A Landsat 7 Scene Selection Strategy for a National Land Cover Database*

Limin Yang, Collin Homer, Kent Hegge, Chengquan Huang, Bruce Wylie and Bradley Reed
Raytheon ITSS, EROS Data Center, Sioux Falls, SD 57198 USA

Abstract- A strategy for selecting Landsat 7 ETM+ imagery for development of new generation national land cover database of the United States has been developed. This strategy is formulated to target Landsat 7 ETM+ scenes based on land cover and land use, vegetation phenology and image quality (cloudiness, haze).

Criteria based on phenology and scene quality provide a national baseline for acquiring Landsat 7 data. Optimal time periods for discriminating land cover types were identified for each Landsat 7 path-row footprint and each proposed land cover mapping zone (mosaic of several path-rows based on landscape and ecoregion), from which three Landsat scenes were selected. This database of selected scenes is used to guide Landsat 7 data purchasing. This methodology provides a consistent framework for populating Landsat 7 imagery to be used for a new national land cover characterization initiative.

I. Introduction

The Multi-Resolution Land Characteristics 2000 (MRLC 2000) consortium consists of a partnership among several U.S. federal agencies. The major driver of the consortium is the common need for accurate land cover information across variety of spatial and temporal scales [1]. One new initiative of MRLC 2000 is to develop a national land cover database (NLCD 2000) using 2000 vintage Landsat 7 data. Hence, a critical component of the initiative is acquisition of multi-temporal Landsat 7 ETM + imagery. A strategy was needed for nation-wide imagery selection that meets the requirements of the land cover database specifications, and also provides flexibility for other applications. This approach established a scene selection strategy based on vegetation dynamics of target land cover types over a growing season. It assumes that a distinct seasonal trajectory of land cover dynamics can be identified using multi-temporal remote sensing data, and that this information will provide increased land cover identification capability. Scene selection criteria are established using multi-temporal greenness as a surrogate for vegetation phenology for any given spatial unit [2] (e.g. path-row, mapping zones).

This strategy for selecting Landsat 7 imagery based on vegetation phenology and image quality provide a relatively objective method to populate a nation-wide image database in a consistent way.

II. Vegetation Phenology

Information on vegetation phenology was derived from the multi-temporal normalized difference vegetation index (NDVI) data of the conterminous U.S. acquired by the advanced very high resolution radiometer (AVHRR) from 1994-1998. The AVHRR data were radiometrically and geometrically calibrated and then smoothed using a weighted multiple regression technique to further reduce cloud and view angle effect [3]. Metrics of vegetation phenology derived from the smoothed NDVI time series includes magnitude and date of mean NDVI for each biweek period throughout a calendar year as well as those at time of onset-, peak-greenness, and senescence [4], (Fig. 1). All seasonal metrics were computed for each land cover type within each Landsat path-row or mapping zone stratified by the 1992 national land cover data (NLCD 1992) developed by the original MRLC consortium. This results in a database that contains, for each mapping zone (or path/row), mean NDVI time series of each land cover type, value and date of NDVI at onset, peak and senescence for each land cover type, and the total NDVI difference among land cover types. The total difference is defined as:

$$NDVI_{diff} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left| NDVI_i - NDVI_j \right| \qquad (1)$$

where n represents the total number of land cover types within each mapping zone.

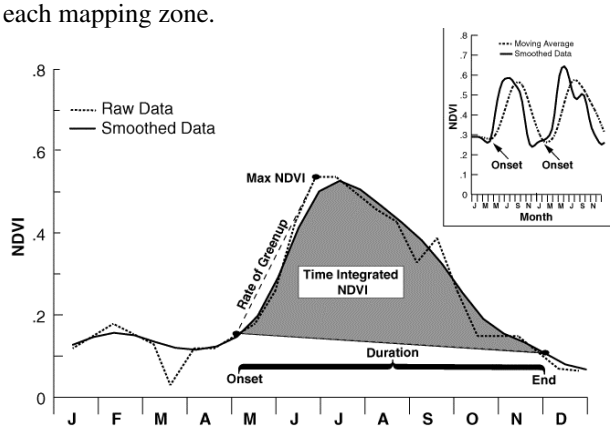

Fig.1. Seasonal characteristics derived from multi-temporal NDVI (from [4]).

III. Scene Selection Criteria Based on Phenology

Recognizing the need for development of a nationally consistent land cover database, it was decided that three Landsat 7 scenes, representing different seasons, would be selected per path-row for discriminating land cover types. The three scenes selected should capture critical stages of vegetation growth such as magnitude, amplitude, and timing of temporal NDVI greenness. Ideally, for areas with a single-peak greenness, the first scene should always be chosen at or near time of the peak-

* Work was performed under U.S. Geological Survey contract 1434-CR-97-CN-40274. This paper is preliminary and has not been edited or reviewed for conformity with U.S. Geological Survey standards or nomenclature.

greenness (maximum NDVI) of the dominant land cover class (es) with second and third scenes selected from pre- and post-peak in the leaf-off season. Hence, the first scene should capture the most rigorous and productive stage of vegetation growth, whereas the second and third scenes capture other vegetative stages and changes over the rest of the growing season. Specifically, three criteria based on NDVI time series were set up as general guidelines for scene selection:

1. Time of peak-greenness is defined as the biweek period when maximum NDVI occurs for the dominant land cover type(s). Once the peak-greenness is determined, persistence of peak-greenness is defined as a time window within which change of NDVI is less than 0.02/biweek prior or after the biweek of maximum NDVI occurrence. This total length (biweek periods of peak and persistence) generally defines the time window for leaf-on scene selection.
2. A second scene should be selected outside the time window of peak greenness as is defined by step 1, and be outside the dormant season defined by when NDVI is below 0.15 (bare soil, snow, clouds and sun angle/illumination effect).
3. The third scene should be selected from the remainder of the growing season based on separability among major land cover classes or separation of land cover classes not well distinguished using the first and second scenes.

Another factor affecting scene selection is the modality of a land cover type observed within the growing season. Modality often indicates the periodicity of the vegetation production. For any given area, there can be either a single or multiple peak greenness depending on land cover/use and geographic location. Multi-modality occurs in areas with double cropping, or with vegetation highly responsive to bi-modal temperature and/or precipitation regime, or with diverse land cover types. For a phenology-based scene selection approach, modality of NDVI greenness within a path/row or mapping zone needs to be identified first to understand if there is: 1) a single peak greenness (referred to hereafter as case A), 2) multi-peak greenness of a single land cover type (referred to as case B), and 3) multi-peak greenness related to several land cover types (referred to as case C). Based on modality and other phenological parameters, we further implemented the following procedures for scene selection.

IV. Scene Selection Procedures

Case A. single-peak greenness

1. For selecting the first scene, identify a biweek period of maximum NDVI greenness for a dominant land cover type. The optimal time window for scene selection covers the biweekly period of peak-greenness plus two adjacent biweekly periods that meets persistence of peak-greenness definition.
2. A second scene will be selected from non-peak growing season at the time period when difference of greenness among major land cover types is distinct (maximum $NDVI_{diff}$). The period should also be outside the dormant season (i.e., NDVI is less than 0.15).
3. The third scene will be selected from the remainder of the growing season based on maximum separability of temporal NDVI greenness of all major land cover types.

***Case B. Bi-modal (two-peak) greenness of a single land cover type*:**

1. Select the first and second scene from the two time periods of peak greenness (excluding dormant season). The step is similar to step 1 in case A.
2. Select the third scene from leaf-off season when maximum difference in NDVI greenness among land cover types occurs. Similar to step 2 in case A.

Case C. Peak greenness at different times for different land covers

1. select the first scene at a time when the most dominant land cover type reaches its peak greenness. Similar to step 1 in case A and B.
2. select the second scene at a time when the second dominant land cover type reaches peak greenness.
3. select the third scene from leaf-off season based on maximum greenness difference among all cover types or separation of land cover classes not well distinguished from the first and second scenes.

V. Examples of Scene Selection

Selection based on phenology only

Three mapping zones are selected to illustrate how scene selection is implemented. Mapping zone 1 (southern Idaho and northern Nevada) is an area predominantly of shrubland, grassland, and agricultural land with a single peak-greenness (case A). Mapping zone 2 (Arizona and New Mexico Mountains) is mostly covered by shrubland, grassland and forest, with two peaks of greenness (case B). Mapping zone 3 (eastern Kansas and Nebraska) has the dominant land cover of cropland, hay/pasture, grassland, and some deciduous forest, with two peak of greenness occurring at different time associated with different land cover types (case C).

Mapping zone 1 (a single peak greenness, Fig. 2)

First scene (leaf-on): Maximum NDVI of dominant land cover (shrubland) occurred at 13th biweek period (June 19 to July 2). The persistence of peak greenness lasted from the 12th to 14th biweek (June 5 to July 16). The best time to select the first image is the 13th to 14th period (June 19 to July 16) based on greenness difference among major land cover types.

Second scene (leaf-off): look for period with NDVI > 0.15 and maximum difference in NDVI among major cover types. The 19th to 20th biweek period (Sep. 12 to Oct. 9) was selected.

Third scene, (leaf-off): time period with NDVI > 0.15 and distinct difference in NDVI among land cover types in the reminder of the growing season. The 8th to 9th biweek period (Apr. 9 to May 7) was selected.

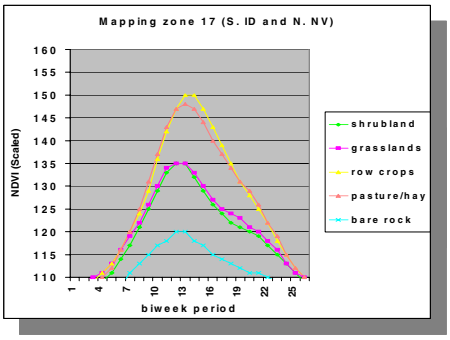

Fig. 2. Temporal NDVI greenness of mapping zone 1.

Mapping zone 2 (case B, two peaks of greenness related to a single land cover, Fig. 3)

First scene (leaf-on): Peak greenness of dominant land cover (forest) occurred at the 13th biweek period (June 19 to July 2). The persistence of peak greenness lasts from the 12th to 14th

biweek (June 5 to July 16). The 12th and 13th biweek is the best time to select the first scene considering greenness difference among cover types.

Second scene (leaf-on): Second peak of greenness occurred at 20th and 21 St. biweek period (Sep. 26 to Oct. 23), and is the best time for second scene selection based on NDVI difference.

Third scene (leaf-off): time period with NDVI > 0.15 and maximum difference of NDVI among cover types from the remainder of the season. The 7th to 9th biweek period (March 26 to May 7) meets the criteria with 8th to 9th being the best time based on NDVI difference.

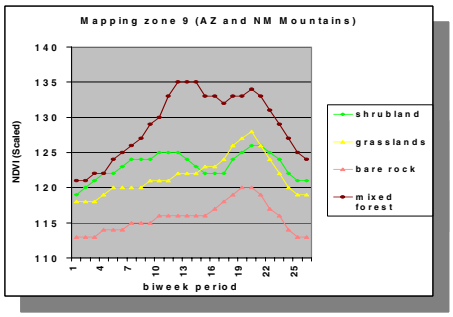

Fig. 3. Temporal NDVI greenness of mapping zone 2.

Mapping zone 3 (Case C. Peak greenness occurs at different time for different land cover types, Fig. 4)

First scene (leaf-on): The peak-greenness of dominant land cover (Grassland) occurred at the 16th biweek period (July 31 - Aug. 13) and is also the best period in terms of land cover separability. The time window can be extended based on persistence of peak greenness, ranging from the14th through 17th biweek period (July 3 to Aug. 27).

Second scene (leaf-on, second peak): the second peak of greenness associated with winter wheat occurred at the 10th biweek period (May 8 to May 21). The 9th and 10th biweek is recommended.

Third scene (leaf-off): from time period with NDVI > 0.15 and large difference of NDVI among cover types in the remainder of the growing season. The 6th to 8th biweek period (March 12 to April 23) meets the criteria. The best time is around the 7th biweek period (March 26 to April 8) based on land cover separability.

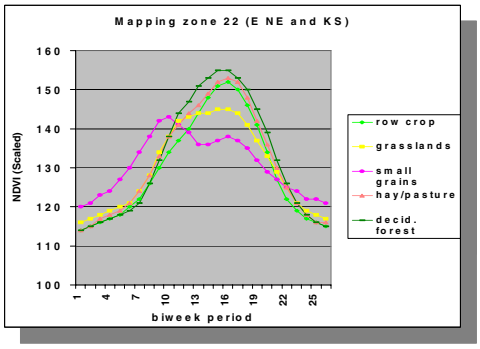

Fig. 4. Temporal NDVI greenness of mapping zone 3.

Selection based on phenology, regional landscape and scene quality

When searching scenes from the current Landsat 7 data archive, the recommended image selection time period based on vegetation phenology only needs to be constrained by scene quality (cloudiness and haze). For MRLC 2000 specification, a scene with less than 20% cloudiness may be considered as a potential candidate.

The phenology-based method presented so far assumes that each mapping zone is completely independent from any other mapping zones. In other words, it does not take into account spatial distribution and complexity of landscape in area adjacent to the mapping zone. While this assumption is reasonable when selecting scenes for a single mapping zone, it is less valid when selecting images for larger region made of several mapping zones. Therefore, for the entire U.S., we need to select scenes based on vegetation phenology, landscape features of a mapping zone and its adjacent areas, and the scene quality. Table 1 illustrates, for sixteen mapping zones, available scenes from the Landsat 7 archives up to 1 January 2001 based on phenology and scene quality.

Mapping Zone	Peak	Spring	Fall
1	8 of 11 (73%)	0 of 11 (0%)	7 of 11 (64%)
3	12 of 12 (100%)	6 of 12 (50%)	10 of 12 (83%)
5	13 of 13 (100%)	8 of 13 (62%)	11 of 13 (85%)
6	12 of 12 (100%)	7 of 12 (58%)	10 of 12 (83%)
10	8 of 14 (57%)	0 of 14 (0%)	14 of 14 (100%)
11	16 of 18 (89%)	6 of 18 (33%)	18 of 18 (100%)
15	7 of 15 (47%)	13 of 15 (87%)	14 of 15 (93%)
16	8 of 12 (67%)	5 of 12 (42%)	11 of 12 (92%)
17	10 of 12 (83%)	9 of 12 (75%)	12 of 12 (100%)
18	12 of 15 (80%)	7 of 15 (47%)	15 of 15 (100%)
19	7 of 15 (47%)	1 of 15 (7%)	12 of 15 (80%)
23	6 of 11 (55%)	3 of 11 (27%)	8 of 11 (73%)
37	5 of 21 (24%)	14 of 21 (67%)	21 of 21 (100%)
41	7 of 20 (35%)	8 of 20 (40%)	11 of 20 (55%)
60	5 of 12 (42%)	7 of 12 (58%)	10 of 12 (83%)
61	0 of 12 (0%)	8 of 12 (67%)	7 of 12 (58%)
Totals	136 of 215 (63%)	102 of 215 (47%)	191 of 215 (89%)

Table 1. Scene availability for 16 mapping zones of U.S.

VI. Conclusions

A Landsat 7 ETM+ scene selection strategy for large area land cover characterization was established. The strategy is formulated to target the best scenes available from the current Landsat 7 data archive and from the on-going data acquisition, based on land cover characteristics, vegetation phenology and image quality.

The optimal time periods for land cover characterization based on vegetation dynamics over a growing season are identified for each Landsat 7 path-row footprint and for each proposed land cover mapping zone. Criteria established based on phenology, regional landscape and scene quality (cloudiness, haze) provides general guidance to searching the Landsat 7 data archive. This methodology provides a consistent framework to select satellite imagery for the new national land cover characterization initiative. The method developed in this study can be readily adopted to aid scene selection of other remote sensing data by sensors aboard various satellites and based on user specific requirements.

REFERENCES

[1] Loveland, T.R., and Shaw, D.M., 1996, Multi-resolution land characterization--building collaborative partnerships, in Scott, J.M., Tear, T.H., and Davis, F.W., eds., GAP analysis--a landscape approach to biodiversity planning: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, p.79-85.

[2] McGwire, K, 1998, Improving Landsat scene selection systems. Photogrammetric Engineering & Remote Sensing, Vol. 64, No.7. 717-722.

[3] Swets, D.L., B.C. Reed, J.R. Rowland, S.E. Marko, 1999. A weighted least-squares approach to temporal smoothing of NDVI. In 1999 ASPRS Annual Conference, From Image to Information, Portland, Oregon, May 17-21, 1999, Proceedings: Bethesda, Maryland, American Society for Photogrammetry and Remote Sensing, CD-ROM, 1 disc.

[4] Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., and Ohlen, D.O., 1994, Variability of land cover phenology in the United States. Journal of Vegetation Science, 5, 703-714.