Documents-GLCF Reports

Proposal Type 2-ESIP

University of Maryland, College Park

A Landcover Earth Science Information Partnership

Contents

Introduction

The University of Maryland at College Park (UMCP) is proposing to establish a Land Cover ESIP. Development and operation of this ESIP is a logical development of the substantial terrestrial research activities which have been accomplished on this campus over the last decade. In 1986 Dr. Goward established the Laboratory for Global Remote Sensing Studies (LGRSS). Shortly thereafter Drs. Prince and Dubayah joined the Geography faculty and Dr. John Townshend was recruited as Department Chair. From the beginning the major emphasis of our research has been to address critical global and regional scale terrestrial Earth Systems Science issues that are central to NASA's Mission to Planet Earth. This is indicated in our publications since the mid 1980s (e.g., Goward et al., 1985; Townshend and Tucker 1984; Justice et al., 1985; Prince and Tucker, 1986.)

Our explorations of the value of available regional and global-scale satellite observations have caused us to pay considerable attention to various efforts to process and analyze AVHRR and Landsat observations (Goward et al., 1990; Goward et al., 1991; Goward et al., 1993; Townshend et al., 1995; Townshend 1989; Townshend 1995; Prince and Goward 1995; Goward 1989; Goward et al., 1990.) We further have examined other observation sources when they appear to provide important additional information (e.g., Dye 1992; Tucker 1989.) Our attention to data quality has always been driven by our science needs. Each of the LGRSS researchers has pursued specific aspects of Earth Systems Science (Townshend et al.,1987; Goward 1989; Goward and Kazantsev 1991; Prince 1991; Townshend et al., 1991; Dubayah 1992; Goward et al., 1994; Goward and Prince 1995; Prince and Goward 1995; Prince et al., 1995.) Recognition of our attention to both science and data quality is found in our current service to the NASA EOS program, as members of the MODIS team, the Team

Leader for Landsat-7, IDS membership, our participation in the Landsat Tropical Deforestation Pathfinder project and our leadership in the AVHRR Land Pathfinder for both versions 1 and 2, as well as leadership in the International Geosphere Biosphere Programme (IGBP) and the Global Climate Observing System (GCOS).

In the early 1990s we concluded that typical, existing computer facilities and data handling procedures were insufficient to address the large volumes of data needed for our research. In 1993, the Department of Geography joined forces with the University of Maryland Institute for Advanced Computer Studies (UMIACS) to explore use of advanced computational approaches to attack this problem (Davis and Townshend, 1993; Liang, 1995 #81.) This resulted in the awarding of a National Science Foundation Grand Challenge grant to UMIACS and Geography to resolve the data volume issue in regional and global-scale studies of land cover and other critical Earth Systems Science issues as well as the need for greatly increased computational speeds. This work has led to remarkable computational advances. For example, atmospheric clearing of Landsat TM images which would normally take 3-4 hours on a typical UNIX workstation we now accomplish in less than 10 minutes. We are also re-examining the basic approaches to remote sensing data storage and retrieval for global scale research (Schock et al., 1995.) We now understand these advanced scaleable computer technologies sufficiently to propose their use in Working Prototype Earth Science Information Partnership.

We believe that our long experience in processing and analyzing regional and global-scale terrestrial remote sensing measurements, combined with our recent progress in employing advanced computer technologies to solve basic data handling and processing problems, provides an exceptional opportunity to develop and exploit the ESIP concept. The recent founding of the Earth System Science Interdisciplinary Center as a collaborative activity between the University and the Goddard Space Flight Center, with its emphasis on inter-disciplinary research contributing to the program of Mission to Planet Earth (MTPE) is another key argument for locating the proposed ESIP at UMCP.

There is no single path to addressing the data handling and processing issues posed by the NASA Earth Observing System and MTPE goals. This University of Maryland team has significant experience in EOS/MTPE science and data handling, internationally recognized expertise across the wide range of skills needed to address the problems from a fundamental perspective, and enthusiasm for the advances in Earth System Science that can come from focused collaboration on developing the new data, information and analytical systems that are now possible.

UMCP Research Activities Relevant to Land Cover

UMCP has an extremely broad range of research capabilities associated with land cover research using remotely sensed data. We propose an ESIP to facilitate the distribution of data products and provision of analytical services based on the extensive program of on-going research on this topic at and research in computer science at UMCP. Current funds support either research leading to the creation of algorithms or, in some cases protype products and services, but we lack funds to enable sustained support of their provision to the broader user communities.

We envisage an expanding array of products and services as the ESIP evolves, based on the continuing research that we are carrying out and eventually also on collaboration with other research groups concerned with land cover. We are able to specify in some detail the initial products and services to be made available, but for later periods we deliberately give only general indications of the new products and services to be provided, since we wish to be able continuously to introduce innovation into all aspects of our work. The initial and subsequent products and services will, nevertheless, all conform to certain archiving and retrieval standards that we propose as fundamental innovations in this ESIP.

The principal land cover research projects from which we will draw are outlined below for the proposed ESIP. They clearly demonstrate the breadth and depth of our research in land cover and the leadership we have in this field. This perspective gives us a unique overview of the current community of land cover researchers and will facilitate our efforts under this ESIP to encourage community participation in the development and identification of new products and services.

Landsat Pathfinder Humid Tropical Deforestation Project
(PI John Townshend)

The NASA Landsat Pathfinder Humid Tropical Deforestation Project is a collaborative effort between the UMCP Geography Department, the University of New Hampshire Complex Systems Research Group, and NASA-Goddard Space Flight Center GIMMS Group. The project is focusing on the three regions where most of the tropical deforestation in the world has occurred--the Amazon Basin, Central Africa, and Southeast Asia (Townshend et al., 1995.) UMCP has responsibility for the non-Brazilian Amazon Basin (also known as the Pan-Amazon) and Central Africa. Results of the country-wide mapping for Bolivia and Peru are available through our homepage.

The prime application for the data set is the identification of changes in carbon stock for global modeling. But many other applications are developing, including monitoring of fragmentation of habitats; identification of new areas of future development for planning purposes; assessment of forestry resources; and land use changes in relation to National Parks and other protected areas.

Coastal Marsh Loss Mapping Project
(PI Michael Kearney)

The purpose of the Coastal Marsh Project, funded by a NASA Applications Center grant, is to analyze the surface condition of coastal marshes and detect areas that are at risk from rapid loss of land area (Kearney and Rogers 1995.) Thematic Mapper (TM) satellite imagery and National Wetlands Inventory (NWI) classification maps are combined in our study using mixture modelling procedures. Marshes are categorized into one of four different classes: (1) healthy, (2) moderately, (3) heavily, and (4) completely deteriorated, based primarily on the percentage of the marsh surface that is covered with water detected using spectral mixture modeling. So far, we have analyzed satellite images covering the east coast of the United States from Massachusetts to Georgia, for a time period from the early 1980s to 1990s.

High Performance Computing For Land Cover Characterization
(PI Larry Davis)

UMIACS, in conjunction with the Department of Geography, holds an NSF Grand Challenge grant for the application of high performance computing to land cover characterization. Among the many results from this work has been the development of procedures allowing very rapid atmospheric correction of complete scenes of Thematic Mapper data using the dark target approach (Fallah-Adl et al., 1996; Fallah-Adl et al., 1996,) generation of bidirectional reflectance distribution function products at a global scale, (Kalluri et al., 1997) and several new classification algorithms and a range of high performance computing tools for dealing with problems associated with the necessarily very high I/O rates.

Global hierarchical indexing system
(PI Samuel Goward)

Also as part of the NSF Grand challenge grant, we have developed a prototype global hierarchical indexing scheme that will optimize access and processing of remotely sensed data and higher level data sets (Schock et al., 1995.) The goal is to allow data sets to be stored, accessed and processed in an unsampled form so as to preserve data integrity. The procedure means that data sets can be fused in various ways without having to suffer from multiple reprojections and resampling. Currently we have implemented the system for a global AVHRR data set.

Global Land Cover Classification Using Coarse Resolution Data
(PIs Ruth DeFries and John Townshend)

NASA's Terrestrial Ecology Program provided funding from 1991 through August 1997 (to be continued by NASA's Land Cover land Use Change (LCLUC) program) for mapping global land cover from satellite data, with the aim of improving the accuracy of land cover data sets and to parameterize global change models. These efforts have produced a global land cover data set at one-degree spatial resolution (Defries and Townshend, 1994; Defries et al., 1995) based on a maximum likelihood classification of annual NDVI temporal profiles. This data set has been made available at a web page. An 8-km data set has subsequently been created and is ready for distribution.

Improved AVHRR Data Sets
(PIs Eric Vermote and John Townshend)

Over many years, since the inception of global remotely sensed data sets, we have contributed significantly to the development and improvement of algorithms and processing chains to advance their capabilities. This effort has been especially strong for AVHRR data where members of the Department of Geography took leadership roles both in the specification of the NOAA/NASA Pathfinder AVHRR Land data set (James and Kalluri 1994) and in the specification of the IGBP-DIS global 1 km AVHRR data set (Townshend et al., 1994.) More recently, as part of our NASA-funded second phase of Pathfinder activity, new algorithms and processing procedures are being implemented to create much improved land products from the AVHRR data stream.

Monitoring Global Primary Production
(PI Stephen Prince)

Net Primary Production (NPP) is a key ecological parameter that describes the fixation of atmospheric carbon by plants, and thus measures the potential supply of carbon to humans and animals for food fuel and fiber (Field et al., 1995.) Our GLO-PEM model (Global Production Efficiency Model) represents the first attempt to infer plant production and also the environmental factors that determine the actual growth (Prince and Goward, 1995; Prince et al., 1995.) Our goal is to estimate annual NPP from the entire AVHRR data set which starts from 1981 and continues to the present and to examine the variations in NPP from year to year associated with land cover changes, interannual vegetation dynamics and climate fluctuations.

Retrieval Of Atmospheric And Biospheric Variables from Remote Sensing Data

(PI Samuel Goward)

Accurate high spatial resolution meteorological and surface variables are needed to model land-atmosphere interactions in ecosystem and energy balance studies. We have developed and tested several algorithms to derive these parameters from satellite data during our participation in large scale field experiments (such as FIFE, HAPEX-Sahel and BOREAS) and through the Pathfinder 2 project. Surface temperature, air temperature, atmospheric perceptible water, and vapor pressure deficit were derived from AVHRR measurements alone (Prince et al., 1997.). Net solar radiation has been estimated using transportables oftware using Landsat Thematic Mapper corrected for topographic and atmospheric effects (Dubayah 1992.)

Principal Investigatorship Of The Vegetation Canopy Lidar
(PI Ralph Dubayah).

Dr. Ralph Dubayah is the principal investigator for the recently-announced NASA-funded Vegetation Canopy Lidar (VCL) instrument. The VCL is an active laser remote sensing system whose principal objective is the three-dimensional characterization of the vertical structure of the earth, including canopy top heights, vegetation canopy vertical distribution, and sub-canopy topography. The information provided by this instrument is unique, and is expected to provide information on global vegetation biomass and topography that is highly relevant to understanding biogeochemical and hydrological cycling, land use and land cover change.

Memberships of MODIS Team
(PIs John Townshend and Eric Vermote).

John Townshend is a member of the MODIS Team generating algorithms for the creation of a 250m resolution land cover change indicator product, which in a simplified form can also be used as a land cover change alarm product to assist acquisitions strategies data from fine resolution systems such as Landsat 7. Algorithms to create 500m and 250m land cover products depicting continuous fields of land cover components are also being produced.

The work of our second MODIS team member, Dr. Eric Vermote, is on the atmospheric correction of the MODIS visible bands. Previous operational correction schemes have assumed a standard atmosphere with zero or constant aerosol loading and a uniform, Lambertian land surface. The MODIS operational atmospheric correction algorithm will use aerosol and water vapor information derived from MODIS data, correct for adjacency effects and take into account the directional reflectance properties of the observed surface.

Membership of Landsat 7 Science Team
(PI Samuel Goward)

Dr. Samuel Goward is currently a team member and Team Leader of the MTPE Landsat Science Team. He is designing an end-to-end Landsat data acquisition, handling and processing system which will achieve current MTPE science objectives in global-scale land cover/land use assessment. The critical contribution of Landsat to MTPE land cover assessment is that it provides the combination of spectral integrity, spatial detail, and temporal repetitiveness (SST) which captures and delineates the spatio-temporal dynamics of the Earth's land area in sufficient detail to assess location and cause of lands cover change.

Land Cover and Land Use Change in Central and Southern Africa

(PI Stephen Prince)

As part of NASA's LCLUC program Dr. Stephen Prince is devloping an integrated land degradation and deforestation detection system for the Southern African Development Community (SADC) region plus southern Zaire. The scale for the inventory and monitoring will be 1 km^2. Using our past experience with degradation studies in Africa, the analysis of very large volumes of Landsat data, socio-economics of land and fuelwood, and inference of biophysical variables for large areas from remotely sensed measurements, we will map land cover and biophysical properties of the land surface related to degradation, thus moving beyond classification of land cover to monitor the processes involved. Socio-economic drivers of land cover change as well as biophysical factors will be employed to select processes that can be expected to cause degradation and to choose representative study areas. Radar and optical methods will be implemented to measure biomass. Primary productivity of crops, rangelands and forests will be monitored using models driven by remotely sensed data. Soil moisture and runoff will be derived from surface water and energy balance models also driven with remotely sensed data.

Data Sets and Products To Be Provided By the Land Cover ESIP

Selection of the specific data sets and products for any given year will be carried out through the process described in Section 5 (Management). The basic rationale behind our approach involves careful internal and external review of the results of both our own existing research projects and those of others to assess whether the results have a maturity appropriate for product and service distribution by the Land Cover ESIP. The review process will identify the highest priority products and services arising from the land cover research activities. Those selected will then be incorporated into the ESIP and the ESIP will then be the principal mechanism for their distribution to the external user community.

We are proposing both improved remotely sensed data products which can be used for land cover research, advanced land cover products and in some cases highly portable software for users to apply the procedures at their own facilities (all software will be made available to users requesting it)

YEAR 1: Data sets and Products to be provided (1997-98)

Based on an evaluation of the results of our current research and scientific priority we have identified a number of products from our research projects, which have reached a level of maturity and timeliness such that they are suitable for immediate distribution through our proposed Land Cover ESIP. A final decision on the release of these would only be made following input from the Strategic Advisory Team described in section 5. Details of the supporting research on which the products are based is given in section 2.1.

Atmospherically corrected Thematic Mapper data on demand for users.

Using the procedures developed under our NSF Grand Challenge grant (Fallah-Adl, et al., 1996; Fallah-Adl et al., 1996) involving implementation of the dark target approach (Kaufman and Sendra 1988) we will correct Thematic Mapper images for aerosol and other atmospheric effects. This method can do much to remove the effects of thin clouds as well. Users will provide their own data sets unless we have the data in-house. Users will receive advice on the appropriateness of the approach for their data, since in the absence of suitable dark targets within the scene the procedure is not effective. Users will receive full resolution data sets with all the reflective bands corrected for these effects. We will also provide our software for this correction so that users can carry out their own corrections if they so wish.

Coastal wetlands health products.

We propose providing products from our coastal wetlands research; these are completed for much of the coastal eastern US. The data layers will include 1) visually enhanced National Wetlands Inventory products, 2) enhanced Thematic Mapper data - (depending on date and copyright restrictions, charges may have to be made for this layer to the copyright owner) and 3) a layer in which the marshes are characterized according to health categories. The products will involve provision of data in flexible units of aggregation including Thematic Mapper scenes, US map quads or by local government unit.

Tropical humid rain forest land cover

Wewill use the Land Cover ESIP to operationalize the distribution of the completed products from our work on the Landsat Pathfinder Humid Tropical Rain Forest project for those areas for which UMCP has responsibility; namely, the pan-Amazon (Bolivia, Peru, Ecuador, Colombia Venezuela and the Guianas) and central Africa (principally the Republic of Congo, Democratic Republic of the Congo and Cameroon). Products showing the distribution of closed tropical rain forest, and areas of clearance and regrowth will be made available at the full resolution of the product (128m). Data will be provided on flexible units of aggregation, principally by Landsat scene, multiple scenes from a Landsat row, or by country. Comprehensive meta-data describing the classificationr esults will be provided with each product. This will include information on product quality including information and commentaries on the analyses from in-country experts and the published documents used to generate the products. Change products between the 70s and 80s and the 80s and 90s will also be available at a resolution of 16 km. (The lower resolution relates to the need to reduce errors introduced by spatial misregistration).

We regard this effort as being highly complementary to the proposed ESIP of Dr. David Skole and others, which is concerned with tropical rain forests in Brazil and south east Asia.

Coarse resolution land cover products.

Over the last three years we have generated global land cover products at resolutions of 1 degree and 8 km using the IGBP-DIS land cover classification scheme, based on AVHRR data. They are widely requested and we will make these available through the proposed Land Cover ESIP. We anticipate they will be of particular use in modelling work at regional and global scales. Associated with these global products are the training and testing data sets based on Landsat data used in their production. Partial sets of these Landsat data have already been published on the Web. The additional resources available through the Land Cover ESIP will allow a much more comprehensive set of data for training and testing to be released. We will make available the original Landsat data (none of these data has copyright restrictions), and the data layer identifying the land cover types present. The meta-data describing the source of information used in identifying the cover types will also be provided. All these data sets will be provided in a format that is registered to the Pathfinder 8-km data set in its Goode projection.

YEAR 2: Data Sets and Products to be provided (1998-99).

In subsequent years we envisage adding products to those described above. The identity of these will be dependent on the progress of our research and collaborations with other land cover researchers, who may wish to collaborate with us in distributing their products through the Land Cover ESIP. Therefore, the following products and services should be considered preliminary and subject to the review and evaluation process described in Section 5.

Production of user-defined 1 km AVHRR products for the land.

We will use code developed as part of our Pathfinder 2 project to allow users to specify the type of product they want derived from 1 km AVHRR data, based on choices of map projection, type of compositing procedure and choice of atmospheric correction algorithm. We intend to acquire from EDC all of the Level 1B composited orbits of the global 1 km data set and stage them to allow users access to this type of processing for any region. We will offer this flexible processing possibility **only** for options other than the standard processing being offered through EDC. Currently EDC offers a product based solely on the initial IGBP recommendations. It is likely they will wish to offer a version 2 product based on our research under Pathfinder 2. It is our understanding that this product will consist of a single non-optional processing chain without the flexibility we propose to offer. Initially we will only make our products available for regional subsets, since provision of global data sets based on multiple options could easily overwhelm our proposed capacity.

We will also consider providing products based on other 1 km data sets, assuming computing resources are available, since the EDC 1 km data set only represents a fraction of the 1 km data available. We will entertain requests from users to process other 1 km data and will use the various management structures described in section 5 to decide on whether to respond positively to requests for such processing. For a maximum specified computing load (provisionally set at approximately a 1000 by 1000 pixel image for a year), we will carry out on-demand processing of products and supply these in near-real time.This will form an important prototypE activity in assessing the extent to which non-standard products are required and the logistics involved in supplying these to the user community.

Atmospherically corrected products from Landsat 7.

We will provide atmospherically corrected Thematic Mapper products on a similar basis to that proposed for those from previous Landsats (see *Year 1*). Additionally we intend to use products from MODIS, with its channels designed specifically to assist atmospheric correction, to carry out improved correction of Thematic Mapper data. This will be facilitated by the similar overpass times of the two platforms.

Improved global land cover products.

Our Pathfinder AVHRR project has proposed a number of substantial improvements in the processing of the GAC and LAC/HRPT data sets. We anticipate that these data sets will be available by the second year. Using these we will generate revised global land cover products using these data sets. We will be incoporating advanced decision-tree procedures and object-based algorithms derived from our NSF Grand Challenge project.

<u>*YEAR 3: Data sets and Products to be provided (1998-99).*</u>

In subsequent years, we envisage several other products' becoming available, dependent in part upon our current research and collaborations ith other groups. For example, we currently have bids for resources to allow us to carry out large scale integration of multiple remotely sensed data sets, including Landsat and MODIS. We would anticipate using the ESIP to assist in the distribution of the resulting advanced products.

Processing of data through the Global Hierarchical Indexing System.

We propose to begin handling, processing and (as desired) delivering data sets in the Global Hierarchical Indexing System structure developed as part of our NSF Grand Challenge project, in which data sets are stored in unresampled forms and are located through an advanced indexing system. Use of this system means that the full integrity of the data is preserved. Such an approach will be of particular value where it is necessary to fuse two or more data sets with inherently different internal geometries. We anticipate that this approach will eventually form the procedure for handling all data in the ESIP. A working prototype already exists, so it may be possible to advance this capability to the second year of ESIP operation.

Advanced Land cover Products from MODIS and Landsat.

For both MODIS and Landsat projects described in the section 2.1, we have already defined procedures for the creation of standard land cover products and distribution through EOS-DIS and existing DAACs. However, continuing research is leading to the creation of more advanced products. As these products mature, we will use the ESIP to allow early production and release of prototype data sets from this work

Additionally, we currently have under consideration bids for a major enhancement of computing resources for a Remote Sensing Center of Excellence that will allow us to carry out large-scale integration of multiple remotely sensed data sets including Landsat and MODIS. We would use the Land Cover ESIP for the distribution of the resulting advanced products, whereas we anticipate that the standard EOS products we are contracted to generate would be distributed through EOSDIS as planned.

Degradation early warning system (DEWS)

The aim of Dr. Stephen Prince's work on a Degradation Early Warning System (DEWS) is to develop a prototype that can be applied to the whole of southern and central Africa and provide a pattern for similar areas worldwide. By the third year of the ESIP we anticipate that we will have established a prototype DEWS. FEWS has a routine protocol for famine early warning; we will propose a parallel protocol for a DEWS. The DEWS will contain a range of techniques developed in the project from which an operational DEWS can select. The protocol will be designed so that it can respond to existing and new threats to the environment, optimize the use of a range of remotely sensed and human indicators and add value to the available information through the use of degradation models.

Annual Assessment of the State of Global Land Cover Change

Several of our currently funded projects, notably Landsat and MODIS, are concerned with the generation of land cover change. Using this information, we intend to provide an annual assessment of the state of the Earth's land cover transformations starting in Year 3. This assessment will contain indicators of the distribution of the main types of land cover change and will provide estimates of the intensities of change. Embedded within the text of the assessment will be hyperlinks to the Landsat and MODIS products and data sets used in deriving the assessments.

Services to Support Earth System Science Research

Our Land Cover ESIP will provide the following main types of service to user communities:

1. *Supply of standardized products derived from existing and on-going land cover projects (see Section 2.2).* Through a process of selection involving internal and external advisors, as described below, we will provide products of high scientific priority in user-friendly forms.

   On-demand processing of certain data sets to create advanced products according to options selected by the user (see section 2.2). For selected products we will allow substantial user input in specifying their characteristics.

   High level global assessment of land cover change with active links to the data bases on which the assessment is based.

   Supply of transportable computer code for selected products and services so that users can apply the code to use on other computing systems. This capability is an important one that will promote more rapid progress in the development of advanced products.
2. *A web interface that will allow users to browse meta-data and data sets to assess how their needs can best be met.* This interface will assist users in making best judgments about whether any particular products will assist their goals. Provision of interfaces to meta-data and carefully designed browse products will substantially reduce the user effort needed to identify the required products.
3. *High level science support of the products* (see section 2.5). Many levels of user support will be provided. To ensure the highest possible levels of scientific support we will provide a hierarchy of support levels as described in section 3.
4. *Documentation of data products, software and services.* Information on these topics is necessary to provide users with best opportunities to assess and use the provided products. They will be provided both as web-based information, as hard copy and through several other mechanisms including workshops and seminars.
5. *Provision of advice and cooperation with other members of the federations* (see section 3).

All data sets and products supplied by the Land Cover ESIP will be accompanied by metadata, browse data sets, documentation and science support. Information about the services of the ESIP will also be publicized by a wide variety of means for example on the Web, by presenting papers at professional meetings organized for example by the AGU (American Geophysical Union), IGARSS (IEEE Geoscience and Remote Sensing Symposium), and ASPRS (American Society for Photogrammetry and Remote Sensing) and through journal publications.

Innovative Approaches for Data Processing, Archiving, Distibtution and User Support

Standard products inevitably pose problems for many users who need data customized in specific formats to analyze them in their own computing environments. In addition to providing easy and efficient access to the multiple sets of the data products described in section 2.2, we intend to offer the possibility for user-specified products starting with regional samples of the AVHRR 1 km data set. Reprocessing of large volumes of data sets will be scheduled to reduce impact on the overall operation of our Land Cover ESIP. A process will be set up to handle such requests from the users. For some products, the processing of relatively limited data sets will be permitted on demand.

To enable the generation of customer-specified data products and efficient access to the AVHRR and Landsat data products, we propose a robust and cost-effective architecture that is scaleable in terms of size and functionality of EOSDIS. This architecture consists of:

- A balanced, high performance, integrated compute- and data-server capable of 1 GFLOP of peak performance, with a 10 TB tape subsystem and a 160 GB of disks with a transfer rate up to 40 MB/s for tertiary storage.
- A software infrastructure that offers a distributed programming environment, automatic data management between the different levels of the storage hierarchy controlled by an object-relational database, and a novel Web-interface design, all tailored for easy integration into the Working Prototype ESIP Federation.

We have had considerable experience in using, developing and enhancing such technologies through the UMIACS Laboratory for Parallel Computing, which currently houses a 16-node IBM SP2 system with 240 GB of disks and 3-TB of tape subsystem, and an ATM Cluster of 10 DEC SMP Alpha nodes, each node containing 4 Alpha processors, and through the Human-Computer Interaction Laboratory (HCIL). The current UMIACS SP2 system will be used to develop prototypes of many of the data products.

UMIACS has been selected as an Earth System Science data repository partner of the recently announced NSF award to the San Diego Supercomputer Center, and therefore will be able to transfer many of the high performance technologies for data-intensive computing and for handling distributed digital libraries directly into the proposed ESIP and the WP-Federation. In addition, we have an excellent networking infrastructure that includes a campus-wide high speed ATM network (currently managed by UMIACS), an OC3 connectivity to NASA at Goddard Space Flight Center through the ATDNet, and, soon to be established, an OC3 connectivity to the NSF-funded vBNS network.

Data Processing and Archiving Hardware

Our intent is to set up an integrated high performance data processing and archiving system that provides a distributed, scaleable, and open computing and data access environment. The data products and information services of our proposed Land Cover ESIP impose several requirements in terms of storage capacity, I/O throughput, computational speed, and operating environment. We consider these requirements first in terms of raw storage capacity and then in terms of total system throughput in the expected operating environment.

The activities conducted by our Land Cover ESIP require adequate storage to handle both a large volume of "raw" (level 1) remotely-sensed data sets as well as standard and user-defined products. Based on the AVHRR and Landsat raw data currently available to us and the data products that we intend to provide, we have set the storage requirement to 10 TB of tertiary storage and 160 GB of disks. With respect to the throughput requirement of such a storage system, consider, for example, the production of user specified 1 km AVHRR data products. Some studies require applying a specific set of processing algorithms to multiple years of level 1B AVHRR data. It is clear that such processing requires fast access to and manipulation of large amounts of data residing in tertiary storage. We have set a transfer rate of 20 MB/s for tertiary storage and 200 MB/s between the disks and the main memory, rates which will allow a 10-TB archive to be rotated onto the disks in approximately a week.

The computational heart of the system must strike a compromise between several disparate requirements. We expect heterogeneous usage and must support co-existing batch and interactive jobs as well as sequential jobs. A wide variety of users also demands a flexible and robust operating environment. The compute engine should not only be computationally fast , but should co-exist well with other computing resources. We have selected IBM's SP-2 platform as an integrated compute and data server. UMIACS researchers and technical staff already have extensive experience with the UMIACS SP2, currently supporting a wide variety of users. The SP-2, with its native software suite, has proven effective at handling concurrent heterogeneous processing tasks. The software configuration for the SP includes all of the necessary elements for system management, monitoring, job scheduling, and development. Several years after its installation, its user community has grown to include a wide variety of applications users. We will use our current SP2 to prototype many of the products to be developed under the proposed Land Cover ESIP, which will then be transferred to the compute/data server system requested under this proposal.

The proposed SP frame houses five nodes connected through a high performance data switch. The switch allows nodes to exchange data at 100MB/sec. The five nodes are divided into one "high" node and four "super thin" nodes. Each thin node contains a single 120 Mhz POWER2 processor, achieves 234 MFLOPS, and is configured with 128 MB of main memory to allow large data sets to be processed completely in core. The high node contains two 112-Mhz 604 processors and 256 MB of main memory and, while available for computation, is primarily intended as a data server. The combined system will deliver a peak of about 1 GFLOPS. The 160 GB of fast disk will be distributed across the nodes using 4 of IBM's SSA adapters per node. This yields sustained throughput of 35MB/sec per SSA card for a nominal aggregate disk throughput of 700MB/sec. We expect the actual throughput to be around 400 MB/sec.

Finally, we have selected StorageTek's Infinity 9710 Cartridge Tape Library equipped with eight Quantum DLT7000 drives for tertiary storage. Each drive is capable of 5MB/sec sustained data throughput, and each DLT cartridge holds 35 GB of data with no compression. We chose DLT because it is a common format and the large capacity of an individual cartridge reduces latency from tape swapping. The unit will hold 588 DLT cartridges for an overall capacity in excess of 20 TB. Aggregate tape throughput is 40MB/sec. The tape library will be connected to the SP's high node using an SCSI-3 interface rated at 40MB/sec. We will acquire the ADSTAR Distributed Storage Manager (ADSM) software to support our tape library, which will allow the users to readily access data from tertiary storage.

Storage Organization and Management

We have set the following ambitious but achievable goals for our hierarchical storage management:

1. Efficient search across multiple data sets by queries on metadata attributes through a Web interface.
2. Ease of incremental evolution, derivation, and integration of metadata and data sets.
3. Ability of cross-correlation and knowledge discovery through data mining.
4. Automatic data handling and movement between the different levels of the storage hierarchy.
5. Ease of integration into the ESIP WP Federation.

These goals impose a number of challenging performance requirements on the file management system, metadata modeling and management, and database management system. The issues related to Web and interaction technologies will be addressed separately. We focus here on the storage management and access issues. Our overall approach will be based on an object-relational database management system and a hierarchical indexing scheme that is tailored for efficient access and manipulation of remotely sensed data at different levels.

A hybrid object-relational database will support declarative, multi-resolution and temporal queries on top of relational systems, with the more interactive, navigational aspects of object-oriented database (OODB) systems. Such a hybrid database management system (DBMS) will allow the creation of a "meta-catalog" that integrates metadata from heterogeneous databases and other data systems. We will use reference types to link object classes and collection types to support multi-valued attributes, special purpose methods and "overloaded" operators. We will use Illustra as the underlying commercial DBMS, since it offers the advantages of OODB modeling and the ease of writing SQL queries. It supports SQL3 OO features such as inheritance, polymorphism, and user-defined data types. One of the most important features of Illustra is its extendible data types that are crucial for implementing complex searches and in particular, its spatial data type with its underlying R-tree indexing. Illustra also allows access to relational systems (such as Sybase) through special type and access method extensions called datablades.

Our UMCP Data Base group has built a similar system for astrophysics meta-data catalogs which became operational a year ago (Cheung et al., 1995.) The AMASE prototype development is a joint effort between the ADF at GSFC and the University of Maryland. The development was funded by the NASA Space Science Technology and Information Systems Office from 1995 to 1996 and currently renewed for another three years of funding. In AMASE we concentrated on building the fundamental capabilities that allow multispectral queries and have populated the database with data from the ROSAT and IRAS missions.

Encouraged by the success of AMASE, we propose to further exploit the OODB technology for accessing and managing multi-temporal, multi-resolution, and multi-mission earth system science data archives. We will capitalize on the "richness" and "extensibility" of OODB in order to create a "meta-catalog" that truly integrates metadata from heterogeneous databases and/or systems. We propose to develop a query engine for scientific search and discovery which supports object classification, object cross-identification, and knowledge discovery through data mining.

We will use emerging computer science technologies to provide scientists and other users with timely and efficient access to large data sets. The methodologies that will be used for the engine include positional queries, positional and attribute-based cross-identification of objects, multiple and user-defined views, and automated meta-data capture that are directly applicable to many scientific disciplines.

If the WP federation is to have a significant impact on the scientific community, the core of the effort should be devoted to data integration. NASA and several other agencies have created many data centers and accumulated very large amounts of data, but the level of integration is very limited. The reason for this is that most centers concentrate on data entry, usually with very restricted data integration. Data entry just collects data as a byproduct of missions rather than taking a "data-integration-centric view" that recognizes the reason for these missions is the discovery of the information and that integration and linking with the previous discoveries is the goal, rather than just collecting unlinkable data sets. Therefore, this proposal departs from the previously "mission-centric view" and treats the data integration and products as the prime goal.

In the AMASE project, 85% of the effort was devoted to the loading and linking of data sets. Data cleansing and data linking were the most time- and resource-consuming tasks. For this reason, we propose to concentrate the core of the effort in formalizing and creating appropriate tools for managing the evolution, derivation, and integration of data sets. The tools that we propose to build include bulk incremental loaders, pipelined data loaders, data derivation capture managers, and other pre-processing for integration and linking (Xue 1997.)

The OODB approach reflects the "data-integration-centric" view, as loading and linking prepares the data for the scientist, as opposed to the data-centric one of the Relational Database (RDB) which leaves all the linking and cross-referencing to the scientist. Among features of the OODB, we will use reference types to link object classes and collection types to support multi-valued attributes, special purpose methods and "overloaded" operators. These will allow us to take advantage of the hierarchy of classes and its "property inheritance."

The storage and data movement will be supported by the ADSM file management system that will allow transparent access to the tape subsystem and, in particular, will allow applications to store and retrieve files in a secure manner without human intervention. In addition, we will develop a hierarchical indexing scheme that will optimize access and processing of level 1B and higher level data sets. The goal is to allow data sets to be stored in an unsampled form so as to preserve data integrity.

Processing Software and Data Production

We will provide a distributed and open, heterogeneous computing environment that will support both batch and interactive jobs as well as sequential and parallel jobs. The basic software configuration will include the necessary elements for system management and monitoring, job scheduling, and parallel application development and execution. The IBM AIX provides the low level functionality such as network I/O, file system management, backup and restore capabilities and other basic administration utilities. Other high level packages installed include the LoadLeveler software to handle job management such as allocating and maintaining system resources, load balancing, and scheduling for batch and parallel jobs. The application environment will include all the necessary elements for software development, including FORTRAN, C, HPF, MPI, and PVM. It will also include the tools for code development, debugging and profiling.

In addition to the basic software infrastructure, we will transfer some of the software and tools developed through our NSF-funded Grand Challenge project on the development of high performance computing technologies for land cover dynamics. One type of software encompasses processing algorithms and their efficient implementations on current and emerging high performance computing platforms such as the system requested under this proposal. Algorithmic efficiency and the optimization of data movement between the different levels of the storage hierarchy are of utmost importance, especially given that we will provide custom-tailored products and on-demand processing.

Our past work (Fallah-Adl et al., 1996) has shown several orders of magnitude improvements in execution times (including I/O time) compared to traditional methods. As an illustration, we have recently developed an optimized global BRDF (bidirectional reflectance function) retrieval software and applied it to the Pathfinder AVHRR Land data covering a period of four years (1983 through 1986), using three widely different models, namely the modified Walthall model (Nilson and Kuusk

1989) , the Rahman model (Rahman et al., 1991), and a new temporal model developed by S. Liang (Liang and Townshend 1997.) For example, the global BRDF retrieval of 30 GB of data using Walthall's model took less than four minutes on our SP2 system, including I/O time (Kalluri et al., 1997.) Another example is the atmospheric correction of complete Thematic Mapper scenes (Fallah-Adl et al., 1996) based on Kaufman's dark target approach (Kaufman and Sendra 1988), which we can complete in approximately 10 minutes per scene, including I/O.

Development and Application of Web Technologies and Visualization Tools

Query languages are effective for frequent users who are also willing to take training, but more robust querying interfaces are needed for a system that supports intermittent users, from scientists to high school teachers, with very varying levels of computer experience. Users will come to the WP-Federation information systems with problems that vary according to several factors, including how well-defined is the solution to the problem and how well-defined is the problem in the information seeker's mind. Thus information systems must accommodate different levels of experience with the data, with the information system itself, and with information seeking in general. Classic form fill-in interfaces for EOSDIS exist, but zero-hit queries are a problem as well as queries returning too many hits. It is difficult to estimate how much data are available on a given topic and how to refine or relax a query to return a useful set of hits.

In the past two years, the Human-Computer Interaction Laboratory (HCIL) at the University of Maryland has been developing prototypes of highly interactive query interfaces for EOSDIS using the principles of dynamic queries and query preview. Dynamic queries have been shown to be an effective technique to browse information and encourage exploration, as well as to find patterns and exceptions. Dynamic queries involve the interactive control by a user of visual query parameters that generate rapid (100 ms update), animated, and visual displays of database search results (Ahlberg and Shneiderman 1994.) We have developed novel algorithms and data structures that guarantee real time interaction (less than 100 ms updates) for 100,000 records and 10 attributes (Tanin et al., 1996.) To overcome the problems of slow networks and data volume, we have also developed a two-phase approach to query formulation using query preview and query refinement, both following dynamic query principles (Doan et al., 1997.) Our work has been very well received by the EOSDIS community and has reached a level of maturity such that a working prototype is now being developed for an operational system (Global Master Change Directory). We believe that preview interfaces will be a major component of successful networked information systems interfaces (North et al., 1996.)

For this project, we propose to work on the following tasks for the proposed Land Cover ESIP:

- Refinement of the dynamic query and query preview.
- User interfaces for on-demand data processing.

We will integrate visualization tools with our Web interface coupled with the Geographic Information System (GIS) ESRI ARC-Info for which the University has a site license. In-situ spatial information will be handled by the GIS and correlated with the remotely sensed data.

Dynamic query and query preview interface for the UMCP ESIP

Our Human-Computer Interaction Laboratory has developed prototypes of dynamic queries and query previews for EOSDIS. Those prototypes were well received by EOSDIS users during a prototype workshop organized by Hughes in Landover. Our prototype interface works as follows:

Users first select rough ranges for a few attributes (time, location and parameter) in the query previewer (Figure 2). The impact of their selections is shown on the preview bars, which are dynamically updated to reflect the number of data sets available: e.g., when a user selects North America (Figure 3), the preview bars reflect the distribution of data sets for North America. The query preview interface makes use of data set counts maintained by providers about their holdings, and downloaded when users initiate a query session.

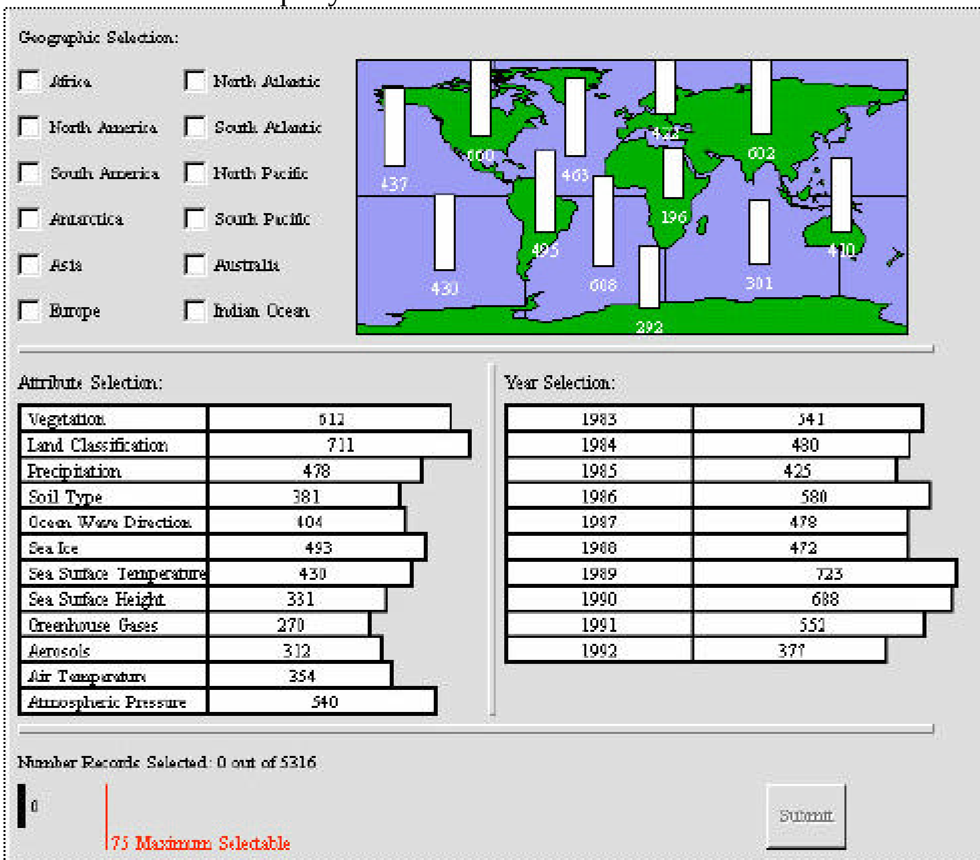

Figure 2(a): The Query Provider displays on preview bars aggregate data about all EOSDIS datasets. Users learn about the holdings of the collection and make rough selections over a few parameters (here locations, parameter and time). The preview bars are updated immediately. The result bar at the bottom shows the total number of selected datasets.

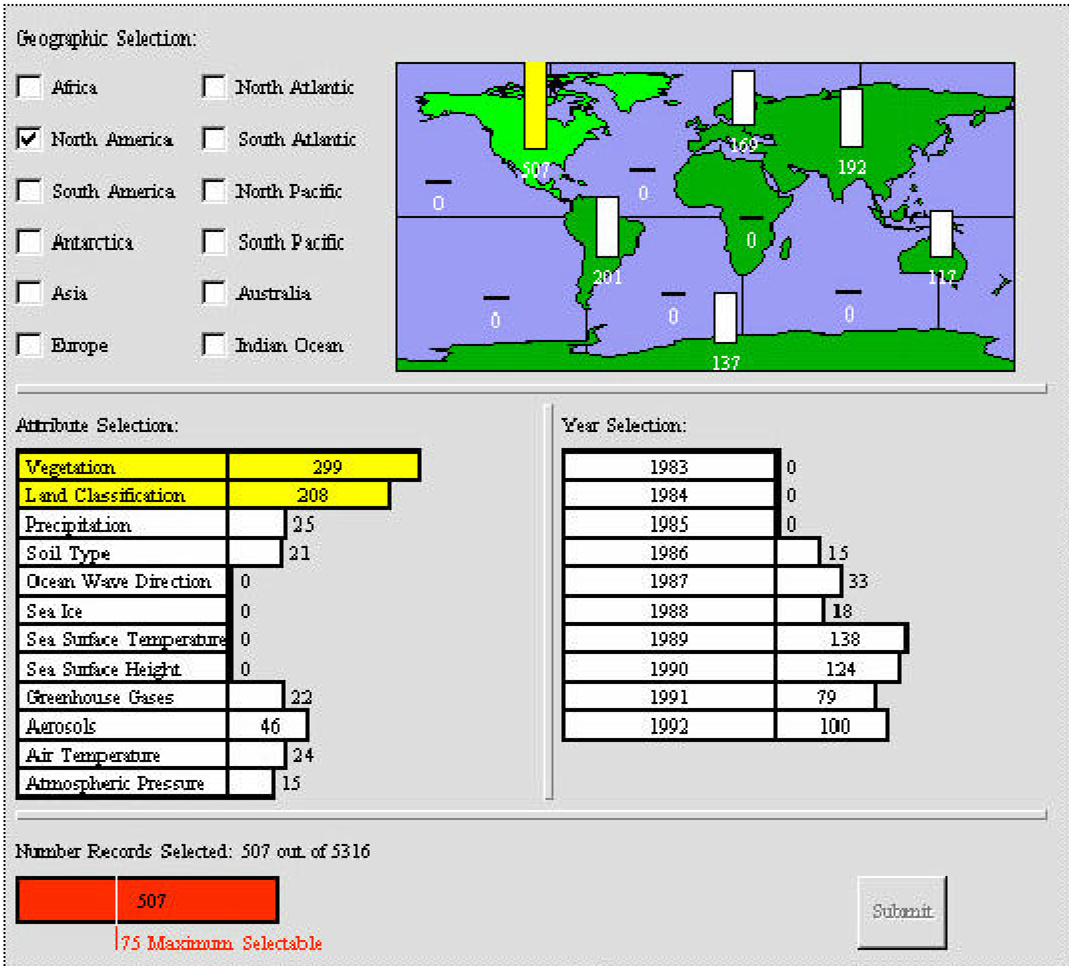

Figure 2(b): North America and two parameters are selected. Next, years will be selected and the query submitted to request more details about the datasets.

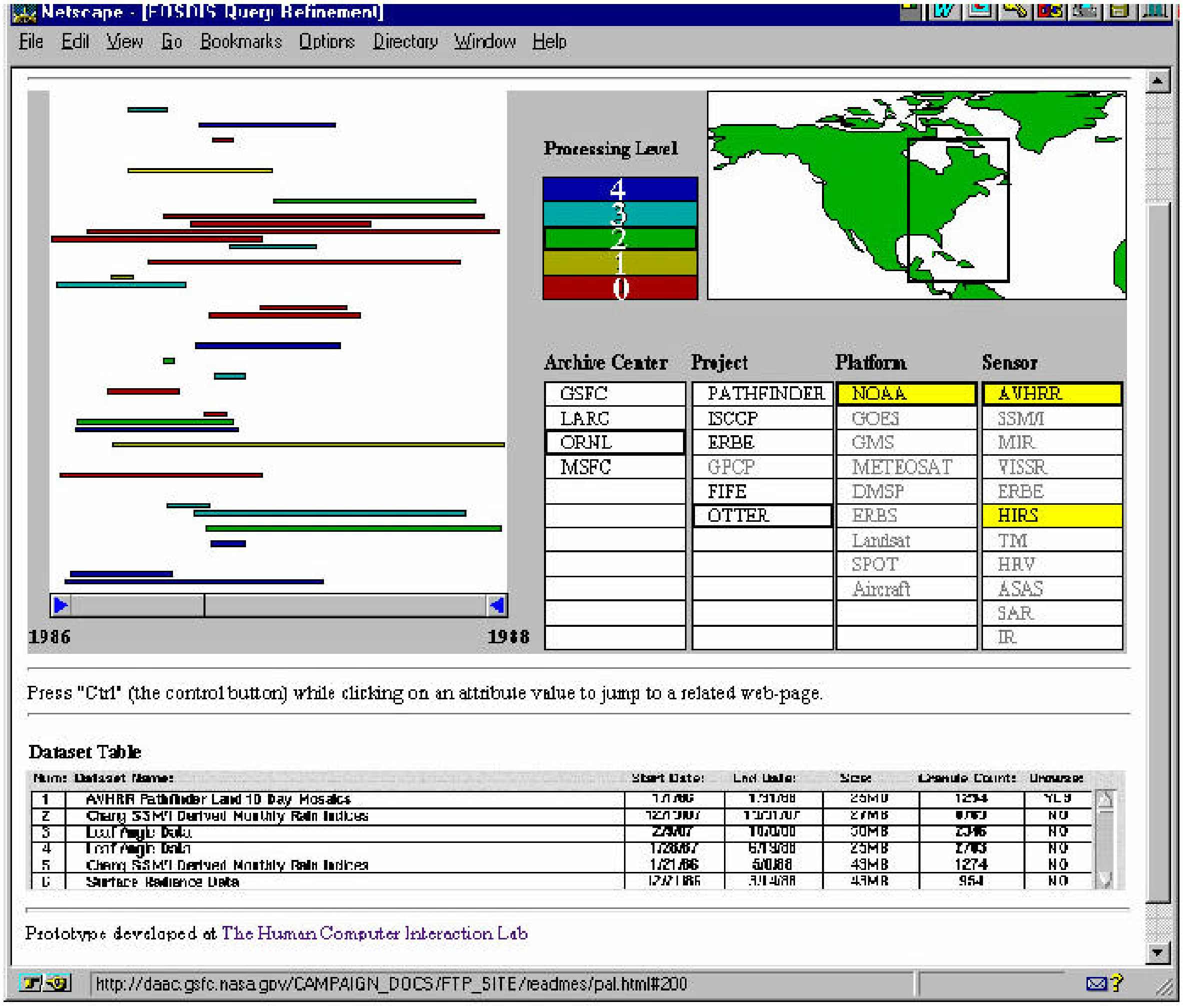

Figure 3: In the Query Refinement users can browse all the information about individual datasets. The result set is narrowed again by making more precise selections on more attributes. Sample data can be viewed before the long ordering process.

When the number of data sets is small enough, the metadata corresponding to the query preview is downloaded for further exploration in the query refinement phase. A second dynamic query interface allows users to specify precise values for more attributes and further filter the result set. The timeline shows the coverage of the data sets, already zoomed on the years selected in the query preview. Large data sets appear at the top, small ones at the bottom, color coded by processing level. An active cursor highlights the corresponding attribute values such as location, sensors, campaign, data center. A recent user study confirmed the benefits of query preview for imprecise queries as the information about the holdings helped users rapidly refine their query (Tanin et al., 1996.)

Refinements are currently being made in conjunction with an operational EOSDIS directory information system (Global Change Master Directory.) In the fall of 1997, we will have a working web prototype available for user testing. But current EOSDIS directory systems deal with a few thousand data sets (about 5000 for GCMD) while inventory systems (i.e. systems allowing queries at the granule level such as expected at the Land Cover ESIP) will have to deal with a much larger number of records and a larger number of attributes. The maintenance of the query preview tables also becomes an issue. We will refine our algorithms but we are confident that we can support highly interactive query interfaces for such inventory systems.

User Interfaces For Earth System Science Data Processing

The current interfaces of EOSDIS have been limited to the searching and ordering of data sets. For our Land Cover ESIP, the users will not only want to retrieve data but also require or apply different types of processing for calibration, projection, atmospheric and topographic correction, classification, etc., so that multiple data samples can be processed to produce a new product satisfying users' needs.

Typically, those operations have been performed with either *ad hoc* programming tools or with complex science data toolkits that are not widely available to the public. Therefore, there is a need for simple but powerful data processing services available to the broader user communities. For the more expert user, this will be accessed from a flexible and powerful web user interface to specify the operations needed and to maintain an history of the operations performed. For the occasional users, such as students and teachers, this implies an intuitive visual interface making clear what operations are possible to combine data of interest, and what standard combination of operations is required in common situations.

We propose to start by focusing on AVHRR data processing and identifying the spectrum of possible operations likely to be requested. Then we will prepare prototypes of visual interfaces that allow both occasional and expert users to specify a series of operations and repeat those operations on sets of data sets. A powerful history and macro facility will be provided that allows users to record a set of base operation and then apply variations of those operations on each selected data sets. This will also lead to a better understanding of the requirement of such adaptable interfaces that can both satisfy the needs of occasional and expert users. As the project progresses, we will adapt and extend our technique to other data types.

Our prototypes will be tested with a series of usability studies. First, informal studies will identify missing functionalities and implementation problems; later, usability tests will verify that users can accomplish their data processing tasks rapidly without help and feel confident that the data was processed as they requested.

JAVA-Based Architecture for Interoperation

We plan to integrate Java and the Java Runtime Environment with the static DBMS operations found in conventional database execution engines in order to allow "code shipping" in addition to "data shipping." Code shipping will allow us to transfer small inter-platform code to be executed at remote data storage facilities where the bulky data are stored, e.g. export of functions to other sites of the WP-Federation systems. In this way special purpose queries and/or processing modules that have been developed in one site can be exported to a remote site and filter most of the data before transferring back either much smaller amounts of it (decision support aggregates) or simply control.

Java provides an object-oriented and architecture-independent programming paradigm which allows developers to create modular applications that can be run on any Java-enabled platform. Using the object-oriented paradigm on the web and the relational support of Illustra we will develop component-based software by combining several independent software components. The advantage of this approach is the modularity and reusability of the code that is written,and the ability to change parts of an application in a plug-and-play manner. It also enables the inclusion of third party software components so as to use the most suitable tools available. Another advantage is the ability to make applications interact with each other and with the host operating system in an easy and efficient manner by using the proper component interface.

With the full support of classes in Java, we will create just one version of the software that will run on many different machines. For example, a calibration module will be transportable to a remote site and run on the raw data without having to ship and import the data. Having access to a remote object-oriented DBMS, the module will be able to store the new data product at a remote site as well. This Java-based architecture will provide the infrastructure for the WP-Federation and the interoperation of its constituent systems. Within this framework, we can integrate heterogeneous data storage facilities, web access to them, and visualization tools.

User Need Assessment

The diversity of the products we create will also attract many other users, as have our activities in this arena in the past. We envisage that the main user communities above and beyond the purely scientific will include the undergraduate and graduate students, the upper levels of the K-12 community, and state and local authority users concerned with natural resources and land management. Additionally, there will likely be a demand from commercial users. We anticipate these users based on analyses of current substantive acces ses to our current Web pages. Moreover, we believe that many of their needs can be met by a careful design of our Web pages and other information and delivery systems we will be providing, but additional outreach activities will likely be needed.

We will institute a monitoring system to assess the characteristics of our users. This information will be used to adapt the operation of the Land Cover ESIP, its services and products through the advice of our Strategic Advisory Team. We will combine a variety of techniques to assess the needs of the user community targeted by the Land Cover ESIP.

First, we will document automatically, using Web tools that record the characteristics of user interactions with the Land Cover ESIP, the typical number of parameters used, and the type of operations most commonly applied to the data sets. We will ensure that the usage data of every Land Cover ESIP data set or tool is adequately recorded and analyzed. If available, we will also study the data collected by other data centers.

Second, we will gather information from current and potential users of the Land Cover ESIP regarding the type of query and data processing they usually perform or that they wish to perform. This will be conducted with one or two small workshops and an e-mail survey broadcast on appropriate mailing lists. Statistics will be generated every month about the access and distribution volume of every product from the ESIP to determine the popularity and usage of the services. By monitoring the customers, we will be able to determine the distribution of users in educational institutions, government, industry and private sector. This information will help in targeting specific user communities to increase the usage of our services. We plan to hold an internal meeting at least once every two months to discuss, improve and monitor the services of the ESIP.

Third, we will conduct interviews of representative users, discuss their use of existing on-line data centers and more traditional data gathering techniques. Those interviews will be important to identify the needs for community support tools. We will produce and update a report regarding user assessment needs and make it available on the Web.

Community Building and Collaboration Tools

The Land Cover ESIP will be used to foster the development of a scientific community that truly merits the title Earth System Information Partnership We will provide information on other users who ordered the same data, performed similar operations, or published papers using the same data, subject to necessary privacy constraints.

This implies building an archive center that not only delivers data and information services, but nurtures the community of users by facilitating access to other users. The first step is to keep track of how the data are used. Summary data will be made available about popular data sets and a reward system can recognize the most productive data sets. We will build a collaborative environment that allows the users to ask for the documents which refer to a data set of interest and find colleagues with similar specialist interests.

Costs and benefits of application of new technology

All the technologies proposed for our ESIP are based on commodity off-the-shelf hardware and software components that are supplied by a number of vendors. We believe that an integrated processing and data server hardware architecture based on commodity high-performance processors with a high-performance interconnect is the most cost-effective architecture; that it will allow efficient processing/re-processing/on-demand processing, archiving, and access to large amounts of multi-sensor data. In fact, it is the only scaleable architecture for use in large science data centers.

The specific vendors selected were based on the best performance/cost quotes obtained from among a number of vendors contacted. In addition to the ADSM software, we have considered the IBM HPSS, which is perhaps an emerging standard for a high-performance storage management system that allows parallel I/O, but the cost was prohibitive. The ADSM software available at no cost to us is appropriate for our purposes, and will allow easy access for tertiary storage by concurrent users as we have configured our hardware. As for the users, we are only assuming that they have access to low-cost platforms with Web browsers.

Other specific aspects of the proposal leading to costs and benefits include the following:

1. *On-demand production of custom products.* On-demand production dramatically reduces the volume of data the system needs to archive on tertiary storage. In addition, custom processing allows an "infinite" number of products to be created from every L1B scene, giving far more value. While this does incur an added CPU load on the processor, we believe that our SP2 system will reduce operational costs through parallel processing. Part of our investigation will seek to quantify the cost/benefits of this "lazy processing" architecture, as opposed to the current EOSDIS "eager processing" architecture, in which all data products are precomputed and stored. We will also examine the trade-off between the higher upfront costs associated with the installation of the powerful SP2 system, and the expected lower operational costs associated with faster processing.
2. *Global Hierarchical Indexing System.* Storing projected data sets both corrupts the data and incurs excess costs. For example, the Interrupted Goode Hemispheric projection used by the 8km AVHRR Land product contains ~20% blank space part from the oceans. This space represents useless data that must be stored on tape or disk. Our proposed spherical databases make subsetting a routine part of query resolution, giving users only the data they require, thus dramatically reducing the network data transfer load.

3. *User interface*: Our HMI Lab has already developed EOSDIS prototype interface. This interface supports dynamic query preview, a mechanism which allows users to locate more quickly the data that is of interest to them.
4. *Interoperability concept*: Our interoperability architecture advocates keeping inter-site as simple as possible. While users can find data at any ESIP through a single interface, advanced services (e.g. scheduling, security) are not included in this infrastructure. This intentional simplicity should reduce costs associated with running Distributed Computing Environment (DCE) cells, as EOSDIS currently does.
5. *Development Costs* . There will be considerable cost-saving as a consequence of the extent to which we can and will leverage off already-completed research and prototype implementations as part of our NSF-funded Grand Challenge high-performance and EOSDIS prototyping.
6. *Platform specific development.* Given our approach and required commitments, for most software we only need to develop for the SP2, and not a multiplicity of other platforms.such as SGI, Sun, HP as ECS does.

Earth System Science User Support

A crucial goal of the Land Cover ESIP will be to reach out to the potentially large community of users and involve them from the outset in the development of various products, assessment of their needs and services required, and provision of related services. Our approach will focus on three issues (1) target community, (2) user support, and (3) data distribution.

Target community

Our main user community will be scientists who need land cover products to further their Earth Science System goals. This is by no means a homogeneous group. Their heterogeneity must be addressed in the ways that the ESIP operates. Some of the diversity which needs to be addressed can be described as follows:

Science users do not form a homogeneous community but vary across several dimensions:

- from users who will be intensive users of remote sensing data to those who will need access only to high level products,
- from users who will deal with extremely large volumes of data to those who will use only very low volumes,
- from users concerned mainly with the biophysical environment to those interested in socio-economic applications,
- from users interested mainly in scientific applications to those assessing natural resources,
- from users who will apply sophisticated models to those who will search for empirical trends in the data.

The problems they will examine will be very diverse and likely will include the following:

- Global scale modeling of the climate system, e.g. for climate prediction.
- Global scale observation and modeling of components of the biospheric systems, such as net primary productivity on land.
- Global modeling of the carbon cycle.
- Regional observation and modeling of hydrological processes and biogeochemical cycling.
- Monitoring of climate change to assess the extent of global and regional scale changes of

climate elements.

- Exchange of heat, momentum, water and carbon dioxide and trace gases across the land-air interface.
- Assessment of the impacts of climate change on vegetation cover.
- Impact of climate change on land use, agricultural production.
- Monitoring and modeling of land cover and land use change especially global deforestation.
- Assessment of loss of biodiversity.
- Mapping and quantifying land degradation/desertification
- Impacts of land cover changes on freshwater and coastal waters.
- Early warning assessments on the state of the environment.

Although our principal customer is likely to be the scientific user as indicated in section 2.4.8 we anticipate many other types of users as well, including K-12 students, undergraduates, state, local and commercial and many other users. We believe it is crucial to the success of the Land Cover ESIP (and for that matter to the whole of MTPE) that we address their needs by providing comprehensible, highly informative Web pages and documents, as well as sample data sets illustrating key issues in Earth System Science and their broader societal significance. The integrated research and educational mission of the newly created joint UMCP/NASA-GSFC (Earth System Science Interdisciplinary Center will be especially valuable in fostering and developing these activities

Data Distribution

The primary method of data ordering will be through the Internet. We will also use the Global Change Master Directory and the Advertising Service provided by EOSDIS to announce the availability of our products and services. In addition to the Internet, orders for products may be placed by conventional means, such as telephones and faxes. For large data sets and images, reduced-resolution browse images will be made available on-line for visual inspection by the users before they place an order. The data will be distributed either by magnetic media (tapes) or by anonymous ftp, depending upon the volume of the data. Users will be given a choice of formats in which they can order the data (e.g., 8mm, 4mm, tar, dd, etc.) Some of the frequently-requested data will be provided on CD-ROMs. Besides providing the products as flat binary files, some of the products will also be distributed as ARC/Info files, so that they can be readily ingested into a GIS. If the products received from the Land Cover ESIP are published by the users in their research, the users would be requested to cite the source of the data.

User Support

User support has to be an integral part of the work of the ESIP. We are proposing an integrated user support system that involves several different types of activity. In this way we intend to give a higher level of support than conventionally has been achieved and at the same time provide a degree of protection for scientists having to respond to basic questions.

Web based and hard copy support: Our experience indicates that the majority of information that users need can be provided through well-designed scientifically and technically authoritative web pages and corresponding hard copy support. We will publish and disseminate a brochure that will detail all the products and services provided by our ESIP. Our Web page will have the following information:

1. Information about the Land Cover ESIP (project objectives, personnel, services, user statistics etc.)
2. On-line documentation and publications about the data and services provided by the ESIP. This documentation will be very detailed, and will include a description of the algorithms, accuracy and limitations of the products, references and bibliography, product status, and frequently asked questions (FAQ) for each product. The documentation will be in a platform independent form such as PDF or postscript.
3. Data browse and order.
4. Operational status and schedule.
5. Software tools.
6. News and announcements.
7. User feedback and comments.
8. Customer support
9. Links to other ESIPs and data services.

Technical support through members of the user support group: User-support personnel will continue to be an important part of the support that we provide, since there will inevitably be queries about product creation, availability and delivery, not dealt with through publication on the web or through hard copy means. Users can contact the Land Cover ESIP either by e-mail or by other means such as voice and fax. Questions regarding the status of a specific order by the users will be dealt with by the operations personnel; however, more specific queries about the algorithms and processing methods will be referred to an appropriate member of the science team or the technical team. Feedback from the users will be utilized to define and implement new products.

Workshops and seminars focusing on new products and services and the underlying science: The underlying philosophy behind our ESIP is to bring forward reliable novel products and services for land cover. To aid that goal we will be running regular seminars, workshops, and special conference sessions to discuss current and planned products in a very timely fashion. The ESIP will provide a channel through which land cover scientists can much more effectively provide products to a broad user community.

Land Cover ESIP Interoperability

As participants in a Federated EOSDIS, the Land Cover ESIP will adopt the system-wide interoperability layer (SWIL) chosen by the WP-Federation. For the purposes of this proposal, we will outline our interoperability philosophy (based on the data needs of the terrestrial science community), and present a provisional plan to implement these ideas.

User Services

The concept of "one-stop shopping" for data and services has been a long-standing need for EOSDIS, so that users might access any data provider from a single user interface. In fact, system-wide interoperability is actually a spectrum, offering increasingly powerful service levels to users at the cost of increased complexity (Figure 4). The most basic service level is that of "collection discovery" (e.g. catalog or directory search), which enables users to locate data sets (collections) that physically reside at disparate sites around the world. The Global Change Master Directory (GCMD) and the EOSDIS Advertising Service represent examples of this service level. More tightly coupled systems allow distributed search and order of specific data granules (e.g. images), typically in response to a set of user-provided attributes (e.g. "find all MODIS NDVI images for Missouri during May, 2000"). This service level requires metadata commonality, or translation software that can convert a system query into a local query. The most sophisticated service level provides remote invocation of a wide array of software methods across a distributed computing environment. Currently, only ECS is scheduled to provide this level of interoperability.

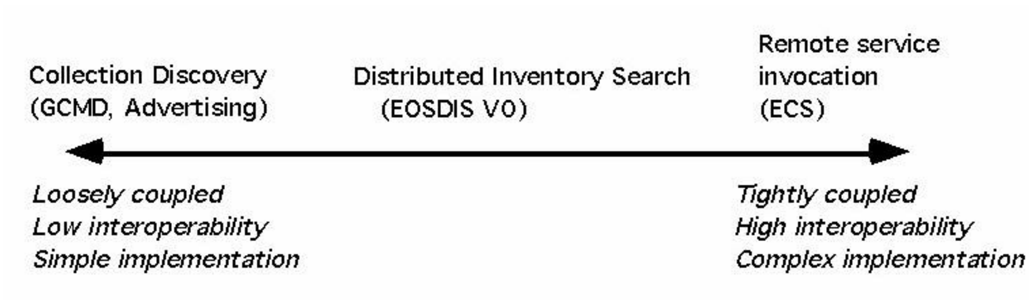

Figure 4: Spectrum of user service associated with "one-stop shopping".

Unfortunately, high levels of interoperability come at the price of either (1) reduced site autonomy or (2) increased system complexity. For example, implementing production scheduling across a distributed environment either requires identical software at all sites, or a complex distributed computing architecture. We feel reducing site autonomy compromises the purpose of the WP-Federation, while distributed computing lacks the maturity needed to produce a robust, low-cost system. Therefore, the University of Maryland advocates a level of interoperability which supports "one-stop" collection discovery, granule search/order, and a *few* key services (e.g. subsetting). More advanced services must be accessed by users through local client-server systems.

This conceptual system architecture is illustrated in Figure 5. Queries for individual granules or data sets can be launched through the distributed Federation of ESIPs (path 1) using a Federation client. These queries will perform collection and granule searches, using a generic set of metadata terms. Translator software (T) can translate between local and distributed protocols. This system-wide interoperability may be thought of as a "least-common denominator" version of one-stop shopping. More sophisticated, site-specific services can be invoked through the use of custom client software (Path 2), which in turn can be accessed from the Federated client via the World-Wide Web (WWW). Finally, we anticipate that both routine and on-demand production will require automated data "feeds" from other ESIPs (Path 3).

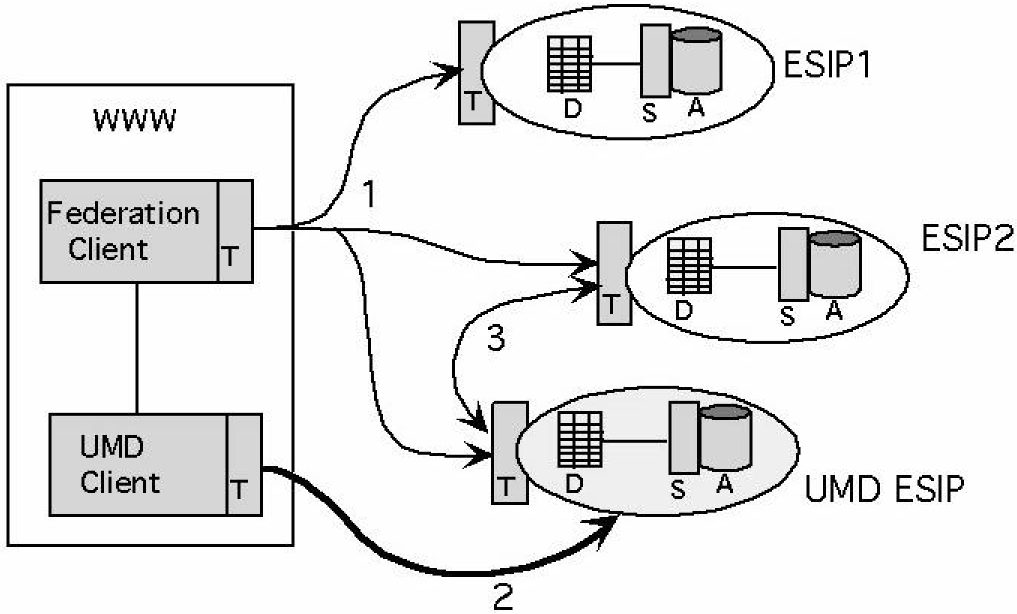

Figure 5: Conceptual interoperability architecture for the WP-Federation. Key: T=translation software (different at each site); D=database; S=server software; A=data archive.

The simplest existing solution that meets the immediate needs of our user community is the EOSDIS V0 IMS. The V0 system offers support for directory search using the GCMD Directory Interchange Format (DIF) standard, and inventory search by sensor, data set name, geophysical parameter, site, geographic location, and temporal range. In addition, planned extensions to the V0 system will allow a subsetting specification (coordinate list) to be passed to a V0 server. The V0 system is accessible via WWW, which will provide easy access to our custom WWW-based client or other local WWW tools. Finally, users of the V1 EOSDIS will be able to search V0 data sets through the 2-way V1/V0 Interoperability Gateway.

Specifically, our Landcover ESIP proposes to set up a V0 IMS server, database management system, and translation service, the latter necessary for translating V0 queries into our local metadata specification. Users will be able to:

- find data sets resident at UMCP;
- search for existing granules based on sensor, parameter, space, and time;
- view browse images;
- subset data products using a rectangular region-of-interest;
- order data products.

Certain key services will not be available via "one-stop shopping." For example, on-demand production of custom data sets, geographic projection, and file format conversion services cannot be currently supported via the V0 system. For these services, users will need to link to the Land Cover ESIP WWW client. While the V0 system does not currently support coincident search, routing of queries through an intermediate query optimizer/join function could facilitate this service.

The Committee on Earth Observation Satellites (CEOS) Catalog Interoperability Protocol (CIP) provides a somewhat more flexible and sophisticated solution based on the popular Z39.50 protocol. The CIP relies on a three-part interaction between client, server, and a middleware layer, the CIP Retrieval Manager, which allows for greater autonomy for both client and server. The presence of the Retrieval Manager also facilitates system-wide distributed services, such as query optimization and coincident search. In addition, the CEOS system organizes data sets into flexible, hierarchical collections, which can be defined according to user specifications.

Despite these advantages we feel that the CEOS CIP has not been prototyped sufficiently for use in the early stages of our Land Cover ESIP. We will, however, follow development of the protocol closely, and could migrate our system within the three years of prototype support. In this case, we would reuse our metadata model and database, but change the translation service to make it compatible with CIP syntax. We may also install a Retrieval Manager locally to allow hierarchical organization of data for the land processes community.

Metadata and Data Format

A key aspect of interoperability is the choice of metadata to be associated with archived products. This choice should be driven by (1) the needs of the targeted user community (e.g. terrestrial science in this case) and (2) FGDC compliance. Although the V0 IMS requires only a handful of attributes for distributed query, we intend to archive considerably more attributes. Storing additional attributes both empowers our custom client software, and also provides "hooks" for the WP-Federation SWIL.

To meet these goals, we intend to use a select subset of the EOSDIS Core System (ECS) science data model attributes as our base metadata model. ECS has assimilated metadata models from a diverse range of data providers for their current (B.1) science data model. While not all of these attributes are appropriate for our needs, many can be reused, and doing so guarantees FGDC compliance. In addition, we intend to add attributes specific to land processes data which might currently be Product Specific in ECS. For example, solar zenith angle, WRS coordinate, and view angle are all critical parameters for terrestrial remote sensing, and will be included in the Landcover ESIP archive.

Participation in the WP-Federation

The goal of the Federation, namely to provide user-friendly and seamless access to Earth System Science data that reside on widely distributed and heterogeneous systems, is a challenging task. This task will only be accomplished by a visionary intellectual leadership coupled with active participation of users from the very beginning. The evolution of the Federation governance structure will be a critical first step. The Land Cover ESIP intends to fully participate in the governance of the Federation and in setting the process for Federation-wide system interoperability and standards, user participation, and coordination in user services.

Our philosophy regarding SWIL was outlined in Section 2.6. Other systems issues that users need to be addressed at the Federation level include metadata, data integration and interface design. As we indicated in Section 2.4.2, we believe that object-oriented data base systems will allow the creation of a "meta-catalog" that integrates metadata from heterogeneous data bases across the Federation. The prototype object relational DBMS to be installed at our ESIP can be easily scaled to handle data access and integration across the different ESIPs of the Federation. The arguments presented in Section 2.4.2 in favor of such an approach carry over to the Federation level. Regarding the interface design, we believe that the development of the different ESIP interfaces should be coordinated from the start so that all the interfaces will be consistent, and have the same look and feel. A group of human-computer experts from the various ESIPs should be set up to coordinate the development of various user interfaces and collaborate on generating the minimum acceptable features. Given that the group led by Professor Shneiderman is one of the most prominent HCI groups in the nation, we believe that we can play an important role in helping the Federation in setting the design specifications of the user interface and in conducting related usability studies.

A crucial component of the success of the Federation is the speed of access, bandwidth, and the reliability of the networking infrastructure linking the different ESIPs. Downloading large files on the Internet is currently a frustrating experience which is likely to become worse before it improves. We suggest two concurrent approaches to deal with this issue. The first is to minimize the amount of unnecessary data transfers by designing an interface based on a two-phased approach (query preview and query processing) as we described in Section 2.4.5, and by using a Java-based architecture that will allow code shipping to where the bulky data is in the Federation. The second is to require enhanced participating ESIP sites to plan for enhanced connectivity. In our case, the University of Maryland is an active participant in the Internet II project and we are aiming to set up a "gigapop" for Internet II, in addition to our OC3 connectivity to vBNS and ATDNET

Management Plan

The overall management of the project will be the responsibility of the two Project Leaders, Drs. John Townshend (Project Leader) and Joseph J□□(Co-Project Leader). They will develop strategic directions and goals, set implementation plans, assess project progress and establish collaborative mechanisms with the WP-Federation. They will be assisted by a Strategic Advisory Team (composed largely of external members) and Technology and Science Teams. Implementation will be through a project manager who will be responsible for day-to-day operations of the two working teams responsible for Data Processing and Archiving, and User Support. The overall management structure and the line management are shown in Figure 6.

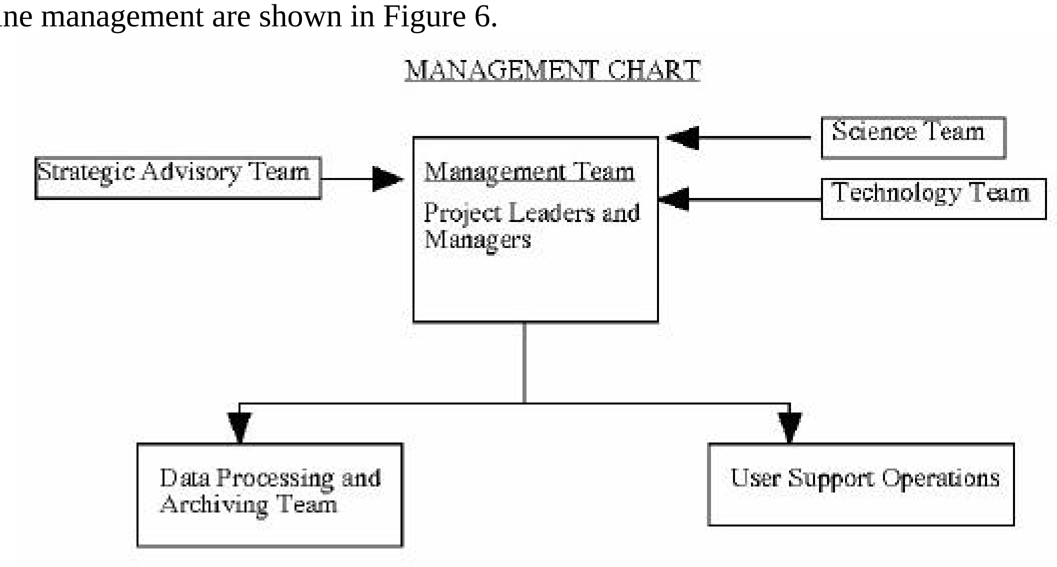

Figure 6. Project management and operations tasks.

The Science Team (see below) will meet on a monthly basis and will make recommendations regarding data products and related algorithms, processing and search requirements and capabilities, and plans for close interactions with the user community. The Science Team includes Dr. C.J. Tucker from NASA/GSFC, whose group interacts closely with members of the Department of Geography; as well as Dr. Eric Vermote, a UMCP Associate research scientist based at NASA/GSFC. The Technology Team (see section 6), will review the recommendations of the Science Team and develop strategies for hardware/software acquisitions, software development for data production, data storage and archiving, and user interactions. In addition, the Technology Team will set and review goals for system interoperability with the WP-Federation. The Technology team includes Dr Nazmi El Saleous, a member of the UMCP research faculty based at NASA/GSFC.

As demonstrated in Section 2.1 of this proposal, we are currently carrying our a wide range of research activities, all of which potentially can contribute useful products and services contributing to the mission of MTPE. We intend to carry out a rigorous internal and external review process with the help of the Strategic Advisory Team to establish which of these many products and services should be developed in consultation with research Pis. The ESIP Management Team will identify in a preliminary fashion which of these should be developed. The Strategic Advisory Team will then consider these proposals on an annual basis and then advise the ESIP Management Team on the activities of the ESIP for the following year.

We will use the following basic criteria to assess the priorities for products and services:

1. Scientific priority.
2. Novelty of services and products, especially in relation to those provided by other members of the Federation.
3. Thoroughness of validation of products.
4. Computing loads.

The Project Manager will be in charge of day-to-day management of the ESIP. He will oversee the implementation plans in each of the working groups, review their progress, coordinate between them, and meet with the Project Leaders on a weekly basis to review overall operations and development activities. He will also be involved in coordination with the WP-Federation and attend related meetings as appropriate.

The Data Processing and Archiving Working Team will consist of a database manager, a database administrator, three programmers, one systems staff, and two graduate students. This team will be responsible for systems support, setting up the file and database management systems, developing the tools to handle incremental evolution, derivation, and integration of metadata, as well as implementing processing algorithms and generating various data products.

The User Support Working Team will consist of two people for customer services, two programmers for the Web-interface, visualization tools and human-computer interaction, and one graduate student. This team will be responsible for handling all the user services and setting and updating the Web-interface and visualization tools, as well as conducting the user need assessment.

Interactions with the WP-Federation will be the prime responsibility of Drs. J□□and Townshend and the Project Manager. Other members of the management team will participate as needed and as deemed appropriate by the Project Leaders. We have provisionally identified Mr. J. Sobieski of UMIACS and Ms. Vivre Bell from Geography to fill the main management positions.

The Strategic Advisory Team will provide overall advice on all aspects of the work of the Land Cover ESIP including advice on the inclusion of new products, especially from research groups outside of UMCP, the relative allocation of human financial and computing resources for different tasks. They will consist of leading experts from the fields of Earth System Science and Computer Science. Dr. Christopher Justice of the University of Virginia has agreed to chair this Team.

Personnel

Dr. John Townshend (Project Leader)	Professor, Department of Geography, member of the Institute for Advanced Computing Studies, and Interim Director, Earth Systems Interdisciplinary Center
Dr. Joseph J□¡ (Co-Project Leader)	Director of the Institute for Advanced Computing Studies, and Professor of Electrical Engineering

Management Team

Dr. John Townshend (Chair)	Professor, Department of Geography, and Interim Director, Earth Systems Interdisciplinary Center
Dr. Joseph J□¡	Director of the Institute for Advanced Computing Studies
Mr. Bruce Douglas	Sr. Research Scientist, Department of Geography
Mr. Jerry Sobieski	UMIACS
Ms. Vivre Bell	Project Manager, Humid Tropical Deforestation Project, Department of Geography

Science Team

Dr. Samuel Goward (Chair)	Professor and Chair, Department of Geography
Dr. John Townshend	Professor, Department of Geograp hy, and Interim Director, Earth Systems Sciences Interdisciplinary Center.
Dr. Joseph J□¡	Director of the Institute for Advanced Computing Studies
Dr. Stephen Prince	Professor of Geography, Department of Geography
	Director Laboratory for Global Remote Sensing Systems
Dr. Ralph Dubayah	Associate Professor, Department of Geography and Institute for Advanced Computing Studies
Dr. C. J. Tucker	Biospheric Sciences Branch, NASA Goddard Space Flight Center
Dr. Eric Vermote	Associate Research Scientist, Department of Geography and GSFC
Dr. Satya Kalluri	Assistant Research Scientist, Department of Geography

Technology Team

Dr. N. Roussopoulos (Chair)	Professor, Computer Science and UMIACS
Dr. John Townshend	Professor, Department of Geography, and Interim Director Earth Systems Sciences Interdisciplinary Center.
Dr. Joseph J□¡	Director of the Institute for Advanced Computing Studies
Dr. Ben Shneiderman	Professor, Department of Computer Science and Institute for Advanced Computer Studies
Dr. Larry Davis	Professor, Department of Computer Science and Institute for Advanced Computing Studies
Dr. Nazmi El Saleous	Assistant Research Scientist, Department of Geography
Dr. C. Plaisant-Schwenn,	Acting Director, Human Computer Interaction Laboratory and UMIACS
Dr. Jeffrey Masek	Assistant Research Scientist, Department of Geography

---

APPENDIX A

Bibliographic Details of Team Members

PROJECT LEADER

John R. G. Townshend
Department of Geography,
University of Maryland, College Park, Maryland 20742
Telephone: 301-405-4050.
Fax: 301-314-9299.
E-mail: jt59@umail.umd.edu.

EDUCATION

B.Sc.1967 Geography with Geology, University College London.
Ph.D. 1971 Geomorphology, University College London.

APPOINTMENTS

Professor, Department of Geography, University of Maryland since 1989, member of UMIACS, interim Director of the Earth Systems Science Interdisciplinary Center (until December 1997).

Professor in the Department of Geography, University of Reading, Reading, U.K., and Director, Natural Environment Research Council's Unit for Thematic Information Systems (NUTIS); previously lecturer and reader (1972-89)

Visiting National Research Council Senior Post-Doctoral Research Fellow at NASA's Goddard Space Flight Center, Greenbelt, MD. (1979-80).

RELEVANT PUBLICATIONS

Townshend, J.R.G., Tucker, C. J. and Goward, S. N. 1993. Global Vegetation Mapping. In Gurney, R., Foster, J. and Partensin, C. (Eds.) *Atlas of Satellite Observations related to Global Change* , Cambridge University Press 301-311.

Townshend, J.R.G. and Rasool, S. I. 1993 Global change, In *Codata Bulletin*, 24 (4) (special issue on Data for Global Change), 1-14.

Townshend, J.R.G. 1994. Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction, *International Journal of Remote Sensing* , 15 (17) 3319-3332.

Townshend, J.R.G., Justice C.O., Skole, D., Malingreau, J.-P., Cihlar J., Teillet, P., Sadowski, F. and Ruttenberg, S. 1994. The 1 km resolution global data set: needs of the International Geosphere Biosphere Program, *International Journal of Remote Sensing 15* (17), 3417-3442.

Townshend. J.R.G. and Skole, D. L. 1995. The global 1km AVHRR data set. In P. Matter (editor), *Understanding the Terrestrial Environment, Remote Sensing Data Systems and Networks* , Ed P.M. Mather, John Wiley and Sons, 75-82.

PROFESSIONAL SERVICE

- Chairman International Geosphere Biosphere Program's Working Group on Land Cover Change (1988- 93)
- Member and Chair IGBP-DIS Standing Committee (1990-1996)
Chairman NASA/NOAA AVHRR Land Pathfinder Working Group (1991-1993)
- Member Editorial Board of the International Journal of Remote Sensing (1981-1995)
- Member National Academy of Science's Committee on Geophysical and Environmental Data (1992-5)
- Chair of the Joint Scientific and Technical Committee of the Global Climate Observing System (1995-present)
- Member of the Planning Committee of the Global Terrestrial Observing System (1993-1995).

C0-PROJECT LEADER

Joseph J□¡
Director
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742
Tel: 301-405-6722
Fax: 301-314-9658
E-mail: joseph@umiacs.umd.edu

EDUCATION

PhD 1977, MS 1976, Applied Mathematics, Division of Engineering and Applied Physics, Harvard University.

APPOINTMENTS

Director, Institute for Advanced Computer Studies, University of Maryland, 1994- present.

Electrical Engineering, University of Maryland, 1983- present.

Computer Science, Pennsylvania State University, 1977-1983.

HONORS AND AWARDS

IEEE Fellow Award, 1996.

Maryland Industrial Partnerships Award of Excellence, 1992.

BOOKS

An Introduction to Parallel Algorithms, Addison-Wesley, 1992.

RELEVANT PUBLICATIONS

H. Fallah-Adl, J. J□□ S. Liang, Y. Kaufman, and J. Townshend, Fast Algorithms for Removing Atmospheric Effects from Remotely Sensed Imagery, *IEEE Computational Science and Engineering*, 66-77, 1996

H. Fallah-Adl, J. J□□ and S. Liang, Estimating Aerosol Optical Depth and Correcting Thematic Mapper Imagery, *The Journal of Supercomputing*, 10, 315-329, 1996.

S. Liang, H. Fallah-Adl, J. J□□ S. Kalluri, Y. Kaufman, and J. Townshend, Development of an Operational Atmospheric Correction Algorithm for TM Imagery, *Journal of Geophysical Research-Atmosphere* (to appear).

Retrieval of Bidirectional Reflectance Distribution Function (BRDF) at Continental Scales from AVHRR Data Using High Performance Computing, to be presented at *IGARSS Symposium*, August 1997.

D. Bader, J. J□□ and R. Chellappa, Scalable Data Parallel Algorithms for Texture Synthesis and Compression Using Gibbs Random Fields, *IEEE Transactions on Image Processing*, 4(10):1456-1460, 1995.

PROFESSIONAL SERVICE

- Special Issue of JPDC on Data-Parallel Algorithms and Programming, Guest Editor, 1994.
- Journal of Parallel and Distributed Computing, Subject Area Editor on Parallel Algorithms, 1993-.

- IEEE Transactions on Parallel and Distributed Systems, Associate Editor, 1995-.
- Frontiers on Massi vely Parallel Computation, Program Committee Chair, 1991.
- International Parallel Processing Symposium, Program Committee Chair, 1995.
- Workshop on Parallel Algorithms, Program Chair, 1996.

PROJECT MEMBER

Samuel N. Goward
Professor and Chair
Department of Geography
University of Maryland, College Park
College Park, MD 20742
Voice: (301) 405-4050
FAX: (301) 314-9299
E-mail: sg21@umai l.umd.edu

EDUCATION

Ph. D.,1979. Indiana State University, Department of Geography and Geology
M. A., 1974. Boston University, Department of Geography
A. B., 1967. Boston University College of Liberal Arts. Major: Geography

RELEVANT PUBLICATIONS

Goward, S. N. and Dye, D. G. 1997, Global biospheric monitoring with remote sensing. The Use of Remote Sensing in Modeling Forest Productivity at Scales from Stand to the Globe. (H. L. Gholz, K. Nakane and H. Shimoda, Ed.), New York, Kluwer Academic Publishers, pp. 241-272

Prihodko, L. and Goward, S. N., 1997, Estimation of air temperature from remotely sensed observations, *Remote Sensing of Environment* (in press).

Czajkowski, K. P., T. Mulhern, S. N. Goward, and J. Cihlar, 1997, Validation of the geocoding and compositing system (GEOCOMP) using contextual analysis for AVHRR images, *International Journal of Remote Sensing*, (in press).

Goward, S. N. and D. Williams, 1997. Landsat and Earth Systems Science: Development of Terrestrial Monitoring, *Photogrammetric Engineering and Remote Sensing* (in press)

Prince, S. D. and Goward, S. N., 1995, Global primary production: a remote sensing approach, *Journal of Biogeography* 22: 815-835.

Prince, S. D. and Goward, S. N., 1995, Evaluation of the NOAA/NASA Pathfinder Land data set for global primary production modeling, *International Journal of Remote Sensing*, 17 (1): 217-221.

PROJECT MEMBER

Ben Shneiderman
Department of Computer Science
Institute for Advanced Computer Studies
University of Maryland

College Park, MD 20742
E-mail: ben@cs.umd.edu
Voice: (301) 405-2680
Fax: (301) 405-6707
Personal: http://www.cs.umd.edu/~ben
Lab: http://www.cs.umd.edu/projects/hcil

Ben Shneiderman is a Professor in the Department of Computer Science, Head of the Human-Computer Interaction Laboratory, and Member of the Institute for Systems Research, all at the University of Maryland at College Park. He has taught previously at the State University of New York and at Indiana University.

He regularly gives conference keynote speeches and public talks on "Human values and the future of technology" and "Relate-Create-Donate: Applying educational technology for the post-TV generation." Dr. Shneiderman's professional talks and courses cover user interface design, information visualization, and educational technology. He organizes an annual satellite television presentation on "User Interface Strategies" seen by thousands of professionals since 1987.

Dr. Shneiderman is the author of Software Psychology: Human Factors in Computer and Information Systems (1980) and Designing the User Interface: Strategies for Effective Human-Computer Interaction (1987, second edition 1992, third edition 1997), Addison-Wesley Publishers, Reading, MA. His 1989 book, co-authored with Greg Kearsley, Hypertext Hands-On!, contains a hypertext version on two disks. It was the world's first commercial electronic book and pioneered the highlighted embedded link. This concept was part of the Hyperties hypermedia system, now produced by Cognetics Corp., Princeton Junction, NJ.

Dr. Shneiderman has co-authored two textbooks, edited three technical books, and published more than 180 technical papers and book chapters. His 1993 edited book Sparks of Innovation in Human-Computer Interaction collects 25 papers from ten years of research at the University of Maryland. This collection includes Dr. Shneiderman's seminal paper on direct manipulation, a term he coined in 1981 to describe the graphical user interface design principles: visual presentation of objects and actions combined with pointing techniques to accomplish rapid incremental and reversible operations.

Ben Shneiderman has been on the Editorial Advisory Boards of nine journals including the newly formed ACM Transactions on Computer-Human Interaction and the ACM Interactions. He edits the Ablex Publishing Co. book series on "Human-Computer Interaction." He has consulted and lectured for many organizations including Apple, AT&T, Citicorp, GE, Honeywell, IBM, Intel, Library of Congress, NASA, and university research groups.

Ben Shneiderman received his BS from City College of New York in 1968, his PhD from State University of New York at Stony Brook in 1973. He received an Honorary Doctorate of Science from the University of Guelph, Ontario, Canada in 1996 and was elected as a Fellow of the Association for Computing (ACM) in 1997.

PROJECT MEMBER

Nicholas Roussopoulos
Department of Computer Science
E-mail: nick@cs.umd.edu

Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742
Voice: (301) 405-2687
Fax:(301) 405-6707

APPOINTMENTS

1991-present: Professor, Institute of Advanced Computer Studies, University of Maryland
1991-present: Professor, Department of Computer Science, University of Maryland
1993-1994: Visiting Professor, NTUA Athens, RWTH Aachen Germany, ETH Zurich
1985-1990: Associate Professor, Institute of Advanced Computer Studies, University of Maryland
1985-1990: Associate Professor, Department of Computer Science, University of Maryland
1985-1987: Associate Chairman for Administration and Research, Department of Computer Science, University of Maryland
1981-1984: Assistant Professor, Department of Computer Science, University of Maryland
1977-1980: Assistant Professor, Department of Computer Science, University of Texas at Austin
1976-1977: Research Scientist, IBM Research Lab, San Jose, California

FIELDS OF SPECIALIZATION

Fast OLAP Querying and Indexing of data aggregates, Client-Server Database Architectures, Distributed Database Systems, Geographic Information Systems, Spatial Indexing, High Performance Database Systems, Scallable Parallel I/O Systems, Parallel and Scalable I/O, Network Management Information Systems, Database Design and Management, Expert Database Systems, and Artificial Intelligence Databases.

RELEVANT PUBLICATIONS

N. Roussopoulos, Y. Kotidis, M. Roussopoulos. Cubetree: Organization of and Bulk Updates on the Data Cube,*ACM-SIGMOD International Conference on Management of Data*, Tucson, Arizona, May 13-15, 1997.

N. Roussopoulos, C.M. Chen, S. Kelley, A. Delis and Y. Papakonstantinou. The ADMS Project Views 'R' Us, *IEEE Data Engineering Bulletin*, June 1995.

N. Roussopoulos, S. Kelley and F. Vincent, Nearest Neighbor Queries. *ACM-SIGMOD International Conference on Management of Data*, San Jose, California, May 22-25, 1995.

A. Delis and N. Roussopoulos. Management of Updates in the Enhanced Client-Server DBMS. (Best Paper Award), *IEEE/CS 14th International Conference on Distributed Computing Systems*, Poznan, Poland, June 1994.

C. Cheung, N. Roussopoulos, D. Leisawitz, S. Kelley, G. Reichert, D. Silberberg and J. Wang. AMASE: An Object-Oriented Metadatabase Catalog for Accessing Multi-Mission Astrophysics Data. Science Information Systems Interoperability Conference, College Park, Nov. 6-9, 1995.

OTHER PROJECT MEMBERS

Dr. Stephen Prince	Professor of Geography, Department of Geography; and Director, Laboratory for Global Remote Sensing Systems
Dr. Ralph Dubayah	Associate Professor, Department of Geography and Institute for Advanced Computing Studies
Dr. C. J. Tucker	Biospheric Sciences Branch, NASA Goddard Space Flight Center
Dr. Eric Vermote	Associate Research Scientist, Department of Geography and GSFC
Dr. Satya Kalluri	Assistant Research Sci entist, Department of Geography
Dr. Larry Davis	Professor, Department of Computer Science and Institute for Advanced Computing Studies
Dr. Nazmi El Saleous	Assi stant Research Scientist, Department of Geography
Dr. C. Plaisant-Schwenn	Acting Director, Human Computer Interaction Laboratory and UMIACS
Dr. Jeffrey Masek	Assistant Research Scientist, Department of Geography
Mr. Bruce Douglas	Sr. Research Scientist, Department of Geography

REFERENCES

Ahlberg, C. and B. Shneiderman (1994). "Visual Information Seeking: Tight coupling of dynamic query filters with starfield displays." Readings in Human-Computer Interaction: Toward the Year 2000. R. M. Baeker, J. Gruden, W. A. S. Buxton and S. Greenberg, Morgan Kaufmann Pubs., Inc: 450-456.

Cheung, C., N. Roussopoulos, D. Leisawitz, S. Kelley, G. Reichert, D. Silberberg and J. Wang (1995). AMASE: An Object-Oriented Metadatabase Catalog for Accessing Multi-Mission Astrophysics Data. Science Information Systems Interoperability Conference, College Park, MD.

Davis, L. D. and J. Townshend (1993). High Preformance Computing for Land Cover Dynamics. UMIACS/University of Maryland.

Defries, R. S., M. Hansen and J. R. G. Townshend (1995). "Global descrimination of land cover types from metrics derived from AVHRR Pathfinder data." Remote Sensing of Environment 54: 209-222.

Defries, R. S. and J. R. G. Townshend (1994). "NDVI-derived land cover classifications at a global scale." International Journal of Remote Sensing 15(17): 3567-3586.

Doan, K., C. Plaisant, B. Shneiderman and T. Burns (1997). Query Previews for Networked Information Systems: A Case Study with NASA Environmental Data. College Park, Department of Computer Science, University of Maryland.

Dubayah, R. (1992). "Estimating net solar radiation using Landsat Thematic Mapper and digital elevation data." Water Resources Research 28(9): 2469-2484.

Dubayah, R. (1992). "Estimating regional variabiity in net solar radiation using Landsat Thematic Mapper and digital elevation data." Water Resources Research (In press).

Dye, D. G. (1992). Satellite estimation of global distribution and variability of incident photosynthetically active radiation Geography, University of Maryland at College Park.

Fallah-Adl, H., J. Ja□a□ S. Liang, J. R. G. Townshend and Y. J. Kaufman (1996). "Fast algorithms for removing atmospheric effects from satellite images." IEEE Computational Science and Engineering 3(2): 66-77.

Fallah-Adl, H., J. JaJa and S. Liang (1996). "Estimating aerosol optical depth and correcting thematic mapper imagery." The Journal of Supercomputing.

Field, C. B., J. T. Randerson and C. M. Malmstrom (1995). "Global net primary production: Combining ecology and remote sensing." Remote Sensing of Environment 51: 74-88.

Goward, S. G. and S. D. Prince (1995). "Transient effects of climate on vegetation dynamics: Satellite observations." Journal of Biogeography 22: 549-563.

Goward, S. N. (1989). "Landsat 1989: Remote sensing at the crossroads." Remote Sensing of Environment 28: 3-4.

Goward, S. N. (1989). "Satellite bioclimatology." Journal of Climate 7(2): 710-720.

Goward, S. N., D. G. Dye, W. Dulaney and J. Yang (1990). "Critical assessment of the NOAA Global Vegetation Index data product." Workshop on Use of Satellite-Derived Vegetation Indices in Weather and Climate Prediction Models, NOAA World Weather Building, Camp Springs, Maryland, U.S. Department of Commerce, National Oceanic and Atmospheric Administration.

Goward, S. N., D. G. Dye, S. Turner and J. Yang (1993). "Objective assessment of the NOAA Global Vegetation Index data product." International Journal of Remote Sensing 14(18): 3365-3394.

Goward, S. N. and N. N. Kazantsev (1991). "Monitoring environmental change: Geographic information systems and remote sensing." Global Change: Geographical Approaches. J. R. Mather and G. V. Sdasyuk. Tucson, University of Arizona Press: 183-198.

Goward, S. N., B. Markham, D. G. Dye, W. Dulaney and J. Yang (1991). "Normalized difference vegetation index measurements from advanced very high resolution radiometer." Remote Sensing of Environment 35(2&3): 259-279.

Goward, S. N., C. J. Tucker and D. G. Dye (1985). "North American vegetation patterns observed with the NOAA-7 advanced very high resolution radiometer." Vegetatio 64: 3-14.

Goward, S. N., R. H. Waring, D. G. Dye and J. Yang (1994). "Ecological remote sensing at OTTER: Macroscale satellite observations." Ecological Applications 4(2): 322-343.

James, M. E. and S. N. V. Kalluri (1994). "The Pathfinder AVHRR land data set: an improved coarse resolution data set for terrestrial monitoring." International Journal of Remote Sensing 15(17): 3347-3364.

Justice, C. O., J. R. G. Townshend, B. N. Holben and C. J. Tucker (1985). "Analysis of the phenology of global vegetation using meteorological satellite data." International Journal of Remote Sensing 6(8): 1271-1381.

Kalluri, S. N. V., Z. Zhang, S. Liang, J. JaJa and J. R. G. Townshend (1997). "Retrieval of bidirectional reflectance distribution function (BRDF) at continental scales from AVHRR data using high performance computing." IGARSS, Singapore.

Kaufman, Y. J. and C. Sendra (1988). "Algorithm for automatic atmospheric corrections to visible and near-IR satellite data." International Journal of Remote Sensing 9(8): 1357-1381.

Kearney, M. and A. Rogers (1995). "Measurement of percent water in coastal marsh pixels using mixture modeling." ACSM/ASPRS annual convention, Charlotte, NC.

Liang, S. and J. R. G. Townshend (1997). "Angular signatures of NOAA/NASA Pathfinder AVHRR Land data and applications to land cover identification." IGARSS, Singapore.

Nilson, T. and A. Kuusk (1989). "A reflectance model for homogeneous plant canopies and its inversion." Remote Sensing of Environment 27: 157-167.

North, C., B. Shneiderman and C. Plaisant (1996). "User Controlled Overviews of an Image Library: A Case Study of the Visible Human." Proc. of the 1st ACM International Conference on Digital Libraries, Bethesda, MD.

Prince, S. D. (1991). "A model of regional primary production for use with course resolution satellite data." International Journal of Remote Sensing 12(6): 1313-1330.

Prince, S. D., S. J. Goetz, R. O. Dubayah, K. P. Czajkowski and M. Thawley (1997). "Inference of surface and air temperature, atmospheric precipitable water and vapor pressure deficit using AVHRR satellite observations: comparison with field observations." Journal of Geophysical Research (Submitted).

Prince, S. D., S. J. Goetz and S. N. Goward (1995). "Monitoring primary production from earth observing satellites." Water, Air and Soil Pollution 82: 509-522.

Prince, S. D. and S. N. Goward (1995). "Evaluation of the NOAA/NASA Pathfinder AVHRR Land Data Set for global primary production." International Journal of Remote Sensing 17: 217-221.

Prince, S. D. and S. N. Goward (1995). "Global net primary production: A remote sensing approach." Journal of Biogeography 22: 815-835.

Prince, S. D. and S. N. Goward (1995). "Global primary production: a remote sensing approach." Journal of Biogeography 22: 815-835.

Prince, S. D., Y. H. Kerr, J.-P. Goutorbe, T. Lebel, A. Tinga, P. Bessemoulin, J. Brouwer, A. J.

Dolman, E. T. Engman, J. H. C. Gash, M. Hoepffner, P. Kabat, B. Monteny, F. Said, P. Sellers and J. Wallace (1995). "Geographical, biological, and remote sensing aspects of the Hydrologic Atmospheric Pilot Experiment in the Sahel (HAPEX-Sahel)." Remote Sensing of Environment 51: 215-234.

Prince, S. D. and C. J. Tucker (1986). "Satellite remote sensing of rangelands in Botswana, II, AVHRR and herbaceous vegetation." International Journal of Remote Sensing 7(11): 1555-1570.

Rahman, H., M. M. Verstraete and B. Pinty (1991). "Coupled surface-atmosphere reflectance model (CSAR). 1. Model description and inversion against synthetic data." Journal of Geophysical Research 98: 20791-20801.

Shock, C. T., L. Davis, S. N. Goward and A. Sussman (1995). A global-scale remotely sensed image data base.

Tanin, E., R. Beigel and B. Shneiderman (1996). "Incremental Data Structures and Algorithms for Dynamic Query Interfaces." Proceedings of ACM SIGMOND 25: 21-24.

Townshend, J. R. G. (1989). "Special issue: Comparison of passive microwave with near-infrared and visible data for terrestrial environmental monitoring." International Journal of Remote Sensing 10(10): 1575-1691.

Townshend, J. R. G. (1995). "Global data sets for land applications from the Advanced Very High Resolution Radiometer." International Journal of Remote Sensing 15(17): 3319-3332.

Townshend, J. R. G., V. Bell, A. C. Desch, C. Havlicek, C. O. Justice, W. E. Lawrence, D. Skole, W. W. Chomentowski, B. Moore III, W. Salas and C. J. Tucker (1995). "The NASA Landsat Pathfinder Humid Tropical Deforestation Project." Land Satellite Information in the Next Decade, Tysons Corner, Virginia, American Society of Photogrammetry and Remote Sensing.

Townshend, J. R. G., C. O. Justice and V. T. Kalb (1987). "Characterization and classification of South American land cover types using satellite data." International Journal of Remote Sensing, 8: 1189-1207.

Townshend, J. R. G., C. O. Justice, W. Li, C. Gurney and J. McManus (1991). "Global land cover classification by remote sensing: present capabilities and future possibilities." Remote Sensing of Environment 35: 243-255.

Townshend, J. R. G., C. O. Justice, D. Skole, J. P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski and S. Ruttenberg (1994). "The 1 km AVHRR global data set: needs of the International Geophere Biosphere Programme." International Journal of Remote Sensing 15(17): 3417-3441.

Townshend, J. R. G. and C. J. Tucker (1984). "Objective assessment of AVHRR data for land cover mapping." International Journal of Remote Sensing 5: 492-501.

Tucker, C. J. (1989). "Comparing SMMR and AVHRR data for drought monitoring." International Journal of Remote Sensing 10: 1663-1672.

Xue, X. (1997). Developments in AMASE: An Object-Oriented Astrophysics Database. Department of Computer Science. College Park, University of Maryland.