Geocover[tm]-Ortho and Geocover[tm]-Lc: Orthorectified Landsat Tm/Mss Data and Derived Land Cover for The World

G. T. Koeln J. D. Dykstra D. J. Cunningham[1]

Earth Satellite Corporation, EarthSat
[1]6011 Executive Blvd., Suite 400, Rockville, MD, USA 20852-3804
Tel: 301-231-0660, Fax: 301-231-5020
E-mail: gkoeln@earthsat.com, jdykstra@earthsat.com, dcunningham@earthsat.com,

ABSTRACT Sponsored under NASA's Scientific Data Buy program, the GeoCover™-Ortho program has created a geodetically accurate digital database of Landsat TM and MSS multispectral imagery covering the earth's land mass. The GeoCover™-Ortho coverage is comprised of over 14,000 Landsat images that have been photogrammetrically adjusted and digitally orthorectified to create a seamless global coverage of multispectral digital imagery with 50 meter (RMS) geodetic positional accuracy. The Landsat source images have been hand picked from the Landsat archives of both the EROS Data Center and international Landsat foreign ground receiving stations, and represent the highest image quality and lowest cloud cover available for the specified time period. GeoCover™-Ortho provides readily available, affordable, and accurate Landsat TM and MSS multispectral imagery which not only can be used as a geodetically accurate base map, but also provide an excellent digital source for multispectral image processing and analysis. Working initially with NASA and currently with NIMA, EarthSat has developed a set of procedures and processes to produce a landcover analysis for all land areas of the world using Landsat TM data rectified under the GeoCover™-Ortho program. The 13 classes of land cover being mapped include: deciduous forest, evergreen forest, scrub/shrub, grassland, barren/sparsely vegetated, urban/built-up, rice agriculture, other agriculture, permanent herbaceous wetlands, mangroves, water, and ice/snow. The land cover from the first 1200 Landsat TM scenes covering the Middle East, East Africa, and North Africa will be available in May 2000. EarthSat owns the landcover product and the U.S. Government has purchased a government-wide license for the first 1200 scenes. Land cover from all 7,600 Landsat TM scenes should be completed in 2003. The land cover will be distributed in both raster and vector formats. The raster products will be available for each of the Landsat TM scenes and by 1-degree by 1-degree blocks. The vector products will be available by 1-degree by 1-degree blocks as ARC/INFO coverages.

KEY WORDS Orthorectified Landsat Images, Landsat Image Mosaics, Controlled Image Base, Global L
and Cover

1. Geocover[tm]-Ortho Program

Earth Satellite Corporation (EarthSat) was one of five commercial firms selected to provide the National Aeronautics and Space Administration (NASA) with data to support Phase II of NASA's Scientific Data Buy (SDB) Program. NASA's SDB program is a demonstration program developed in response to the President's Space Policy, which directed NASA to purchase remote sensing data from the private sector. The SDB was initiated in fiscal year 1997 and was funded under the Earth Science Enterprise (ESE) program (formally known as Mission to Planet Earth) to provide scientific data to the ESE community. NASA's hope was that the private sector could provide images for scientific purposes in a cost effective manner. This $50 million program is an opportunity to advance global-systems research, to strengthen the U.S. economy through development of remote sensing technologies, and to test a new way of doing business.

The NASA Commercial Remote Sensing Program (CRSP) Office at the John C. Stennis Space Center in Mississippi is managing the SDB program. This program is working to expand the resources available to the ESE science community in its quest for knowledge about the Earth and its changing environment. Under this program, EarthSat is providing historical orthorectified Landsat images and mosaics covering global land areas for a period of two decades.

As part of the ESE program, NASA estimates that ESE scientists will study at least 70,000 Landsat scenes. It is estimated that without EarthSat's GeoCover™-Ortho product, 4 to 7 staff-hours would be required to rectify each of these scenes, and the existing maps used to rectify the images will likely have widely varying quality and accuracy. Using the GeoCover™-Ortho product as an image base, each of these 70,000 Landsat scenes can be rectified in less than 1 staff-hour. This and other benefits of EarthSat's GeoCover™-Ortho product will save NASA over $50 million. For further information on NASA's SDB and ESE programs, visit:

www.crsp.ssc.nasa.gov/databuy/OVERVIEW.HTM and www.earth.nasa.gov/visions/index.html.

1.1. GeoCover™-Ortho PRODUCTS

GeoCoverTM-Ortho products are being produced for all land areas of the world which requires approximately 7,600 Landsat TM scenes. The scenes are being acquired from EROS Data Center (EDC) when quality scenes for a particular path/row are available, when quality images are not available from EDC scenes from foreign ground receiving stations are reviewed and purchased. The target collection date for the TM scenes is 1990, but images within three years of 1990 are obtained when high quality images for 1990 are not available. The two TM image-based products being delivered to NASA are individual orthorectified images (by path/rows) and in 5 degree by 6 degree mosaics.

The individual TM scenes include all 7-bands and are in UTM coordinates based upon the WGS84 datum. The images are resampled to 28.5 meters using a nearest neighbor approach to preserve the original data values. The positional accuracy is 50 m (RMS). Each of the 7 TM bands are delivered in GEOTIFF format, so there is one GEOTIFF file for each band. A typical GEOTIFF file is usually less than 60 megabytes. Most of the scenes have less than 10 percent cloud cover.

The mosaic product is a 3-band mosaic covering approximately 5 degrees of latitude by 6 degrees of longitude. There are usually 15 to 20 scenes required to cover an area that size. The mosaics are created using TM bands 2, 4, and 7. To create a mosaic that is as seamless as possible, the histograms of the individual scenes are matched and the overlapping areas between scenes are feathered. These data are projected to UTM coordinates based upon the WGS84 datum and resampled to 28.5 meters using a cubic convolution process to enhance the visual quality of the mosaic. The mosaics are distributed as uncompressed 3-band GEOTIFFS (approximately 1.5 gigabytes) or as MrSid compressed files (approximately 50 megabytes).

A Landsat MSS product is also being produced. This MSS product is co-registered to the Orthorectified TM data and is created from images acquired in the mid-1970's. Like the TM data, the MSS data are projected to UTM, WGS84 using nearest neighbor techniques. All 4 of the MSS bands (bands 4, 5, 6, and 7) are provided as separate GEOTIFF files with a pixel size of 57.0 meters (twice the size of the TM products). The position accuracy of the MSS product is 100 m (RMS).

NASA has an unrestricted right to copy, use, and distribute all of the GeoCover™-Ortho products described above. EarthSat may also distribute these data and produce additional derived products from the data provided to NASA. Questions concerning the distribution of the GeoCover™-Ortho products by NASA should be directed to:

Mr. Fritz Policelli
Bldg. 1100
Stennis Space Center, MS 39529-6000
Telephone: 228-688-2042
Fax: 228-688-7455
E-Mail: fpolicel@ssc.nasa.gov

Question concerning the distribution of these data or the development of other derived products by EarthSat should be directed to:

Dr. Jon Dykstra
Earth Satellite Corporation
6011 Executive Blvd., Suite 400
Rockville, MD 20852-3804
Telephone: 301-231-0660
Fax: 301-231-5020
E-Mail: jdykstra@earthsat.com

1.2. Image Selection Criteria

To aid in the selection of scenes used for the GeoCoverTM–Ortho products, a preliminary image selection algorithm was developed. This algorithm narrows the set of images that are individually inspected based on the digital archive metadata from EDC and the foreign ground receiving stations. The algorithm ranks all available scenes for a specified path/row by: 1) cloud cover (minimum cloud cover was the highest objective), 2) image quality (selected scenes must have top 10 percent quality rating), 3) phenology (ideally, scenes should be acquired at or just past the peak of greenness), 4) collection year (scene should be collected in 1990 or within three years of 1990), and, when possible, 5) same date run (for image mosaics and landcover processing, there are advantages to selecting scenes within a path that were acquired on the same day). Parameters within each of the above criteria and a set of weights of importance (i.e. minimum cloud cover was the most important objective) between criteria were established. Each potential scene was evaluated by the algorithm outlined above and assigned a score that was used to select the top ranking scenes in each path row, these top ranking scenes were then further evaluated by visual inspection of the available thumbnail prints.

1.3. Orthorectification Process

The geodetic control and digital elevation models used in the photogrammetric block adjustments and orthorectification process were provided by the US

Government and represent the best available control for Landsat data processing. In order to accomplish a geodetically accurate Landsat adjustment on a global scale, a new approach, called "MosPoly" (Mosaicking Polynomials), was developed by EarthSat. (This technique is patent pending.) The MosPoly technique provides the unique capability to simultaneous adjust large blocks of Landsat data (approximately 400 scenes in one block) with less than 15% of the scenes containing ground control points. The software uses pixel correlation to acquire tie-points between adjacent images in an automated manner. A similar pixel correlation technique is used to pass control from the orthorectified Landsat TM imagery to the uncorrected Landsat MSS imagery. This control is then used, along with the same digital elevation models used for the TM processing, to orthorectify the Landsat MSS imagery.

1.4. Application Of Earthsat'S Geocover^{tm}-Ortho Products

The Landsat TM products described above have positional accuracies better than most existing 1:250,000- and 1:100,000-scale maps and provide an attractive option to serve as the cartographic standard at these scales. These products are ideally suited to be used in analysis of land cover and landcover change between the mid-1970's and 1990 and can readily be used as the image base from which Landsat 7 data can be orthorectified. Recent uses of EarthSat's GeoCoverTM-Ortho products include:

- UN/FAO to orthorectify TM data for their AFRICOVER Program
- Mapping changing shoreline of the Caspian Sea
- UN's use for Ethiopean/Eritrean conflict resolution
- Disaster response to Turkey's earthquake
- Conflict resolution of Algeria/Morocco border dispute
- NASA's orthorectification of PATHFINDER data sets.

Currently the GeoCover^{TM}-Ortho products are available for the Middle East, Africa, Mexico, eastern North America, and parts of Central America (see www.geocover.com for an up-to-date availability status). By the end of the year 2000, these products should be available globally.

2. GeoCoverTM-LC Program

GeoCoverTM-LC is EarthSat's global land cover and landcover change product derived from Landsat Thematic Mapper (TM) and Multispectral Scanner (MSS) data created under EarthSat's GeoCoverTM-Ortho program. The U.S. Government has currently funded the derivation of land cover from the TM data. Sources of funding for developing the landcover change product are being solicited. The 1990 landcover product will be derived from 7,600 GeoCoverTM-Ortho Landsat TM scenes (with mean acquisition date of 1990). Each of the products described below will be produced for the area covered by a single Landsat TM path/row footprint. EarthSat will also provide the vector landcover data in 1-degree by 1-degree cells.

2.1. Landcover Categories

The landcover categories in the GeoCoverTM-LC product follow closely to those defined by Anderson et al. in his 1976 publication, "A land use and landcover classification system for use with remote sensor data" (U.S.G.S. Professional Paper 964). Brief descriptions of the 13-landcover categories are shown below:

Forest

Deciduous – includes woody vegetation greater than 3 meters in height that lose their leaves periodically due to changing seasons or drought. The deciduous forest category will generally have greater than 35% canopy closure. Also included in this category are areas commonly referenced as "swamp" or forested wetland if dominated by a deciduous canopy.

Evergreen – includes woody vegetation greater than 3 meters in height that retain their leaves through various seasons. The evergreen forest category includes both needle leaf and broad leaf species. Some tree plantations, such as palm or date plantations may be included in this class. Evergreen forests will generally have greater than 35% canopy closure. Mixed evergreen and deciduous forest will be classified as evergreen forests if 25% or more of the trees are evergreen

Scrub/Shrub – includes woody vegetation less than 3 meters in height, both open and closed canopies. Minimum scrub/shrub ground cover is 10% for this category.

Grasslands – includes grass and herbaceous areas. It may at times include herbaceous

wetlands if the image was collected during dry periods. Landcover types often referred to as savannas or open savannas are included in this category.

Barren/Sparsely Vegetated – includes sand dunes, deserts, rocks, gravel, bare soil, and sparsely vegetated areas of grass or shrubs.

Urban/Built Up – includes cities, towns, wide roads, airports, or other developed areas. Areas of non-urban cover within the urban fringe are only separated from the urban category if they exceed 25 ha in size.

Agriculture

Agriculture, Other - includes all non-rice agricultural fields (with crops or in fallow), highly managed pastures and haylands (but not grasslands commonly referenced as "rangeland"), and complex mosaics of natural vegetation and croplands. Some orchards and tree plantations, such as palm or date plantations may be included in this class.

Rice Fields – includes paddy agricultural fields, mainly rice, that are seasonally inundated with water. Depending upon the date of the imagery, this category may be included in the Agriculture, Other class if the paddies are not inundated when the image was acquired.

Wetlands

Permanent Herbaceous Wetlands – includes emergent herbaceous wetlands, as well as other irregularly inundated areas that may not be vegetated including: mud flats, saltpans, and playas. Many of these permanent herbaceous wetlands are referred to as marshes. Areas typically called swamps, including forested wetlands, are not included in the wetland class (they are included in the forest class), with the exception of Mangroves that are included in the wetland, mangrove class.

Mangroves – includes regularly inundated coastal areas that are covered by mangroves.

Water – includes all type of water bodies such as oceans, seas, lakes, large ponds, reservoirs, and wide rivers. This classification does not separate these various types of water bodies.

Ice/Snow – includes only those areas that are covered by ice or snow at the time of image acquisition

No Data/Clouds/Shadows/Data Drops – Areas that are masked by clouds, cloud shadows, or smoke or areas that are otherwise obscured are categorized as No Data. This class also includes areas for which imagery is missing due to either sensor errors or data transmission failures.

2.2. GeoCover^{TM}-LC Products

There are four GeoCoverTM-LC products being created:

1) Landcover Raster File with No Minimum Mapping Unit in ERDAS IMAGINE format

This landcover file has the full resolution of the TM data (28.5 m) and has not been filtered or processed to a user-specified minimum mapping unit (MMU). Consequently, it will have inherent speckle and noise of raster data categorizations from remotely sensed data as well as artifacts from raster editing. The speckle and noise is removed when this data set is clumped to the minimum mapping unit. This UNCLUMPED raster file will be an 8-bit ERDAS IMAGINE file with cell values that indicate the landcover codes.

2) Landcover Raster File with a 5 ha Minimum Mapping Unit in ERDAS IMAGINE format

This landcover file will be processed to remove much of the speckle and noise associated with landcover categorizations based upon Landsat TM data. The only difference between this file and the previous file is the incorporation of EarthSat's CLUMP and ELIMINATE routines to generalize the land cover to a user specified minimum mapping unit. The CLUMP routine is used to define the size of all contiguous features with the same land cover code. The ELIMINATE routine removes any of these clumps that were smaller in size than the user-specified minimum mapping unit and the eliminated areas are filled by the predominate land cover surrounding the eliminated clump. The MMU may vary depending upon the feature (water vs. vegetation). In general, EarthSat's policy is that the minimum mapping unit should not be less than an area measuring 2 mm by 2 mm on map at the appropriate scale for these data. Landsat TM data are typically mapped at scales ranging from 1:100,000 to 1:250,000. Therefore, EarthSat generally uses a minimum mapping unit from 4 ha (200 m by 200 m if square) to 25 ha (500 m by 500 m if square). Note that the procedure employed

using CLUMP and ELIMINATE will not remove narrow linear features with areas exceeding the user-specified minimum mapping unit that would be removed if standard filtering procedures were used to create the minimum mapping unit. Currently, EarthSat is using a MMU of 5.0 ha for all categories except water. For water, a 0.2-hectare MMU (approximately 30m by 30m) is being used. This landcover raster with a minimum mapping unit will be an 8-bit ERDAS IMAGINE file.

3) TM Land Cover of the MMU Raster File in ARC/INFO GRID Format

The ERDAS IMAGINE minimum mapping unit raster files described above are converted into ARC/INFO GRID format. The raster files generated for each Landsat TM image are mosaicked and converted into 1-degree by 1-degree ARC/INFO GRID files. The ARC/INFO GRID is vectorized to create the ARC/INFO vector coverage of the landcover categorization as derived from the TM data for the 1990's era.

4) Vector Coverage of the 1990's Era TM Land Cover in ARC/INFO Vector Format

ESRI's ARC/INFO software is used to convert the GRID file as described above to an ARC/INFO vector coverage. GRIDPOLY is used is used to create the vector coverage. This coverage is reprojected from UTM to geographic coordinates and clipped precisely by a 1-degree by 1-degree polygon. This process creates 1-degree by 1-degree tiles of landcover data in geographic coordinates that are edge matched.

2.3. Geocover^{tm}-LC DATA PROCESSING STEPS

A brief description of the processes used to create EarthSat's GeoCoverTM-LC data base is presented below. The process uses EarthSat's IPS software for orthorectification of the TM imagery, ERDAS IMAGINE for the extraction of land cover from the TM imagery, and ARC/INFO to vectorize the derived land cover. Many of the steps utilize ERDAS IMAGINE Models that were created by EarthSat specifically for the production of this GeoCoverTM-LC product.

1) Image Selection – Landsat TM images selected were from those images available from the USGS EROS Data Center and various foreign ground receiving stations as described above for the GeoCover^{TM}-Ortho product. The same images selected for the GeoCover^{TM}-Ortho product are used to create the GeoCover^{TM}-LC product
2) Image Orthorectification - the images were orthorectified using EarthSat's IPS software. Elevation data used for orthorectification was LEVEL 1 DTED (supplied by the U.S. Government) or GTOPO30 elevation data. A sparse network of highly accurate ground control points were provided by the U.S. Government.
3) Ground Truth Collection - a variety of maps, data bases, ground photos, IGBP 1 km landcover data, and CORONA images have been collected and are the source of ground truth used for the landcover classification. Ground truth input for other sources is being sought and is welcomed.
4) Image Prints - for each TM scene, image prints at the scale of 1:250,000 are produced. Bands 4, 5, and 3 are printed to red, green, and blue. In mostly barren areas, band 5 is replaced by band 7 to reduce over saturation of barren features.
5) Ground Truth Delineation - on each image print, ground truth information is delineated onto the image prints. Polygons of contiguous samples of a single land cover are delineated on the 1:250,000-scale image prints and labeled with the appropriate landcover code.
6) Digitizing Ground Truth - the ground truth delineated on the image prints are digitized to create a raster file of the ground truth data.
7) Generate Spectral Classes - ISODATA is used to generate up to 240 spectral classes for each Landsat TM scene. Parameters used for ISODATA include:
   Maximum Iterations =20
   Convergence Threshold = 0.999
   X and Y Skip Factors =3
8) Maximum Likelihood Classifier - Supervised classification is used to create the 8-bit 240 spectral class files from the signature files created with ISODATA. The supervised classification parameters used are:
   Non-parametric rule: None
   Parametric Rule: Maximum Likelihood
9) Dendrogram - a dendrogram is created to graphically display the spectral class groupings and to aid in assigning the spectral classes to the required informational classes.
10) Summary Command with Ground Truth - the SUMMARY command is used with the ground truth file and the 240 spectral class file to identify the informational class to be initially assigned to the spectral classes and to identify spectral classes having high confusion.

11) Summary Command with Adjacent Scenes - the SUMMARY command is also executed with the finished classification file of adjacent scenes to help reduce differences in the classification from one scene to the next.
12) Assign Informational Classes - based upon the information provided in steps 9 - 11, the landcover classes are assigned to each of the 240 spectral classes.
13) Raster Edit - clouds, cloud shadows, terrain shadows, developed areas, and other areas of spectral confusion are manually edited.
14) Review of Land Cover - the derived land cover is reviewed by EarthSat's Quality Control (QC) Team assigned to the project. The QC Team may identify additional raster edits.
15) CLUMP and ELIMINATE - CLUMP and ELIMINATE models are run to create the final landcover classification file with a minimum mapping unit of 5 hectares for all categories except water. For water a 0.2 ha MMU is used. Any small group of contiguous pixels of a single landcover class with an area of less than the MMU is eliminated and replaced with the predominant land cover surrounding this small group of pixels. It should also be noted that islands, regardless of size are retained, and islands that are less than the minimum mapping unit are given the value of the majority landcover on that island.
16) MOSAIC TM Land Cover to Create 1-degree Cells – The raster data produced in step 15 are mosaicked into large blocks to cover many 1-degree by 1- degree cells.
17) Convert to ESRI GRID Format – The large mosaics of landcover are clipped into tiles with each tile being slightly larger than 1-degree by 1-degree cell. This ERDAS IMAGINE file is converted to ARC/INFO GRID format to enable vectorization of the landcover data.
18) Raster to Vector Conversion – ESRI's ARC/INFO software is used to convert the GRID file as described above to an ARC/INFO vector coverage.
19) Clip Vectors – the vector coverage produced in step 18 is reprojected from UTM to geographic coordinates and precisely clipped to 1-degree by 1-degree cells.

2.4. Application Of Earthsat'S Geocover^{tm}-Lc Products

The production of EarthSat's GeoCoverTM-LC products began in July 1999. The first 1200 of the 7600 TM scenes required to cover all land areas of the world will be completed in May 2000. The first 1200 scenes cover the Middle East, East Africa, and North Africa. It is anticipated that the global land cover product from TM data should be completed in 2003.

The U.S. Geological Survey's (USGS) Earth Resources Observation System (EROS) Data Center, the University of Nebraska-Lincoln (UNL) and the Joint Research Centre of the European Commission have generated a 1-km resolution global landcover characteristics data base for use in a wide range of environmental research and modeling applications. (See edc www.cr.usgs.gov/landdaac/glcc/globdoc1_2.html for more information on the 1-km resolution global landcover data base.) The GeoCover^{TM} -LC product provides global land cover at a spatial resolution 30 times better than the 1-km resolution global landcover database described above.

The issue of land cover change is becoming an important national and international policy question. International environmental treaties relating to global warming issues, carbon budgets, and carbon credits require detailed landcover data at the global scale. The GeoCover^{TM} -LC products will be a great aid in addressing these issues. Through the application of change detection methodologies, such as EarthSat's patented Cross-Correlation Analysis technique, various rates of landcover change can quickly be calculated.

For further information on EarthSat's GeoCover^{TM} -LC products contact:

Dr. Greg Koeln, David Cunningham, or Thom Jones
Earth Satellite Corporation
6011 Executive Blvd., Suite 400
Rockville, MD 20852-3804
Telephone: 301-231-0660
Fax: 301-231-5020
E-Mail: gkoeln@earthsat.com, dcunningham@earthsat.com, or tjones@earthsat.com