International Journal of Remote Sensing
Vol. 30, No. 8, 20 April 2009, 1913–1940

Effects of satellite image spatial aggregation and resolution on estimates of forest land area

M. D. NELSON*†, R. E. MCROBERTS†, G. R. HOLDEN‡ and M. E. BAUER§
†US Department of Agriculture, Forest Service, Northern Research Station, St Paul, MN, USA
‡US Department of Agriculture, Forest Service, Cibola National Forest, Albuquerque, NM, USA
§University of Minnesota, College of Food, Agricultural and Natural Resource Sciences, St Paul, MN, USA

(*Received 22 May 2007; in final form 4 March 2008*)

Satellite imagery is being used increasingly in association with national forest inventories (NFIs) to produce maps and enhance estimates of forest attributes. We simulated several image spatial resolutions within sparsely and heavily forested study areas to assess resolution effects on estimates of forest land area, independent of other sensor characteristics. We spatially aggregated 30 m datasets to coarser spatial resolutions (90, 150, 210, 270, 510 and 990 m) and produced estimates of forest proportion for each spatial resolution using both model- and design-based approaches. Average-based aggregation had no effect on per-image estimates of forest proportion; image variability decreased with increasing spatial resolution and local variability peaked between 210 and 270 m. Majority-based aggregation resulted in overestimation of forest land in a heavily forested landscape and underestimation of forest land in a sparsely forested landscape, with both trends following a natural log distribution. Of the spatial resolutions tested, 30 m was superior for obtaining estimates using model-based approaches. However, standard errors of design-based inventory estimates of forest proportion were smallest when accompanying stratification maps which were aggregated to between 90 and 150 m spatial resolutions and strata thresholds were optimized by study area. These results suggest that spatially aggregating existing 30 m land cover datasets can provide NFIs with gains in precision of their estimates of forest land area, while reducing image storage size and processing times; land cover datasets derived from coarser spatial resolution sensors may provide similar benefits.

1. Introduction

Satellite remote sensing provides image data for mapping and estimating land cover characteristics. While estimates of forest composition and structure may be derived directly from image-based datasets, it is difficult to determine the precision of those estimates. Therefore, national forest inventories (NFIs) of many countries use design-based approaches based on samples of field data for estimating forest attributes. Because natural variability in the resource and budgetary constraints limit the sufficiency of sample sizes for some variables, simple random sampling

*Corresponding author. Email: mdnelson@fs.fed.us

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160802545631

(SRS) techniques may be insufficient for achieving NFI precision standards. Various forms of stratification are used by NFIs to reduce the variance of estimates produced using design-based approaches. Stratified sampling provides for greater precision than post-sampling stratification approaches, but does not serve the needs of NFIs that remeasure permanent plots and produce estimates for a wide variety of forest attributes. Stratified estimation may be used with post-sampling stratification or post-stratification to reduce uncertainty of design-based estimates, even when sampling designs are predetermined. Increased precision occurs when (1) stratum weights are unbiased, (2) within-strata variances are smaller than the overall variance, (3) strata with smaller weights capture plots with more uncertainty, or (4) georectification of imagery and horizontal accuracy of plot location coordinates lead to correct assignment of plots to strata (McRoberts *et al.* 2002, 2006).

Stratifications based on 30 m spatial resolution Earth Observing System Landsat 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) imagery are effective as ancillary data in reducing uncertainty of NFI estimates of forest land area (McRoberts *et al.* 2002, 2006). However, classification of TM or ETM+ imagery across large geographic regions, temporally coincident with NFI reporting intervals, requires substantial cost and poses considerable data management challenges. In addition, Landsat 5 has already exceeded its projected mission lifespan and is expected to cease operations in the near future. The ETM+ imaging sensor onboard Landsat 7 has encountered technical problems with the scan line corrector system, resulting in gaps of missing data on each ETM+ image. As a result, there is a need to assess the potential utility of satellite imagery from alternative satellite-borne imaging sensors.

Coarser spatial resolution satellite imagery, such as that using the 250–500 m Moderate Resolution Imaging Spectroradiometer (MODIS), 1 km MODIS and 1 km Advanced Very High Resolution Radiometer (AVHRR), is receiving increased consideration for mapping and stratifying regions encompassing large geographic extents because such products can be produced more easily and with lower cost than for datasets derived from finer spatial resolutions. In a comparison of stratification maps derived from 30 m TM, 500 m MODIS and 1 km MODIS products, Liknes *et al.* (2004) reported the smallest standard errors for stratified estimates of forest land area with 30 m pixel resolution stratification maps, although their study did not separate the effects of nonspatial sensor characteristics (e.g. spectral, temporal and radiometric resolution; classification or estimation models) from the effects of spatial resolution.

In this study, only spatial resolution is evaluated; the effects of spectral, temporal and radiometric resolution are held constant within study areas of fixed extent. We assessed the effects of simulated spatial resolutions on the potential utility of satellite imagery for providing estimates of forest land area derived from (1) estimates obtained using model-based approaches, that is directly from satellite image-based datasets, and (2) design-based inventory estimates stratified using the same satellite image-derived maps for post-sampling stratification. We simulated effects of spatial resolution by aggregating a 30 m pixel dataset of forest proportion derived from an ETM+ dataset to six coarser spatial resolutions. The simulation involved two case studies that cover relatively small geographic extents of north central USA (figure 1). The results of these tests are intended to provide insight into the effects of spatial aggregation and resolution on estimates of forest land over large geographic extents, such that NFIs may better address the issues of spatial resolution when selecting satellite remotely sensed imagery and derived datasets.

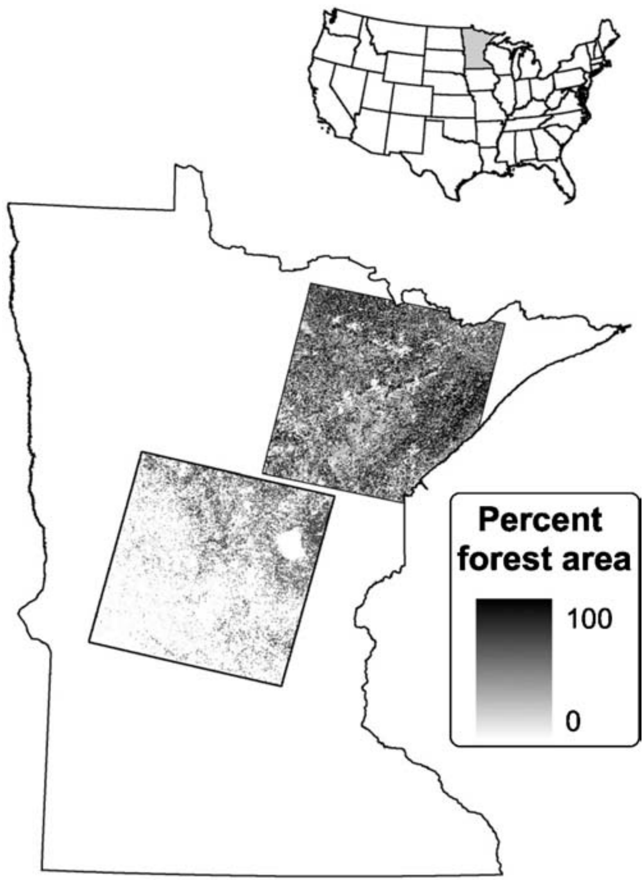


Figure 1. Study areas: WRS-2 Path27/Row27 and Path28/Row28, northeastern and central Minnesota, USA, respectively.

1.1 *Spatial resolution and aggregation*

The following literature review is designed to define terms and summarize effects of spatial resolution and spatial aggregation, representing a variety of land cover conditions, image sensor characteristics and geographic extents, such that results of our study may be viewed within a wider context. Of particular interest is the assessment of the effects of spatial resolution and spatial aggregation on image-derived datasets, such that the subsequent use of these datasets as stratification layers for improving precision of estimates from design-based approaches may be better understood.

Both grain and extent comprise the spatial scale of remote sensing data, where grain is described by Turner *et al.* (1989) as 'the resolution of the data, i.e., the area represented by each data unit' and extent as 'the overall size of the study area'. In the context of cartography, the term 'scale' typically refers to the ratio of the distance between two points on a map or aerial photograph to the distance between the same two corresponding points on the Earth's surface (Goodchild and Quattrochi 1997). Hence, large-scale maps display more spatial detail over smaller geographic extents. Except for references to the literature, we refrain from using the term 'scale'; instead, we use the term 'spatial resolution'. Example definitions of spatial resolution include 'the size of the area on the ground from which the measurements that compose an image are derived' (Hay *et al.* 1997); and 'the dimension in metres of the ground-projected instantaneous field of view (IFOV)' (Jensen 1996). Similarly, Joseph (2000) referred to the instantaneous geometric field of view (IGFOV) as 'the geometric size of the image projected by the detector on the ground through the optical system, called often pixel footprint'. We refer to spatial resolution both as IFOV and for describing spatial aggregations of pixels or derived datasets (e.g. per-pixel proportion forest cover or per-pixel thematic classification).

Descriptors such as 'fine' or 'coarse' resolution have no numerical definition, but may be useful when applied in a relative sense (Forshaw *et al.* 1983). Within the range of spatial resolutions assessed in our study, we refer to relatively smaller pixels as having finer spatial resolution whereas relatively larger pixels have coarser spatial resolution. Strahler *et al.* (1986) proposed an explicit framework for describing remote sensing models. H-resolution models are characterized by resolution cells smaller in size than scene objects, whereas L-resolution models have resolution cells larger than scene objects. H-resolution models of empirical relationships between sensor measurements and cover types provide the basis for land cover classification, while L-resolution models provide the basis for canopy models and for estimating structural parameters (Strahler *et al.* 1986). Strahler *et al.* (1986) suggest using images at varying spatial resolutions for inquiry into the transition between H- and L-resolution.

Tian *et al.* (2002) used the term scaling to describe the process by which biophysical attributes derived from coarse spatial resolution sensor data are established as equal to the arithmetic average of those same attributes derived from finer spatial resolution data. Hay *et al.* (1997) used the term scaling synonymously with resampling and the term upscaling to describe 'resampling techniques designed to transform an image collected at a high spatial resolution (Strahler *et al.* 1986) to a lower spatial resolution representation of the same image'. Atkinson and Curran (1995) used the term 'regularization' as increasing the size of pixels with resulting spatial resolution becoming coarser. Justice *et al.* (1989) described aggregation as degradation of finer spatial resolution satellite image data for simulating and assessing data from coarser spatial resolution sensors. We use the term 'spatial aggregation' synonymously with upscaling, resampling, regularization and degradation.

Ju *et al.* (2005) described two variations of spatial aggregation of attributes from fine spatial resolution pixels to successively coarser spatial resolution pixels. The first, spectral aggregation, entails moving window averaging of spectral measurements from which thematic labels subsequently may be assigned to pixels of coarser spatial resolutions. The second, label aggregation, entails assigning existing thematic class labels from finer spatial resolution pixels to pixels of coarser spatial resolutions. We refer to spectral and label aggregation as average-based and majority-based aggregation, respectively, and broaden the definition of average-based aggregation to include per-pixel estimates of continuous variables (e.g. proportion forest cover) from which moving window averages can be calculated.

Marceau and Hay (1999) discussed the modifiable areal unit problem (MAUP) as a heuristic method for reviewing the contributions of remote sensing to the scale issue. They described MAUP as 'the sensitivity of analytical results to the definition of data collection units', for example aggregation and spatial resolution of remote sensing pixels. They concluded: 'it is of considerable importance that the effects of the MAUP in remote sensing be fully understood to avoid arbitrary and erroneous analytical results'. In the section that follows we review MAUP effects on remote sensing of forested landscapes.

Continuous estimates of forest cover-like attributes can be obtained using optical sensors of any spatial resolution (Lefsky and Cohen 2003). However, areal estimates from coarse spatial resolution satellite image-derived thematic land cover maps (e.g. MODIS, AVHRR) tend to underestimate forest area where forest cover is more fragmented, and overestimate forest area where forest cover is less fragmented (Kuusela and Päivinen 1995, Mayaux and Lambin 1995). Nelson (1989) observed that AVHRR classifications tend to overestimate forest cover except when the cover is <20%. Similarly, Tian *et al.* (2002) reported that AVHRR per-pixel estimates of Leaf Area Index (LAI) show increasing error with increasing heterogeneity of forest cover. Atkinson and Curran (1995) demonstrated that the spatial resolution of satellite remote sensing pixels had an important influence on the precision of estimates of mean percentage vegetation cover.

Estimation of land cover area with coarse spatial resolution imagery is affected by mechanisms associated with mixed pixels, including: (1) nondetection of land cover fragments much smaller than pixels, and (2) multiple detections of the same land cover fragment across adjacent pixels, when that fragment is similar in size to the pixel size (Hlavka and Livingston 1997). In a study of burn scars and wetlands, fragment sizes followed lognormal ($\log_{10}$) distributions, both before and after degrading pixel spatial resolutions, with the exception of an overabundance of tiny fragments (Hlavka and Dungan 2002). By applying the lognormal transformation of fragment sizes, Hlavka and Dungan (2002) were able to estimate the area of small fragments not detected within pixels of coarse spatial resolution, thereby adjusting their estimates of total area.

Nelson *et al.* (2005) reported area-weighted root mean square deviations (RMSDs) from statewide estimates of conterminous United States (CONUS) forest land area derived from forest inventory reports versus estimates derived from a 1 km AVHRR-based Forest Type Groups dataset (2.5% RMSD) (Zhu and Evans 1994), a 500 m MODIS-based Vegetation Continuous Fields (VCF) percentage tree canopy cover dataset (10.7% RMSD), and a 30 m TM-based National Land Cover Dataset (NLCD) (5.8% RMSD). These observations reveal no trend with pixel spatial resolution but several investigators have reported decreasing classification accuracies for forest environments with decreasing pixel spatial resolutions finer than 60–80 m (Sadowski *et al.* 1977, Latty and Hoffer 1981, Markham and Townshend 1981). Reese *et al.* (2002) reported lower accuracies for continuous estimates of forest volume, age and biomass when using 20–25 m (0.4–0.0625 ha) pixels (58–80% RMSE) than when spatially aggregating pixels over 19 ha (17% RMSE) and 100 ha (10% RMSE).

In a forested study area of northern California, local variance, which is a first-order texture measure computed as the statistical mean of the standard deviations of pixel values over a moving window, typically 3×3 pixels, increased for pixel spatial resolutions between 0.75 and 6 m, peaked at 6 m, decreased between 6 and 120 m, and reached an asymptote at about 150 m (Woodcock and Strahler 1987). Because the terms local variance and image variance in fact refer to measures of standard deviation (the positive square root of variance), we substitute the more generic terms local variability and image variability. Although a second peak related to stand size was expected, none was observed for their study area, probably because the forest stands were not well defined and varied so widely in size that there was no typical stand size occurring with sufficient frequency to cause a second peak in image variability. Local variability is directly related to the relationship between the size of the objects in a scene and the spatial resolution of a sensor, with peaks in local variability occurring at spatial resolutions one-half to three-quarters the size of scene objects (Woodcock and Strahler 1987). Ferro (1998) also described spatial resolution effects on texture, edge and land cover separability.

Morisette *et al.* (2003) summarized the relationship between image variability (statistical variance of pixel values over an entire image) and spatial resolution. The relationship led the authors to suggest that ETM+ imagery is suitable for determining the scale of variation in biophysical properties for their study area in Sevilleta National Wildlife Refuge, New Mexico, USA. Semivariogram ranges decreased from about 800 to 500 m as image pixel spatial resolution was degraded from 4 to 512 m, and indicated that spatial support of imagery has some effect on the observed scale of spatial variation. Cohen *et al.* (2003) stated that spatial complexity of land cover types is captured by 30 m ETM+ data but not by 500 m MODIS data. Despite the authors' expectation that lower image variability of coarser spatial resolution MODIS data would lead to lower spectral sensitivity of LAI, they observed improved relationships between spectral data and LAI when moving from 30 m ETM+ data to 500 m MODIS data, and realized further improvements with 1 km MODIS data.

Moran *et al.* (1997) observed negligible error in aggregation of remotely sensed surface temperature and reflectance over a wide range of conditions. Thomas *et al.* (1996) reported that spatial aggregation of pixel spectral values improved the fit of mixture models and resulting per-pixel area estimates, for example forest proportion. Hay *et al.* (1997) discussed spatial scale and the methods for estimating image properties at new scales. Townshend and Justice (1988) resampled red and near-infrared image bands using a technique designed to simulate these same bands if obtained from coarser spatial resolution sensors ranging from 125 to 4000 m. The authors tested the effectiveness of coarse spatial resolution data for detecting change in landscape types by deriving the Normalized Difference Vegetation Index (NDVI) from the resampled imagery and calculating differences in the NDVI between various spatial resolutions. They reported better representation of land cover conversions with finer spatial resolutions, but analyses using spatial resolutions as coarse as 500 m or 1000 m are capable of detecting conversions of land cover occurring over relatively short periods. Moody and Woodcock (1994) aggregated a 30 m TM image-derived map of land cover to coarser resolutions ranging from 90 to 6000 m. The authors hypothesized that estimates of the proportion of land cover class are affected by spatial resolution, area of land cover within each class, and spatial arrangement of the classes. Specifically, the proportions of smaller, more fragmented classes decrease with coarser spatial resolution while the proportions of larger, more clumped classes increase with coarser spatial resolution.

Ju *et al.* (2005) assessed the effects of spatial aggregation approaches on estimates of simulated land cover fraction (conifer, hardwood, brush and grass) from TM 30 m imagery of a state forest near Boston, MA, USA. Following aggregation of TM spectral values to coarser spatial resolution imagery, the authors used a statistical finite mixture model to classify land cover classes and the resulting land cover fractions. Mean absolute errors (MAEs) of these land cover fractions were compared with MAEs from majority-based aggregations of the 30 m land cover classification. The authors determined that the finite mixture model approach resulted in consistently lower MAEs than the majority-based aggregations, with the lowest MAEs at spatial resolutions between 60 and 150 m, and that the 240 m resolution MAE was similar to that of the 30 m MAE.

Forest patch size, number of patches and other landscape metrics are also affected by spatial resolution. Saura (2004) assessed the effects of spatial resolution on six fragmentation indices used by the third Spanish NFI. They determined that power scaling laws were effective in predicting the effects of spatial resolution on estimates of number of patches, mean patch size and edge length, but not for estimates of largest patch index, landscape division and patch cohesion. Wu (2004) spatially aggregated 30 m land cover pixels to assess effects of grain on landscape metrics. Number of patches, patch density, total edge, edge density and landscape shape index exhibited consistent, robust scaling relationships that were characterized by a power law, while other metrics exhibited less robust relationships or unpredictable scaling behaviour. However, García-Gigorro and Saura (2005) suggest that the previously reported accuracy and utility of power scaling laws may have been overestimated. Incorporating sensor point spread function into their aggregation procedure better simulated actual sensor radiation at different resolutions and improved the comparability of forest fragmentation indices.

As forest patch size changes, the cell size differentiating H- from L-resolution also changes. Within the confines of H-resolution, Atkinson and Curran (1995) reported increasing precision in estimates of percentage vegetation cover as cell (e.g. pixel) sizes increased. Pax-Lenny and Woodcock (1997) reported that, for coarser spatial resolutions, small patch sizes of reclaimed agricultural fields led to poorer map accuracies whereas large patch sizes of old agricultural lands led to better map accuracies and more accurate area estimates. Colombo *et al.* (2004) rescaled a 30 m Brazilian tropical forest/nonforest classification to coarser spatial resolutions of 160 and 1000 m, corresponding to image resolutions of the Russian remote sensing satellite Resurs MSU sensor (160 m) and the European Space Agency's Along Track Scanning Radiometer sensor (1000 m). From their variographic analyses, the authors observed that rescaling to coarser spatial resolutions retained the general structure and spatial distribution of forest cover in areas of low fragmentation but not in areas with higher fragmentation. Colombo *et al.* (2004) observed that semivariogram range parameters increased at coarser spatial resolutions. However, in areas of medium or low fragmentation, range values decreased for 160 m spatial resolution data and increased for 1000 m spatial resolution data. Millington *et al.* (2003) assessed the effects of spatial resolution on landscape metrics by spatially degrading a 30 m TM image-derived classification of Bolivian forest/nonforest to 80, 125, 500 and 750 m spatial resolutions. The authors observed varying responses to decreases in spatial resolution: increases in total core area, mean patch size and total edge length; decreases in core area density, patch density and edge density; no change in number of patches, several shape indices and several diversity indices; and complex patterns of change in the number of core areas. Spatial degradation resulted in patch merging and elimination and produced a landscape classification with fewer large patches and increased distances between patches (Millington *et al.* 2003).

Marceau *et al.* (1994*a*) stated 'There is no unique spatial resolution appropriate for the detection and discrimination of all geographical entities composing a complex natural scene such as a forested environment'. However, the authors proposed that, logically, there is an optimal spatial resolution for each type of entity that corresponds with spatial and spectral characteristics intrinsic to that entity of interest. Hay *et al.* (1997) suggested a multiscale approach for detection and analysis of image-objects with different optimal image spatial resolutions. Marceau *et al.* (1994*a*) analysed local variability and spectral separability versus spatial resolutions (5, 10, 20 and 30 m) to derive optimal spatial resolutions for classifications of nonforested areas (5 m optimal spatial resolution) and natural forest (20 m optimal spatial resolution). Marceau *et al.* (1994*b*) determined optimal spatial resolutions for discrimination of each of 14 forest classes consisting of combinations of four conifer tree species with varying levels of stand density, tree height and tree groupings (e.g. plantations, natural stands). Among the resolutions the authors tested (0.5–29.5 m), optimal spatial resolution ranged from 2.5 to 21.5 m and was defined as the pixel size resulting in the minimal within-class variance for the largest number of spectral bands. Although not investigated in their study, Marceau *et al.* (1994*b*) suggested that a methodology similar to theirs could be applied at a variety of resolutions, for example using satellite imagery to determine optimal spatial resolutions for different types of forest or for an entire forest corresponding to a particular ecosystem.

Jarnagin *et al.* (2004) compared 'per-pixel' and 'whole-area' measures of accuracy of 30 m TM-based estimates of subpixel proportion impervious surface. The authors obtained whole-area means and coefficients of variation for various samples of pixels, for all pixels within the study area, and for spatial aggregations of 3×3, 5×5, 9×9, 15×15 and 25×25 pixel blocks. The results from Jarnagin *et al.* (2004) indicated that accuracies of pixel-based estimates were consistently less than for whole-area estimates, and accuracies increased as pixel block sizes increased except for 5×5 pixel blocks (150 m spatial resolution), which had lower accuracies than both 3×3 and 9×9 pixel blocks.

To summarize the preceding literature review, spatial resolution refers to the size of the area on the ground from which measurements that compose an image are derived. In this study, we extend this term to refer to image-derived datasets as well. Fine spatial resolution refers to smaller pixels, while coarse spatial resolution refers to larger pixels; both terms are relative. The size and variability of objects being imaged, among other factors, affect the suitability of an image's spatial resolution. Examples of several metrics are reported for quantifying the size and variability of image objects and the relationship of these metrics to spatial resolution. Spatial aggregation refers to processes that change the size of image pixels or derived datasets from relatively finer to relatively coarser spatial resolution. Various aggregation approaches are reported that enable us to test for the effects of resulting spatial resolutions. Knowing these effects allows us to better understand the potential utility of stratification layers derived from these image-based datasets. The objective of our study was to integrate and adapt several of the approaches discussed into two spatial aggregation approaches, and to test for the effects of resulting spatial resolutions on estimates of forest land proportion.

1.2 *Forest Inventory and Analysis (FIA): the NFI of the USA*

The FIA Program (http://fia.fs.fed.us/) of the United States Department of Agriculture (USDA) Forest Service conducts detailed surveys of forests across all land ownerships, serving as the USA NFI. FIA's four regional units report statewide estimates of forest land area. FIA defines forest land as commercial timberland, some pastured land with trees, forest plantations, unproductive forested land, and reserved, noncommercial forested land. The FIA definition of forest land also requires the following: a minimum area of 0.405 ha (1 acre), a minimum continuous canopy width of 36.58 m (120 ft), and 10% minimum stocking level (USDA Forest Service 2003). The minimum stocking requirement is modified for several western woodland types where stocking cannot be determined and where at least 5% tree canopy cover occurs (or has occurred in the past). The National Resources Inventory of the USDA Natural Resources Conservation Service (NRCS) interprets 10% stocking as canopy cover of at least 25% when viewed from a vertical direction (Lessard *et al.* 2003). This interpretation is corroborated by Nelson *et al.* (2005), who reported that a 25% tree canopy cover minimum threshold resulted in map-based estimates similar to FIA inventory estimates of USA and CONUS forest land area when using 500 m MODIS VCF data as the map source. FIA estimates of forest land area are obtained by multiplying total land area inventoried by the mean proportion of forest land, estimated from FIA plot observations. National FIA precision standards limit the allowable error associated with estimates of forest land area to ± 3–5% per 404 700 ha (million acres) of timberland and $\pm 10\%$ per 404 700 ha (million acres) of all other forest land.

2. Data and methods

2.1 *Study area*

The study was conducted in two distinct landscapes in Minnesota, USA, that differ substantially in forest distribution and composition (figure 1). The two landscapes correspond with TM images and are labelled by the Worldwide Reference System (WRS-2) Path and Row numbers of their corresponding images. The first landscape, designated P27R27, encompasses the Minnesota portion of Path 27, Row 27 (excluding Lake Superior), encompasses approximately 3.1 million ha in north-eastern Minnesota, and is dominated by forest cover (75% forested) and open water. P27R27 forest land is primarily Aspen–Birch and Spruce–Fir associations. The second landscape, designated P28R28, includes the entire extent of Path 28, Row 28, encompasses approximately 3.3 million ha in central Minnesota, is dominated by agricultural land use, and includes patches of forest land comprising 25% of the area. Forest land in P28R28 is dominated by Aspen–Birch, Oak–Hickory and Maple–Beech–Birch associations.

2.2 *Estimates obtained using model-based approaches*

Per-pixel predictions of P27R27 and P28R28 proportion forest cover were obtained from studies by McRoberts (2002, 2006), who used a logistic regression estimation approach and 30 m ETM+ satellite imagery acquired during spring, summer and leaf-off seasons between 1999 and 2001. Although mean predicted pixel proportion forest values for both areas are within 2–5% of FIA design-based inventory estimates, in this study it is the subsequent effect of spatial aggregation and resulting spatial resolutions and not the accuracy of the original datasets that are of interest. For each area, two 30 m thematic maps were produced from the continuous per-pixel predictions by assigning each pixel to either a ‘forest’ or a ‘nonforest’ class, hereafter referred to as F/NF maps. For each thematic map, pixel assignments were based on one of two thresholds of proportion forest. For the first approach, termed ‘standard’ threshold, pixels with forest proportion $\geqslant 0.25$ were assigned to the forest class, while pixels having forest proportion <0.25 were assigned to the nonforest class. This standard threshold of 0.25 proportion forest represents 25% tree canopy cover for the 30 m forest proportion dataset used in this study, a canopy cover value that has been suggested as a minimum threshold for defining forest land (Lessard *et al.* 2003). In the second approach, 'optimal' thresholds were determined for each of the two study areas based on the minimum per-pixel proportion forest threshold that resulted in map-based mean predicted proportion forest most similar to FIA design-based estimates of proportion forest. This approach is similar to the VCF percentage tree canopy cover threshold approach described by Nelson *et al.* (2004, 2005). At 30 m spatial resolution, these optimal forest proportion thresholds were 0.72 and 0.45 for P27R27 and P28R28, respectively. Within each study area, optimal thresholds were determined for each spatial resolution and one standard threshold was applied to all spatial resolutions. The utility in identifying an optimal threshold is for subsequently producing a stratification layer whose stratum weights are unbiased, resulting in a potential increase in the precision of stratified estimates.

We tested for effects of spatial aggregation and resulting spatial resolutions on estimates of proportion forest land using spatial averaging and block majority filtering (figure 2). Both forms of spatial aggregations were performed using block functions in ArcInfo™ Grid software (Environmental Systems Research Institute, ESRI®). The block functions work by clumping adjacent pixels into equal-size, square blocks (jumping windows), summarizing those pixel values within a block, and assigning the summary value to every pixel in that block. In a subsequent processing step, each block was converted to a single pixel of a spatial resolution equivalent in spatial extent to the input block.

Spatial aggregation was performed on both the 30 m continuous forest proportion datasets and the 30 m F/NF datasets. For this study, pixel values from the 30 m data were summarized within block sizes of 3×3, 5×5, 7×7, 9×9, 17×17 and 33×33 pixels, with output pixel spatial resolutions of 90, 150, 210, 270, 510 and 990 m, respectively (table 1). Figure 3 shows examples of spatial aggregation for a subset of coarser spatial resolutions.

2.2.1 Average-based aggregation. For the continuous forest proportion dataset, pixel prediction values were spatially aggregated based on the arithmetic average (AVG) of forest proportion within each jumping window, similar to the approaches used by Morisette *et al.* (2003) and Su *et al.* (1999). With the AVG approach, the average forest proportion within a block window of pixels (e.g. 3×3, 5×5) was assigned to a single output pixel of a spatial resolution comparable in area to the block window size (e.g. 90 m, 150 m).

Three variance measures of forest proportion were calculated for each image and each spatial resolution following average-based aggregation: (1) image variability, (2) local variability, and (3) comparative variability. Image variability is defined as the standard deviation over all pixel values within an image, while local variability is defined as the mean of standard deviations from all 3×3 pixel moving windows over an image; equivalent to local variance from Woodcock and Strahler (1987). We defined comparative variability as the mean of standard deviations from preaggregation pixel moving windows having window extents corresponding to the spatial extents of local variability windows. For example, for pixels of 150 m spatial resolution, a local variability window of 3×3 pixels encompasses a spatial extent of 450 by 450 m. A corresponding comparative variability window also encompasses 450 by 450 m, but includes 15×15 pixels of 30 m spatial resolution. In this study, local variability and comparative variability are identical at 30 m spatial resolution.

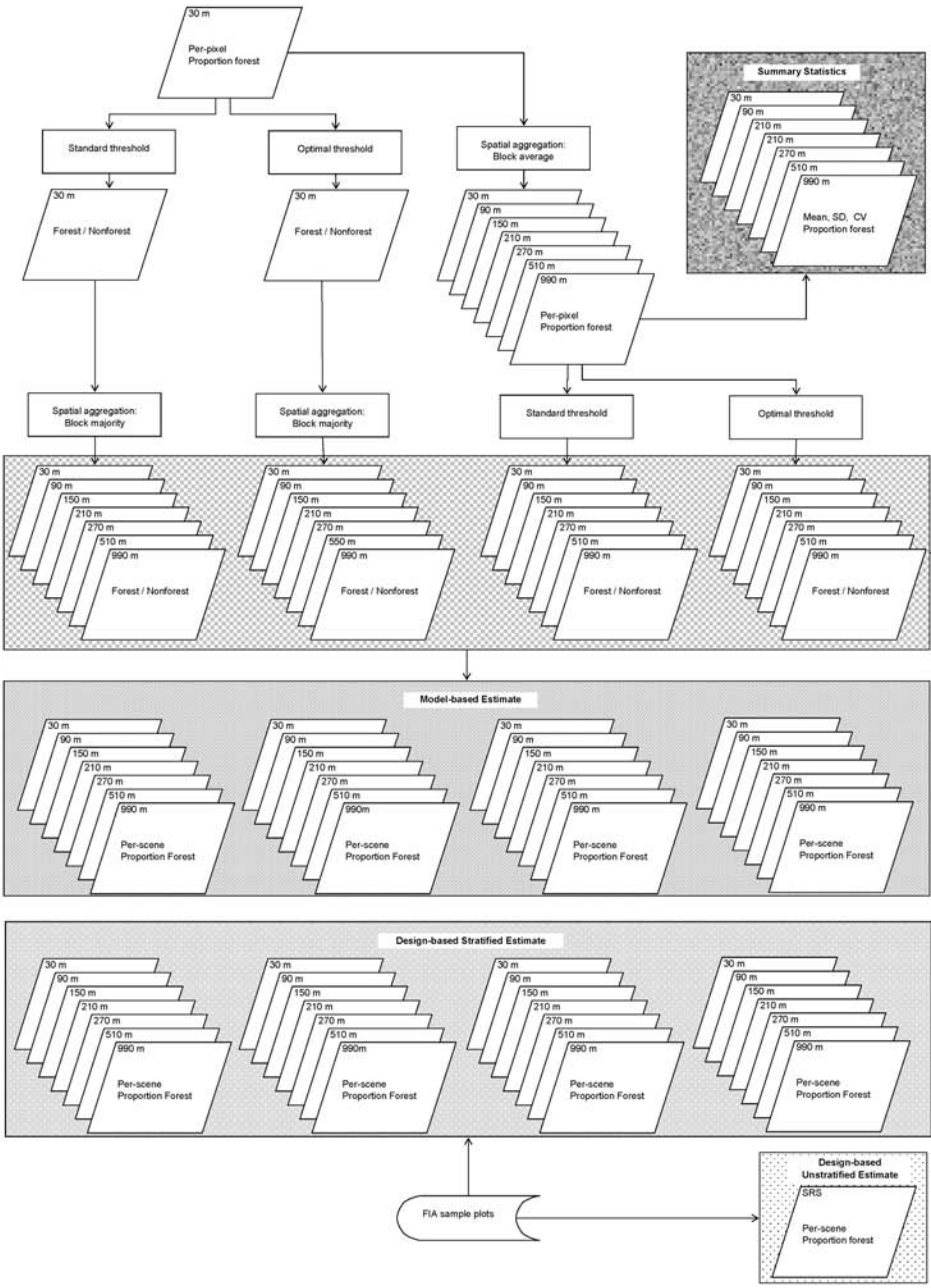


Figure 2. Flowchart of spatial aggregation processing steps.

Semivariograms were used to compare the spatial variability structure of forest proportion aggregated to multiple spatial resolutions (Colombo *et al.* 2004). Also called a variogram, a semivariogram is computed as half the average squared difference between paired data values (Isaaks and Srivastava 1989). An unbiased estimator of semivariance γ at lag distance h, $\gamma(h)$ is calculated as:

Table 1. Image mean, standard deviation (SD, image variability), coefficient of variation (CV), and optimal threshold of proportion forest land following spatial averaging to coarser spatial resolutions, P27R27 and P28R28 study areas, Minnesota, USA.

Window size (pixels)	Spatial resolution (m)	P27R27				P28R28			
		Mean	SD	CV	Threshold	Mean	SD	CV	Threshold
1 × 1	30	0.7496	0.3182	0.4245	0.72	0.2168	0.319907	1.4755	0.45
3 × 3	90	0.7496	0.2885	0.3848	0.69	0.2167	0.295697	1.3646	0.42
5 × 5	150	0.7496	0.2718	0.3625	0.68	0.2167	0.281936	1.3010	0.41
7 × 7	210	0.7496	0.2595	0.3461	0.67	0.2167	0.271988	1.2550	0.40
9 × 9	270	0.7496	0.2499	0.3333	0.67	0.2167	0.264333	1.2196	0.39
17 × 17	510	0.7496	0.2231	0.2976	0.68	0.2168	0.243745	1.1243	0.38
33 × 33	990	0.7496	0.1949	0.2600	0.68	0.2168	0.223671	1.0318	0.37

$$\gamma(h)=\frac{1}{2n(h)}\sum_{i=1}^{n(h)}\left(z_i-z_{i+h}\right)^2 \tag{1}$$

where $n(h)$ is the number of pairs of pixel observations (z_i, z_{i+h}) separated by Euclidean distance within a neighbourhood of h, measured in geographic space (m).

Semivariogram parameters are obtained by fitting a model to the experimental semivariogram. Parameters include the range, which defines the geographic distance at which numerical observations stop correlating with each other; the sill, which is the semivariance at which the range is achieved; and the nugget effect, which represents the random component of the numerical observations and is identified as

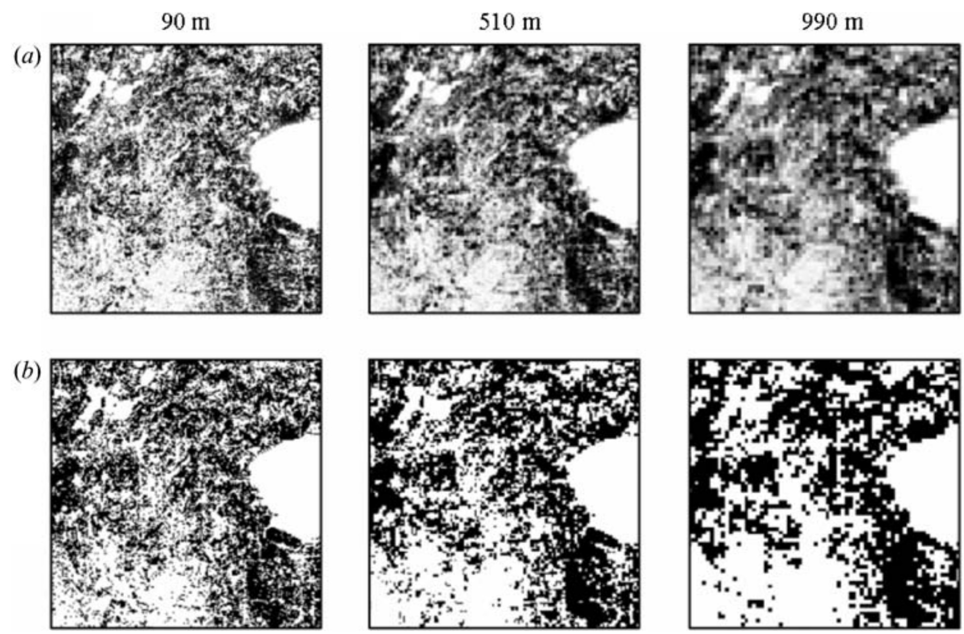


Figure 3. (*a*) Average-based and (*b*) majority-based spatial aggregation of 30 m pixels into 90, 510 and 990 m spatial resolutions, northeastern P28R28, Minnesota, USA. Geographic extent is 69 × 69 km.

the semivariance at distance 0. A variation in the definition of range is the lag (geographic distance) at which 95% of the sill is reached (Deutsch and Journel 1997). In this study, semivariograms were rescaled to allow for direct comparison of imagery aggregated to coarser spatial resolutions. Variographic analyses were performed using the gstat computer program (Pebesma 2004) as an open source add-on package to the 'R' implementation of S language (R Development Core Team 2004).

Each ETM+ image contains millions of pixels – a number of observations that exceeds the current capability of desktop computing platforms for estimating semivariance. Therefore, random samples of pixels were selected from images prior to conducting detailed variographic analysis. Pixel sample intensity varied with spatial resolution and ranged from 0.001 proportion of image pixels at 30 m spatial resolution to 1.0 proportion of image pixels (all pixels) at 990 m spatial resolution. Resulting samples comprised approximately 30 000–40 000 pixel observations per image, which is a sample size within the software's capability for calculating semivariance. In addition to performing omnidirectional variographic analyses, directional semivariograms were produced at azimuths of 0° (N–S), 45° (NE–SW), 90° (E–W) and 135° (SE–NW) to assess directional effects on spatial variability of forest proportion. Semivariograms were compared between study areas and spatial resolutions.

2.2.2 Majority-based aggregation. We created two thematic F/NF classifications from continuous per-pixel estimates at each resulting spatial resolution, one using the standard threshold and the other using the optimal threshold specific to each area. The 30 m F/NF map pixels were spatially aggregated to coarser resolutions based on the clear majority within a block (MAJ). For example, if 41 or more 30 m pixels within a 9×9 pixel aggregation (81 pixels) were classed as forest, the resulting 270 m resolution block was labelled forest. If fewer than 41 pixels were of forest class, the resulting block was labelled nonforest.

In contrast with methods that allow for blocks with even numbers of pixels (e.g. 2×2, 4×4) (Turner *et al.* 1989), we produced only square blocks with odd numbers of pixels (e.g. 3×3, 5×5) so that resulting blocks have a clear majority. For each spatial resolution tested, the aggregations of continuous forest proportion maps were converted to F/NF maps using both threshold approaches described earlier.

The proportion of forest land was estimated as the ratio of forested pixels to total image pixels for each study area. Estimates were compared within each study area among 28 F/NF maps that were derived from two thresholds, two spatial aggregation approaches, and seven spatial resolutions (figure 2). To assess resampling effects, mean, standard deviation (SD, image variability), and coefficient of variation (CV) of pixel proportion forest land were calculated for each study area, for each spatial resolution following spatial averaging.

2.3 *Estimates obtained using design-based approaches*

Stratified estimates of mean proportion forest land, $\bar{P}$, and estimated variance, $\hat{\mathrm{Var}}(\bar{P})$, are calculated using formulae from Cochran (1977):

$$\bar{P} = \sum_{h=1}^{L} w_h \bar{P}_h \tag{2}$$

and

$$\hat{\mathrm{Var}}(\bar{P}) = \sum_{h=1}^{L} w_h^2 \frac{\hat{\sigma}_h^2}{n_h} \tag{3}$$

where $h=1, \ldots, L$ denotes stratum; w_h is the hth stratum weight, $\bar{P}_h$ is the mean forest land proportion for plots assigned to the hth stratum (F/NF); n_h is the number of plots assigned to the hth stratum; and $\hat{\sigma}_h^2$ is the within-stratum variance for the hth stratum. Variance estimates obtained using equation (3) ignore the slight effects due to finite population correction factors and to variable rather than fixed numbers of plots per stratum.

FIA plot data within P27R27 ($n=1022$) and P28R28 ($n=1071$) were collected between 1999 and 2002. Land area within each study area was stratified using each of the 28 previously defined F/NF image datasets and resulted in 28 stratified estimates of proportion forest land for each study area. Each FIA plot was assigned to one of two strata, 'forest' or 'nonforest', based on the pixel associated with plot centre location. Strata weights were determined as the proportion of pixels within each stratum. To test for effects of assigning aggregated subplot forest proportion values to the single pixel associated with only the central subplot, we compared stratified estimates using only the forest proportion value associated with central subplots to stratified estimates based on forest proportion averaged across all four subplots.

The effect of spatial resolution on stratified estimates was determined by comparing estimates using each of the coarser spatial resolution stratification datasets to stratified estimates obtained using the 30 m stratification datasets. Comparisons also were made with unstratified design-based estimates of mean proportion forest land, that is those calculated under the assumption of SRS.

3. Results

3.1 *Estimates obtained using model-based approaches*

Estimates of 30 m pixel mean proportion forest for P27R27 (0.75) and P28R28 (0.22) were similar to corresponding estimates from design-based approaches; these estimates remained almost constant following spatial averaging for all spatial resolutions, but SD and CV decreased as simulated pixel size increased from 30 to 990 m (table 1). Figure 4 shows examples of P28R28 forest proportion histograms for four spatial resolutions. Visual inspection of the histograms reveals that the frequency distribution of the original 30 m dataset is skewed towards the tails; pixels with forest proportions between 0.25 and 0.75 comprise a minor fraction of the frequency distribution. As is typical for smoothed data, the histograms indicated less variance with coarser spatial resolutions because averaged pixel values were closer to the mean (Bian 1997). Similarly, optimal thresholds decreased in both P27R27 (from 0.72 to 0.68) and P28R28 (from 0.45 to 0.37) following spatial aggregation to coarser resolutions (table 1). After assigning 30 m pixels to F/NF classes, estimates of mean forest proportion using optimal thresholds were about 15% lower than estimates using the standard threshold, for both study areas.

Image variability decreased slightly nonlinearly with increasing spatial resolution, from 0.32 at 30 m to 0.20 at 990 m resolution in P27R27, and from 0.32 at 30 m to 0.22 at 990 m resolution in P28R28. Local variability in P27R27 increased from 0.09 at 30 m to a peak of 0.12 at 270 m, and then decreased to 0.10 at 990 m spatial

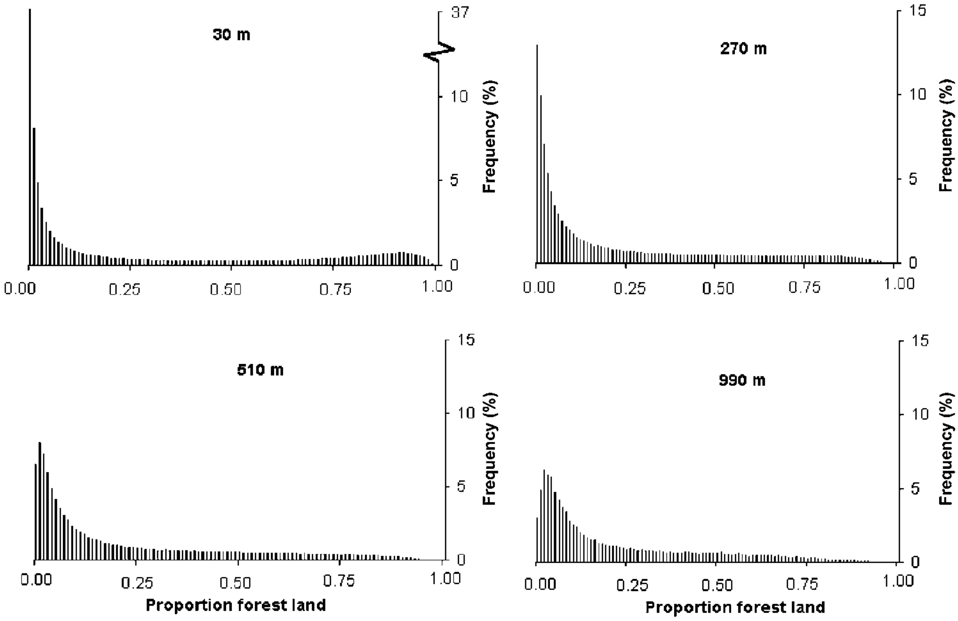


Figure 4. Frequency distribution of forest land proportion at four pixel spatial resolutions (30, 270, 510 and 990 m), P28R28, Minnesota, USA. The upper portion of the y-axis is compressed for the 30 m pixel spatial resolution.

resolution. In P28R28, local variability increased from 0.08 at 30 m to a peak of 0.10 at 270 m, and then decreased to 0.08 at 990 m spatial resolution. Comparative variability increased nonlinearly with increasing spatial resolution, from 0.09 at 30 m to 0.27 at 990 m, and from 0.8 at 30 m to 0.22 at 990 m, for P27R27 and P28R28, respectively (figure 5).

P27R27 directional semivariograms of 30 m forest proportion revealed relatively constant to slightly increasing semivariance for all four azimuths tested (figure 6). P28R28 azimuths of 90° and 135° portray nondescript semivariance, similar to those observed for P27R27. However, P28R28 azimuths of 0° and 45° reveal directional effects of increasing semivariance from 0 to 30 km, relatively constant semivariance from 30 to 60 km, and increasing semivariance from 60 to 100 km (figure 6). For both P27R27 and P28R28, omnidirectional semivariograms show decreasing sills and increasing ranges with increasing spatial resolution following spatial averaging of 30 m pixels (figures 7 and 8).

3.1.1 P27R27. When the standard threshold was used and as pixel size increased from 30 to 990 m, estimates of P27R27 mean proportion forest increased from 0.86 to 0.96, and from 0.86 to 0.93 for average-based and majority-based F/NF maps, respectively. When the optimal threshold was used, mean proportion forest decreased from 0.73 to 0.69 for average-based, but mean proportion forest increased from 0.73 to 0.83 for majority-based F/NF maps as pixel size increased from 30 to 990 m (figure 9). Deviations from 30 m estimates followed a natural log ($\log_e$) distribution for both aggregation and both threshold approaches (figure 9).

3.1.2 P28R28. When the standard threshold was used, P28R28 estimates of mean forest proportion increased with increasing pixel size from 0.28 to 0.33 for

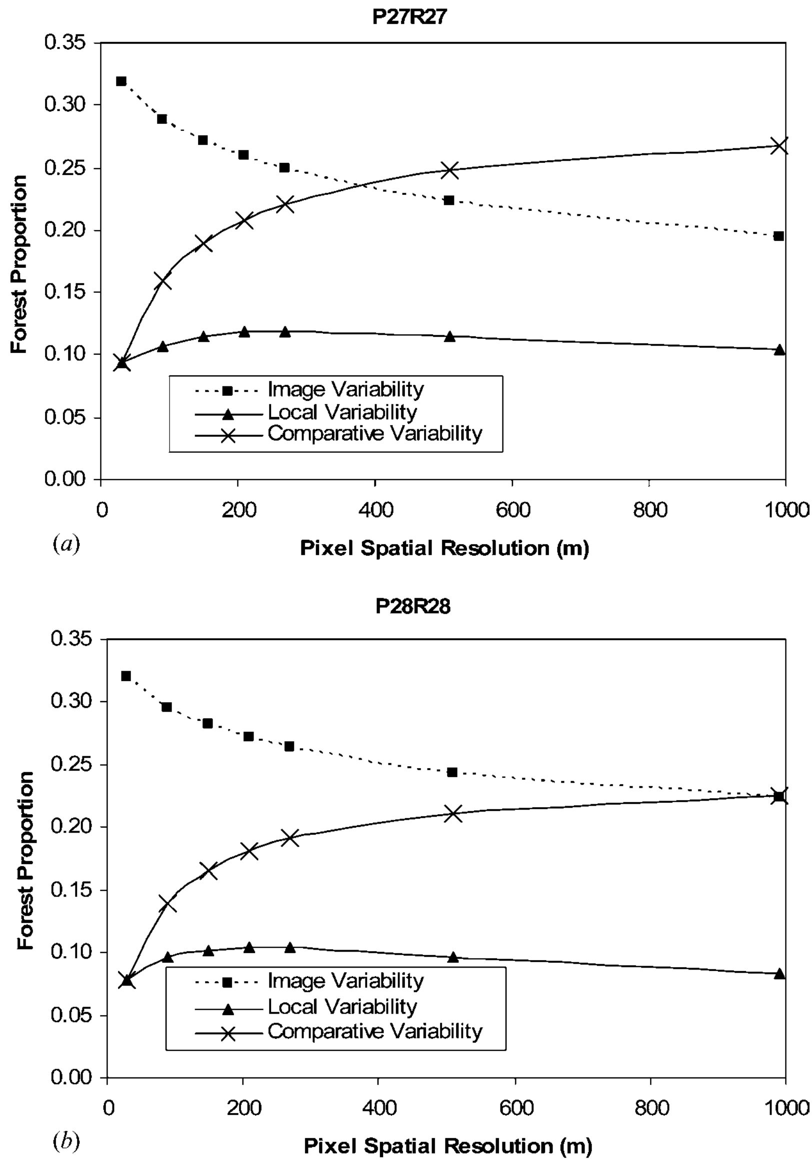
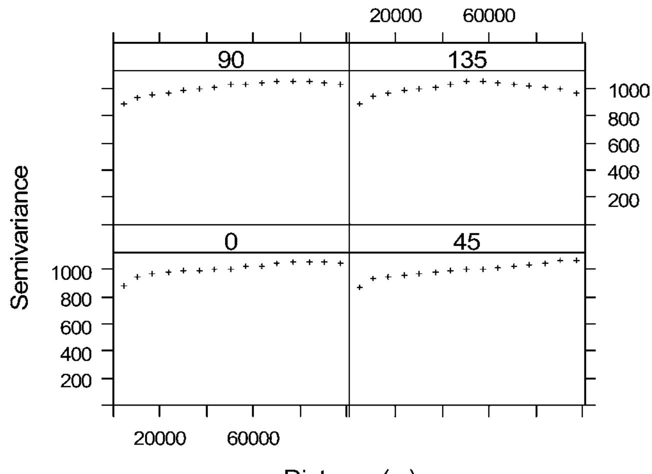


Figure 5. Image variability, local variability, and comparative variability as a function of spatial resolution: (*a*) P27R27 and (*b*) P28R28, Minnesota, USA.

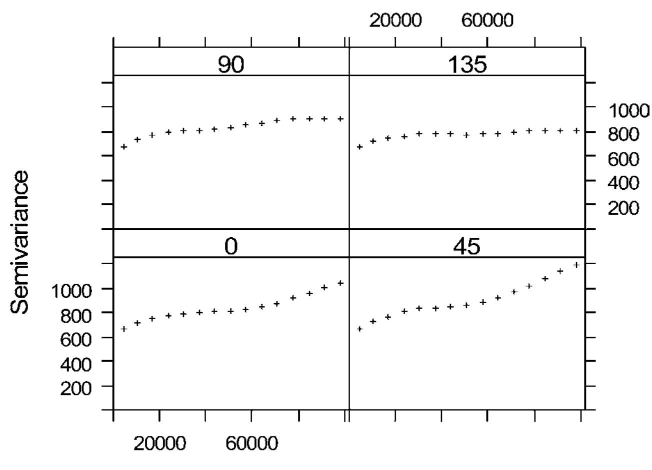
average-based F/NF maps but decreased from 0.28 to 0.24 for majority-based F/NF maps. Mean forest proportion decreased from 0.23 to 0.18 for both average- and majority-based F/NF maps when the optimal threshold was used (figure 9). For both aggregation and both threshold approaches, deviations from 30 m estimates followed a natural log ($\log_e$) distribution (figure 9).

3.2 *Estimates obtained using design-based approaches*

3.2.1 P27R27. Under the assumption of SRS, the design-based estimate for P27R27 proportion forest land was 0.74 with a standard error of 0.01. Design-based



(*a*)



(*b*)

Figure 6. Forest proportion 30 m pixel directional semivariograms for azimuths of 0° (N–S), 45° (NE–SW), 90° (E–W) and 135° (SE–NW): (*a*) P27R27 and (*b*) P28R28, Minnesota, USA.

stratified estimates of proportion forest land ranged from 0.74 to 0.76 across all spatial resolutions for average- and majority-based stratification maps and for standard and optimal thresholds (figure 9). When the standard threshold was used, standard errors of stratified estimates ranged slightly, from 0.009 to 0.012 for both average- and majority-based stratifications. Standard errors ranged from 0.009 to 0.011 for both average- and majority-based stratifications with the optimal threshold (figure 10).

3.2.2 P28R28. Under the assumption of SRS, the design-based estimate of proportion forest land for P28R28 was 0.23 with a standard error of 0.012. Stratified estimates of proportion forest land ranged from 0.22 to 0.25 across all spatial resolutions for both average- and majority-based stratification maps using both standard and optimal thresholds (figure 9). For both average-based and

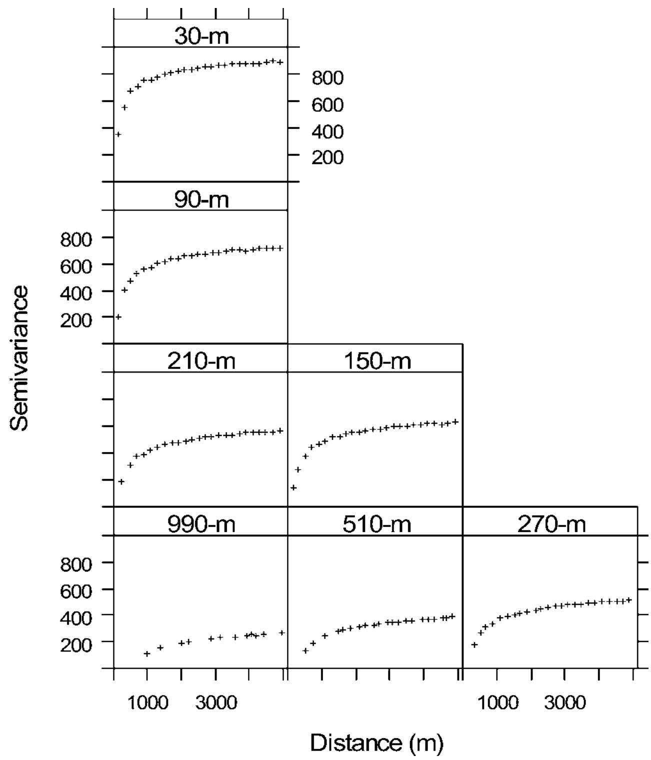


Figure 7. Forest proportion omnidirectional semivariograms for seven spatial resolutions, P27R27, Minnesota, USA.

majority-based stratification maps, standard errors ranged slightly from 0.008 to 0.010 when the standard threshold was used and from 0.007 to 0.010 with the optimal threshold (figure 10).

3.2.3 Subplot aggregation. Compared with estimates of mean forest proportion that aggregate all four FIA subplots, estimates using only central FIA subplots were similar for both P27R27 (0.74 for four subplots, 0.75 for central subplots) and P28R28 (0.23 for four subplots, 0.23 for central subplots). Standard errors of SRS estimates were also similar for both P27R27 (0.013 for four subplots, 0.014 for central subplots) and P28R28 (0.012 for four subplots, 0.013 for central subplots).

4. Discussion

4.1 *Estimates obtained using model-based approaches*

For both study areas, spatial averaging had almost no effect on pixel mean forest proportion, but image variability decreased as spatial resolution became more coarse. This result is similar to the findings of Henderson-Sellers and Pitman (1992), Marceau *et al.* (1994*a*,*b*) and Su *et al.* (1999), who reported near constant mean spectral values and decreasing variances following spatial averaging to coarser pixel resolutions within forested pixels. Trends in image variability, local variability and comparative variability of forest proportion were similar between P27R27 and

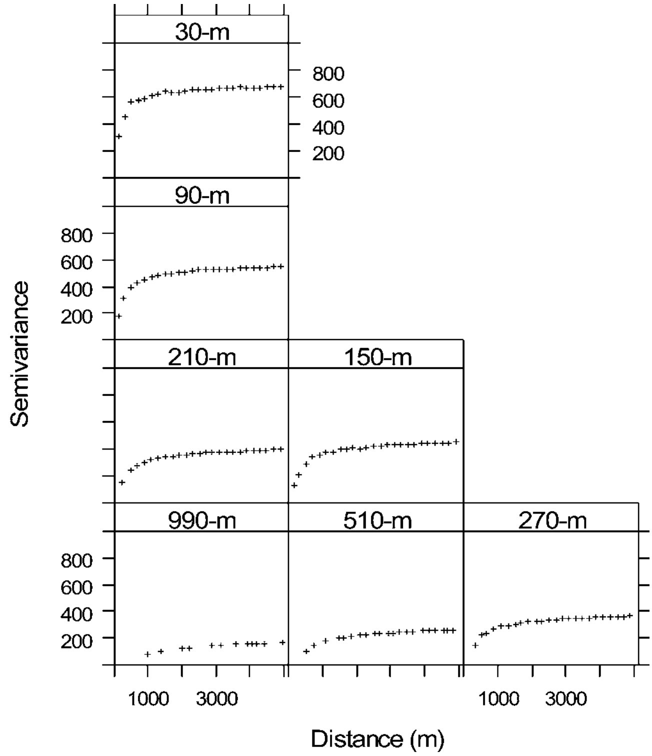


Figure 8. Forest proportion omnidirectional semivariograms for seven spatial resolutions, P28R28, Minnesota, USA.

P28R28. There were slight peaks in local variability of forest proportion at 270 m spatial resolution in both study areas, suggesting that typical image objects (e.g. forest patches) are between 360 by 360 m and 540 by 540 m (Woodcock and Strahler 1987). However, these peaks are much less distinct than the peaks in local variability of pixel spectral reflectance within the forested study areas of Woodcock and Strahler (1987), suggesting that patch sizes are more variable within our image extents or that forest proportion varies less than spectral reflectance, or both.

By contrast, comparative variability continued to increase with coarser spatial resolution from 30 to 990 m for both study areas, although an asymptote at a slightly coarser resolution appears likely. Thus, variability of 30 m pixel forest proportion continued to increase for moving windows of 1000 ha or larger. The results of directional semivariance analyses (figure 6) confirm the visually evident pattern of relatively dense forest across P27R27, and decreasing forest cover from northeast to southwest P28R28 (figure 1). Omnidirectional semivariograms reveal expected patterns of increasing range and decreasing sill with coarser spatial resolutions. These results for continuous forest proportion are not comparable directly with those of Colombo *et al.* (2004), who estimated semivariance for forest/nonforest classifications.

Optimal thresholds were defined such that resulting estimates from model-based approaches for each spatial resolution would correspond to estimates from design-based approaches. At the resolution of the original pixel dataset (30 m), estimates using optimal thresholds were about 15% lower than estimates using the standard

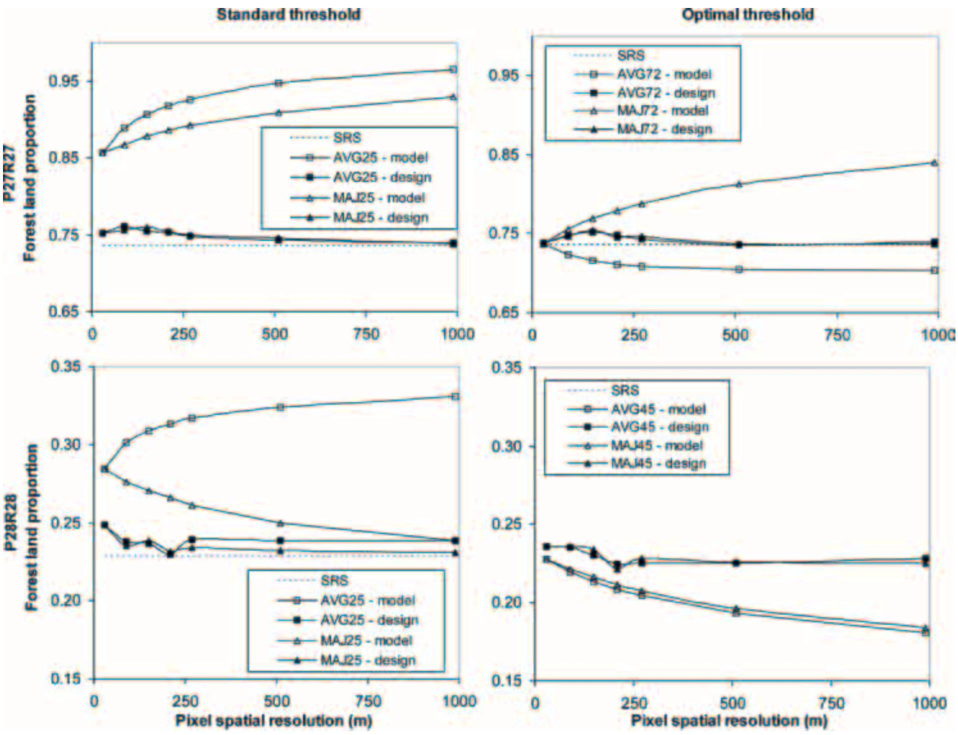


Figure 9. Estimates of forest land proportion calculated under (1) the assumption of simple random sampling (SRS), and (2) model-based and design-based approaches using average-based (AVG) and majority-based (MAJ) aggregations, with standard (0.25) and optimal (0.75 or 0.45) thresholds for P27R27 and P28R28, Minnesota, USA.

threshold, for both study areas (figure 9). This difference calls into question the validity of 0.25 proportion forest as a 'standard' threshold, which was based on a definition that at least 25% tree canopy cover is required to be forest land.

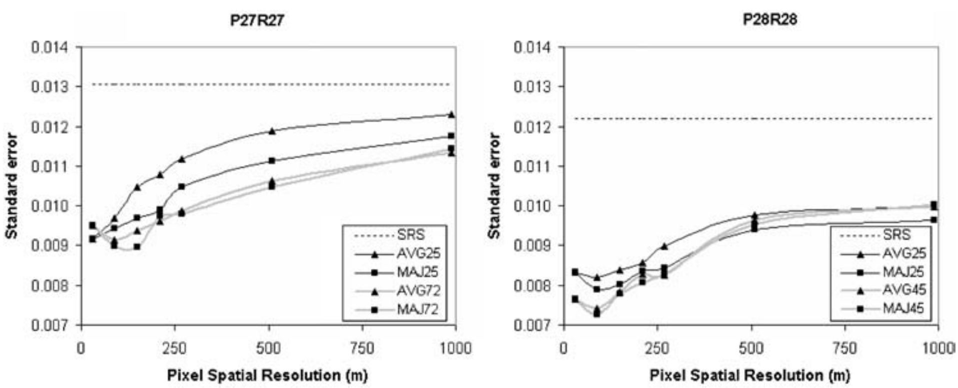


Figure 10. Standard errors of design-based forest proportion estimates calculated under the assumption of simple random sampling (SRS), and standard errors derived using post-sampling stratification with average-based (AVG) and majority-based (MAJ) aggregations, and standard (0.25) and optimal (0.72 or 0.45) forest proportion thresholds for P27R27 and P28R28, Minnesota, USA.

At least two possible reasons may explain the difference between estimates using optimal and standard thresholds. First, 30 m pixels having forest proportion values between 0.25 and 0.75 were infrequent, with a majority of pixels having values nearer 0.0 or 1.0. The 30 m pixel estimates of forest proportion used in this study were produced with a logistic modelling approach, in part because the field training observations were composed primarily of field observations of either 0.0 or 1.0 proportion forest on each field plot, with few intermediate observations. A possible artefact of this modelling approach is that intermediate values of proportion forest (i.e. those associated with standard and optimal thresholds) may not represent the population adequately, even though image-wide mean values of pixel forest proportion corresponded closely to estimates from design-based approaches. This factor has not been tested. A second factor possibly influencing the discrepancy between estimates from standard and optimal thresholds is a difference in the definitions between image pixels and plots. FIA's definition of forest land requires a minimum patch area, minimum patch width, and exclusion of treed lands developed for nonforest uses. Although based on FIA plot observations for training data, the 30 m pixel estimates of forest proportion were not filtered to comply with FIA's definition, resulting in inclusion of isolated pixels in patches too small or too narrow, or pixels with tree cover in land uses not permitted in FIA's definition. Thus, use of the standard threshold would result in overestimates of forest proportion when a number of these pixels occur in conditions not permitted by FIA's definition.

Optimal thresholds of proportion forest decreased nonlinearly as spatial resolution became coarser, much more so for sparsely forested P28R28 (18% decrease) than for heavily forested P27R27 (6% decrease). By contrast, Pax-Lenny and Woodcock (1997) reported that spatial resolution of NDVI datasets calculated from spatially aggregated TM spectral data had a negligible effect on the maximum NDVI threshold used to distinguish productive and nonproductive agricultural lands. For P28R28, the 30 m resolution optimal threshold (0.45) was substantially larger than both the mean pixel proportion (0.217) and the design-based estimate of forest proportion (0.228). In addition to comprising a minority of the area, forest land in P28R28 also appears to be more patchy or fragmented than in P27R27, where the 30 m resolution optimal threshold (0.72) was similar to both the mean pixel proportion (0.750) and the design-based estimate of forest proportion (0.736). Gains in precision of stratified estimates resulted from using an optimal threshold over a standard threshold for certain spatial resolutions. Although constructing these stratification categories from models fitted to the sample data (endogenous post-stratification) violates standard post-stratification assumptions, the probable practical effect is negligible (Breidt and Opsomer 2002, McRoberts *et al.* 2006). Additional work is needed to test the utility of resolution-specific thresholds for increasing precision of stratified estimates.

Majority-based spatial aggregation of 30 m F/NF maps produced from optimal proportion forest thresholds resulted in nonlinear increasing deviation from design-based estimates as the resulting spatial resolution became coarser. Specifically, majority-based aggregations led to overestimation of forest land area in heavily forested P27R27 and underestimation of forest land area in sparsely forested P28R28. Our results support those of Nelson (1989), Turner *et al.* (1989), Moody and Woodcock (1994), Kuusela and Päivinen (1995) and Mayaux and Lambin (1995), all of whom reported the effects of majority-based spatial aggregation.

4.2 *Estimates obtained using design-based approaches*

For both study areas, standard errors of stratified estimates of forest proportion showed generally increasing trends with coarser spatial resolution using both average- and majority-based stratifications for both standard and optimal thresholds of forest proportion. The use of optimal thresholds in place of the standard threshold appeared to reduce standard errors of stratified estimates for some spatial resolutions. By definition, the use of optimal thresholds results in F/NF maps more like an FIA plot-derived map, leading to more FIA forest plots being correctly assigned to the 'forest' stratum, which leads to increased precision of stratified estimates. The fact that this improvement appears greater for P27P27 than for P28R28 may relate to the greater absolute difference between optimal and standard thresholds in P27R27 (0.75 *vs*. 0.25) than in P28R28 (0.45 *vs*. 0.25).

The standard errors of forest proportion stratified estimates for both study areas were smallest at spatial resolutions between 90 and 150 m and not at 30 m, which was the finest resolution tested. This result was similar to those of Sadowski *et al.* (1977), Latty and Hoffer (1981), Markham and Townshend (1981) and Ju *et al.* (2005). This may relate to the fact that 30 m TM/ETM+ pixels are larger than individual tree canopies, smaller than forest patches, and smaller in area (0.09 ha) than FIA's minimum forest area definition (0.405 ha). Hansen and Wendt (2000) and McRoberts *et al.* (2002) observed that standard errors of stratified estimates based on 30 m stratification maps can be reduced by first clumping and eliminating small clusters of fewer than four forest or nonforest TM pixels (<0.36 ha), equivalent in area to a 60×60 m pixel.

Per-plot proportion forest was based on the average forest proportion across four 0.017 ha circular subplots. The per-plot forest proportion was then assigned to a single stratum based on the single pixel associated with the central subplot location. Pixels associated with the locations of the outer three subplots were not considered when assigning plots to strata. FIA plot Global Positioning System (GPS) coordinate location error ranges from 10 to 15 m; pixel locations have an associated image registration error of about 15 m; and the four subplots within each FIA plot are encompassed by a 3×3 window of TM pixels (90 m). By combining these factors, the image area associated with each FIA plot could be characterized as roughly 120×120 m, which is within the range of our observed optimal resolution (90–150 m).

Spatial resolution-dependent trends in standard errors of design-based stratified estimates were nearly identical using only central subplots versus aggregating four subplots for both P27R27 and P28R28 majority-based aggregations with optimal thresholds (figure 10). From this we infer that, for both P27R27 and P28R28, aggregating forest proportion from four subplots and assigning plots to strata based only on their central subplot locations does not affect bias or precision of forest proportion estimates or the spatial resolution for which the smallest standard errors were observed.

We surmise that reductions in standard errors of stratified estimates following spatial aggregation of stratification maps between 90 and 150 m or clumping and eliminating smaller pixel clusters result from related spatial effects. For both study areas, the largest standard errors of stratified estimates resulted from stratification maps of 510–990 m spatial resolution, but even these estimates exhibited gains in precision over estimates calculated under the assumption of SRS. Although we did not test stratification maps at resolutions coarser than 1000 m, we expect the trend of increasing standard errors with coarser spatial resolutions to reach an asymptote beyond 1000 m, as is suggested by the semivariograms in figures 7 and 8.

Although not identical to the spatial resolutions of satellite image sensors from MODIS (250 m, 500 m, 1 km) and AVHRR (1.1 km), our simulations at 270, 510 and 990 m provided for approximate comparisons with these sensor resolutions. However, we acknowledge that aggregations from 30 m imagery registered to within about 15 m to coarser spatial resolutions are likely to benefit from much better image registration characteristics than would occur from registration of coarser spatial resolution imagery (Townshend and Justice 1988). In several cases, the precision of design-based estimates using stratification maps of approximately 250 m spatial resolution was comparable with or only slightly lower than that for 30 m stratification maps. This is corroborated by the results of Ju *et al.* (2005), who observed that MAEs of estimates of land cover fraction following spatial aggregation to 250 m were similar to the MAEs of their 30 m dataset. Although MODIS VCF data on percentage tree canopy cover are available only at 500 m spatial resolution, an anticipated 250 m VCF dataset holds promise for stratifying inventory estimates of forest land area. In addition, calibration of continuous forest or tree canopy cover geospatial datasets holds potential for portraying spatial distributions and area estimates comparable with forest inventory estimates.

The distribution of a nationwide 30 m percentage forest canopy map as a component of the 2001 NLCD (Homer *et al.* 2004) provides a potential source of stratification maps for producing stratified inventory estimates of forest land area. Spatially aggregating the 30 m NLCD 2001 percentage forest canopy dataset to moderately coarser spatial resolution (e.g. 90–150 m) and identifying optimal thresholds of forest proportion may result in simultaneous gains in precision and more efficient management of stratification datasets. We simulated effects of spatial resolution by aggregating a 30 m pixel dataset of forest proportion derived from an ETM+ dataset to six coarser spatial resolutions. In this study, only spatial resolution was evaluated; the effects of spectral, temporal and radiometric resolution were held constant within study areas of fixed extent. We compared estimates of forest land area among and between the resulting forest/nonforest maps and design-based stratified forest inventory estimates using stratifications derived from the maps. The results of these tests are intended to provide insight into the effects of spatial aggregation and resolution on estimates of forest land such that NFIs may better address the issues of spatial resolution when selecting satellite remotely sensed imagery and derived datasets. We are pursuing further investigations in this area.

5. Conclusions

The following conclusions can be drawn from this study.

First, for average-based spatial aggregation of per-pixel forest proportion, local variability showed a slight peak between 210 and 270 m, suggesting that typical forest patches in the study areas measure about 300–500 m per side (9–25 ha), although the weakness of the peak suggests that patch sizes vary somewhat. Thus, use of image-based products with pixel spatial resolutions coarser than 30 m appears to be justified for portraying F/NF in the study areas. Image variability decreased nonlinearly as spatial resolution became coarser, and mean forest proportion did not vary.

Second, majority-based spatial aggregation, following assignment of pixels to F/NF classes based on a minimum threshold of forest proportion, resulted in overestimation of forest proportion in a heavily forested area and underestimation of forest proportion in a sparsely forested area. Deviations from original pixel estimates were nonlinearly related to spatial resolution. Optimum thresholds varied with spatial resolution and should be calculated specific to the spatial resolution of interest.

Third, spatial aggregation had a minimal effect on design-based stratified estimates of forest proportion but standard errors of estimates varied considerably with aggregation approach, threshold and spatial resolution of stratification maps. For some but not all spatial resolutions, reductions in standard errors of stratified estimates of forest proportion resulted from stratifications based on landscape-specific optimal thresholds rather than a standard threshold of minimum forest proportion.

Fourth, aggregation of a 30 m dataset to six successively coarser spatial resolutions resulted in generally increasing standard errors of stratified estimates, although standard errors were lowest for spatial resolutions between 90 and 150 m and resolutions near 250 m produced standard errors of stratified estimates only slightly larger than those from 30 m stratification datasets. Standard errors were largest for stratifications at the coarsest spatial resolutions tested (510–990 m), but even these were smaller than for estimates calculated under the assumption of SRS. Determining optimal spatial resolution and aggregation of satellite image-derived stratification maps can lead to improved stratification efficacy of design-based estimates of forest land area. Although these results are not based on tests of moderate resolution imagery, they do suggest that moderate resolution image products, such as a forthcoming global 250 m MODIS VCF dataset, may provide precision benefits to NFIs similar to those of finer spatial resolution 30 m Landsat imagery but without the associated limitations of temporal and radiometric resolution, greater cost of data acquisition and processing, drawbacks of technical failures with current sensors, and uncertainty of future Landsat missions. Additional research is needed for upscaling *in situ* NFI data to moderate resolution satellite image data.

Spatial resolution is but one of several characteristics affecting satellite image utility. By itself, degradation of spatial resolution results in loss of local specificity. However, enhanced spectral, radiometric and temporal resolutions in some sensors with moderate to coarse spatial resolution (e.g. MODIS) provide advantages over 30 m Landsat sensors. The interactions of sensor characteristics are becoming better understood through upscaling of field data, subpixel modelling, multisensor/multiresolution image processing, and other comparative studies. Continuing assessments of satellite image characteristics, both generic and sensor-specific, are required to address the evolving needs of NFIs.

Acknowledgements

We thank the editor and both anonymous reviewers for prescribing substantial improvements to the original manuscript.

References

Atkinson, P.M. and Curran, P.J., 1995, Defining an optimal size of support for remote sensing investigations. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, pp. 768–776.

BIAN, L., 1997, Multiscale nature of spatial data in scaling up environmental models. In *Scale in Remote Sensing and GIS*, D.A. Quattrochi and M.F. Goodchild (Eds), pp. 13–26 (Boca Raton, FL: CRC Lewis).

BREIDT, F.J. and OPSOMER, J.D., 2002, *Endogenous Post-Stratification in Surveys: Classifying with a Sample Fitted Model* (Ames, IO: Department of Statistics, Iowa State University).

COCHRAN, W., 1977, *Sampling Techniques*, 3rd edn (New York: John Wiley & Sons).

COHEN, W.B., MAIERSPERGER, T.K., YANG, Z., GOWER, S.T., TURNER, D.P., RITTS, W.D., BERTERRETCHE, M. and RUNNING, S.W., 2003, Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*, **88**, pp. 233–255.

COLOMBO, S., CHICA-OLMO, M., ABARCA, F. and EVA, H., 2004, Variographic analysis of tropical forest cover from multi-scale remotely sensed imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp. 330–341.

DEUTSCH, C.V. and JOURNEL, A.G., 1997, *GSLIB: Geostatistical Software Library and User's Guide*, 2nd edn (New York: Oxford University Press).

FERRO, C.J.S., 1998, Scale and texture in digital image classification. MA thesis, West Virginia University, Morgantown, WV.

FORSHAW, M.R.B., HASKELL, A., MILLER, P.F., STANLEY, D.J. and TOWNSHEND, J.R.G., 1983, Spatial resolution of remotely sensed imagery: a review paper. *International Journal of Remote Sensing*, **4**, pp. 497–520.

GARCÍA-GIGORRO, S. and SAURA, S., 2005, Forest fragmentation estimated from remotely sensed data: is comparison across scales possible? *Forest Science*, **51**, pp. 51–63.

GOODCHILD, M.F. and QUATTROCHI, D.A., 1997, Scale, multiscaling, remote sensing, and GIS. In *Scale in Remote Sensing and GIS*, D.A. Quattrochi and M.F. Goodchild (Eds), pp. 1–11 (Boca Raton, FL: CRC Lewis).

HANSEN, M.H. and WENDT, D.G., 2000, Using classified Landsat Thematic Mapper data for stratification in a statewide forest inventory. In *Proceedings of the First Annual Forest Inventory and Analysis Symposium*, 2–3 November 1999, San Antonio, TX, R.E. McRoberts, G.A. Reams, and P.C. Van Duesen (Eds), pp. 20–27, General Technical Report NC-213 (St Paul, MN: US Department of Agriculture, Forest Service, North Central Forest Experiment Station).

HAY, G.J., NIEMANN, K.O. and GOODENOUGH, D.G., 1997, Spatial thresholds, image-objects, and upscaling: a multiscale evaluation. *Remote Sensing of Environment*, **62**, pp. 1–19.

HENDERSON-SELLERS, A. and PITMAN, A.J., 1992, Land-surface schemes for future climate models: specification, aggregation and heterogeneity. *Journal of Geophysical Research*, **97**, pp. 2687–2696.

HLAVKA, C.A. and DUNGAN, J.L., 2002, Areal estimates of fragmented land cover: effects of pixel size and model-based corrections. *International Journal of Remote Sensing*, **23**, pp. 711–724.

HLAVKA, C.A. and LIVINGSTON, G.P., 1997, Statistical models of fragmented land cover and the effects of coarse spatial resolution on the estimation of area with satellite sensor imagery. *International Journal of Remote Sensing*, **18**, pp. 2253–2259.

HOMER, C., HUANG, C., YANG, L., WYLIE, B. and COAN, M., 2004, Development of a 2001 national land-cover database for the United States. *Photogrammetric Engineering and Remote Sensing*, **70**, pp. 829–840.

ISAAKS, E.H. and SRIVASTAVA, R.M., 1989, *Applied Geostatistics* (New York: Oxford University Press).

JARNAGIN, S.T., JENNINGS, D.B. and EBERT, D.W., 2004, A technique for assessing the accuracy of sub-pixel impervious surface estimates derived from Landsat TM imagery. In *Remote Sensing and GIS Accuracy Assessment*, R.S. Lunetta and J.G. Lyon (Eds), pp. 269–280 (Boca Raton, FL: CRC Press).

JENSEN, J.R., 1996, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 2nd edn (Upper Saddle River, NJ: Prentice Hall).

JOSEPH, G., 2000, How well do we understand Earth observation electro-optical sensor parameters? *ISPRS Journal of Photogrammetry and Remote Sensing*, **55**, pp. 9–12.

JU, J., GOPAL, S. and KOLACZYK, E.D., 2005, On the choice of spatial and categorical scale in remote sensing land cover classification. *Remote Sensing of Environment*, **96**, pp. 62–77.

JUSTICE, C.O., MARKHAM, B.L., TOWNSHEND, J.R.G. and KENNARD, R.L., 1989, Spatial degradation of satellite data. *International Journal of Remote Sensing*, **10**, pp. 1539–1561.

KUUSELA, K. and PÄIVINEN, R., 1995, On the classification of ecosystems in boreal and temperate forests. In *Proceedings of an International Workshop Designing a System of Nomenclature for European Forest Mapping*, 13–15 June 1994, Joensuu, Finland, P.J. Kennedy, R. Päivinen, and L. Roihuvuo (Eds), pp. 387–393. Report EUR 16113 EN (Ispra, Varese, Italy: Institute for Remote Sensing Applications, Joint Research Centre, European Commission and the European Forest Institute).

LATTY, R.S. and HOFFER, R.M., 1981, Computer-based classification accuracy due to the spatial resolution using per-point versus per-field classification techniques. In *Proceedings of the 7th International Symposium on Machine Processing of Remotely Sensed Data*, 23–26 June 1981, West Lafayette, IN, P.G. Burroff and D.B. Morrison (Eds), pp. 384–393. LARS Technical Note 062281 (West Lafayette, IN: LARS/Purdue University).

LEFSKY, M.A. and COHEN, W.B., 2003, Selection of remotely sensed data. In *Remote Sensing of Forest Environments: Concepts and Case Studies*, M.A. Wulder and S.E. Franklin (Eds), pp. 13–46 (Boston: Kluwer Academic Publishers).

LESSARD, V.C., HANSEN, M.H. and NELSON, M.D., 2003, Comparing Minnesota land cover/use area estimates using NRI and FIA data. In *Third Annual Forest Inventory and Analysis Symposium*, 17–19 October 2001, Traverse City, MI, R.E. McRoberts, G.A. Reams, P.C. Van Deusen, and J.W. Moser (Eds), pp. 36–43, General Technical Report NC-230 (St Paul, MN: US Department of Agriculture, Forest Service, North Central Research Station).

LIKNES, G.C., NELSON, M.D. and MCROBERTS, R.E., 2004, Evaluating classified MODIS satellite imagery as a stratification tool. In *ACCURACY 2004: The Joint Meeting of the Sixth International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences and the Fifteenth Annual Conference of TIES*, The International Environmetrics Society, 28 June–1 July 2004, Portland, Maine, USA, H.T. Mowrer, R.E. McRoberts, and P.C. Van Deusen (Eds), pp. 1–8 (St Lucia, Queensland, Australia: International Environmetrics Society).

MARCEAU, D.J. and HAY, G.J., 1999, Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing*, **25**, pp. 357–366.

MARCEAU, D.J., HOWARTH, P.J. and GRATTON, D.J., 1994a, Remote sensing and the measurement of geographical entities in a forest environment. 1: The scale and spatial aggregation problem. *Remote Sensing of Environment*, **49**, pp. 93–104.

MARCEAU, D.J., HOWARTH, P.J. and GRATTON, D.J., 1994b, Remote sensing and the measurement of geographical entities in a forest environment. 2: The optimal spatial resolution. *Remote Sensing of Environment*, **49**, pp. 105–117.

MARKHAM, B.L. and TOWNSHEND, J.R.G., 1981, Land cover classification accuracy as a function of sensor spatial resolution. In *Proceedings of the 15th International Symposium on Remote Sensing of Environment*, Ann Arbor, MI, pp. 1075–1090 (Ann Arbor, MI: Environmental Research Institute of Michigan).

MAYAUX, P. and LAMBIN, E.F., 1995, Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation. *Remote Sensing of Environment*, **53**, pp. 1–15.

McRoberts, R.E., 2002, Using maps to increase the precision of forest inventory estimates. In *Accuracy 2002: Proceedings of the 5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 10–12 July 2002, Melbourne, Australia, G. Hunter and K. Lowell (Eds), pp. 456–463 (Melbourne: Department of Geomatics, The University of Melbourne).

McRoberts, R.E., 2006, A model-based approach to estimating forest area. *Remote Sensing of Environment*, **103**, pp. 56–66.

McRoberts, R.E., Holden, G.R., Nelson, M.D., Liknes, G.C. and Gormanson, D.D., 2006, Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the USDA Forest Service. *Canadian Journal of Forest Research*, **36**, pp. 2968–2980.

McRoberts, R.E., Wendt, D.G., Nelson, M.D. and Hansen, M.H., 2002, Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, **81**, pp. 36–44.

Millington, A.C., Velez-Liendo, X.M. and Bradley, A.V., 2003, Scale dependence in multitemporal mapping of forest fragmentation in Bolivia: implications for explaining temporal trends in landscape ecology and applications to biodiversity conservation. *ISPRS Journal of Photogrammetry and Remote Sensing*, **57**, pp. 289–299.

Moody, A. and Woodcock, C.E., 1994, Scale-dependent errors in the estimation of land-cover proportions: implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 585–594.

Moran, M.S., Humes, K.S. and Pinter, P.J., Jr., 1997, The scaling characteristics of remotely-sensed variables for sparsely-vegetated heterogeneous landscapes. *Journal of Hydrology*, **190**, pp. 337–362.

Morisette, J.T., Nickeson, J.E., Davis, P., Wang, Y., Tian, Y., Woodcock, C.E., Shabanov, N., Hansen, M.C., Cohen, W.B., Oetter, D.R. and Kennedy, R.E., 2003, High spatial resolution satellite observations for validation of MODIS land products: IKONOS observations acquired under the NASA Scientific Data Purchase. *Remote Sensing of Environment*, **88**, pp. 100–110.

Nelson, M.D., McRoberts, R.E. and Hansen, M.C., 2004, Forest land area estimates from Vegetation Continuous Fields. In *Remote Sensing for Field Users: Proceedings of the Tenth Biennial Forest Service Remote Sensing Applications Conference*, 5–9 April 2004, Salt Lake City, UT, J.D. Greer (Ed.), pp. 1–6 (Bethesda, MD: American Society for Photogrammetry and Remote Sensing).

Nelson, M.D., McRoberts, R.E. and Lessard, V.C., 2005, Comparison of U.S. forest land area estimates from Forest Inventory and Analysis, National Resources Inventory, and four satellite image-derived land cover datasets. In *Proceedings of the Fifth Annual Forest Inventory and Analysis Symposium*, 18–20 November 2003, New Orleans, LA, R.E. McRoberts, G.A. Reams, P.C. Van Duesen, and W.H. McWilliams (Eds), pp. 129–137, General Technical Report WO-69 (Washington, DC: US Department of Agriculture, Forest Service).

Nelson, R., 1989, Regression and ratio estimators to integrate AVHRR and MSS data. *Remote Sensing of Environment*, **30**, pp. 201–216.

Pax-Lenny, M. and Woodcock, C.E., 1997, The effects of spatial resolution on the ability to monitor the status of agricultural lands. *Remote Sensing of Environment*, **61**, pp. 210–220.

Pebesma, E.J., 2004, Multivariable geostatistics in S: the gstat package. *Computers and Geosciences*, **30**, pp. 683–691.

R Development Core Team, 2004, *R: A Language and Environment for Statistical Computing* (Vienna: Foundation for Statistical Computing).

Reese, H., Nilsson, M., Sandström, P. and Olsson, H., 2002, Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture*, **37**, pp. 37–55.

SADOWSKI, F.G., MAILILA, J.E., SARNO, J.E. and NALEPKA, R.F., 1977, The influence of multispectral scanner spatial resolution on forest feature classification. In *Proceedings of the 11th International Symposium on Remote Sensing of Environment*, Ann Arbor, MI, pp. 1279–1288 (Ann Arbor, MI: Environmental Research Institute of Michigan).

SAURA, S., 2004, Effects of remote sensor spatial resolution and data aggregation on selected fragmentation indices. *Landscape Ecology*, **19**, pp. 197–209.

STRAHLER, A.H., WOODCOCK, C.E. and SMITH, J.A., 1986, On the nature of models in remote sensing. *Remote Sensing of Environment*, **20**, pp. 121–139.

SU, Z., PELGRUM, H. and MENENTI, M., 1999, Aggregation rules for surface parameters in global models. *Hydrology and Earth System Sciences*, **3**, pp. 549–563.

THOMAS, G., HOBBS, S.E. and DUFOUR, M., 1996, Woodland area estimation by spectral mixing: applying a goodness-of-fit solution method. *International Journal of Remote Sensing*, **17**, pp. 291–301.

TIAN, Y., WANG, Y., ZHANG, Y., KNYAZIKHIN, Y., BOGAERT, J. and MYNENI, R.B., 2002, Radiative transfer based scaling of LAI retrievals from reflectance data of different resolutions. *Remote Sensing of Environment*, **84**, pp. 143–159.

TOWNSHEND, J.R.G. and JUSTICE, C.O., 1988, Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. *International Journal of Remote Sensing*, **9**, pp. 187–236.

TURNER, M.G., O'NEILL, R.V., GARDNER, R.H. and MILNE, B.T., 1989, Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology*, **3**, pp. 153–162.

USDA FOREST SERVICE, 2003, *Forest Inventory and Analysis National Core Field Guide, Volume 1: Field Data Collection Procedures for Phase 2 Plots, Version 2.0* (Washington, DC: US Department of Agriculture Forest Service, Forest Inventory and Analysis).

WOODCOCK, C.E. and STRAHLER, A.H., 1987, The factor of scale in remote sensing. *Remote Sensing of Environment*, **21**, pp. 311–332.

WU, J., 2004, Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, **19**, pp. 125–138.

ZHU, Z. and EVANS, D.L., 1994, U.S. forest types and predicted percent forest cover from AVHRR data. *Photogrammetric Engineering and Remote Sensing*, **60**, pp. 525–531.