A Comparison of Available National-Level Land Cover Data for National Park Applications

Mike Story, Leona K. Svancara, Thom Curdts, John Gross, Shepard McAninch

Abstract

Four national scale land cover data sets were compared for areas including seven national park units in order to better understand the similarities and differences between the land cover products. We first review important factors that are likely to influence or determine the data source most appropriate for a specific analysis. These include data vintage, resolution, internal consistency, and accuracy. The conclusions that were drawn from the analysis of these comparisons were then used to make recommendations for a variety of National Park Service (NPS) applications.

Introduction

Maps depicting land cover and land use are fundamental to nearly all resource planning, management and monitoring programs at local, regional and national levels. The usefulness, efficiency, and cost-effectiveness of any land cover map is dependent on factors such as the scope and scale of questions being addressed, the spatial and temporal scale of imagery from which the landcover map was derived, the comprehensiveness of field and classification methods employed, and the level of accuracy desired. In other words, no single mapping strategy will work best in all locations for all applications.

Data on land cover and land use are available from multiple sources at a broad range of spatial and temporal scales. Those consistently available at regional and national (US) extents include the National Land Cover Data (NLCD), LANDFIRE, GAP/ReGAP, and NatureServe landcover from LandScope. These data were each developed with specific goals and objectives, using select standards and methodologies at particular spatial, temporal, and thematic scales. Just as their inherent characteristics differ (Table 1), so will their overall applicability to any given situation.

Key Considerations for Selecting a Data Source

If these data are to provide information or metrics that reflect the current conditions in and around a park unit, a user will want to weigh the vintage of the data, the relative accuracy, and the consistency of the data in deciding which source to use. Most of this type of information can be found in the metadata records that accompany the land cover data or in the reports that describe the details of the projects.

Vintage. Consider the vintage of the land cover data. The vintage should be related to the dates of the imagery and plot/field data that were used to support the effort, not the timeframe that the data were developed or released. GAP data for instance can have a wide range of source dates depending upon when a particular state initiated the process. In our examples, Arkansas data was developed primarily from TM imagery acquired in 1992. The imagery used for the North Dakota GAP products was acquired between 1992 and 1998. The vintage of the data must be appropriate for the questions being asked.

Minimum mapping unit (mmu). The minimum mapping unit was used to create the land cover data products will also affect their utility. The minimum mapping unit (mmu) for most of the NPS Inventory and Monitoring Program's (I&M) Vegetation Maps is 0.5 ha. Some of the mmus for the sources being considered can be as large as 100 ha (Arkansas GAP). Even though the data are presented as 30 meter pixels, some of the data have been smoothed to eliminate the "speckle effect" that raster data can show. The mmu for NLCD is 0.45 ha, or a group of five pixels. Southwest GAP used a "clump and eliminate" process that resulted in a mmu of approximately 4 pixels (0.36 ha). As the mmu increases, the ability to differentiate land cover classes that typically occupy "smaller" areas becomes less likely. Small, unique "communities" may be eliminated and larger polygons can become even larger as the small areas are filled in. The use of a mmu thus tends to produce maps that under-represent land cover or habitat types that are typically small, like bogs and wetlands, and overestimate the area of more extensive land cover types.

Consistency. If the analysis zone that a park unit is using extends across multiple states or other mapping units, the consistency of the data across state lines (or mapping units) may vary and thus influence the selection of a data source. Inconsistencies can result from using data of different vintages, different classification systems, different mapping techniques, or just different personnel. GAP data that has resulted from one of the several ReGAP processes has been "normalized" so that the data are consistent within the limits of the ReGAP project. Earlier GAP efforts that were produced in a state by state manner are more likely to exhibit inconsistencies across state boundaries.

Accuracy. Another consideration for selection criteria is the accuracy or reliability of the data. Several of these national land cover projects have included an accuracy assessment phase that has resulted in measures of accuracy or agreement with some reference data. Consider whether inaccuracies in the map product might impact the results of your analysis. If you find that the data includes consistent errors related to the identification of tree species but you are only interested in a more coarse "forest" designation, these errors will not likely impact your analysis.

Deciding which data source to use, and why, is challenging even for those familiar with the intricacies of their development. NLCD2001, LANDFIRE, and ReGAPare based on the same original satellite information; LANDFIRE, ReGAP, NatureServe share a comparable thematic classification, and NatureServe is a combination of the others (LANDFIRE, ReGAP) (summarized in Table 1). Of these sources, only NLCD allows for assessment of temporal changes at this time (Fall 2009). LANDFIRE is in the process of developing and implementing an "Operations and Maintenance" phase that will result in updates to that data set on a two year basis, but at this time only the original map is available. For any particular region of the country

and specific application of the data, any one of these sources is likely to be more accurate and more appropriate than the others. The challenge is to identify which one, and why. A comparison of GAP and NLCD data for the state of Kansas found that most of the disagreement between the two occurred within the grassland and cropland classes (Wardlow and Egbert, 2003). A comparison of NLCD data with more localized land cover data from Rhode Island and Massachusetts (Hollister et al , 2004) found that NLCD data provided reliable area estimates for dominant classes and those classes that exhibit unique spectral signatures. We are unaware of any other comparisons of the available data sets that provide a credible basis for determining if, or where, a particular data set is more appropriate for addressing NPS monitoring questions and objectives.

Our overall objective in this effort was to compare NLCD, LANDFIRE, GAP/reGAP, NatureServe (LandScope), and NPS vegetation maps in and around select NPS I&M parks. These data sources are most likely to be used for broad-scale characterization of landscapes and for monitoring landscape change to inform NPS decisions. Specifically, we assessed the area mapped in each land cover class by each source and developed a land cover application matrix identifying the source likely to be most applicable for particular issues in various parts of the country. Additionally, images from each of the sources were created for the study areas (and focus areas within the parks) to better understand the differences in the spatial distribution of the land cover classes. Several of the park units were recently mapped in detail as part of the NPS I&M Program Vegetation Map Inventory. Where these I&M data were available, they were compared to the national level data from each of the other four sources.

Methods

Example Parks

We identified and selected a set of 7 parks with significant natural resources for evaluating differences in the land cover data sets: Sequoia / Kings Canyon NP (SEKI), Crater Lake NP (CRLA), Theodore Roosevelt NP (THRO), Great Smoky Mountains NP (GRSM), Rocky Mountain NP (ROMO), Zion NP (ZION), and Buffalo National River (BUFF)(Figure 1). These parks include a wide range of habitats and represent a cross-section of the geographical, geophysical, and ecological variability in the national park system. Many of these parks also have completed local NPS vegetation maps.

Data Sources

We focused our comparison on four datasets consistently available at regional and national extents: the National Land Cover Data (NLCD), LANDFIRE, GAP/ReGAP, and NatureServe Landscope. Each of these datasets were acquired from their respective web sites, reprojected as necessary, and archived for future use. In addition, we downloaded local NPS vegetation maps from the NPS/USGS Mapping project. Additional information on each data set can be obtained from the associated data documentation available online from each of the sources. There is a vast overlap among the data sources in terms of imagery, dates, and sometimes even project staff (i.e., GAP processed some of the NLCD data in the SW and SE, LANDFIRE processed some of the NLCD data in other parts of the country). However, what differs are the questions being addressed by each data source and the methods used to answer those questions.

NLCD. The NLCD has competed three major successful data releases under the Multi-Resolution Land Characteristics (MRLC) Consortium, a group of Federal agencies working together to obtain imagery and derive land cover data for the Nation. These products include a 1992 conterminous U.S. land cover dataset with one thematic layer of land cover (NLCD 1992); an updated 50-state/Puerto Rico land cover database with three thematic layers including land cover, percent imperviousness, and percent tree canopy (NLCD 2001); and an NLCD 1992-2001 land cover change product that was specially designed to identify land cover change between the two datasets (NLCD 1992 and 2001). An effort is well underway for a 2006 version of NLCD. These data currently serve numerous Federal, State, local, and non-Governmental programs that seek to understand the impact of land cover on both ecological and social systems.

The 21 land cover classes (modified from Anderson et al. 1976, Level II) that make up the NLCD were derived from consistent analysis of multi-date Landsat Thematic Mapper (TM) imagery. Different methodologies were used for the 1992 and the 2001 products (Vogelman et al. 2001, Homer et al. 2004). Given the different methodologies used in NLCD 1992 and 2001 products, as well as inherent differences in source image seasonality and georegistration, classification accuracy, and map legends, direct comparison of the two datasets is subject to error. To optimally compare the products, the USGS NLCD design team created the NLCD 1992-2001 Retrofit Change product (Fry et al 2009). Accuracy assessments of NLCD 1992 (Stehman et al. 2003, Wickham et al. 2004) suggested measurements of composition and pattern should be based on eight generalized categories (Riitters et al. 2006). Similar assessments for NLCD 2001 are being completed at this time. Similarly, the NLCD 1992-2001 Retrofit Change product is available in provisional status until a formal accuracy assessment can be completed (Coan et al)

LANDFIRE. The Landscape Fire and Resource Management Planning Tools Project (LANDFIRE), is a five-year, multi-partner project producing consistent and comprehensive maps describing vegetation, wildland fuel, and fire regimes across the U.S. Developed products include existing vegetation type, biophysical setting, fire regime groups, and fire regime condition class departure index and are designed to facilitate regional- and national-level strategic planning and reporting of wildland fire management activities.

The LANDFIRE Existing Vegetation Type (EVT) layer represents the vegetation currently present at a given site. Existing vegetation is mapped through a predictive modeling approach using a combination of field reference information, Landsat imagery, and spatially explicit biophysical gradient data. Existing vegetation map units are based on NatureServe's Ecological Systems classification, which is a nationally consistent set of mid-scale ecological units (Comer et al. 2003). LANDFIRE has implemented an Operations and Maintenance Strategy, the first product of which is known as "Refresh." Refresh involves a systematic critique of existing data products and a consistent update of vegetation layers according to changes resulting from disturbance processes. In addition, products will be updated to correct some mapping inaccuracies that were detected after the initial mapping effort as part of a long-term endeavor to continually improve LANDFIRE data.

GAP. The Gap Analysis Program (GAP), conceived and launched in the late 1980s, represented the first comprehensive effort to map elements of biodiversity (land cover and species

distribution) over large areas and identify those which were under-represented (i.e., 'gaps') in the network of conservation lands. Actual (existing) land cover is one of the three main data layers developed and used by GAP to a) determine species and natural community representation within areas being managed for biodiversity conservation; b) provide spatial data that can be used to model wildlife habitat distributions; and c) provide a single temporal data set for monitoring trends (GAP 2007). As part of the MRLC, GAP works with federal partners to obtain imagery and derive land cover products for the nation.

Since the initial GAP project in Idaho (1989), a wide range of tools and procedures have become available including standardized vegetation classifications, consistent sets of satellite images, and vastly improved methodologies. Because of this, inconsistencies may arise when comparing across state lines. Recently, the GAP program has initiated regional efforts (ReGAP) to update and extend the analyses across broader areas. For example, the Southwest ReGAP effort encompassed 5 states (NV, UT, CO, AZ, and NM).

GAP does not require rigid compliance with a specific methodology or imagery source, but rather implements standards for products. GAP land cover maps meet or exceed the following specifications: 1) a nominal map scale of 1:100,000, 2) thematic classes based on the National Vegetation Classification System (NVCS) alliances or aggregations of alliances (though more recent efforts have used Ecological Systems, Comer et al. 2003), 3) complete metadata, and 4) Landsat TM imagery as a primary source of data with the base year for all TM scenes less than 3 years old at the start of the project. While initial state GAP efforts had a mmu of 100 ha for uplands and 40 ha for special features such as riparian areas, more recent efforts typically use 2 ha and 0.5 ha, respectively. Lastly, GAP requires a standardized method for assessing the accuracy of all land cover maps (for additional information see GAP 2007).

NatureServe(Landscope). Recently, NatureServe developed a national composite map of ecological systems based on LANDFIRE and ReGAP data sources. In areas with ReGAP data, these were used as the base map. In areas without ReGAP, NatureServe used LANDFIRE data to produce their land cover map. Additionally, NatureServe made extensive edits on parts of the data in an effort to improve the accuracy and usefulness of the data for their purposes. The thematic classes are ecological systems (Comer et al. 2003) and are attributed with the upper levels of the 2008 National Vegetation Classification System (Class, Subclass, Formation, Division). An additional level ("macrogroup") between the 44 Divisions and 512 ecological systems is being developed by NatureServe. As a composite product, these data reflect the same spatial and temporal scale, as well as inherent errors and assumptions, as do the original source data. The data are provided at 30 m (0.09 ha.) spatial resolution. This landcover set is available through the Landscope web site (www.landscope.org).

NPS. The NPS vegetation mapping program is one of the first of several inventories initiated by the NPS Inventory & Monitoring program. These efforts are typically initiated by detailed interpretation of high resolution aerial photography supplemented by extensive field sampling in order to define the vegetation communities and assess the accuracy of the final products. These projects use the National Vegetation Classification System as the basis of the map classifications. (needs similar level of detail to other sources- what is mmu, level of classification, source and vintage of imagery).

Land Cover Comparison

All land cover data sources were cross-walked to a comparable classification (Anderson Level I) prior to analysis. We used ArcGIS (9.2) to estimate and summarize the number of cover types, area of each type, and percent cover at the extent of park boundary, the park plus a 3 km buffer, and the park plus a 30 km buffer. The results for the park unit and 30 km buffer were used for this comparison (table 2). In addition, thematic maps were created for each source covering the park and the buffer areas. These maps provided a graphic/visual comparison of the different sources of land cover data (figure 3). By comparing these thematic maps, we could identify the types and locations of differences in classifications as well as classes that are consistently mapped for a particular area.

For the parks that had I&M Vegetation Map products, we rasterized those products into 30 m cells and cross-walked the I&M classification into the same Anderson Level 1 as the four national data sets. Using ERDAS Imagine (9.3), we created contingency tables that directly compared (on a pixel by pixel basis) the I&M classification with each of the four national data sets. The I&M data were also summarized by calculating the number of acres and percent of each cover type in the park. We compared the areas of each cover type in each NPS vegetation map to the four national data sets.

Results

We intended to identify relationships between these land cover data sets based upon consistent observations across the Park study areas. For instance, one might expect a particular data soruce to consistently indicate more forest than other data sets due to a focus on forests by the data set developers.

Unfortunately, comparing these land cover data was a complex process. We encountered difficulties to directly compare land cover classes because of the different classification systems used, differences in the focus of the developers, the different regional development schemes, and the variety of contractors and cooperators with different techniques for mapping.

As an example, the LANDFIRE, and GAP data often use the NatureServe Ecological Systems classification. This system is not a strict hierarchical system and sometimes combines dissimilar vegetation types into one class. A good example is the series of classes that fall under “Mixed Upland and Wetland” classification. This is a combination of the fundamentally different, very broad Uplands and Wetlands classes.. These classes can contain Forests, Shrublands, Grasslands, Floodplains and Riparian systems. Because they contain so many dissimilar elements, we were unable to compare them to the Anderson Level 1 classes in a one to one relationship.

Recently harvested forest around Crater Lake National Park in Oregon serve as a good example of different interpretations of land cover. These areas no longer support a dense population of tree species but are more likely to contain disturbed ground, sparse vegetation, herbaceous cover or shrubs and small trees (depending on the age of the cut). These areas were typically classified as Shrub/Scrub by NLCD, Barren/Transitional by GAP and NatureServe, and as Forest by LANDFIRE. This might well reflect the different focus of the different programs. NLCD is

classifying the area as the vegetation exists at the time of the mapping, GAP and NatureServe data are indicating that the area is in a transitional state while the focus of LANDFIRE in this area is towards fire fuels associated with forest cover. These interpretations are not correct or incorrect, but reflect the specific application(s) that were the focus of the development teams. Nonetheless, the interpretations have a strong influence on the potential use of the data.

We made these observations from our comparison of the four national land cover data sets. They are based on comparisons of areas within parks and a 30 km buffer around the park boundary, Our our recommendation are based on these results (table 3).

Overall,

- For several of the data sets (GAP, LANDFIRE, NatureServe), the classification tends to be fairly consistent within the mapping regions (or state boundaries) but not as consistent between mapping regions. This results in the potential lack of consistency across the conterminous U.S. and makes broad 'one-size-fits-all' conclusions difficult. Conversely, the focus for NLCD is for a consistent classification nationwide.
- For NLCD, Calculated Impervious Surface is used along with masks of buffered roads, census data and NOAA night sky data to derive "Built Up" classes. This has the effect of showing more roads. (Colin Homer, personal communication). Our comparisons validated this. Visually, the roads within the NLCD data appear to be more consistent.

Wetlands,

- National Wetland Inventory (NWI) data were used (when available) to train the "Wetlands" classes for the NLCD. But, wetland classes were not "aggressively" mapped. Therefore, NLCD typically has much less "Wetlands" classes than the others. (Colin Homer, personal communication). This was validated by our comparison. For five of seven parks that were compared, NLCD classified less Wetland area than all of the others. In many cases, much less. For example, NLCD data indicates from 7% to 81% less area of Wetlands than the source with the next highest estimates for Wetlands. The source with the largest total area of Wetlands indicated 46% to 94% more Wetlands than NLCD.
- NatureServe data tend to estimate more area of Wetland class than the other three in four of the seven parks. In one of the exceptions, the NatureServe, GAP and LANDFIRE data were all very close in area estimates (within 6 hectares). NatureServe actually uses a flow accumulation model to help depict riparian and wetland areas. This is the likely reason that the NatureServe data indicate more wetland area (Pat Comer, personal communication). Visually, the wetland class in many of the NatureServe examples resembles a map of stream channels.
- When the I&M data were compared to the four land cover sets for the Wetland class, the GAP data tended to match the area estimates best for most of the parks with the NatureServe data having the closest estimates in one of the parks.

Forests,

- LANDFIRE data typically indicate more Forest area than the other three. For our comparison, that was true for six of seven park areas. The exception was at THRO where NLCD indicated more forest and LANDFIRE indicated more Grassland than the others. Possibly, LANDFIRE tends to focus on the predominant fire fuel in an area resulting in an estimate of that land cover (fuel) type that is greater than the other sources. LANDFIRE is more focused on fire fuel types than proper species identification and should not be used for more rare types. GAP would likely be more appropriate for species level information. (Jim Voglemann MRLC meeting March 2009)
- When the I&M data were compared to the four land cover sets for the Forest class, the GAP data tended to provide the best match for area estimates.

Transitional,

- The GAP data estimated more Barren/Transitional area than the other three in four of seven cases. In one of the exceptions, the GAP data were within 25 hectares of the highest estimate. In the other exception (BUFF) the state based GAP data indicated no Barren/Transitional area.
- When the I&M data were compared to the four land cover sets for the Barren/Transitional class, there were no close matches. All four of the national land cover data sets estimated more (sometimes much more) area of Barren/Transitional than the NPS I&M data. This may well be related to the different base image and analysis techniques used in these programs. The I&M effort relied upon high resolution aerial photography interpreted by specialized analysts and supported by extensive field data. Some of the area that was spectrally classified as Barren could have been classified as sparse vegetation by a photo interpreter.

Water

- All four land cover sets consistently mapped the Water class. Due to the spectral uniqueness of water, this should not be a surprise. When compared to the I&M data, all four national sets typically indicated less water than the I&M maps. This is possibly due to the difference in resolution and scale of the products. Interpretation of high resolution photos allows the analyst to map more narrow streams and detailed shorelines.

Grassland/Shrub

- Two of the level 1 classes, Grassland and Shrub/Scrub were typically inconsistently mapped by these four efforts. NLCD designated more grasslands in four of the test parks, Gap showed more in two of the parks while LANDFIRE indicated more grasslands in one (THRO). For each park, the source that indicated the most grassland area typically indicated 1.25 to 2 times more than the next source.
- I&M data typically indicated less grassland than the four national sets.
- All four national land cover data sets were relatively consistent when mapping the Agriculture category. An exception was noted at ROMO where there seemed to be some

confusion between Agriculture, Grasslands and Wetlands. All four of the national data sets indicated some presence of Agriculture in ROMO in an area that consists of several large wet meadows.

Summary of data source accuracy assessments

NLCD

NLCD 1992 accuracy was assessed at 80.4% for Level I and 55.7% for Level II. NLCD 2001 accuracy was assessed at 84.6% for Level I and 78.3% for Level II. The accuracy is lowest in Southeast mapping zone (78%) and highest in Rocky Mountains mapping zone (90%) (Wickham, 2009).

- A small gain in accuracy when lumping from Level II to Level I (78.3% to 84.6%) indicates a significant portion of the misclassification is across rather than within the thematic classification hierarchy. (Wickham, 2009).
- At least 5% of total map error is confusion between upland and wetland classes (regions with lowest Level II accuracy -7 & 9 – are wet) (Wickham, 2009).
- Overall there is a high positive correlation between the relative abundance of a class and accuracy. (Wickham, 2009).
- For the NLCD 2001 data "A significant portion of error was due to difficulty in distinguishing the context of grass; that is confusion between 21, 71, 81" (Open urban, Grassland, Agriculture (pasture)). (Wickham, 2009).
- Accuracies for urban classes are not as good as the map "feels." Accuracies generally increase as percentage impervious increases. Open urban (21) "looks" like its surroundings. **(**Wickham, 2009**).**
- Preliminary results from the NLCD 2001 Alaska efforts indicate an overall accuracy of 72.8% for the northern half of the state. The southern half will be assessed in 2009. (Selkowitz, 2009).

LANDFIRE

- LANDFIRE data were evaluated for accuracy in the west only (at this time). Overall accuracy values for the Ecological Systems range from a low of 32% in the Northern Rockies mapping zone to a high of 50% for the Great Basin mapping zone.
- The accuracies of the LANDFIRE data used in this comparison were found to be: 38.4% for zone 28 which includes ROMO, 75% for zone 16 which includes ZION, 41.4% for zone 6 which includes SEKI and 35% for zone 7 which includes CRLA. LANDFIRE accuracy measures were not available yet for the remaining study areas.
- Many of the state based GAP efforts have also been assessed for accuracy. Users can view these results in the final project reports. The Arkansas GAP data indicate an overall accuracy of 92% at level 1 dropping to 69% at level 2.

Recommendations:

Whatever national land cover data set a park chooses to use for its applications, we strongly suggest that the analyst follow the suggestions for appropriate use offered by the developers of those data (Appendix 1). "For most uses, it is unlikely that GAP will provide the only data needed, and for uses with a regulatory outcome, field surveys should verify the result. In the end, it will be the responsibility of each data user to determine if GAP data can answer the question being asked, and if they are the best tool to answer that question."(Appropriate and Inappropriate Use of GAP Data, appendix A). And… "Managers and planners must evaluate LANDFIRE data according to the scale and requirements specific to their needs (for example, habitat requirements for the species being considered or requirements by community leaders and interagency partners). LANDFIRE products are not intended to replace local products, but rather serve as a back-up by providing wall-to-wall cross-boundary products. It is the responsibility of the user to be familiar with the value, assumptions, and limitations of LANDFIRE products." (Scale and Use of LANDFIRE Products, appendix A).

To compare land cover features over time and identify change, the logical choice for a data source is the National Land Cover Data (NLCD). Of the four land cover data we compared, the NLCD is the only one that has been repeated. NLCD data are available from a 1992 vintage, 2001 vintage and 2006 is schedule for completion in 2010. Furthermore, the 1992-2001 change product is already available at Anderson Level 1, and this is appropriate for many applications. In the future, LANDFIRE plans to update its product every two years. When these new LANDFIRE products are available, they should be evaluated and compared with the NLCD change product.

To evaluate more detailed information related to ecological systems or communities, and where the area of analysis is within a single Re-GAP zone, we suggest using Re-GAP data. The reasonable agreement between I&M vegetation maps and Re-GAP supports the use of Re-GAP data to extend vegetation a map beyond its geographical extent. If the analysis area extends beyond a single Re-GAP zone, the recommended source for data is the NLCD.

For applications related to wildland fire, we recommend use of the LANDFIRE data since that is the stated focus of those data.

The NatureServe LandScope land cover data may prove to be a valuable resource since it seems to take the "best" of the LANDFIRE and Re-GAP data and further refine this mosaic. At this point in time, however, the specific edits that were used to create the LandScope data are not known. Also, these data have not been assessed for accuracy at this time. As a result, the maps are not reproducible and there may be large and undeterminable inconsistencies in mapping accuracy that affect use of LandScope data for specific purposes.

Our recommendations for uses of land cover data for specific applications is summarized in Table 3..

None of these national land cover programs are static. They continue to evolve and improve their methods and products. It will be important to stay up to date on these improvements as they may further influence these recommendations.

Acknowledgements

Special thanks to Collin Homer (USGS), Jim Wickham (EPA), Pat Comer (NatureServe) and Jim Voglemann (USGS) for personal insight, review and comments.

Literature Cited

Coan, M., Fry, J., Homer, C., Meyer, D.K., Larson, C. & Wickham, J. Completion of the National Land Cover Database 1992/2001 Change Product. Online at http://www.mrlc.gov

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague. 2003. *Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems*. NatureServe, Arlington, Virginia.

Gap Analysis Program (GAP) (2007). Retrieved August 4, 2007 from: http://gapanalysis.nbii.gov/

Hollister, Jeffrey W., M. Lillana Gonzalez, John F. Paul, Peter V. August, and Jane L. Copeland. 2004, Assessing the Accuracy of National Land Cover Dataset Area Estimates at Multiple Spatial Extents. *Photogrammetric Engineering and Remote Sensing*, Vol 70, No. 4, April 2004, pp.405-414.

Homer, Collin, Chengquan Huang, Limin Yang, Bruce Wylie, and Michael Coan, 2004, Development of a 2001 National Land-Cover Database for the United States. *Photogrametic Engineering and Remote Sensing* Vol70, No. 7, July 2004, pp829-840.

Riitters KH, Wickham JD, Wade TG (2006) Evaluating ecoregions for sampling and mapping land-cover patterns. *Photogrammetric Engineering and Remote Sensing* 72:781–788

Selkowitz, David, 2009, NLCD Alaska Accuracy Assessment, Update Strategy, and Other Products, Presentation at the March 2009 MRLC Annual Partners Meeting.

Stehman, S., Wickham, J., Smith, J. Yang, L, 2003. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. *Remote Sensing of Environment* 86, 500-51

Vogelmann, J.E, S.M. Howard, L.Yang, C.R. Larson, B.K. Wylie, and J.N. Van Driel, 2001. Completion of the 1990's National Land Cover Data Set for the conterminous United States, *Photogrammetric Engineering and Remote Sensing* 67:650-662.

Vogelmann, James, 2009, LANDFIRE and GAP: 2008-9 Highlights and Future Plans, Presentation at the March 2009 MRLC Annual Partners Meeting.

Wardlow, Brian and Stephen L. Egbert. 2003 A State-Level Comparative Analysis of the GAP and NLCD Land-Cover Data Sets. *Photogrammetric Engineering and Remote Sensing* Vol 69, No. 12, December 2003, pp. 1387-1397.

Wickham, J.D., S.V. Stehman, J.H. Smith, L. Yang. 2004. Thematic Accuracy of the 1992 National Land-Cover Data for the Western United States. *Remote Sensing of Environment*, Vol. 91, pp. 452-468.

Wickham, James. 2009, NLCD 2001 Accuracy Assessment, Presentation at the March 2009 MRLC Annual Partners Meeting.

Table 1. Developmental characteristics of the four national land cover data sets assessed plus the NPS I&M Vegetation Map.

	National Land Cover Data (NLCD)	**LANDFIRE**	**GAP / ReGAP**	**NatureServe LandScope**	**NPS Vegetation Map**
Agency	USGS, MRLC	USGS, USFS, DOI	USGS	NatureServe	NPS
Extent	Conterminous US	Conterminous US	State / Region	Conterminous US	Park
Source Imagery and Dates	Landsat, 2001 and mid-1990s	Landsat 2001	Varies, most Landsat late 1990s-2002	Composite of LANDFIRE and ReGAP	Varies
Thematic Resolution	21 classes (Anderson Level II)	400-450 classes (ecological systems)	Hundreds (NVCS alliances or ecological systems)	512 classes (ecological systems)	Varies, target is for NVCS associations
Spatial Resolution, Minimum Mapping Unit	30m, 0.45 ha	30m,	30m, MMU varies	30m, dependent on source data	0.5 ha
Temporal Change	Yes, using the 1992/2001 Retrofit Product	No	No	No	No
Overall Objective	Regional to national level assessment focused on land cover and land use (urban and agriculture)	Regional to national level assessment focused on wildland fuel and fire regimes	State to regional level assessment focused on biodiversity and conservation	Regional to national level assessment focused on ecological systems	Park level assessment focused on vegetation communities
Accuracy Assessment Complete?	Yes, by mapping zone	Partially, western states done by mapping zone	Yes, by mapping zone or state	No	Yes

Figure 1. Location of sample parks used in the comparison

Table 2. Example of Class Summaries for ROMO

ROMO + 30 Km Land Cover Comparison

LEVEL I		NLCD 2001		LANDFIRE		ReGAP		NatureServe	
Class #	Class Name	Area (Ha.)	Percent	Area (Ha.)	Percent	Area (Ha.)	Percent	Area (Ha.)	Percent
11	Open Water	7,119	0.91	9,311	1.19	7,780	1.01	7,662	1.02
12	Perennial Ice/Snow	34,806	4.45	35,834	4.58	173	0.02	173	0.02
2	Developed	12,136	1.55	13,865	1.77	3,684	0.48	3,580	0.48
3	Barren/Quarries/Transitional	27,537	3.52	29,546	3.78	48,960	6.34	48,783	6.48
4	Forest	509,209	65.07	555,403	70.98	548,608	71.08	529,816	70.34
5	Shrub/Scrub	87,885	11.23	75,078	9.59	81,829	10.60	80,277	10.66
7	Grassland/Herbaceous	85,330	10.90	25,194	3.22	47,143	6.11	47,109	6.25
8	Agriculture	16,905	2.16	35,037	4.48	27,538	3.57	27,092	3.60
9	Wetlands	1,604	0.21	3,214	0.41	6,078	0.79	8,694	1.15
	TOTALS	**782,532**	**100.00**	**782,481**	**100.00**	**771,793**	**100.00**	**753,186**	**100.00**

Figure 2. Comparison of NLCD, NatureServe, LANDFIRE, and GAP land cover data in a portion of ROMO.

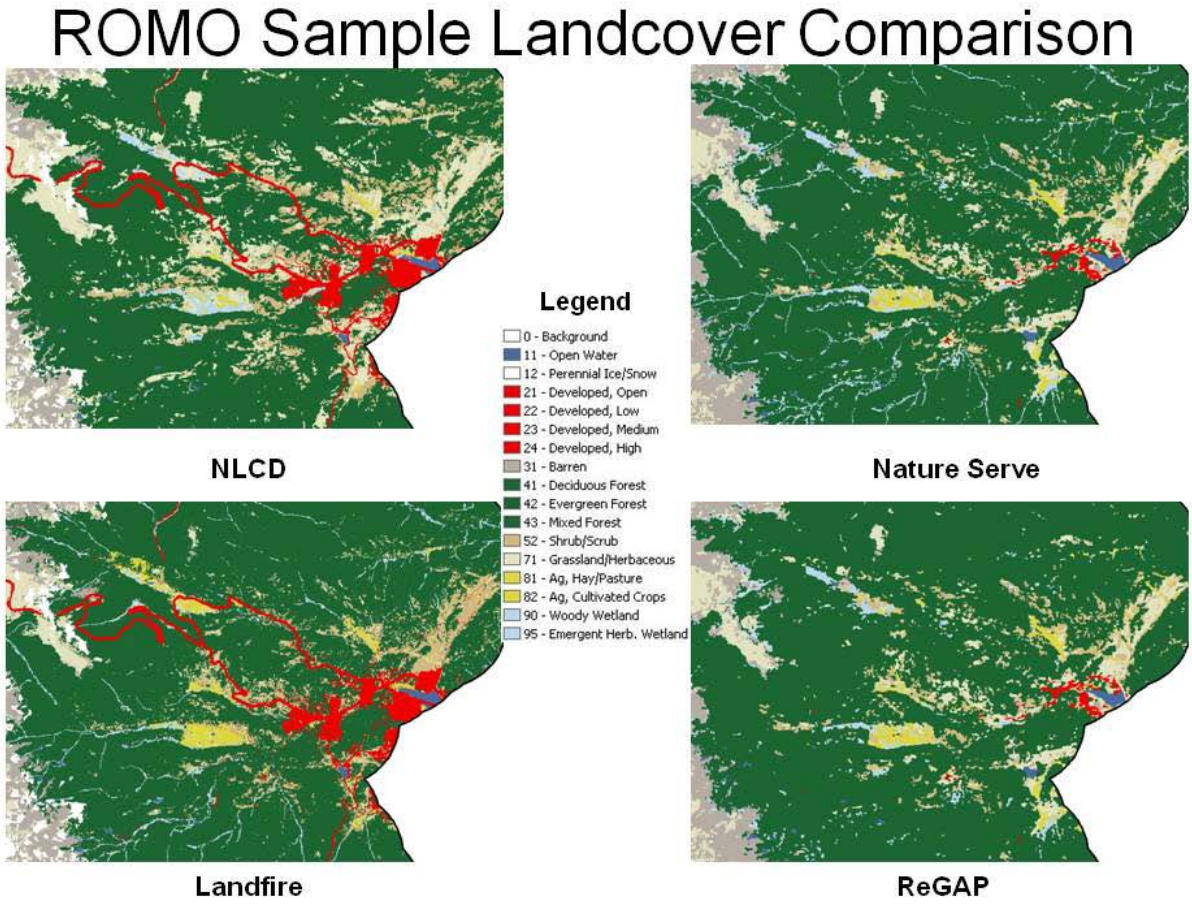

Table 3. Application Matrix

Potential Applications for National-extent Land Cover Products

Where 1 = Best Choice, 2 = Second Best Choice, 3 = Third Best Choice, and NA = Not Appropriate

Applications/Questions	NLCD	LANDFIRE	NatureServe	ReGAP/GAP	Your Local Veg Map(s)
Coarse Spatial and Thematic Scale					
Assessing change in land cover types (e.g., forest to agriculture)	**1**	**NA**	**NA**	**NA**	**NA**
Assessing fire fuel potential	**3**	**1**	**2**	**2**	**NA**
Assessing broadscale wildlife habitat needs	**3**	**2**	**1**	**1**	**NA**
Assessing distribution of invasive species (e.g., annual vs perennial grasses)	**NA**	**1**	**2**	**2**	**NA**
Assessing distribution in broad land cover types (e.g., forest vs nonforest, natural vs converted)	**1**	**2**	**2**	**2**	**NA**
Assessing pattern in broad land cover types (e.g., forest vs nonforest, natural vs converted)	**1**	**2**	**2**	**2**	**NA**
	NLCD	**LANDFIRE**	**NatureServe**	**ReGAP/GAP**	**Your Local Veg Map(s)**
Moderate Spatial and Thematic Scale					
Assessing change in land cover types (e.g., woodland to forest)	**1**	**NA**	**NA**	**NA**	**?**
Assessing fire fuel potential	**3**	**1**	**2**	**2**	**2**
Assessing moderate scale wildlife habitat needs	**3**	**2**	**1**	**1**	**2**
Assessing distribution of invasive species	**NA**	**2**	**2**	**2**	**1**
Assessing distribution in moderate land cover types (e.g., conifer forest vs deciduous forest)	**1**	**1**	**1**	**1**	**2**
Assessing pattern in moderate land cover types (e.g. conifer forest vs deciduous forest)	**1**	**1**	**1**	**1**	**2**
	NLCD	**LANDFIRE**	**NatureServe**	**ReGAP/GAP**	**Your Local Veg Map(s)**
Fine Spatial and Thematic Scale					
Assessing change in land cover types (e.g., ponderosa pine to douglas fir)	**NA**	**NA**	**NA**	**NA**	**?**
Assessing fire fuel potential	**NA**	**2**	**3**	**3**	**1**
Assessing fine scale wildlife habitat needs	**NA**	**3**	**2**	**2**	**1**
Assessing distribution of invasive species	**NA**	**NA**	**NA**	**NA**	**1**
Assessing distribution in fine land cover types (e.g., oak hickory forest vs poplar forest)	**NA**	**3**	**2**	**2**	**1**
Assessing pattern in fine land cover types (e.g. oak hickory forest vs poplar forest)	**NA**	**3**	**2**	**2**	**1**

Appendix A. Appropriate and Inappropriate Data Use Descriptions (from sources)
Each of these data sources (GAP, NLDC and LANDFIRE) contains information (either in the metadata or on the web sites) relating to the suggested use of the data. These instructions are included in this appendix and the potential users of these data are encouraged to read these warnings. No such information was provided by LandScope.

GAP Data
(http://gapanalysis.nbii.gov/portal/server.pt/gateway/PTARGS_0_2_16708_1483_5122_43/http%3B/public-content%3B7087/publishedcontent/publish/gap/public_sections/maps__data_and_reports/appropriate_uses/more_content.html)

Appropriate and Inappropriate Use of GAP Data

All information is created with a specific end use or uses in mind. This is especially true for GIS data, which is expensive to produce and must be directed to meet the immediate program needs. For GAP, minimum standards were set (see Standards Chapter in the GAP Handbook) to meet program objectives. These standards include: scale or resolution, accuracy, and format. The standards are constantly evolving, and therefore projects begun previous to the date of the current standards may not meet those standards.

Recognizing, however, that GAP would be the first, and for many years likely the only, source of statewide biological GIS maps, the data were created with the expectation that they would be used for other applications. Therefore, we list below both appropriate and inappropriate uses. This list is in no way exhaustive but should serve as a guide to assess whether a proposed use can or cannot be supported by GAP data. For most uses, it is unlikely that GAP will provide the only data needed, and for uses with a regulatory outcome, field surveys should verify the result. In the end, it will be the responsibility of each data user to determine if GAP data can answer the question being asked, and if they are the best tool to answer that question.

Scale
First we must address the issue of appropriate scale to which these data may be applied. The data were produced with an intended application at the ecoregion level, that is, geographic areas from several hundred thousand to millions of hectares in size. The data provide a coarse-filter approach to analysis, meaning that not every occurrence of every plant community or animal habitat is mapped, only larger, more generalized distributions. The data are also based on the USGS 1:100,000 scale of mapping in both detail and precision. When determining whether to apply GAP data to a particular use, there are two primary questions: do you want to use the data as a map for the particular geographic area, or do you wish to use the data to provide context for a particular area? The distinction can be made with the following example: You could use GAP land cover to determine the approximate amount of oak woodland occurring in a county, or you could map oak woodland with aerial photography to determine the exact amount. You then could use GAP data to determine the approximate percentage of all oak woodland in the region or state that occurs in the county, and thus gain a sense of how important the county's distribution is to maintaining that plant community.

Appropriate Uses
The above example illustrates two appropriate uses of the data; as a coarse map for a large area such as a county, and to provide context for finer-level maps. Following is a general list of applications:

- Statewide biodiversity planning
- Regional (Councils of Government or ecoregional) planning
- Regional habitat conservation planning
- County comprehensive planning
- Large-area resource management planning
- Coarse-filter evaluation of potential impacts or benefits of major projects or plan initiatives on biodiversity, such as utility or transportation corridors, wilderness proposals, regional open space and recreation proposals, etc.
- Determining relative amounts of management responsibility for specific biological resources among land stewards to facilitate cooperative management and planning
- Basic research on regional distributions of plants and animals and to help target both specific species and geographic areas for needed research
- Environmental impact assessment (EIS) for large projects or military activities
- Estimation of potential economic impacts from loss of biological resource based activities
- Education at all levels and for both students and citizens

Inappropriate Uses
It is far easier to identify appropriate uses than inappropriate ones, however, there is a "fuzzy line" that is eventually crossed when the differences in resolution of the data, size of geographic area being analyzed, and precision of the answer required for the question are no longer compatible. Examples include:

- Use of the data as a "content" map for small areas (less than thousands of hectares), typically requiring mapping resolution at 1:24,000 scale and using aerial photographs or ground surveys.
- Combining GAP data with other data finer than 1:100,000 scale to produce new hybrid maps or answer queries resulting in precise measurements.
- Generating specific areal measurements from the data finer than the nearest thousand hectares (minimum mapping unit size and accuracy affect this precision).
- Establishing exact boundaries for regulation or acquisition.
- Establishing definite occurrence or nonoccurrence of any feature for an exact geographic area (for land cover, the percent accuracy will provide a measure of probability).
- Determining abundance, health, or condition of any feature.
- Establishing a measure of accuracy of any other data by comparison with GAP data.
- Altering the data in any way and redistributing them as a GAP data product.
- Using the data without acquiring and reviewing the metadata and this report.

NLCD (taken from the NLCD 1992 metadata)

While we believe that the approach taken has yielded a very good general land cover classification product for a large region, it is important to indicate to the user where there might be some potential problems. The biggest concerns are listed below:

1) Some of the TM data sets are not temporally ideal. Leaves-off data sets are heavily relied upon for discriminating between hay/pasture and row crop, and also for discriminating between forest classes. The success of discriminating between these classes using leaves-off data sets hinges on the time of data acquisition. When hay/pasture areas are non-green, they are not easily distinguishable from other agricultural areas using remotely sensed data. However, there is a temporal window during which hay and pasture areas green up before most other vegetation (excluding evergreens, which have different spectral properties); during this window these areas are easily distinguishable from other crop areas. The discrimination between hay/pasture and deciduous forest is likewise optimized by selecting data in a temporal window where deciduous vegetation has yet to leaf out. It is difficult to acquire a single-date of imagery (leaves-on or leaves-off) that adequately differentiates between both deciduous/hay and pasture and hay-pasture/row crop.

2) The data sets used cover a range of years (see data sources), and changes that have taken place across the landscape over the time period may not have been captured. While this is not viewed as a major problem for most classes, it is possible that some land cover features change more rapidly than might be expected (e.g. hay one year, row crop the next).

3) Wetlands classes are extremely difficult to extract from Landsat TM spectral information alone. The use of ancillary information such as National Wetlands Inventory (NWI) data is highly desirable. We relied on GAP, LUDA, or proximity to streams and rivers as well as spectral data to delineate wetlands in areas without NWI data.

4) Separation of natural grass and shrub is problematic. Areas observed on the ground to be shrub or grass are not always distinguishable spectrally. Likewise, there was often disagreement between LUDA and GAP on these classes.

LANDFIRE Products (http://www.landfire.gov/documents/Scale_and_Use_of_LF_Data.pdf)

Scale and Use of LANDFIRE Products

LANDFIRE data products facilitate national- and regional-level strategic planning and reporting of wildland fire management activities. LANDFIRE products are designed to be used at a landscape-scale in support of strategic vegetation, fire, and fuels management planning to evaluate management alternatives across boundaries.

Although LANDFIRE National products are delivered as 30-meter pixels, they should not be used at the individual pixel level or on small groups of pixels. LANDFIRE National products were designed to support 1) national (all states) strategic planning, 2) regional (single large states or groups of smaller states), and 3) strategic/ tactical planning for large sub-regional landscapes and Fire Management Units (FMUs) (such as significant portions of states or multiple federal administrative entities). The applicability of LANDFIRE National products to support fire and land management planning on smaller areas will vary by product, location, and specific use. Further investigation by local and regional experts should be conducted to inform decisions regarding local applicability. However, it is the responsibility of the local user, using LANDFIRE metadata and local knowledge, to determine if and/or how LANDFIRE can be used for particular areas of interest.

Managers and planners must evaluate LANDFIRE data according to the scale and requirements specific to their needs (for example, habitat requirements for the species being considered or requirements by community leaders and interagency partners). LANDFIRE products are not intended to replace local products, but rather serve as a back-up by providing wall-to-wall cross-boundary products. It is the responsibility of the user to be familiar with the value, assumptions, and limitations of LANDFIRE products.