LCLUC- Breakout Group #2 Summary Report

Charge: Assessment of Immediate Data Needs

Chairs: Ronald Rindfuss & Mark Harmon

Our working group focused on general issues concerning data rather than specific data needs. For the most part the remote sensing needs of the group were being met, but there were problems related to linking the remotely sensed data to biophysical and social science data. The key challenge facing the LCLUC community was seen as linkage across different types of data sets. Finally, there was a consensus that the LCLUC community must be involved in the discussion of the sensors being planned and the acquisition strategies used for remote sensing data.

Our discussion was driven by three assumptions:

1. Significant progress in this arena depends on how well land cover is linked to land use. This indicates that, while improvements in remote sensing technology are welcome, significant steps need to be made to move from an assessment of land cover (provided by remote sensing) to exactly how that land cover functions biophysically and in a social context. Specifically we need to be able to link existing and planned biophysical and social data sets to the land cover data derived from satellite data. This will involve coordination between the remote sensing community and those who collect the other data sets. This will be difficult because many of the critical actors are not part of the USGCRP.

2. Carbon dynamics of ecosystems is a very important issue for the LCLUC community, however, it is part of the larger issue of use of land. The latter can be viewed very broadly to include not only human uses, but also use by plants, animals, and microbes. It is therefore essential that LCLUC not be converted into a carbon assessment program- maintaining a diversity of issues is very important.

3. An important step has been made by introducing land cover and land use into the global change discussion. In the next few years, however, we must move beyond this to a consideration of how land use and climate interact to control ecosystem function (sensu latu). Several of the briefing sessions contained examples of how land use coupled with climate dramatically changed ecosystem function. For example, clearing in tropical forests lead to increased drying, which lead to increased deforestation as fires set in pastures spread into normally fire proof forests. An example from the Middle East was the spread of barley farming onto marginal lands in years when abundant water is expected. These interactions are clearly not unidirectional; causality runs both ways. General themes that might drive the examination of land use-climate interaction might be why, when, and where these interactions occur. Another might be the degree to which current land use strategies (recall the African example) are dislocated spatially as climate change proceeds.

Problems and Issues Concerning Data Use

While collection of new data is crucial, our group found that key data may already exist, but is not being effectively used. Although this data may eventually be used by the LCLUC community, specific problems and issues are limiting their use in the short-term. A concerted effort should be made to eliminate or reduce these barriers to progress as soon as possible. One reason they may be difficult to eliminate entirely is that they often involve lapses of inter- agency, institutional, or disciplinary coordination. Given the limits on budgets, groups are going to have to learn dual use of data. Specific examples include:

1. Georeferencing of traditionally non-spatial data. There is a need to be able to match what has been traditionally viewed as non-spatial data to as specific a location as possible. Problems include confidentiality (i.e., promises made to respondents during the data collection process), matching ownership to land, and convincing data gathers that location is an important data attribute. This issue exists for social and biophysical data. An issue more specific to the social sciences is an understanding of how to match people to the land they control. This will require methodological development as the solution is dependent on the region and owner (cluster villages versus systematic dispersion, resident versus absentee owners).

2. Temporal referencing of data. This is really the temporal analog of georeferencing of data. It is not clear the degree to which this problem exists given that many human and natural resources are censused or inventoried on 10 year cycles.

3. Historical data recovery. Given that historical perspective on LCLUC is essential to interpreting the consequences of current pattern and process, more attention must be placed on recovering historical data. In particular, NASA needs to continue to focus on recovering old missions as well as planning new ones. This is probably the one area where current research needs are not being met; essential data exist at foreign ground stations, but can not currently be used because of cost, lack of equipment, or other problems.

4. Image quality and consistency. Given that the main task of LCLUC in a remote sensing context is often change detection, it is essential that as much extraneous noise as possible be removed from the data. Atmospheric corrections remain a major problem. Another problem is the lack of georeferencing of historical data. The latter may also be a problem for LANDSAT 7 as registration to earth coordinates may not have been addressed adequately. A major question that needs to be reconsidered is the degree to which it is NASA's job to make these corrections versus the user communities. Passing on these costs to the users makes a good deal of sense given that one may not always be able to predict how the data will be used. On the other hand it also introduces a lack of consistency that may ultimately limit the degree results can be compared.

5. Developing a strategy of remote sensing data collection that meets the needs of the LCLUC community. The current refresh or sampling strategy of LANDSAT 7 is based on a 4 month interval. While this may be ideal for phenological sampling, it may not be appropriate for LCLUC questions. In the latter context it may be better to have wall to wall coverage at certain specified time intervals. More frequent sampling could be done within certain regions expected to undergo high rates of change. Stratification of the sampling could be driven by other remote sensing data (e.g., use AVHRR to set up an ecosystem type screen) or non-remote sensing data on records of fire, land clearing, logging, or economic activity.

6. Development of comparable metadata. It is one thing to provide access to data. It is quite another to provide it in an understandable format. The LCLUC community needs to start thinking about this issue and take advantage of the work other communities (e.g., NSF's Long Term Ecological Research Network has developed a policy on data sharing). One challenge is to develop similar policy and metadata standards for spatial databases. Another is to compile and inventory existing data in a way that helps provide general access. The latter activity should certainly be done for each LCLUC study, but regional compilations would also be extremely valuable.

7. Software issues. The cost, ease of use, and access to algorithms can be the limiting factor in using data effectively. This is going to become more and more an issue as remote sensing passes from the domain of specialists to generalists. On the other hand, there is a real danger of providing access to the untrained (or unwashed). A more problematic issue is providing usable software to scientists in underdeveloped regions where budgets are limited. The involvement of these individuals may be critical to provide truly global coverage of LCLUC analysis.

Immediate Date Needs

Although few immediate data needs were described, there was a need to consider certain types of data in the near term. Two questions need to be considered in selecting data priorities:

1. Are the data required to ask new questions or seek more detail within a region? Or

2. Are the data required to move the results from a singular case study to a global level? The latter data could be primary data describing pattern or process, or could be secondary data that allow extrapolation of results to new regions. For example, if excellent relationships between temperature and decomposition rates exist, then it should be possible to predict the latter from a spatial database of the former.

Examples of the data needed to improve LCLUC studies are:

1. Temporally harmonized data at the global or local levels. Many census data and resource inventory data are collected periodically. It would be extremely helpful if

the period that these data were collected were synchronized and if they would match global level remote sensing products.

2. Cross-scale products. It is essential that information gathered at the finer spatial scale be incorporated into broader scales. Currently these efforts are not well connected and are proceeding on separate fronts. Two alternative strategies were discussed:
   1) synthesis of local and regional information/processes into the broader-scale
   2) provide global coverage of local or regional information or processes. These two approaches need not be mutually exclusive.

   An example of the former approach might be the development of regimes of land use and/or disturbances.

3. Data that adds a historical dimension to LCLUC analysis.There are several cases where historical perspectives may be essential to understand/or predict current processes or patterns. For example, human activity on most landscapes predates the availability of satellite data. A historical perspective may also be essential to understanding the function of certain ecosystem components. The flux of carbon from soil and detritus has a strong historical component (e.g., a 1 year-old pasture does not behave the same as a 100 year-old one). Development of a 25 year old history of land cover will depend on recovery of achieved MSS and TM data. Development of a longer-term historical record will involve a synthesis of many sources of data including maps, historical and management records, event history data (in essence, retrospectively asking about various events, a procedure widely used in the social sciences), and other data. Efficient development of the latter requires that a sensible strategy be developed (i.e., what level of detail is sufficient to answer questions) and how far back does the history extend? The latter question might be driven by an event (e.g., the onset of the industrial revolution) or the time required for the impact of an event to diminish.

4. A one time global database of forest location based on AVHRR data. This database would provide a consistent reference point for studies at many spatial scales and could serve as the basis for stratifying finer resolution sampling.

5. Wall to wall fine resolution data on land cover at a preset temporal sampling interval. This would be based on TM and or ETM. For many natural disturbances and regions with "normal" rates of land use change a 5 year interval should prove adequate. Although the group discussion centered on future data products, there is no reason this 5 year data product could not be made retroactive.

6. Location of fires to indicate location of deforestation events. Night time data may be of real use here, if the location of fixed cities and towns could be factored out. Probably would need to be TM or ETM data.

7. Climate data. While there has been great progress in synthesizing historical climate data at the global level, the resolution of this data is of marginal value to more local studies. A database that allowed access of daily data from specific climatic stations

would allow users to develop regional databases at the appropriate time and spatial scales for their analyzes.

8. Urban sites. The projects funded by NASA-LCLUC are focusing on rural areas. There is a need to include urban sites into the present mix of studies. One question raised was the degree to which the new Urban-LTER sites (Baltimore and Phoenix) are including remote sensing into their projects.

9. Vegetation Canopy Lidar. Access to VCL data needs to be prioritized for LCLUC use. It is anticipated the these data can be used to corroborate ecosystem model predictions or to initialize models themselves.

10. The group thought thermal data of all types would be important, but there was no time to develop any specifics on this issue.