Landsat Thematic Mapper Registration Accuracy and its Effects on the IGBP Validation

Gregory J. Husak, Brian C. Hadley, and Kenneth C. McGwire

Abstract
Research associated with the International Geosphere/Biosphere Programme Data and Information System Cover (DISCover) validation exercise revealed possible registration errors in the high-resolution data employed. Systematically corrected National Landsat Archive Production Systems (NLAPS) Thematic Mapper (TM) data used in the validation of the DISCover dataset was compared with precision registered Multi-Resolution Land Characteristics (MRLC) TM data available only for the conterminous United States. A consistent offset in the systematically corrected data was discovered to be approximately 1 km to the east and south. The impacts of this bias on the validation of the IGBP DISCover dataset are examined and show that possibly 20 percent of the IGBP pixels for the conterminous United States could be adversely affected by this registration error. Fractal dimension is shown to have a close relationship with the effect of the offset on a class, and may be used as a predictive tool for areas outside the conterminous United States.

Introduction
This paper examines the misregistration of Level 1P Landsat Thematic Mapper (TM) imagery, and its associated impacts on accuracy assessment for the International Geosphere/Biosphere Programme (IGBP) Global One-Kilometer Land Cover Dataset. This land-cover dataset, named DISCover (Data and Information System Cover), was developed by researchers at the U.S. Geological Survey's (USGS) Earth Resources Observation Systems (EROS) Data Center and the European Joint Research Center (JRC; Ispra, Italy) from a one-year time-series of full resolution Advanced Very High Resolution Radiometer (AVHRR) imagery (Loveland *et al.*, 1999). The DISCover product is a raster dataset that maps the global distribution of 17 land-cover classes. In addition to being the highest resolution dataset of its type ever developed, the DISCover product is the first thematic map of global land cover to have a valid statistical accuracy assessment (Scepan, 1999, in this issue).

While the issue of thematic accuracy assessment has been well-studied (Congalton, 1991; Stehman, 1996), several challenges were encountered when applying state-of-the-practice techniques and methodologies to the development of a global-scale one-kilometer resolution thematic dataset. These challenges included, but were not limited to, data availability, database management, and economic feasibility. Manual image interpretation of Landsat Thematic Mapper and SPOT (Systeme Probatoire d' Observation de la Terre) imagery was used to validate the land cover at randomly sampled locations. While one obvious aspect of data availability was the existence of appropriate Landsat images around the world, a less obvious issue arises from the lack of precise geometric correction in standard high resolution data products such as Landsat TM.

The validation effort had two significant constraints that affected many aspects of the project. The first constraint was the absence of a data budget for the validation. This required the donation of all high-resolution imagery by various organizations. With precision corrected data costing nearly twice as much as systematically corrected data, the latter was chosen in order to assure the necessary global coverage of samples. The second constraint was the proposed timeline for validation activities. High-resolution imagery was to be delivered within two months of the sample selection, leaving inadequate time to either acquire precision data or to precision register all systematic data.

In order to perform the DISCover product validation in a cost-effective and timely manner, Landsat TM imagery obtained from the EROS Data Center and SPOT multi-spectral imagery were employed as high-resolution ground truth. SPOT data are included in this paper in order to detail its processing for the validation. However, due to time and data limitations, no registration tests were performed for SPOT data. The National Landsat Archive Production System (NLAPS) processed the TM data, and images were geometrically registered using the image corner coordinates calculated by EROS Data Center using satellite ephemeris data. The procedures that were used to register TM imagery to the DISCover product are documented in the background section that follows. While it was thought beforehand that positional error from ephemeris-based registration would be random, during the execution of this project it was determined that a significant bias exists in the corner coordinates that are provided by the NLAPS system. This paper documents the nature of this bias for the conterminous United States. Also discussed are the potential impacts this bias may represent with respect to the DISCover accuracy assessment. The importance of this bias extends well beyond this study area and, indeed, beyond the scope of this validation activity.

Background
The IGBP's global land-cover dataset represents a joint effort from international organizations to map the globe at a one-kilometer level. Data for creating the DISCover map were acquired from the AVHRR sensor that provides daily global coverage at a 1-km spatial resolution (IGBP, 1996). The AVHRR data used in the production of the DISCover dataset were projected into the Interrupted Goode Homolosine projection (Eidenshink

G.J. Husak and B.C. Hadley are with the Remote Sensing Research Unit, Department of Geography, University of California, Santa Barbara, CA 93106 (husak@geog.ucsb.edu).

K.C. McGwire is with the Desert Research Institute, Biological Sciences Center, Reno, NV 89512.

Photogrammetric Engineering & Remote Sensing
Vol. 65, No. 9, September 1999, pp. 1033–1039.

0099-1112/99/6509–1033$3.00/0

and Faundeen, 1994). This projection is beneficial when working with a global dataset for two reasons. First, this projection is an equal area projection, which is ideal for calculations of areal land cover and facilitates visual understanding by minimizing areal distortions. Second, the Goode's projection easily partitions the globe into 12 regions, allowing for the processing of data in parts rather than having to process the entire global dataset at once (Steinwand, 1994; Eidenshink and Faundeen, 1994). The accuracy of registering the AVHRR data to Digital Chart of the World (DCW) was larger than 1.0 pixel root-mean-square (RMS), or one kilometer (IGBP, 1996; Eidenshink and Faundeen, 1994). Because the temporal AVHRR data are a series of monthly composites, it is important to make sure that all of the dates are registered relative to one another. This relative registration is important because each pixel's signature needs to be traced throughout the year to accurately classify its land cover. If the relative registration is inaccurate, then the classifier analyzes different areas through time. It was shown that these layers are registered to one another at better than 1.0 RMS pixels (Eidenshink and Faundeen, 1994; IGBP, 1996).

The time series AVHRR dataset was classified into the IGBP classification scheme using an unsupervised cluster classification. This classification algorithm analyzed the temporal signatures of the NDVI, and pixels with similar temporal profiles for NDVI were grouped together through an iterative process (IGBP, 1996; Loveland *et al.*, 1999). Once the classification was complete, an accuracy assessment was performed (Scepan, 1999, in this issue). This validation effort used both TM and SPOT imagery as a high-resolution data product that substituted for "ground truth." Each of these data sources has inherently different characteristics and formats, so the registration routines were adjusted to take different input data formats and generate compatible output formats.

The TM data used for the DISCover validation were NLAPS level 1P. The NLAPS data were selected as the primary ground truth due to their extensive global coverage. The spatial resolution of NLAPS data is 25 meters. Radiometric and geometric corrections are performed as part of the preprocessing of NLAPS data.

Geometric correction removes systematic distortions by modeling satellite and sensor properties. Satellite imagery may contain many distortions arising from effects such as nonlinear mirror scanning velocity, varying average mirror speed between scans, sequential detector sampling, terrain, and detector offsets in the focal plane. Systematically corrected data account for these and other variations, such as irregularities in the orbit or attitude of the satellite. The model used to correct NLAPS data also accounts for the rotation and curvature of the Earth. The results of the model yield coordinates for the corner points of the TM imagery that can be used in a geometric transformation to arrive at locations for the entire image surface.

A program was developed in the C language to speed the extraction and reprojection of imagery from the original CD-ROM media and to improve the efficiency in generating manual interpretation products from TM and SPOT datasets. This program performed a number of functions, including

- reading image corner coordinates from the image header file,
- extracting a subarea from the CD-ROM approximating the desired region in order to reduce disk space requirements,
- transforming corner coordinates into the Interrupted Goode Homolosine projection,
- creating a file of synthesized control points to warp the image subset to the Goode's projection, and
- generating polygons in an Arc/Info format indicating the 1-km^2 test areas used in the validation activities.

Transformation of image corner coordinates to the Interrupted Goode Homolosine projection was accomplished by compiling portions of the General Cartographic Transformation Package (GCTP) into the program. The GCTP subroutine library was developed by the USGS and is available on the EROS Data Center FTP site **ftp://edcwww.cr.usgs.gov/pub/software/gctpc**. None of the image processing or GIS software packages available at the time of this project supported the Goode's projection. ERDAS Imagine software was used in this effort to warp images to map coordinate systems using polynomial transformations. In testing the transformation of TM images from their original Universal Transverse Mercator (UTM) projection to Goode's, it was found that the distortions involved in warping between these two projections were too great to be modeled adequately with a first-order polynomial, generating RMS errors much greater than 1.0 pixel. However, the five points provided in the image headers (four corners and center point) were not sufficient to allow a higher order polynomial. To correct for this, synthetic control points were generated at equal intervals throughout the UTM projected images. UTM coordinates for these synthetic points were calculated using a first-order polynomial based on coordinates in the image header. Control point files were then created that projected these synthetic points into the Goode's projection. Approximately 36 points were generated per scene, which, when used in a second-order polynomial transformation, typically resulted in RMS errors on the order of 0.02 pixels. Thus, the geometric error contributed by transforming from UTM to the Goode's projection with a second-order polynomial was practically insignificant.

Testing Geometric Accuracy

Validation tests performed in association with the DISCover validation activity indicated that there could be inaccuracies in the NLAPS registration. The locational accuracy of the NLAPS data was tested against the Multi-Resolution Land Characteristics (MRLC) dataset, a set of precision registered TM products that are available for the conterminous United States. This precision dataset undergoes extensive preprocessing before being made available to users. MRLC data are precision registered using ground control points (GCPs) taken from 1:100,000-scale Digital Line Graphs (DLGs) (Loveland and Shaw, 1996). These DLGs are acquired from the National Digital Cartographic Data-Base (NDCDB) which contains transportation, hydrographic, hypsographic, and political boundary information. The DLGs are overlaid on the imagery to aid in correlating features between the two datasets. The latitude and longitude of the feature from the DLG data were then matched with the line and sample of the pixel in the file. Each image contained 12 to 15 control points distributed throughout the image. These control points are then used to calculate a first-order polynomial model that was used when reprojecting the raw data into a UTM projection. Elevation values from a 1:250,000-scale digital elevation model (DEM) were used to correct control point coordinates and to perform corrections for relief displacement on a pixel-by-pixel basis. MRLC data are projected into the UTM coordinate system using the 1983 North American Datum (NAD83). Each MRLC image is tested to verify that the RMS error is within $\pm$ 1 pixel (30 meters).

MRLC data were compared directly with the NLAPS data to determine the difference in registration between the two. Because the MRLC registration errors are quantified to within 30 meters, it was used as the true ground location, and any difference between the two datasets was assumed to be an error resulting from the NLAPS preprocessing. In order to compare the two datasets, the first 25 validation sites for each class in the conterminous United States were selected and MRLC data were ordered to match the NLAPS data. In all, 23 scene pairs were collected. Because the methods for comparison required accurately identifiable image features, we were careful to select a band that would provide sharp contrast between bodies of water and land surfaces, and at cross-sections of linear shaped urban features (e.g., highways and road intersections). Band 5 (1.55 to 1.75 μm) was selected because it is effective at detecting land/water boundaries and it provided good visual contrast.

While the MRLC data were georeferenced using ground control points from USGS cartographic datasets and corrections for topographic relief (Loveland and Shaw, 1996), NLAPS data are simply processed using a geometric model to determine the image location, and from this the corner points of the image are calculated. Both datasets were projected in UTM coordinates with matching scene pairs having identical UTM zones. The NLAPS imagery is projected using the World Geodetic System 1984 (WGS84) datum while the MRLC imagery is projected using the NAD83 datum; however, these two data are identical for the conterminous United States, allowing for direct comparison of feature coordinates (Dana, 1991).

The fact that the NLAPS and MRLC data were resampled to different spatial resolutions did not pose much of an obstacle. The pointing accuracy associated with identifying features in each type of image was very similar. Patterns observed in each image were sufficiently comparable that point features could be confidently located and their coordinates recorded. Errors in cursor placement between the two images was somewhat dependent on the type of feature selected, but in all cases was not expected to exceed $\pm$ 1.0 pixels (Welch *et al.*, 1985).

Coordinates taken from NLAPS images were based on best fit linear equations. This method calculates a first-order polynomial developed from the four UTM corner coordinates in the image header to provide cursor locations in the UTM projection without requiring the images to be resampled. Matching NLAPS/MRLC scene pairs contained sufficient overlap to allow four identical point features to be identified, one near each corner of the image. This placement provided a good spread of coverage for each image, which would hopefully reduce potential biases associated with a single bad corner coordinate. A zoom ratio of 4:1 was used to locate point features, with the images being displayed simultaneously on the computer monitor. This level of zoom provided a broad enough view to give context while allowing precise placement of the cursor on the point features.

Comparison of the NLAPS and MRLC images yielded a systematic offset in both the X and Y directions. Offsets reported here are based on the average for all of the four features that were identified in each image (i.e., single average offset per image). In the east-west direction, features in the NLAPS images were mapped 880 meters to the east of the same feature in the MRLC image on average, with a standard deviation of 140 meters. The minimum shift was 489 meters to the east, while the maximum shift was 1180 meters to the east. The offset in the north-south direction was greater in both magnitude and variance. The average shift of features in NLAPS imagery was 1082 meters to the south of MRLC, with a standard deviation of 492 meters. Offsets ranged from a minimum shift of 116 meters to the south to a maximum shift of 2705 meters to the south. The average combined shift in both axes was calculated to be 1437 meters with a standard deviation of 380 meters. The minimum combined shift was 947 meters and the maximum was 2812 meters.

It should be noted that the within-scene variation of the shift was quite small. All individual corner points were consistent in both the X- and Y-axis shifts, resulting in consistent scene average shifts. To quantify the internal scene variation, a standard deviation for the shift values of each corner point was calculated for each scene. The average of within-scene standard deviations was 54.7 meters. This number reveals that corner points were in agreement with the other points for that scene, and that resulting scene averages were not heavily biased by one wayward point for the scene.

A graphical view of the offsets is shown in Figure 1. This figure depicts the direction and magnitude of the shift for each image as a vector. However, it is important to note that the length of the vectors is not proportional to the scale of the map, and the vectors are grouped into five length classes based on the magnitude of the offset. These cartographic conventions were employed in an attempt to produce a graphic that the reader could more quickly interpret.

Effect on DISCover Validation

The systematic positioning bias associated with the NLAPS processing of TM data discovered during this investigation will affect the overall interpretation of the validation results of the DISCover product. Yet, just what that ultimate effect may be remains unclear. This 1.4-kilometer offset has been measured only for the conterminous United States, so the actual misregistration of TM scenes that were acquired for other parts of the globe is unknown. However, the ubiquitous nature of the offset in this study certainly raises a global concern.

The effect this offset may have on the validation is dependent on the location of the sample point. Figure 2 shows a sample TM image with three pairs of polygons. For each polygon pair, the grey polygon represents the possible location of a sample point used in the validation. The white polygon represents where the polygon would be located after the application of the average shift. The applied shift might not be the exact shift for that scene, but it does give the reader an example of how the shift may affect the validation. For some sample points, the shift would have no effect on the land cover while, for other points, the shift would obviously change the land-cover type.

The impact of a systematic bias on the overall DISCover accuracy assessment will depend on a number of variables, including the size and complexity of land-cover polygons delineated in the DISCover product. For land-cover types characterized by relatively large, homogeneous patches on the landscape, the effect of an offset should be relatively small. For example, the number of pixels in a simple square patch is exponentially related to the length of the sides, while the number of pixels affected by an offset grows linearly. Thus, given similar offsets, a larger square would have proportionally less misregistered pixels than a smaller one. However, this relationship does not necessarily hold for land-cover types that exhibit fractal behavior or, more generally, statistical self similarity (O'Neill *et al.*, 1988). A fractal pattern exhibits a power law relationship between perimeter and area, so the number of boundary pixels that may be affected by misregistration grows at an exponential rate along with area. Thus, land covers with complex, fragmented spatial patterns may be generally affected regardless of patch size.

In an attempt to assess the impact of misregistration for the high resolution images that were used in the DISCover validation, a test was performed for the conterminous United States. An area roughly corresponding to the conterminous United States was subset from the IGBP global dataset. This subset was then offset by the bias found in the study (i.e., 880 meters east and 1082 meters south) and overlaid on the original DISCover dataset. By checking how the values of the overlaid pixel values compared with the original pixel values, statistics could be calculated to estimate the effects the offset would have on each class. This approach assumed that the mean offset of the selected NLAPS/MRLC image pairs was representative of the actual area-averaged effect.

The results of overlaying an offset version of the DISCover dataset for a substantial portion of the conterminous United States on top of the correctly registered DISCover dataset are displayed in Table 1 as a confusion matrix. The water class was excluded from this analysis because the inclusion of open ocean might mask the misregistration effect on overall accuracy. The columns in Table 1 represent the original pixels and rows represent pixels in the shifted data. The overall agreement between the original and offset datasets for land areas of the conterminous United States was 81 percent. This suggests

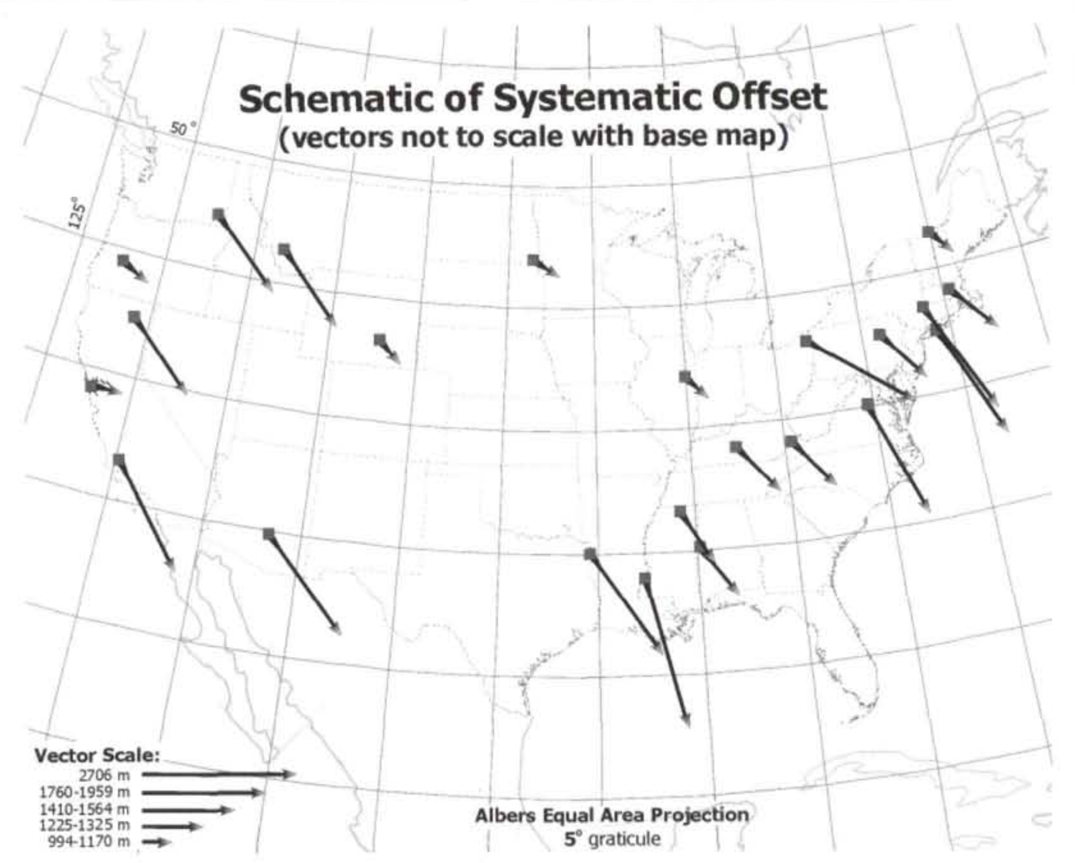

Figure 1. Map of offsets between NLAPS and MRLC images.

that nearly 20 percent of the one-kilometer test pixels in the conterminous United States portion of the DISCover product that were classified by the expert image interpreters would have been incorrectly matched in the product validation (Scepan, 1999, in this issue). Row and column percentages are extremely similar in Table 1. This similarity would be expected because the two products were identical; however, this similarity also confirms that no strong biases between class-pairs arise from the geometry of land-cover patches. Such a bias might have been introduced if patches for a particular class were narrow, linear, and parallel in nature, or if there was a predominant sequence of transitions between any two-class pairs in a given direction.

Classes that show the strongest effects from a systematic offset in the conterminous United States are snow and ice (class 15), urban (class 13), permanent wetlands (class 11), closed shrubland (class 6), and woody savannas (class 8). Figure 3 plots the area of each class versus the agreement between original and offset maps (as indicated by column percentages). There is a moderately strong, negative relationship between the total area and misregistration error for each class ($R^2 = 0.67$). Among classes with the least areal extent, the snow and ice class was more strongly affected than might be expected, while savannas, wetlands, and barren show somewhat less of an effect than expected. It is interesting to note that, because those classes most affected by registration error are of smaller extent and because overall map accuracy must be adjusted to area-weighted representation for each class (Card, 1982), the effect of misregistration on overall accuracy will be less than the effect on individual classes.

To explore the effect of boundary complexity on misregistration errors, the area-weighted patch fractal dimension was calculated for each landcover class. The fractal dimension (δ) measures the geometric complexity of a polygon. This landscape metric is formulated as

$$\delta = \sum_{i=1}^{n} \left[\frac{2 \ln (P_i/4)}{\ln (a_i)} \frac{a_i}{A} \right]$$

where i is the patch number, P_i is the perimeter of patch i, a_i is the area of patch i, A is the total area, and n is the number of patches (adapted from McGarigal and Marks (1995)).

Figure 4 shows the relationship between area-weighted fractal dimension and the agreement in shifted versions of the DISCover subset for each land-cover class. This relationship is stronger than that observed for areal extent ($R^2 = 0.72$). A regression of both area and fractal dimension against the agreement in shifted products yielded an R^2 of 0.75, so the areal extent of each class does not provide much additional explanatory power.

Conclusion

As part of the validation activities for the IGBP DISCover land-cover map, a systematic offset was identified in Thematic Mapper corner coordinates for imagery used in the validation exercise. Within the resources available to this project, at this time, this offset could only be documented using select scenes for the conterminous United States. In North America, it appears that misregistration of the TM imagery may have affected almost 20

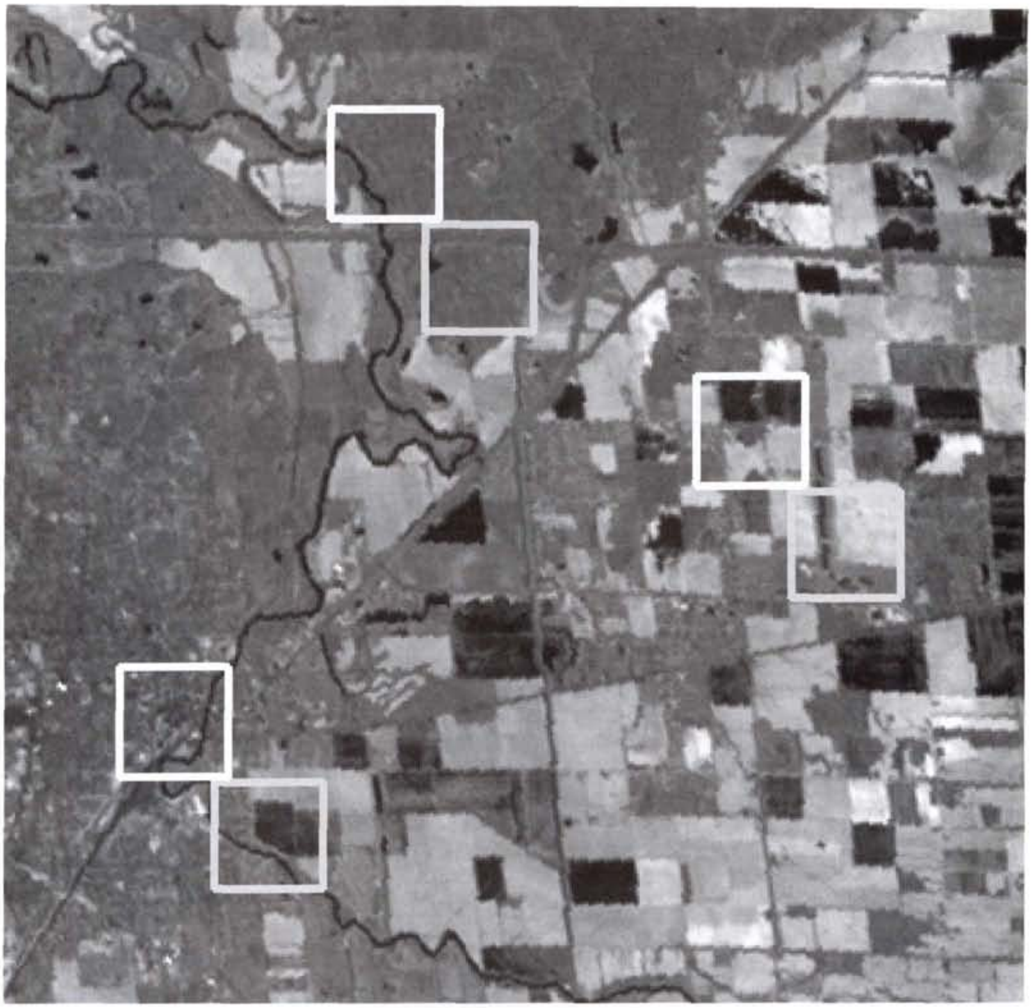

Figure 2. TM scene with 1-km^2 polygon pairs representative of a DISCover sample before (grey) and after the application of the bias (white).

TABLE 1. CONFUSION MATRIX SHOWING EFFECT OF OFFSET ON DISCOVER PRODUCT

	Original pixel															
Shifted pixel	1	2	4	5	6	7	8	9	10	11	12	13	14	15	16	
1 Evergreen Needleleaf	**1113390**	1969	33848	49902	7098	1187	25008	814	32913	99	44525	1970	28102	1	4	83.04
2 Evergreen Broadleaf	1871	**16443**	684	2250	0	4	13	21	443	1	695	3	102	0	1	72.98
4 Deciduous Broadleaf	30877	689	**924296**	54563	1102	701	12809	87	5794	1	26223	3871	36562	0	0	84.21
5 Mixed forests	48035	2292	5654	**627557**	479	1303	3631	990	1543	127	12419	963	35946	1	3	79.34
6 Closed Shrublands	6833	0	1344	483	**26472**	1873	542	117	1158	1	502	330	138	0	0	66.69
7 Open Shrublands	1156	1	644	1494	1837	**1564198**	15156	1167	62222	160	6257	1210	4789	54	5412	93.9
8 Woody Savannas	28289	14	11291	3351	603	12925	**289739**	3373	47931	45	13513	1334	15095	0	2	67.77
9 Savannas	980	29	72	954	0	1143	3371	**24492**	366	19	1671	225	709	1	1	71.97
10 Grasslands	32897	271	4998	1513	1044	64664	47890	444	**1195393**	685	43345	3844	58998	0	11	82.10
11 Permanent Wetlands	70	1	5	121	2	130	36	29	869	**6947**	1079	186	12	5	104	72.39
12 Cropland	44978	691	25584	12558	572	6372	13782	1613	43466	1076	**952766**	7186	103510	2	6	78.47
13 Urban	2053	1	4123	993	335	1259	1352	186	3791	259	7118	**51867**	7149	0	5	64.44
14 Mosaic	29872	108	35705	35299	163	4372	14207	669	59820	6	103423	6742	**862230**	0	1	74.81
15 Snow and Ice	0	0	0	0	0	51	0	0	0	9	6	0	0	**336**	201	55.72
16 Barren	2	1	0	8	1	5652	5	0	7	92	3	11	2	190	**16237**	73.10
	83.01	73.05	84.16	79.33	66.67	93.90	67.77	72.24	82.12	72.92	78.51	65.04	74.76	56.95	73.81	

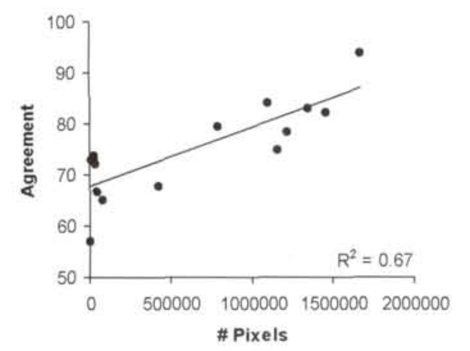

Figure 3. Plot of class area versus agreement in offset maps.

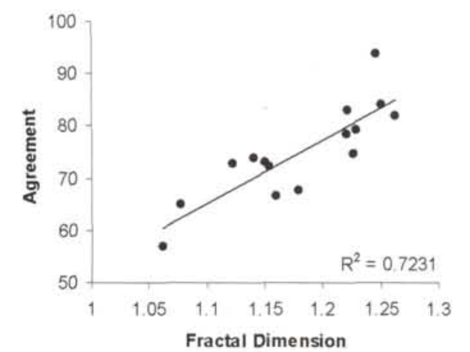

Figure 4. Plot of area-weighted patch fractal dimension versus agreement.

percent of the test samples used in the validation of the DISCover product. It may be useful to try shifting samples used in the IGBP validation in an attempt to correct for this systematic bias. Although every NLAPS scene that was tested against MRLC was offset in a similar direction, simply shifting all NLAPS scenes by the average bias might not fix the problem completely. Unfortunately, in implementing a project of this scale, it was not practical for the DISCover validation effort to precision geo-register each scene in remote parts of the world with local control points.

It is important to stress that the results of this offset test do not mean that the user of the IGBP DISCover product should just add 20 percent to the overall accuracy value for the map. The effect that this offset in the high-resolution data has on the validation accuracy statistics is difficult to ascertain, and work is continuing in this area. It is possible that the distribution of validation samples was not proportional to the 20 percent of changed pixels reported for North America. The possible over- or under-representation of misregistration effects might be approximated by the binomial variance associated with the number of validation sites falling in or out of locations where the offset led to disagreement. Ignoring class-specific variation in area and error rates, we would expect the standard deviation to be the square root of $p * (1 - p)/n$. Given a proportion (p) of 0.814 and 53 samples, this corresponds to a standard deviation of 5.3 percent or 2.8 samples. From this we might infer with 95 percent confidence that at least 39 of the 53 samples were "unchanged" after the shift in the high-resolution imagery. Alternately, with the same confidence interval, as many as 48 of the 53 samples might have been unchanged. However, the effect becomes substantially more complicated if area weighting, unequal error rates, and possible locational dependence of errors were to be factored in. It is unknown if there is any systematic offset in areas outside the conterminous United States, so the effects on the overall accuracy documented here are only known to pertain to validation samples in the study area.

The MRLC dataset provided a convenient point of comparison for the test of registration accuracy because the scenes corresponded closely with the extent of the NLAPS data, and it was easy to identify identical features in the two types of TM imagery. A further test that could be done to measure the accuracy of the registration procedures would be to use independent checkpoints. Testing against a set of independent surveyed checkpoints would quantify the registration errors more precisely and would allow the test to include any possible problems arising from the transformation to the Goode's projection. However, such an approach would not be any more extensible to handle the global extent of the validation than that tested here. Acquiring and testing reliable checkpoints for the entire globe could prove to be difficult due to lack of existing, high-resolution, positional datasets for parts of the globe and the number of TM and SPOT images to be tested.

For small study areas requiring the registration of only a few NLAPS scenes, there are options. One would simply register each scene manually using available maps, surveys, or Geographic Positioning System coordinates. However, studies of large areas requiring the processing of many NLAPS scenes may necessitate a different approach. For a project of this magnitude, it was assumed that there would be too many resources required to independently register each scene. Under these circumstances, correcting the TM to DISCover overlay by the average bias found here should result in imagery that is more accurate overall, though individual scenes may be less accurate.

This study tells us little about possible systematic geometric biases in other parts of the world. Further work is needed to characterize the problem with the NLAPS processing of scene information for the rest of the globe. Work is underway to improve the TM registration methodology by employing National Imagery and Mapping Agency (NIMA) image chips covering locations throughout the world. These chips may provide an accurate benchmark for determining image location and result in a library of accurately registered imagery for the globe. It must be noted that more precise products will come at a higher cost. This higher cost could prove to be prohibitive for projects operating on donations or small budgets like the IGBP validation.

There is much work to be done as an extension of this study. Checking for systematic offsets in other parts of the world to see if they match what has been found for the conterminous United States would be important to the IGBP validation. It is also important to get a more complete understanding of how this offset affects the validation by checking sample points individually to check the exact class-to-class shift for each point. Finally, it would be interesting to develop a confusion matrix, like the one in this study, for other regions of the globe in order to understand how such a shift would affect each class around the world. These global shifts could then be checked against the actual validation results to possibly explain some of the errors in the accuracy statistics from the validation exercise.

Acknowledgments

This work was supported by grants from the National Aeronautics and Space Administration (NAG5-3620 and NAG5-6259). Dr. John Estes provided integral guidance and support during this study. In addition, this research received support from the global land-cover team at the USGS EROS Data Center, Centre

Nacional d'Etudes Spatiales (CNES), and Joint Research Center (JRC) Space Applications Institute. Thanks to Dr. Dar Roberts for reviewing a draft of this paper and to Jeffrey Hemphill for his cartographic assistance.

References

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of the Environment*, 37(1):35–46.

Card, D., 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy, *Photogrammetric Engineering & Remote Sensing*, 48(3):431–439.

Dana, P.H., 1991. *The Evolution and National Selection of Geodetic Datums*, University of Texas, Austin, 134 p.

Eidenshink, J.C., and J.L. Faundeen, 1994. The 1 km AVHRR global land data set: first stages in implementation, *International Journal of Remote Sensing*, 15(17):3443–3462.

International Geosphere-Biosphere Programme, 1996. *The IGBP-DIS Global 1 Km Land Cover Data Set: Proposal and Implementation plans*, IGBP, 66 p.

Loveland, T.R., and D.M. Shaw, 1996. Multi-resolution land characterization—Building collaborative partnerships, *GAP Analysis—A Landscape Approach to Biodiversity Planning* (J.M. Scott, T.H. Tear, and F.W. Davis, editors), American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, p. 79–85.

Loveland, T.R., B.C. Reed, J.F. Brown, D.O. Ohlen, J. Zhu, L. Yang, and J.W. Merchant, 1999. Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data, *International Journal of Remote Sensing*, (in press).

McGarigal, K., and B. Marks, 1995. *FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure*, Gen.Tech. Rep. PNW-GTR-351, U.S. Forest Service, Pacific Northwest Research Station, Portland, Oregon.

O'Neill, R., J. Krummel, R. Gardner, G. Sugihara, B. Jackson, D. DeAngelis, B. Milne, M. Turner, B. Zygmunt, S. Christensin, V. Dale, and R. Graham, 1988. Indices of landscape pattern, *Landscape Ecology*, 1(3):153–162.

Scepan, J., 1999. Thematic Validation of High Resolution Global Land-Cover Data, *Photogrammetric Engineering & Remote Sensing*, 65(9):1051–1060.

Stehman, S.V., 1996. Estimating the kappa coefficient and its variance under stratified random sampling, *Photogrammetric Engineering & Remote Sensing*, 62(4):401–407.

Steinwand, D.R., 1994. Mapping raster imagery to the Interrupted Goode Homolosine projection, *International Journal of Remote Sensing*, 15(17):3463–3471.

Welch, R., T. Jordan, and M. Ehlers, 1985. Comparative evaluations of the geodetic accuracy and cartographic potential of Landsat-4 and Landsat-5 Thematic Mapper image data, *Photogrammetric Engineering & Remote Sensing*, 51(9):1249–1262.