Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forest Ecology and Management 222 (2006) 167–181
Forest Ecology and Management
www.elsevier.com/locate/foreco

Key issues in making and using satellite-based maps in ecology: A primer

Karin S. Fassnacht [1], Warren B. Cohen *, Thomas A. Spies

USDA Forest Service, Pacific Northwest Research Station, Forestry Sciences Laboratory, 3200 SW Jefferson Way, Corvallis, OR 97331, USA

Received 26 June 2005; received in revised form 30 September 2005; accepted 30 September 2005

Abstract

The widespread availability of satellite imagery and image processing software has made it relatively easy for ecologists to use satellite imagery to address questions at the landscape and regional scales. However, as often happens with complex tools that are rendered easy to use by computer software, technology may be misused or used without an understanding of some of the limitations or caveats associated with a particular application. The results can be disappointment when maps are less accurate than expected or incorrect decisions when they are treated as truth. In this paper, we discuss several key issues which are critical to ensuring the effectiveness and value of remote sensing products, but which are also sometimes sources of confusion: (1) direct versus indirect models of land surface properties and processes, (2) differences between class-based and continuous mapping models, (3) scale, and (4) accuracy assessment. We illustrate our points with examples from the application of satellite imagery to forest management issues in the Pacific Northwest, USA. While our examples focus largely on Landsat image data, the issues we discuss have broad relevance across sensor data types, land cover properties, and geographic locations.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Classification; Scale; Accuracy assessment; Remote sensing; Satellite

1. Introduction

In recent years, the use of satellite imagery in ecological research and ecosystem management has increased significantly. For example, 9 of 10 recent issues of the journal *Ecological Applications* contain at least one article in which satellite imagery was used to characterize vegetation or land cover. The widespread availability of imagery and image processing software has made it relatively easy for ecologists to employ satellite imagery to address questions at landscape and regional scales. However, as often happens with complex tools that are rendered easy to use by computer software, technology may be misused or used without understanding some of the limitations and caveats associated with a particular application. The promise of satellite imagery to solve landscape research and management problems has frequently been high (Green, 1994), but some assert that it has not delivered (Meyer and Werth, 1990; Holmgren and Thuresson, 1998). Some of the reasons for "failures" in the application of satellite imagery may derive from overly optimistic promises by remote sensing advocates, and from poor understanding by ecologists of the limits of the technology and the methods used to assess accuracy.

While there is a substantial body of literature exploring the application of remote sensing to ecological endeavors, there is a need for a non-technical overview of fundamental issues that ecologists should consider when using remote sensing. In this paper, we provide such an overview with respect to (1) feature description and spatial representation and (2) accuracy assessment, illustrating our points with examples from recent applications of remote sensing in ecological research and ecosystem management in western Oregon and Washington (Weyermann and Fassnacht, 2001; Cohen et al., 2001; Ohmann and Gregory, 2002; Spies et al., 2002; Nighbert et al., public communication[2]). We chose the Pacific Northwest for our examples because satellite imagery has been widely used in

* Corresponding author. Tel.: +1 541 750 7322; fax: +1 541 758 7760.
E-mail address: wcohen@fs.fed.us (W.B. Cohen).
[1] Present address: Wisconsin Department of Natural Resources, 107 Sutliff Avenue, Rhinelander, WI 54501, USA.

[2] WODIP Guidebook, http://www.or.blm.gov/gis/projects/wodip.asp, visited 28 November 2003.

0378-1127/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.foreco.2005.09.026

ecological research here and is increasingly being used to address ecosystem management problems (e.g., FEMAT, 1993; Weyermann and Fassnacht, 2001; Spies et al., 2002). Furthermore, the region is a center of considerable debate about forest management and biodiversity, creating a real need for landscape-scale information for assessment and policy decision-making.

2. Feature description and spatial representation

2.1. Is remote sensing the right tool?

It may seem trivial, but an important first question to address is: is remote sensing the right tool? When considering this question, the strengths of the technology must be balanced against its limitations. Remote sensing can be particularly useful when there is a need for (1) complete spatial coverage, especially over large areas; (2) monitoring, or frequently repeated measurements; or (3) measurements at inaccessible or sensitive locations. For environmental monitoring and management, Phinn et al. (2000) have identified four broad types of information which can be derived from remote sensing data: landscape composition, landscape pattern, biophysical parameters, and changes in time of these three elements.

The top-down view of remote sensing platforms restricts what can be "seen" by the sensor. Consequently, many important ecological variables measured beneath or within the canopy are not good candidates for direct remote measurement. For example, the Interagency Vegetation Mapping Project (IVMP; Weyermann and Fassnacht, 2001) attempted to map quadratic mean tree stem diameter directly using Landsat Thematic Mapper data with limited success. Overall accuracies for four to six structural classes were 33–44% (USDI, public communication[3]). Another limitation is that information recorded by airborne or satellite sensors is typically restricted to energy returned in one or more wavebands and relative geographic position. As a result, features that cannot be distinguished by "color" and spatial arrangement alone may be confused when using only spectral data.

Fortunately, there are ways to address these shortcomings. If a variable of interest is not visible from overhead (e.g., soil nitrogen mineralization rates), it can be calculated by using remotely sensed values of a visible variable (e.g., foliar lignin content; Wessman et al., 1988) if the two are statistically related. Furthermore, if variables cannot be distinguished by spectral values and spatial location alone, results can be improved by using ancillary data, such as topography, ownership, or soils data (e.g., Hutchinson, 1982; Palacio-Prieto and Luna-Gonzalez, 1996; Eiumnoh and Shrestha, 2000). In addition, multiple image processing techniques (e.g., per pixel classifiers and image segmentation) can be combined with other complementary analyses (e.g., GIS-based ecological modeling) and data sources (e.g., photointerpretation and field data) to enhance the capabilities of remote sensing alone (Franklin and Woodcock, 1997).

It is possible, however, that remote sensing technology may not be able to provide acceptable results for a given objective (Franklin and Woodcock, 1997; Cohen et al., 2003). In these cases, remote sensing may not be appropriate. A framework has been developed which can be helpful in making this evaluation (Phinn et al., 2000, 2003).

2.2. Information classes

2.2.1. Defining a classification scheme

If data will be divided into groups, there are a number of factors to consider in developing a classification scheme. When hard class boundaries are desired, class definitions must be clear, mutually exclusive, and exhaustive to prevent confusion and to avoid unlabelled map regions (Congalton and Green, 1999). Mutual exclusivity is not required, however, if soft class boundaries are to be used (i.e., when features are allowed to belong to more than one class, e.g., a forest with 60% conifer might reasonably be considered to belong to either a "conifer" or "mixed" class). In either case, classes should be detailed enough to be useful, yet not so narrowly defined that they cannot be spectrally separated from other classes (Lillesand et al., 1998). Hierarchical classification schemes can help provide a balance between utility and accuracy by allowing users to define and organize desired detail, while permitting classes to be collapsed as needed if accuracies are not acceptable. For example, using the classification scheme presented in Table 1, Cohen et al. (2001) were able to provide a greater level of detail for forest compared to other land cover classes, and the most detail for closed canopy conifer forests.

If preexisting classification schemes or maps are to be used, it is important to know how the classes were defined in order to evaluate how well the scheme/map matches project objectives

Table 1
Hierarchical classification scheme used by Cohen et al. (2001)

Class
Water
Snow and ice
Cloud and shadow
Serpentine
Urban
Agriculture
Other non-forested
Forested
Open canopy
Semi-open canopy
Closed canopy
Broadleaf
Mixed
Conifer
Young
Mature
Old

[3] IVMP Datasets, http://www.or.blm.gov/gis/projects/ivmp_data.asp, visited 26 February 2005.

(Cohen et al., 2001). For example, if a new study seeks to identify potential habitat for a forest dwelling creature, is a map that defines forest as any area with over 10% canopy closure (Anderson et al., 1976) suitable? It depends on the ecology of the creature being studied. Perhaps a classification scheme that required at least 60% canopy closure (Scepan, 1999) would be more appropriate. Without looking into how classes were defined, it would be easy to assume the map did a poor job of identifying forested areas or to misidentify areas of suitable habitat.

2.2.2. Classes versus continuous representations of information

When features of interest are categorical in nature (e.g., land cover/land use), techniques are typically used which place pixels or polygons into classes with either hard or soft boundaries. Franklin et al. (2003) provide a nice overview of many classification techniques including more traditional methods (e.g., maximum likelihood and minimum distance to mean classifiers) as well as newer methods (e.g., nearest neighbor and artificial neural networks).

When the features of interest are continuous in nature (e.g., leaf area index) or are based on such features (e.g., habitat suitability based on a tree-size threshold), using classes of any kind can be restrictive. Maps produced using classes divide continuous-variable data into arbitrary categories which can be rendered obsolete if class definitions are changed after the maps have been completed. Mapping variables continuously removes artificial class structure and allows multiple projects to use the same initial map by defining class boundaries for secondary maps to meet their individual needs (as in IVMP where four or more end users with still-developing needs would be using the maps being produced; Fassnacht, personal observation). For example, a single continuous map of tree density (Fig. 1a) is easily recoded into two classed maps to accommodate different habitat-suitability requirements for the marbled murrelet (*Brachyramphus marmoratus*) (Fig. 1b; McGrath, unpublished manuscript) and the northern spotted owl (*Strix occidentalis caurina*; Fig. 1c; McGrath, unpublished manuscript). In addition, changes in class requirements can be accommodated by reclassing the original continuous map, avoiding the need to remap the entire area. It should be noted, however, that at times sufficient data are not available to develop continuous models or the form of the relationship between independent and dependent variables is quite complex. Using classification methods may be preferred in these cases (Franklin et al., 2003).

Major projects in the Pacific Northwest are moving from using classification (e.g., Nighbert et al., public communication[2]) to using continuous mapping methods based on regression (e.g., Weyermann and Fassnacht, 2001) and gradient nearest neighbor (Ohmann and Gregory, 2002) analyses. Although earlier studies using continuous methods focused primarily on model development (e.g., Franklin, 1986; Running et al., 1986; Lathrop and Pierce, 1991; Cohen and Spies, 1992), more recent studies have used these models to map forest attributes over large areas (e.g., Cohen et al., 2001; Spies et al., 2002).

Multiple independent regression models, such as those used by Cohen et al. (2001), however, do not preserve the covariance among attributes. Joint variability, particularly of species occurrence and understory and overstory characteristics, is of interest to ecologists. To retain data covariance, Ohmann and Gregory (2002) developed a gradient nearest neighbor (GNN) method, which combines direct gradient analysis and nearest neighbor imputation (Moeur and Stage, 1995). Ground-inventory-based values (e.g., species basal area), satellite spectral data, and ancillary information related to ownership, topography, geology, climate, and location were combined in a canonical correspondence analysis to produce a distribution in eight dimensions. Each pixel was then assigned all the characteristics of whichever ground plot was closest to it in this eight-dimensional space.

2.3. Scale

2.3.1. Definitions

Scale is one of the most important characteristics to consider when matching remote sensing data with natural resource applications. Because the term "scale" has been inconsistently defined (Lam and Quattrochi, 1992; O'Neill and King, 1998), we want to be explicit in its definition here to avoid confusion (Raffy and Blamont, 2003). In this paper, *scale* will refer to the combination of extent, grain, and minimum mapping unit (MMU). Changing the value of any of these elements changes the scale. *Extent* is typically defined as the total physical area to be examined in a project (project extent) or covered by a data source (data extent; Fig. 2a), while *grain* is the smallest resolvable element in the extent (Turner et al., 2001). In a raster image, the grain is the pixel size (Fig. 2b); in a photograph, it is related to the size distribution of silver halide crystals (but is also influenced by atmospheric effects, uncompensated airplane or camera motion, etc.; Lillesand et al., 2003). *Minimum mapping unit*, in contrast, is the smallest area in the extent that will be mapped as a discrete unit (Fig. 2c). Because MMU is often larger than the grain, spatial and/or content information contained in the mapping unit is lost. For example, an aerial photograph polygon labeled as conifer may contain many conifer trees, a few hardwood trees, and a grassy opening. These different stand elements are resolvable because of the grain of the photo, but the analyst decided the latter two were too small to be considered separately for the map being made.

2.3.2. Choosing appropriate values for scale elements

To get the most from remote sensing data, it is important to make conscious decisions about scale (Phinn et al., 2000). Using popular data sources (e.g., Landsat Thematic Mapper (TM); grain = 28.5 m) or the defaults of the data (i.e., MMU = grain) will result in a view of the world that matches the sensor but not necessarily the question being asked. To facilitate decision-making, it is helpful to understand how changing scale elements can affect landscape information. Increasing extent, for example, increases the probability of sampling rare classes (Wiens, 1989), whereas increasing grain size reduces the number of landscape elements that can be

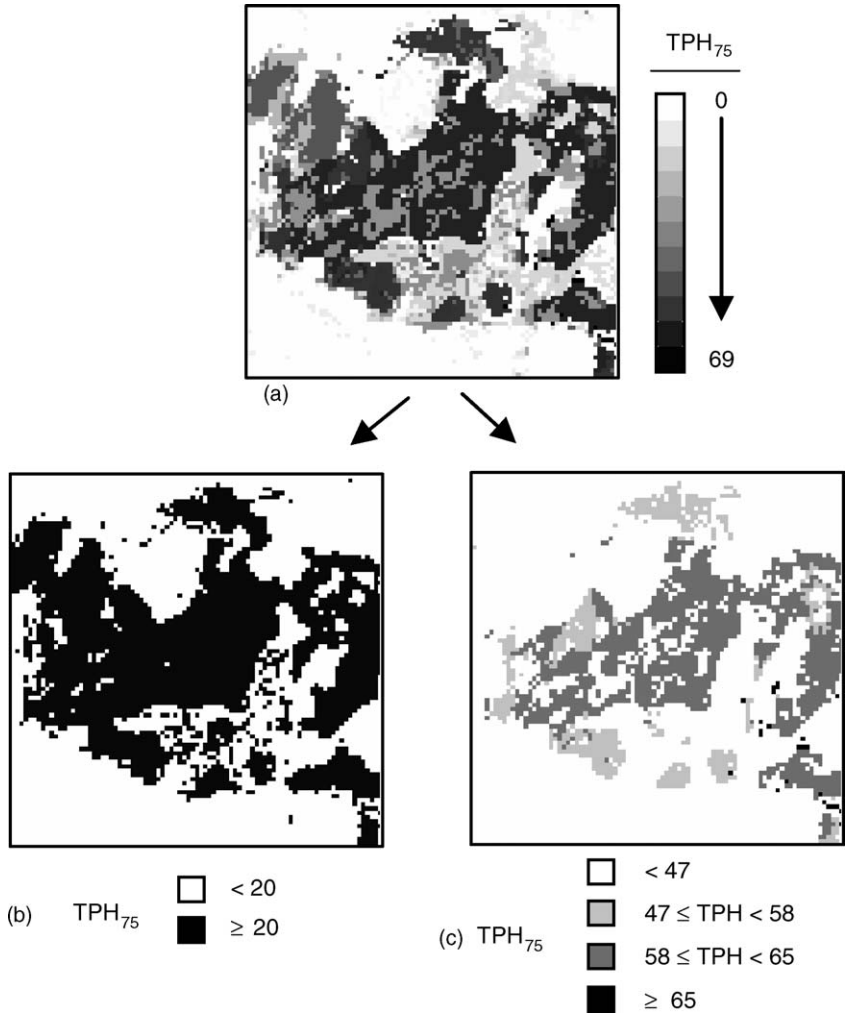

Fig. 1. (a) A continuous map of trees per hectare with diameter at breast height greater than 75 cm (TPH_{75}) used to create classed maps tailored to habitat-suitability requirements of (b) the marbled murrelet (*Brachyramphus marmoratus*) (platform tree index; McGrath, unpublished manuscript) and (c) the northern spotted owl (*Strix occidentalis caurina*) (diameter class 3 index; McGrath, unpublished manuscript). Data from Ohmann and Gregory (2002).

resolved. Changes in MMU affect the spatial expression of patches on the landscape (Fig. 3a–d). Smaller MMUs may result in large within-patch variability, making patches of interest difficult to discern (Fig. 3a) and reducing classification accuracies compared to maps where this variability has been removed (Fig. 3b and c; Toll (1985) and Irons et al. (1985), but see Markham and Townshend (1981) and Hsieh et al. (2001) for a discussion of the compensating effect of fewer mixed pixels). On the other hand, larger MMUs can result in patches of interest being artificially combined with surrounding patches (Fig. 3d).

So what are the optimum extent, grain, and MMU? There is a large body of literature devoted to this question, which speaks to both its importance and complexity. Phinn et al. (2003) have presented a framework for selecting appropriate remote sensing data for management issues which includes tackling the scale issue. In case studies presented (Phinn et al., 2000, 2003), initial grain size has typically been selected to equal the smallest feature that is to be resolved. This size is then evaluated in subsequent exploratory analyses which sometimes find that a different size would work better. Marceau et al. (1994a), in fact, found that the optimum grain size for best classification accuracy was different for different image elements. For example, to most accurately classify maple forest required a grain of 5 m while poplar and birch had optimal grains of 20 and 30 m, respectively. Likewise, at a coarser classification level, optimum grain sizes to classify deciduous, conifer, and mixed forest were 5, 10, and 30 m, respectively. Based on these results, Marceau et al. (1994a) recommend abandoning the common practice of selecting a unique resolution on which to perform analyses, instead pursuing multi-scale or hierarchical approaches which make use of logical decision rules.

In general, the factors involved in selecting optimum grain and MMU are complex. Considerations include not only project goals and the object of interest, but also the spatial arrangement of features on the landscape, the internal structure of objects of interest (e.g., stand characteristics), and the processing techniques to be used (Woodcock and Strahler, 1987). For additional discussion on this topic, see Quattrochi and Pelletier (1991, pp. 67–69), Marceau et al. (1994a,b), Cao and Lam (1997), Hsieh et al. (2001), and Franklin (2001, pp. 97–103).

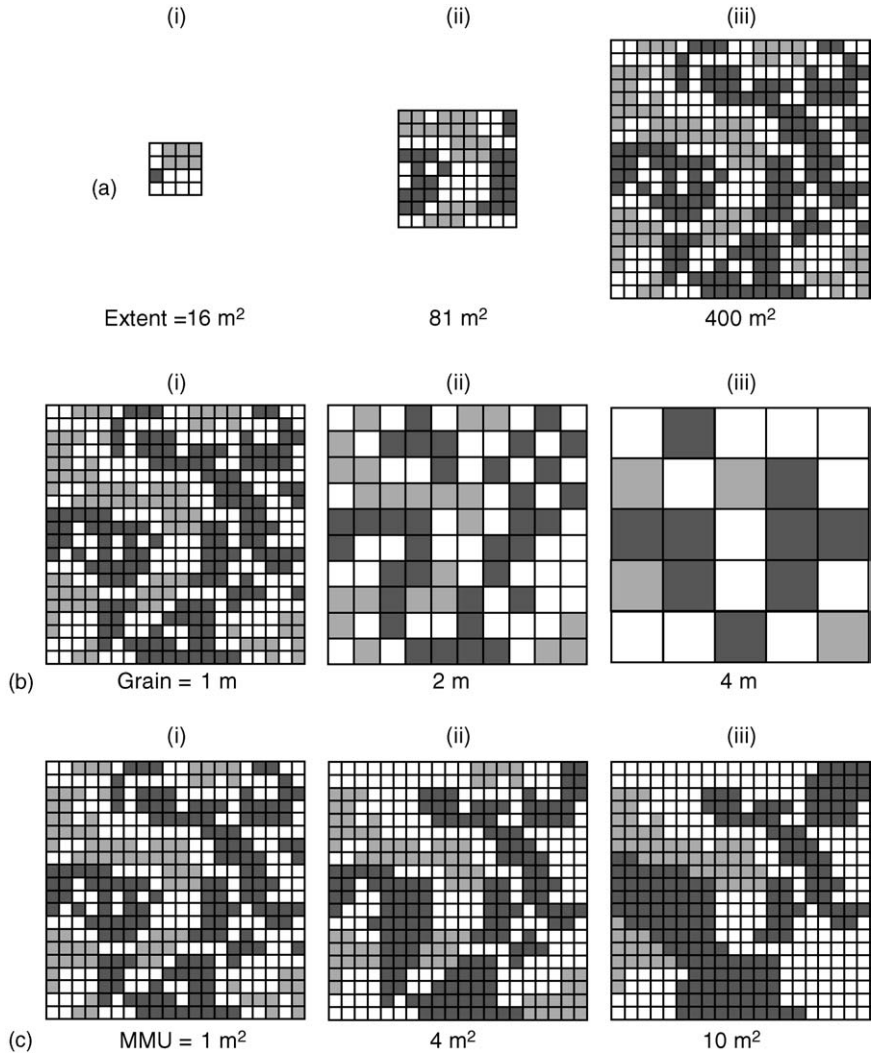

Fig. 2. Examples of increasing (a) extent, (b) grain, and (c) minimum mapping unit (MMU). Values held constant are (a) grain and MMU at 1 m and 1 m^2, respectively; (b) extent at 400 m^2; and (c) extent and grain at 400 m^2 and 1 m, respectively. Note that (b(ii)) and (c(ii)), both have extent = 400 m^2 and MMU = 4 m^2. Both were created from (b(i)) (the same as c(i)) rather than one being created from the other. Figure modified from Turner et al. (1989).

2.3.3. Example

The impacts of decisions about scale can be best appreciated through an example. To look at the effects of grain and MMU on landscape metrics, we will use a small area of classified imagery from the Coast Range of Oregon to compute landscape metrics for determining wildlife habitat suitability. We consider four scenarios: (1) grain = 25 m and MMU = 1/16 ha (i.e., one pixel), created by classifying the raw 25-m-grained imagery (Fig. 4a); (2) grain = 25 m and MMU = 1 ha, created by aggregating the map from scenario 1 to produce patches no smaller than 1 ha (Fig. 4b); (3) grain = 100 m and MMU = 1 ha, created by resampling the classified map from scenario 1 to obtain the desired grain and MMU (Fig. 4c); and (4) grain = 100 m and MMU = 1 ha as in scenario 3, but created by degrading the raw imagery to a grain of 100 m, and then reclassifying the image (Fig. 4d). Diagonals were not considered in determining patches for any scenario.

When the MMU is changed from 1/16 to 1 ha, the number of patches drops, and patch size and core area increase (Table 2), especially if the grain stays small (e.g., scenario 2). For a given MMU, smaller grained images can capture finer details in patch shape, and therefore, include more small areas into a patch. In contrast, by making the grain 100 m (scenarios 3 and 4), a 1-ha unit (pixel) belongs entirely or not at all to a given class, leading to more blocky, smaller patches (Fig. 4c and d; Table 2). When small grain is combined with small MMU (e.g., scenario 1), finer level ‘‘salt-and-pepper’’ variability is captured, leading to consistently higher edge densities and smaller solid blocks of any vegetation class (Table 2). Finally, increasing the grain of the raw imagery prior to classification (scenario 4) fundamentally changed the results, reducing the amount of ‘‘broadleaf’’, and increasing the amount of ‘‘small and medium mixed’’ (Fig. 4d; Table 2). The tendency of broadleaf cover in this image subset to form narrow linear features along rivers and drainage ways made this cover type particularly susceptible to changes in grain (Turner et al., 1989, 2000).

Obviously the four scenarios resulted in different characterizations of the landscape. None of these is wrong. A particular one, however, may be most appropriate given the species of wildlife being considered, the objectives of determining the

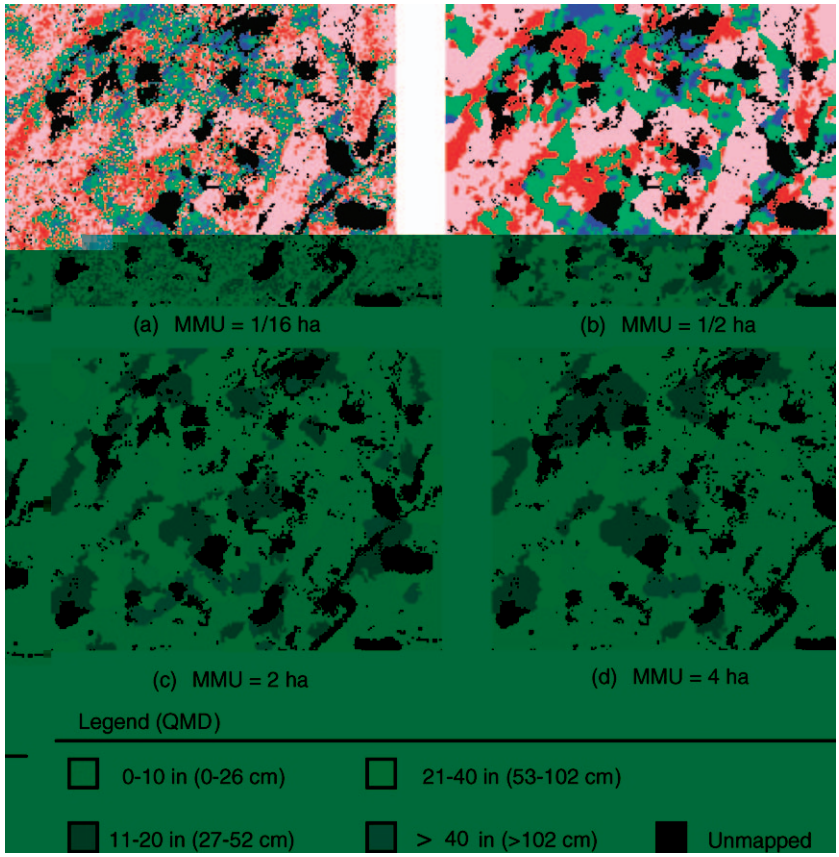

Fig. 3. The effect of increasing minimum mapping unit (MMU) on the expression of patches and heterogeneity on a landscape (a–d). The grain is 25 m in all cases, and the feature mapped is quadratic mean diameter (QMD). Unmapped areas (black) were not included in the increase in MMU. Data are from the Interagency Vegetation Mapping Project (Weyermann and Fassnacht, 2001).

landscape metrics, etc. Also noteworthy is the high degree of sensitivity of some landscape metrics to changes in scale. Such sensitivities have been widely noted in the literature (Turner et al., 1989; Moody and Woodcock, 1995; Wickham and Riitters, 1995; O'Neill et al., 1996; Cain et al., 1997; Wu et al., 2002). Wu et al. (2002) describe three categories of metrics having predictable, less-predictable staircase-like, and erratic responses to changes in scale. Given these sensitivities, they recommend the use of scalograms (which they describe as "response curves of landscape metrics to changing grain size or extent") for describing and comparing landscapes rather than seeking a single "optimum" scale at which to calculate metrics.

2.4. A word about preexisting maps

The decisions discussed in this section are extremely important for setting the foundation of a project and for maximizing the utility of remote sensing data. We recognize, however, that it is not always possible to make a new map for every project. Nonetheless, when using preexisting maps, there are tradeoffs: all the decisions made with respect to information classes, scale, and processing techniques were based on the needs of the original project. The scale selected for mapping owl habitat may not work well for studying carbon flux. Similarly, maps developed regionally may not capture local variation well (Fig. 5).

3. Accuracy assessment

Although error in remote sensing-derived maps is an important issue (Verbyla, 1995; Stehman and Czaplewski, 1998; Lowell and Jaton, 1999), users frequently treat error analysis in a *pro-forma* manner. That is, users will want to know if an accuracy assessment has been conducted but may not use the information in any way beyond establishing their general confidence level in the map. It is advisable, however, to take the time to study a map's accuracy assessment. The assessment provides not only a general idea of how "good" a map is but also details on how error is distributed both spatially and non-spatially, and insight into how further analyses of, or with, the map might be affected by error. Because a variety of accuracies can be obtained for the same map simply by modifying the methods used to calculate those accuracies, it is also worthwhile to understand how accuracies were computed. Below we discuss (1) sources of confusion with respect to accuracy, (2) effects of different methods on reported accuracy, and (3) impacts of error on further analyses.

3.1. Sources of confusion

Problems can arise when there is uncertainty about what accuracies are acceptable and misunderstandings about what accuracies are achievable. The acceptability of map accuracy