Partitioning the Conterminous United States into Mapping Zones for Landsat TM Land Cover Mapping

Collin Homer
Raytheon, EROS Data Center, Sioux Falls, South Dakota
605-594-2714
homer@usgs.gov

Alisa Gallant
EROS Data Center, Sioux Falls, South Dakota
605-594-2696
gallant@usgs.gov

U.S. Department of the Interior
U.S. Geological Survey

1. Introduction

The Multi-Resolution Land Characteristics (MRLC) Consortium was formed in 1992 in order to meet the needs of several federal agencies (USGS, EPA, NOAA, and USFS) for satellite data and for current land-cover information. In 1995, EPA's Office of Research and Development (ORD) and USGS National Mapping Division (NMD at EROS Data Center) agreed that a consistently classified national land-cover data set would be mutually beneficial. A successful mapping of the conterminous United States into the National Land Cover Dataset (NLCD 1992), was completed from circa 1992 Landsat TM at an Anderson et al. (1976) level II thematic detail in September 2000. The increasing need for land-cover and other geospatial data within the federal government again culminated in reconvening the MRLC Consortium in 2000, to plan a second generation national land-cover mapping effort. This data set, called the National Land Cover Dataset 2000 (NLCD 2000), would provide a second-point-in-time land-cover product to complement the original NLCD 1992. In addition, continued technological advancements in remote collection of earth surface data and further development of national-scope ancillary data and image classification algorithms will also lead to more comprehensive and accurate second generation data.

The NLCD 1992 was created at the TM pixel scale (30m), which results in a data layer with approximately 9 billion pixels. The size of this dataset, underscores the complexity and difficulty of land cover mapping at a national scale. NLCD 1992 was mapped in zones determined by EPA administrative boundaries (Fig 1), which were unrelated to the biogeography of land cover. This caused difficulty in staging an effective mosaic based classification of land cover, especially when mosaic boundaries included disparate land cover. This experience led to the recommendation for improved development of mapping units for NLCD 2000.

2. Mapping Zone Background

Because mapping over large landscapes typically involves many satellite scenes, multi-scene mosaicking is often used to group scenes into a single file for classification. This approach optimizes both classification and edge matching (Homer et al 1997; Homer 1996, 1999). However, large multi-scene mosaics create a variety of spectral gradients within the file, and successful spectral classification of land cover requires effective partitioning of these gradients. Spectral gradients typically represent patterns of physiographic, phenologic, solar, atmospheric and instrument influences within and between remotely sensed imagery. The degree to which this variability can be isolated in local context largely determines the success of the classification. A common method of isolating spectral gradients is to stratify landscapes into sub-regions of similar biophysical and spectral characteristics. This process is not new to remote sensing and has been widely used as a method to improve accuracy (Pettinger 1982; White et al. 1995; Lillisand 1996; Homer et al 1997). For example, Bauer et al. (1994) showed overall classification accuracy could potentially be improved by 10 to 15 percent using physiographic regions.

In this paper we outline the development of regions for NLCD 2000 called "mapping zones." The application of mapping zones as a pre-classification stratification method has been developed throughout the western

United States (Homer et al. 1997; Homer 1998, 1999). This research has shown that carefully defined mapping zones can maximize spectral differentiation, provide a means to facilitate partitioning the workload into logical units, simplify post-classification modeling and improve classification accuracy. The underlying concept of mapping zone delineation is a pre-classification division of the landscape into a finite number of units that represent relative homogeneity with respect to landform, soil, vegetation, spectral reflectance, and image footprints at a project scale that is affordable. There are five general concepts that are useful in defining mapping zones:

1.) **Economics of Size -** Mapping zones require independent treatment in classification, training site collection, modeling, assessment and edge-matching, with cost determined by mapping detail, area, and methods. Understanding the project requirements and resources is critical to determine an appropriate number of mapping zones (or independent classification units). Smaller zones may be more homogenous in spectral properties and land cover patterns, but increase individual classification efforts. Conversely, larger zones include more variability in spectral and spatial properties, but reduce the number of individual classification efforts. Once the appropriate delineation scale for the project has been determined, it constrains the subsequent scale and goals of mapping zone boundaries.

2.) **Physiography –** Criteria such as topographic variability, micro-climate (rain shadows), soil patterns, hydrography or other landscape features are important considerations in defining boundaries.

3.) **Landcover Distribution-** The distribution of vegetation and land-use patterns (e.g. agricultural and urban practices) are critical components in mapping zone delineation. Optimal boundaries reduce the total number of target classes for any one zone, thereby reducing confusion in spectral land cover associations and reducing post-classification stratification.

4.) **Spectral Uniformity -** Mapping zones allow spectral variability to be localized within a geographic area. This maximizes the potential association of land cover and spectral patterns, while minimizing potential confusion. Because the zone is mapped as a single unit, spectral uniformity among all the scenes is important. Consideration should be given to identifying zone boundaries that optimize existing scene boundaries wherever possible.

5.) **Edgematching** – Mapping zones provide an opportunity to not only minimize edgematching between individual scenes as they are compiled in the zone, but also provide a proactive selection of optimal overlap areas to edgematch adjacent zones into a single seamless project-wide coverage.

3. NLCD 2000 Mapping Zone Methods

The development of mapping zones across the conterminous United States included an initial review of project scope, which determined that approximately 60-70 zones would be the appropriate grain size for staging NLCD 2000. Initial mapping zone boundaries were based on Omernik (1987). These boundaries were displayed over two principal data layers, NLCD 1992 and AVHRR normalized greenness maps (NDVI) for modification. These data layers provided a landscape overview of both interpreted land cover and gross vegetation phenology patterns, and provided the context to further refine the initial Omernik boundaries on 1:5,000,000 scale paper maps. Paper map boundaries were subsequently crafted into a digital file by on-screen digitizing with 300m-pixel NLCD as the background. The next interpretative stage in this process will be to re-draft boundaries over full-resolution TM data to create local interpretation that can be applied at the single pixel scale. It is during this process that mapping zone concepts of physiography, spectral uniformity and edgematching will be fully applied.

4. Results

The current draft of NLCD 2000 mapping zones is shown in Figure 2. This draft represents only the conceptual scale of the potential boundary lines for the mapping zones. Because the process is iterative, it is anticipated that future updates will be applied as boundaries are interpreted at full TM scales.

5. Discussion

Taking time in the early stages of the NLCD 2000 project to develop well-defined mapping zones is expected to improve image classification, and provide a foundation to organize and prioritize the entire project (Homer 1999). Delineating mapping zones is an iterative process involving input from collaborating participants and refinement using multiple data sources. Because mapping zones boundaries are required to work at the single pixel scale, further local iteration of the lines are often necessary during actual local mapping. Because the definition of mapping zones is a relatively subjective process, it can require a significant amount of personal and collective knowledge. However, the ultimate result is a proactive process that divides the enormously complex task of large area land classification into manageable pieces.

6. References

Anderson, J.R., Hardy, E.E., Roach J.T., and Witmer R.E., 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. USGS Professional Paper 964, U.S. Geological Survey, Reston, VA.

Bauer, M.E., T.E. Burk, A.R. Ek, P.R. Coppin, S.D. Lime, T.A. Walsh, D.K. Walters, W. Befort, and D. F Heinzen. 1994. Satellite inventory of Minnesota forest resources. *Photogrammetric Engineering and Remote Sensing* 60(3):287-298.

Homer, Collin G. 1998. Southern Idaho & Western Wyoming Landcover Classification Using Landsat TM. Technical/Project report, USFS and BRD Contract fulfillment 1999.

Homer, C.G., R.D. Ramsey, T.C. Edwards, Jr., and A. Falconer. 1997. Land cover-type modeling using multi-scene Thematic Mapper mosaic. *Photgrammetric Engineering and Remote Sensing* 63:59-67.

Homer, Collin G. 1996 Nevada Landcover Classification. Technical report, Utah Cooperative Fish and Wildlife Research Unit, Utah State University, Logan, Utah. 180 pp.

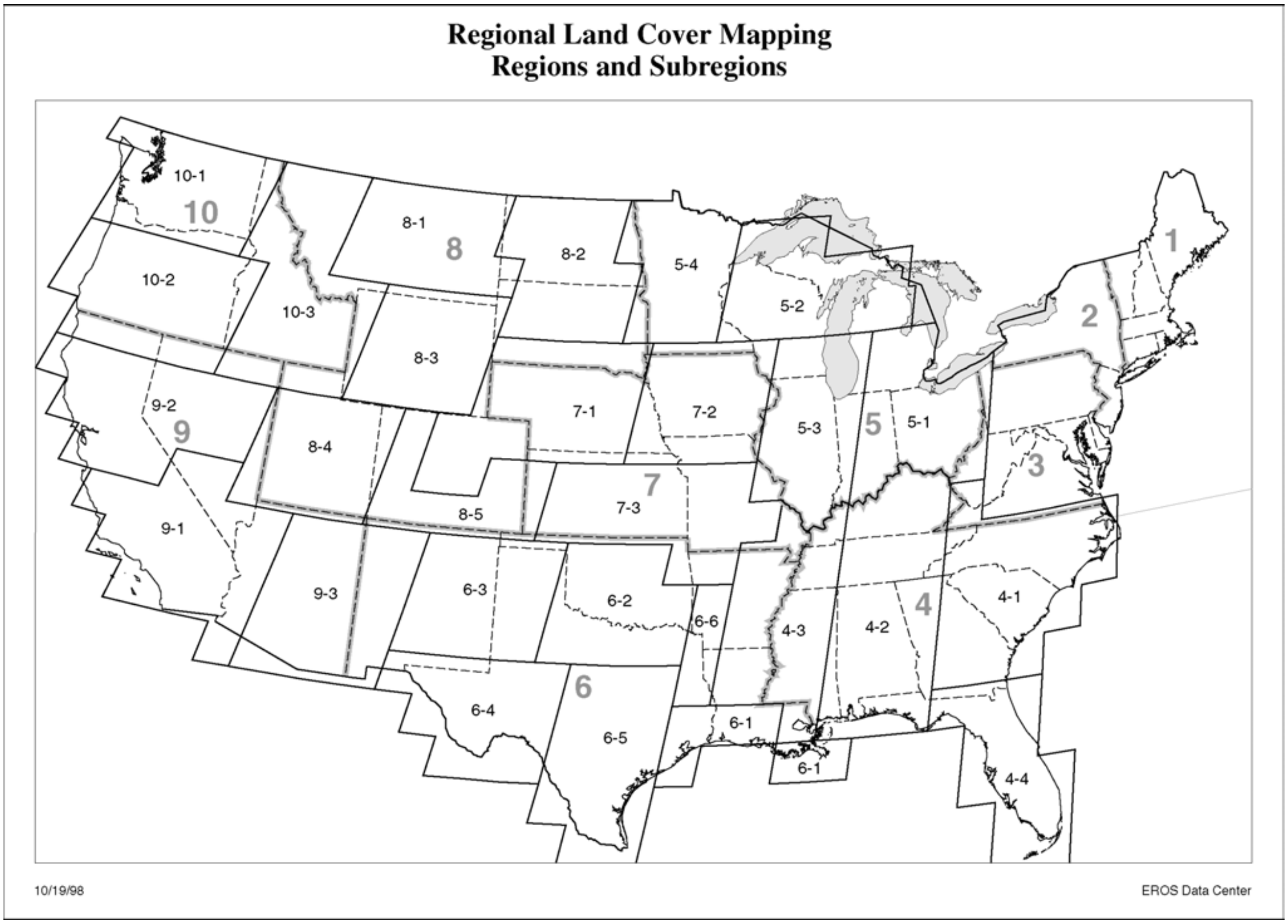
Lillisand, T. M. 1996. A Protocol for Satellite-Based Land Cover Classification in the Upper Midwest. *Gap Analysis: A Landscape Approach to Biodiversity Planning*, editors J. Michael Scott, Timothy H. Tear and Frank W. Davis. ASPRS, 320 pp.

Omernik, J.M. 1987 "Ecoregions of the Conterminous United States." Map (scale 1:7,500,000). Annals of the Association of American Geographers 77(1): 118-125

Pettinger, L.R. 1982. Digital Classification of Landsat Data for Vegetation and Land-Cover Mapping in the Blackfoot River Watershed, Southeastern Idaho.: U. S. Geological Survey Professional Paper.

White, J.D., G.C. Kroh, C. Glenn, and J.E. Pinder 3rd. 1995. Forest mapping at Lassen Volcanic National Park, California, using Landsat TM data and a geographical information system [Feature-Article]. Photogrammetric Engineering and Remote Sensing v 61 Mar P. 299-305.

Figure 1. NLCD 92 original mapping units



USGS
science for a changing world

Figure 2. Initial NLCD 2000 Mapping Zones