Processing of Digital Data for the

Central American Ecosystem Map

Dr. Jeffrey R. Jones
Director, GIS Laboratory,
CATIE
Turrialba, Costa Rica

The processing of digital data for the Central American Ecosystem Map had two phases; an initial phase where data were collected 'in the field' by 7 country teams, and a second phase where the 7 individual country maps were integrated into a final regional map after they had been initially digitized on the basis of satellite imagery. This discussion refers only to the second phase, which was carried out in the GIS Laboratory of CATIE, in Turrialba, Costa Rica. This activity was supported by a number of individuals with the laboratory; Antonio Alcaraz, Sebastian Wesselman, Rogelio Solano and Sergio Velasquez in image rectification and image management, Jeffrey Jones and Alexander Salas in data input and editing.

The processing carried out in the CATIE GIS Laboratory was divided into four distinct activities;
1) Image importation and registration
2) Integration of country vector layers
3) Establishment of a common projection
4) Creation of the 1:250,000 map set

Image importation and registration

A set of imagery was turned over the the GIS Laboratory which came from two sources; four 60 gigabyte backup hard-drives from the WICE offices, and a set of 161 CD-ROMs. This set of images represented some 30 images to cover the Central American Isthmus, with duplicates in some cases where some images suffered from cloud cover. Early on the process, it was recognized that the information from the hard-drives would be extremely difficult to use, because the data had been transformed but not documented, so positive identification of images would in most cases be impossible. Cataloging and and image management then concentrated on the set of 161 images on CD-ROMs which as original images in most cases carried sufficient documentation to determine path, row, and date of the image.

The overall objective of activities with the imagery was the creation of mosaics based on the best available and most recent images of each country. This began with the definition of a catalogue of images, projections, and qualities. Ideally, this activity would have been based on the imagery used by each of the countries in the digitizing of their ecosystems; nevertheless, only Guatemala supplied original imagery, and this converted in 1:250K quads in 3 band tif format. As a result, all imagery was reviewed and re-registered based on information available in the GIS Laboratory; in many cases, 1:50,000 topographic sheets were available from the laboratory archives, and in the worst case, the 1:250,000 aeronautical navigation charts supplied by WICE were used for registration.

Paradoxically, imagery was documented as to projection, but its final registration to the desired projections for the purposes of the Ecosystem Map project was not possible directly on the basis of the projection information. Whether the problem resided with the documentation of the images or a bug in the conversion software in ERDAS was ignored in the rush to complete this activity; each image was registered directly from topographic maps to avoid any future uncertainty relating to the process of conversion. The completed list of documented images is found in Appendix 1.

The purpose of this activity was two-fold. First, the set of images was documented for the purpose of making them available to the Central American research community involved in Biodiversity and Ecosystem research; the mapping activity and monitoring constitute an activity which should be ongoing. The existence of a ready databank of imagery facilitates the involvement of a variety of institutions even in cases where no funding is available to finance work. The expectation is that in the process of carrying out normal activities, such as theses, research or ministry activities, the information gathered can be integrated into the Ecosystem map if the original imagery is available to facilitate the geo-referencing of the ground data.

The second objective of the image registration activity was the creation of country mosiacs to be distributed with the final Ecosystem Map product, on CD-ROM. In order for this activity to be feasible, images were degraded to a resolution of 240m/pixel, so the completed mosaics were between 20 and 80MB each. While these degraded images lack resolution, they provide a guide and backdrop for viewing the completed ecosystem files which help knowledgeable researchers orient themselves with regard to defined categories.

Integration of country vector layers

The integration of country vector layers presented the greatest challenge in the work at the CATIE Lab. While it was expected that data would arrive in a variety of formats, most came as ArcView shapefiles, which minimized problems of data translation. Nevertheless, the overriding problems in integrating the layers were;
1) coordinating legends from the different countries
2) converting semi-documented projections to a common projection
3) making borders compatible in terms of national limits and the continuity of vegetation.

Coordinating legends

At the initiation of the Ecosystem Map project, before the involvement of CATIE, all country teams were assembled and agreed upon a methodology for describing ecosystems. This method was based on the UNESCO classification. The UNESCO classification is a hierarchical system with the possibility of expanding categories with increased detail, avoiding the straightjacket of classes which might not be applicable in highly varied tropical environments. However, as a result, each country team was in a position to introduce new categories of vegetation according to their own experience without consulting with neighboring countries. As a result, each country map had an independent legend, with some common elements, but some unique classes, distinct from those of neighboring countries.

The first task in coordinating the legends was the development of a conversion table between all the different country legends (See Appendix 2). This table was assembled by the WICE team, using descriptions from the individual country reports and legends, and consolidating categories where possible. This table underwent considerable evolution in the course of the project, the underlying problem being how to maintain accurately the relation between coded classes and polygons, while at the same time modifying the coding scheme. The final legend was derived by creating a new code and sub-code for Central American vegetation types, with a suffix indicating specific national or regional variants. As an indication of this evolution, the data presented in Table 2 is named CODES12.DBF. This conversion table allows information from original country maps to be translated to the regional map with the consolidated regional legend codes.

As a final note on the translation of codes, after November of 2000 when a meeting was held to review the regionally integrated map, all editing of legends was done to the country maps as integrated into the regional map. Polygons were altered, codes were modified, and legend categories were subdivided, but only on the regional map. The original country maps and codes were not kept up to date, in the sense that if a new category was developed for Panama, the original country map was not updated. This detail is mentioned to explain the possibility of incompatibilities between the completed regional map, the completed country maps, and the country maps initially presented to the CATIE Laboratory. In the chain of conversions of legend categories and changes in coding schemes, the project finally had to concentrate on the final map products to ensure consistency. The underlying reason for the divergence between the current regional map and the earlier final products of the country teams lies in the process of negotiation required to reach agreements regarding the ecosystem categories. This negotiation took place at a national level, and again at a regional level, with relatively few representatives of the debate at a national level.

Converting semi-documented projections to a common projection

In Central America, each country has developed its own official projection. These projections are officially presented on the cartographic map sets of each country.

In recent years there have been changes in the cartographic systems of the countries, with the introduction of new map sets with slight variations in projection details. The documentation of these details requires that the country Ecosystem map teams document which particular map was used in each case, and the projection parameters of each map. In most cases the country teams overlooked this type of documentation (making the assumption that there was a complete consistency in the national map sets), and there were doubts as to the source maps for georeferencing, the level of error encountered in the georeferencing of the base images used for digitizing, or other source materials.

Table 1

Country	**Projection**
Belize	UTM 16, NAD27 derived datum
Guatemala	UTM 15, NAD27 derived datum
El Salvador	Lambert, NAD27 derived datum
Honduras	UTM 16, NAD27 derived datum
Nicaragua	UTM 16, NAD27 derived datum
Costa Rica	Lambert, Ocotepeque datum
Panama	UTM 17, NAD27 derived datum

As a result, the final adjustment and conversion of the national ecosystem maps to the regional projection, and the coordination of national borders in adjacent countries required a certain amount of research and testing to acheive the best and most consistent fit of data. It should be noted that the typical error encountered in these adjustments was between 200 and 500 meters, errors quite typical of a 1:250,000 map set. While these values in a broad sense were marginally acceptable on their own, when juxtaposed with the adjacent country polygons and borders, the errors became quite visible.

Making borders compatible in terms of national limits and the continuity of vegetation

Another area requring work in the CATIE Lab was the creation of compatible/acceptable border definitions. This problem has two elements; first, the location of the border itself, and second, the continuity of vegetation patterns from one side of the border to the other.

In the case of the locations of borders, final editing was done to resolve more obvious errors, which in many cases were simply questions of which year a river edge was marked as a border; if the river meaner moved in its course, and the adjancent countries used river edges from different years, then an 'error' appeared even though the actual location of the border was clear. In the case of Guatemala and Belize, sandbars appearing in the Sarstun River were mapped as part of both countries, while border disputes between Honduras and El Salvador created areas where the definition of the national boundary was problematic. The regional map of Central America was used as an arbiter whereever possible, although in the end a simple line was drawn indicative of the border location.

It goes without saying that the location of the indicative border between different countries has not been done in this project with the intention of supporting one or the other claim of the border location, and does not indicate acceptance or support of any one competing claim on the part of CATIE or the World Bank, and has been included as a reference to help researchers orient themselves with regard to the vegetation portrayed on the map.

Problems of discontinuities of vegetation arose due to the use of images with cloud cover or images from different dates in adjacent countries. Areas with cloud cover are interpreted on the best available information and the interpretation of the country team. Similarly, drawing of vegetation boundaries for each country was carried out without consulting the adjacent country as to which image was used, or which interpretation and category was being used. These inconsistencies were mapped and printed for a regional meeting in Managua, Nicaragua, in November of 2000, when the different country teams met to agree on consistent boundaries. Once these corrections were agreed upon, the regional and national maps were returned to the CATIE Lab and edited to reflect the changes.

Establishment of a common projection for the Regional Ecosystem Map

The creation of the Regional Ecosystem Map presented an opportunity to enter into a long standing problem for the Central American Isthmus and the process of Central American Integration. If the region is to be presented in a single map, what is the map projection which ought to be used?

Map projections are transformations of data from the spherical earth to flat paper, a process which inevitably introduces distortions into the final map. The selection of a map projection requires the evaluation of the objectives for the use of the final map, and the evaluation of which particular distortion is least damaging. In the case of the Ecosystem Map, it was determined that the most critical criterion for the selection of the projection is that it be 'Equal Area', which is to say that it minimizes distortion in areas between different parts of the map. This characteristic would permit the comparison of areas of different ecosystems between different countries, for the purpose of determining how many hectares of each particular ecosystem existed, how many were protected, and in which countries.

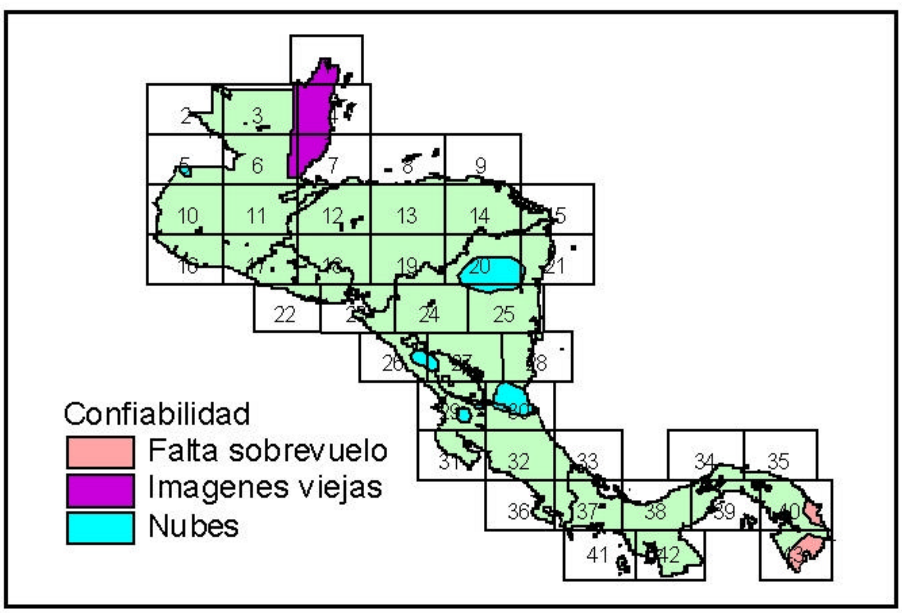
In October of 2000, a formal request for input was circulated to all the heads of the Geographic Institutes in the 7 countries. A number of suggestions were made, including the use of the projection defined by the IPGH (Instituto Panamericano de Geografia e Historia) in the recent regional map. Unfortunately, in the opinion of the president of the Central American IPGH group during the creation of the regional map, the use of that projection without express permission of IPGH would not be advisable.

The projection finally selected was a Lambert Azimuthal, centered at 85 West and 13 North, in the geographic center of Central America (this point is actually inside Nicaragua). To maintain all coordinates as positive, a one million meter offset was incorporated as a false easting and a false northing.

Creation of the 1:250,000 map set

The overall objective of the Central American Ecosystem Mapping project was the creation of a regional map with a scale of 1:250,000. The presentation of the full set of maps on paper requires that the overall map be divided up into a set of individual maps.

Map 1.



The grid for the creation of the individual maps is initially defined arbitarily, on the basis of the size of the final output maps. The borders of the individual maps are then adjusted to avoid the inclusion of empty maps with tiny slivers of map data in a blank frame. The initial 1:250,000 grid included 56 map sheets, but with the careful adjustment of boundaries and the elimination of superfluous coverages, the map set was reduced to 43 maps as illustrated in Map 1. The completed map was then sectioned into rectangles corresponding to the 43 map sheets, and inserted in the documentation frame of the 1:250,000 set, and saved to Windows Metafile format files, to permit electronic distribution of the full map set, and avoid the cost of reproducing full sets in cases where the user would only want the coverage for a particular country.

Data Formats

The consolidated Central American Ecosystem Map is named ECOSISLA, and is the format of an ArcView Shape file, for ease of use in a variety of systems. Viewing of the files requires ArcView or ArcExplorer on the computer of the user.

The fields of the regional database are presented in table 2.

TABLE 2		
Field name	Type	Size
AREA	NUMERIC	12
PERIMETER	NUMERIC	12
ECOSISLA_	NUMERIC	11
ECOSISLA_I	NUMERIC	11
COD14	CHARACTER	16
UNESCO	CHARACTER	16
DESCRIP	CHARACTER	128
CAMCODE	CHARACTER	10
LEYENDA	CHARACTER	128

For purposes of identification, the most important fields are the last 4. ‘Leyenda’ is the legend that appears on the final map, which is a combination of the Central American Ecosystem Code (CAMCODE) and the Description of the ecosystem (DESCRIP). The UNESCO field includes the classification of the ecosystem category according to the Unesco system. Finally, AREA and PERIMETER are expressed in square meters and linear meters. The selection of the Lambert Azimuthal projection for the regional map was based on a need to maintain a consistent basis for measurement of the polygons in all the countries, so the area and perimeter values are 98% accurate.

Conclusion and Recommendations

It should be noted that the activities described could not be carried out precisely in the order presented in this report. The ‘final’ regional map was produced a half dozen times, as a result of editing of country files, changing of legend categories, and finally, as a result of the lack of response from IPGH in the use of their regional projection. After the initial registration of the satelite imagery, the CATIE Lab and the country teams went through an interative process of presenting tentative maps in hard copy and in digital format, discussing and correcting elements, and then editing and reprinting. While in a sense this process was not the most efficient for finishing the map quickly, it was the most effective in guaranteeing the satisfaction of all participants in the process that their observations and corrections had been incorporated.

Although the creation and consolidation of the regional Ecosystem Map represented a huge effort, a nearly equally difficult problem faces the team in the use of the map. For professionals in biodiversity and ecosystems, the use of the map will be obvious and transparent. However, a great number of administrators and policy makers can potentially use the map if its significance and conclusions can be made clear to them.

A first recommendation for the dissemination of the map outside the narrowist community of users is the simplification of the legend. As it currently stands, there are more than 200 legend categories to the map. This richness of detail can be confusing to most users, and may literally not let the user see the forest for the trees. A legend with 10 to 15 categories would be much more understandable, and easier to read and present. The structure of the simplified legend should be defined through consultation with regional users, especially CCAD and CBM (Mesoamerican Biological Corridor).

A second recommendation is the wide distribution of the data through internet, to reduce costs. The continued distribution of the CD-ROM physically incurs costs of time and materials to whoever takes on this task, and it should be more economical to place the data on internet. The exception to this solution are the original satellite images. Since these are each 350MB of data, it is difficult to download these over the internet, especially since there are a total of 30 images.

A third recommendation involves monitoring. Monitoring is an important outcome of the creation of a baseline, such as the 1998-99 Ecosystem map. The process of monitoring would be facilitated by the creation of a formal group in the region; the most effective method of maintaining such a group is to fund it, although that might not be feasible financially. Maintaining the cohesion of this group without funding is a challenge, but it might be acheived through the creation of a mailing list, or an on-line discussion forum, or through periodic meetings. Another option is to locate a very minimal funding source, to manage a 'small grants' program where 1 or 2 projects per year are funded, with the condition that results are shared in a way to promote the cohesion of the group. This is clearly a difficult problem, but may be resolvable through long-term planning. For example, the cohesion of such a group might be guaranteed by the planning of a periodic funded revision of the ecosystem map, which provides an incentive for involvement.

A fourth and final recommendation is the need to explore applications of the map, especially for decision makers regarding biodiversity. Several strategies are possible. One would be to create a mapping module, whose function is only to display the ecosystem map and selected analyses. Such a module could be made very simple and user friendly, but with limited capabilities for GIS professionals; its objective would be to provide an interactive mapping tool to decision makers with little or no GIS training.

A second strategy for reaching decision makers would be the creation of an ecosystem workbook. This would be presented as a training in GIS for biodiversity, and would include a detailed methodology for creating certain analyses. This is clearly a strategy which requires more commitment from the user, but it provides a much more powerful tool.

A third option is to create an ecosystem Atlas. This could be distributed on CD-ROM or on paper, and could feature maps, photos, explanations, etc. This strategy could result in a more colorful and popular result, but its cost of production and distribution would be relatively expensive. Nevertheless, its attractiveness could be a significant diffusion element in its own right, reaching a broad, popular audience through market sales.