INT. J. REMOTE SENSING, 2000, VOL. 21, NO. 6 & 7, 1513–1536

High performance computing algorithms for land cover dynamics using remote sensing data

S. N. V. KALLURI†, J. JÁJÁ‡, D. A. BADER§, Z. ZHANG‡, J. R. G. TOWNSHEND¶ and H. FALLAH-ADL‡

†Department of Geography, University of Maryland, College Park, MD-20742, USA

‡Institute for Advanced Computer Studies, University of Maryland, College Park, MD-20742, USA

§Electrical and Computer Engineering Department, University of New Mexico, Albuquerque, NM-87131, USA

¶Department of Geography and Institute for Advanced Computer Studies, University of Maryland, College Park, MD-20742, USA

Abstract. Global and regional land cover studies need to apply complex models on selected subsets of large volumes of multi-sensor and multi-temporal data sets that have been derived from raw instrument measurements using widely accepted pre-processing algorithms. The computational and storage requirements of most of these studies far exceed what is possible on a single workstation environment. We have been pursuing a new approach that couples scalable and open distributed heterogeneous hardware with the development of high performance software for processing, indexing and organizing remotely sensed data. Hierarchical data management tools are used to ingest raw data, create metadata and organize the archived data so as to automatically achieve computational load balancing among the available nodes and minimize input/output overheads. We illustrate our approach with four specific examples. The first is the development of the first fast operational scheme for the atmospheric correction of Landsat Thematic Mapper scenes, while the second example focuses on image segmentation using a novel hierarchical connected components algorithm. Retrieval of the global Bidirectional Reflectance Distribution Function in the red and near-infrared wavelengths using four years (1983 to 1986) of Pathfinder Advanced Very High Resolution Radiometer (AVHRR) Land data is the focus of our third example. The fourth example is the development of a hierarchical data organization scheme that allows on-demand processing and retrieval of regional and global AVHRR data sets. Our results show that substantial reductions in computational times can be achieved by the high performance computing technology.

1. Introduction

Global change studies through the use of remote sensing techniques require multi-disciplinary research with the fusion of data sets from various sources and instruments. Although multi-temporal high-resolution satellite data have been collected since the early 1970s (e.g. Landsat and Advanced Very High Resolution Radiometer (AVHRR) data sets), the routine use of these data sets in modelling global carbon, biogeochemical, and hydrological cycles and ecosystem response to natural and anthropogenic changes at a global scale is hindered by the requirements of tremendous data storage and high computational complexity.

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2000 Taylor & Francis Ltd
http://www.tandf.co.uk/journals/tf/01431161.html

The acquisition, processing, mapping and conversion of remotely sensed data into science products useful for studying land cover dynamics involves addressing a number of complex modelling and computational problems. The computational tasks involved in a variety of pre-processing or 'conditioning' of satellite data—such as calibration, atmospheric and topographic correction, and identification of clouds—consist of a mixture of simple pixel operations and complex neighbourhood operations.

As typical examples, here we consider Landsat Thematic Mapper (TM) and AVHRR data processing streams to gauge the computational and storage requirements of processing level zero to level two products at a global scale, since data from these two instruments have been widely used to study land cover dynamics for more than 15 years (Townshend 1994, Goward and Williams 1997). We estimate that 4500 TM scenes are required to achieve global coverage of the land surface, and about 237 Giga Floating Point Operations (GFLOPs) are involved in generating a land cover product from a level zero scene. The total storage requirements for such data would be 2.7 Tbytes, requiring 1.06 Peta FLOPs to process. Similarly, processing the entire 17-year archive of Global Area Coverage (GAC) data from the AVHRR would require the processing of 4.65 Tbytes of data, which would take 3.5 years to process on a single processor (Acharya *et al.* 1996). Moreover, significant improvements in processing algorithms could mandate the re-processing of these data sets from time to time to produce enhanced and more accurate measurements of the land surface. It is evident from these examples that high performance computing techniques are needed to achieve the objective of acquiring routine and timely information from Earth orbiting satellites.

Under the sponsorship of the National Science Foundation's Grand Challenge Program, we embarked on a comprehensive research program in 1994 on the application of high performance parallel computing to data and computation-intensive problems in land cover dynamics. Innovative parallel and scalable algorithms are being developed in a heterogeneous distributed computing environment for rapid and accurate processing of large satellite-derived data sets. The processing procedures that have been developed include:

- atmospheric correction of TM data;
- image segmentation;
- retrieval of land surface Bidirectional Reflectance Distribution Function (BRDF), and;
- designing a high performance system for processing, storing and retrieving AVHRR data.

In this paper we present a summary of results from this work. The following sections briefly describe the results from various algorithms that we have successfully implemented using innovative computational techniques.

2. Atmospheric correction of TM data

Remote sensing measurements are contaminated by atmospheric effects such as Rayleigh scattering due to atmospheric molecules, absorption by water vapour, ozone and other gases, and scattering and absorption due to atmospheric aerosols. Unless satellite data are corrected for these affects, large errors in the measurement of the variables required for studying land cover dynamics (Kaufman 1984, Singh and Saull 1988). The atmospheric affect varies both spatially and temporally, and is also dependent upon the wavelength and geometry of observations. It is possible to decouple the effects of various individual components in the atmosphere on the remote sensing signal and then perform selective corrections (Tanre *et al.* 1992). For a plane-parallel atmosphere bounded by a Lambertian surface, the radiance at the top-of-the-atmosphere (TOA) can be expressed as (Chandrasekhar 1960, Fraser and Kaufman 1985):

$$L_\lambda = L_0 + \frac{\rho F_\mathrm{d} T}{\pi(1 - s\rho)} \tag{1}$$

where L_λ is the radiance recorded by the sensor at wavelength λ, L_0 is the upward radiance at the TOA when the surface reflectance (ρ) is zero (path radiance), F_d is the total irradiance at the surface, T is the total transmittance of the atmosphere and s is the spherical albedo of the atmosphere. Molecular scattering in the atmosphere is well understood, and its effects can be easily corrected for (Kaufman and Sendra 1988). Correcting for the aerosol effect is more difficult, since atmospheric aerosols are highly variable in time and space. Atmospheric correction approaches using standard radiative transfer algorithms such as 6S (Vermote *et al.* 1997) and LOWTRAN7 (Kneizys *et al.* 1988) require the input of atmospheric optical depth data from observations. Moreover, these algorithms perform corrections on single-pixel values and are not designed to correct entire scenes.

We have implemented a direct atmospheric correction approach based on the so-called 'dark target method' of Kaufman and Tanre (1996) and Kaufman and Sendra (1988). Two steps are involved in direct atmospheric correction: the estimation of optical properties of the atmosphere from the imagery and the retrieval of surface reflectance. The principle behind this method is to derive the atmospheric properties by inverting equation (1) from measurements of L_λ over targets whose ρ is known. Our algorithm is capable of correcting a whole TM scene at once. A step-by-step approach of the atmospheric correction methodology for TM data is described below (Fallah-Adl *et al.* 1996a, b).

1. For a $w \times w$ window of pixels in the input TM image, dark targets representing dense green vegetation are identified based on a channel 7 reflectance threshold (ρ_7). w typically ranges from 11 to 121, and the default ρ_7 threshold is set at 0.1. Since reflectances in TM channel 7 (2.08–2.35 μm) are least affected by the atmosphere we assume that TOA reflectances in this channel are equivalent to surface reflectances in the same wavelength. The wavelength intervals of different TM bands are given in table 1. Both w and ρ_7 can be changed very easily depending upon the scene conditions.
2. If there are several pixels whose ρ_7 is less than the threshold, then their mean value is computed.
3. Reflectances in TM channels 1 (ρ_1) and 3 (ρ_3) for the dark pixels are estimated as (Kaufman *et al.* 1997) $\rho_1 = \rho_7 \times 0.25$ and $\rho_3 = \rho_7 \times 0.50$.
4. Using a pre-computed look-up table generated by a radiative transfer code (Fraser *et al.* 1992) the aerosol optical thickness (τ) in TM bands 1 and 3 is estimated from the measured $L_{\lambda 1}$ and $L_{\lambda 3}$ for ρ_1 and ρ_3 respectively, i.e. given a surface reflectance and upwelling radiance, τ is estimated. τ_1 should be larger than or equal to τ_3. Otherwise, previous steps are repeated with a smaller ρ_7 threshold until $\tau_1 \geqslant \tau_3$.

Table 1. Landsat Thematic Mapper (TM) bands and their spectral wavelength intervals (Townshend *et al.* 1988).

TM band number	Wavelength interval (μm)
1	0.45–0.52
2	0.52–0.60
3	0.63–0.69
4	0.76–0.90
5	1.55–1.75
7	2.08–2.35
6	10.4–12.5

5. From τ_1 and τ_3, the aerosol optical thickness in TM bands 2, 4, and 5 is computed using the following exponential relationship between wavelength and aerosol optical thickness: $\tau_i = a\lambda_i^{-b}$. τ_i and λ_i are the aerosol optical thickness and central wavelength (in micrometres) in channel i respectively, and $1 \leqslant i \leqslant 5$. The coefficients a and b are derived by fitting an exponential curve to τ_1 and τ_3 at wavelengths λ_1 and λ_3 respectively.
6. Once the aerosol optical thickness is determined in all the TM bands, atmospheric correction is applied to the central pixel of the $w \times w$ window using the same look-up table. From the measured L_λ and retrieved τ_λ, ρ_λ is estimated.
7. The $w \times w$ window moves in single-pixel increments across the image and the above steps are repeated.

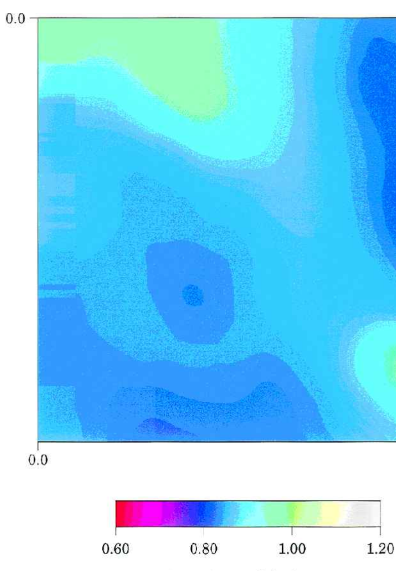
Liang *et al.* (1997) validated the aerosol optical thickness values determined by this algorithm from several TM scenes by comparing them with ground observations during the First ISLSCP Field Experiment (FIFE) (Sellers *et al.* 1992), and the Sulfate Clouds and Radiation Atlantic (SCAR A) experiment (Kaufman and Holben 1996). The correlation between estimated and ground measurements was good ($r = 0.98$). Figure 1(*a*) shows a red–green–blue (RGB) composite of TM bands 1, 2, and 3 from a TM scene over Eastern Maryland, USA, before atmospheric correction, and figure 1(*b*) shows the same scene after atmospheric correction. Most of the atmospheric haze is removed after applying the correction and the image looks clearer than the image without correction. The derived aerosol optical thickness values in TM band 1 (blue) are shown in figure 1(*c*), and the spatial pattern of aerosol optical thickness (figure 1(*c*)) is consistent with haze in the image. The optical thickness is higher in the upper-left and lower-right parts of the image which correspond to areas with dense haze in the original image.

The vertical striping that is observed on either side of the optical thickness image (figure 1(*c*)) is due to the moving window algorithm used here, and can be explained as follows. If we have a $w \times w$ window, then the central pixel within the window will be $w/2$. A window size of 91×91 was used to correct the image shown in figure 1. Thus, a full window of 91×91 pixels can only be created around pixels starting from

Figure 1. (*a*) A red–green–blue (RGB) colour composite of the visible bands (3, 2 and 1) of a TM scene over Maryland, USA, before atmospheric correction. The acquisition date of this scene is 29 August 1993 and the size of the image shown here is 512 pixels by 512 lines. (*b*) A composite of the visible bands after atmospheric correction. (*c*) Band 1 (blue) optical thickness image derived from the atmospheric correction algorithm.

(a)

(b)



(c)

column number 46 to column number n-46, where n is the total number of columns in a two-dimensional image (i.e. we have 45 pixels on either side of pixel number 46). The first 45 pixels around which a complete $w \times w$ window cannot be created are assigned the optical thickness of the nearest window. Therefore, the column of pixels less than $w/2$, and those that are greater than n-$w/2$ present on either side of the image will have a uniform optical thickness which results in striping at the edges.

Our atmospheric correction algorithm was benchmarked on different platforms with serial as well as parallel architectures. Timing results show that it takes 3 h 52 min to correct a full TM scene on an IBM RS6000 machine. We have also tested the performance on an IBM SP2 machine having 16 nodes. Each of these processors on the SP2 have a peak performance of 266 MFLOPS, and are configured with 128 Mbytes of memory and 64 kbytes of cache. All the 16 processors are connected by an Ethernet and a high-performance switch that permits all the processors to exchange messages simultaneously. The parallel version of the code is written in the Single Program Multiple Data (SPMD) model, so that all the processors run the same code, but on different parts of the input image simultaneously. For the parallel algorithm, the image is divided into a number of blocks, and these blocks are distributed equally among all the nodes for processing. The size of the block is dependent upon the memory available to the nodes at the time of processing. On the SP2 configuration described above, atmospheric correction for a full TM scene can be run in 30 min (Fallah-Adl *et al.* 1996a, b), and can in general be shown to achieve a linear speed up as a function of the number of nodes.

These results indicate that a significant reduction in runtime can be achieved by a parallel implementation of the atmospheric correction algorithm, which makes it possible to apply the correction scheme over large areas covered by multiple scenes in a practical manner. Both qualitative and quantitative analyses of several TM scenes corrected by this algorithm show that the results are reliable (Liang *et al.* 1997). The code has been written in a modular fashion, so that individual components of the correction scheme such as instrument calibration, and the empirical functions relating channel 7, 3 and 1 reflectances can be very easily modified.

3. Image segmentation using hierarchical connected components

Segmentation algorithms for remotely sensed imagery cluster pixels into homogeneous regions, which, for example, can be classified into categories with higher accuracy than could be obtained by classifying the individual pixels. Region-growing is a class of techniques used in image segmentation algorithms in which, typically, regions are constructed by an aggregation process that merges pixels to regions when those pixels are both adjacent to the regions and similar in spectral property (see, e.g., Haralick and Shapiro 1985, Westman *et al.* 1990, Chang and Li 1994). Each pixel in the scene receives a label from the region-growing process; pixels will have the same label if and only if they belong to the same region. A segmentation process using region-growing techniques may be used to realize a hierarchy of region-labelled images. At the lowest level of the hierarchy, a region contains the set of connected pixels with strict similarity. As we move up the hierarchy, the similarity criterion relaxes and similar regions merge together.

Image segmentation consists of two steps. The first step involves image enhancement and edge detection, and the second step involves identifying uniform regions and labelling them. In region-growing algorithms a region's border is susceptible to erroneous merging at its weakest point, which can be aggravated by several factors, including noise, blur and lighting. Thus it becomes extremely important to enhance an image before this region-growing process. An ideal image enhancement filter preserves edges as well as smoothes the interior of regions (Bader and JáJá 1996a).

3.1. *Image enhancement*

In remotely sensed imagery, natural regions may have significant variability in each band. Noise, introduced from the scanning of the real scene into the digital domain, will cause single-pixel outliers. Illumination and view geometry can also cause a gradient of grey levels in pixels across the same region due to surface anisotropy. Because of these and other similar effects, it is necessary to pre-process the image with a stable filter, such as the Symmetric Neighborhood Filter (SNF) that smoothes out the interior pixels of a region to a near-homogeneous level using an iterative technique (Harwood *et al.* 1987, Bader and JáJá 1996a). Also, due to the point spread function of the instrument, edges of regions are usually blurred (Poropat 1993, Forster and Best 1994) so that the transition in grey levels between regions is not a perfect step over a single pixel, but ramps from one region to the other over several pixels. Most pre-processing filters will smooth the interior of regions at the cost of degrading the edges or, conversely, detect edges while introducing intrinsic error on previously homogeneous regions. The SNF filter is, additionally, an edge-preserving filter that detects blurred transitions and sharpens them while preserving the true border location as best as possible. Therefore, the SNF is an edge-preserving smoothing filter that performs well for simultaneously sharpening edges and smoothing regions. In addition, it is an iterative filter that also can be tuned to retain thin-image structures in remotely sensed imagery corresponding, for example, to rivers and roads.

The SNF enhancement is a stable filter that is applied either for a fixed number of iterations or until stopping criteria (defined below) are reached, and takes the single parameter ε, as follows. The SNF filter compares each pixel to its 8-connected neighbours. Note that we are using the notion of 8-connectivity, meaning that two pixels are adjacent if and only if one pixel lies in any of the eight positions surrounding the other pixel (Castleman 1996). Figure 2 shows a diagram of a 3×3 neighbourhood centred around a pixel, with the symmetric pairs having the same letter. The neighbours are inspected in symmetric pairs around the centre, that is, top with bottom, left with right, upper-left with lower-right, and upper-right with lower-left. Assume

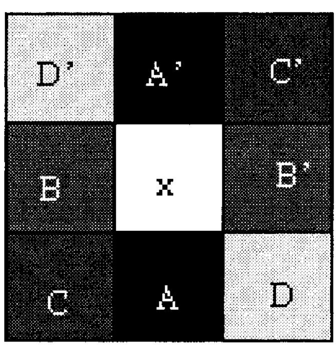


Figure 2. A 3×3 neighbourhood of pixels. The symmetric pairs around the central pixel x have the same letter.

without loss of generality that the pair of pixels have brightness intensities A and A' and that $A > A'$. Using each pair and the centre pixel, four different comparisons are made using the following criteria (figure 3).

- If the centre pixel (with value x) falls within region R_A, that is, $A+A'/2 < x \leqslant A+\varepsilon$, then we select A, where $\varepsilon = \kappa\sigma^*$. σ^* is the median of the standard deviations of all 3×3 neighborhoods centred around each non-border pixel in the image. For satellite images, κ is typically set to 2.
- Likewise, if the centre pixel falls within region $R_{A'}$, that is, $A'-\varepsilon \leqslant x < A+A'/2$ then we select A'.
- If x is midway between A and A', we simply select the average of A and A'.
- Finally, if x is an outlier with respect to A and A' so that $x > A+\varepsilon$ or $x < A'-\varepsilon$, we leave x unchanged.

Thus, four values are computed from the four pairs of pixels surrounding the central pixel in the 3×3 window. In the next step, the mean of these four values is computed. Finally, the central pixel x is replaced by the average of the value computed in the previous step and the centre pixel's original grey level value. This latter average is similar to that of a damped gradient descent, which yields a faster convergence.

The SNF filter is applied three times on the input image with $\varepsilon=0$, $\varepsilon=\kappa\sigma^*$ and $\varepsilon=0$. During the first run with $\varepsilon=0$, the input image is de-blurred. When the SNF is applied with $\varepsilon=\kappa\sigma^*$, it essentially smooths the pixels within a region and makes regions homogenous. During the last run with $\varepsilon=0$, edges are sharpened. During each run several iterations are performed until the pixel values remain unchanged

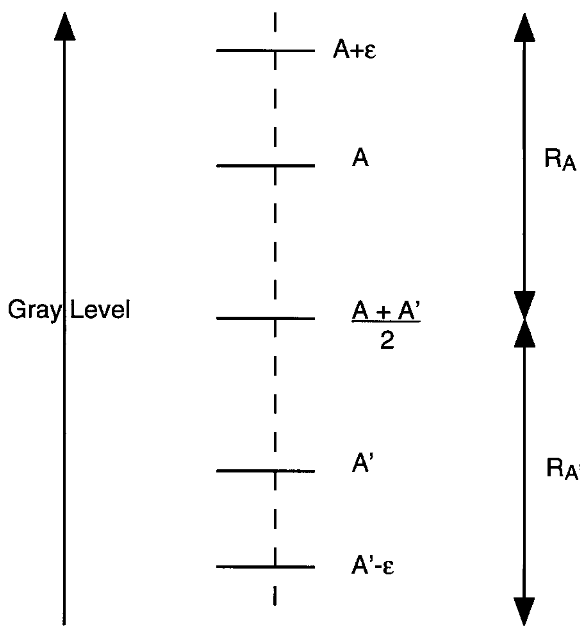


Figure 3. Selection criteria for the SNF filter.

between iterations. The resulting image has near-homogeneous regions with sharp transitions between bordering regions.

3.2. δ-*Connected Components*

A connected component in the image is a maximal collection of pixels with uniform reflectance such that a path exists between any pair of pixels in the component. Each pixel in the image will receive a label; pixels will have the same label if and only if they belong to the same connected component.

It is interesting to note that, in §3.1, we defined connected components as a maximal collection of uniform colour pixels such that a *path* exists between any pair of pixels. The conventional algorithm assumes that there is a connection between two adjacent pixels if and only if their grey level values are identical. We relax this connectivity rule and present it as a more general algorithm called δ-Connected Components. In this approach we assume that two adjacent pixels with values x and y are connected if their absolute difference $|x-y|$ is no greater than the threshold δ. Note that setting the parameter δ to 0 reduces the algorithm to the classic connected components approach. Thus, a series of image segmentations can be performed by varying δ. The lowest level in the hierarchy computes the connected components with $\delta=0$ for labelling the regions of the enhanced image. As δ increases, regions are merged together with respect to each region's decreasing similarity to its neighbouring regions. Typical values of δ are set to $\kappa\sigma^*$, where κ and σ^* are the same as those input to enhancement filters.

To demonstrate the performance of our technique, we applied the image enhancement and segmentation algorithms to an image generated by principal component analysis of the six reflective bands of a Landsat TM image. Figure 4(*a*) shows the first principal component image, which explains 57% of the variability contained in the six reflective bands. Principal component images are commonly used to reduce the dimensionality of the input data (e.g. Singh and Harrison 1985). Image enhancement was performed (figure 4(*b*)) by applying four iterations of the SNF with $\varepsilon=0$ for de-blurring edges, 46 iterations with $\varepsilon=11$ used to flatten interior regions, and 68 iterations again with $\varepsilon=0$ used to sharpen the edges. It can be noticed that the edges are enhanced, and different regions within the input image distinctly stand out. The enhanced image was then segmented using the connected component method with δ set to 11. Figure 4(*c*) shows the segmented image which contains 5476 regions.

Thus, image segmentation is computationally intensive and requires both image enhancement and connected component labelling. Segmenting a single band of a large TM scene (with roughly 17 million pixels) would take roughly 1 h 45 min on a single IBM RS6000 processor. Our research has produced image segmentation algorithms that scale well on a parallel computer. Using high performance computing techniques, the same segmentation of a remotely sensed image requires less than 9 min on an IBM SP2 with 16 processors. (Bader *et al.* 1996, Bader and JáJá 1996b).

Our parallel implementation strategy can be described as follows. The input image, which is an m by n matrix of pixels, is divided in to a number of equal-sized tiles, which are assigned one per processor. A parallel algorithm consists of a sequence of local computations interleaved with communication steps, where computation and communications may overlap. Divide and conquer algorithms typically use a recursive strategy to split problems into smaller sub-problems and, given the solutions to these sub-problems, merge the results into the final solution. In our implementation, the computational tasks (especially the iterations during SNF enhancement phase) are simplified by pre-fetching the neighbourhood cells around the border of each tile in a coordinated fashion from neighbouring processor's tile borders. Instead of communicating requests for individual pixels while computing at the borders of tiles (where the needed pixels are owned by other processors), we created an augmented data structure called 'ghost cells', which buffer the necessary pixels from adjacent tiles. Similarly, the communication between the processors is minimized by scheduling regular communication patterns and overlapping the transfer of border pixels with local computation. The code is written in C with the standard Message Passing Interface (MPI Forum 1995, 1997). The MPI is both portable and efficient on most current high performance parallel machines, for example, the IBM SP2, Cray T3E and Silicon Graphics Origin 2000.

4. Retrieval of the Bidirectional Reflectance Distribution Function (BRDF) from AVHRR data at a global scale

Understanding land surface anisotropy is critical in remote sensing studies because measurements of surface reflectance are dependent upon the view and illumination geometry. The BRDF describes variations in reflectance with illumination and view geometry, and can be mathematically described by Nicodemus *et al.* (1977):

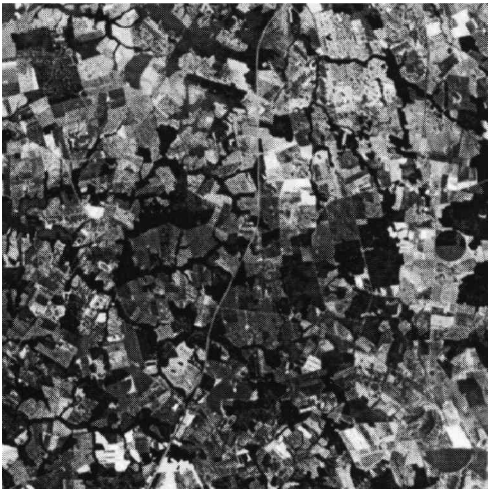
$$f_{\mathrm{r}}(\theta_{\mathrm{s}},\phi_{\mathrm{s}};\theta_{\mathrm{v}},\phi_{\mathrm{v}};\lambda)=\frac{\mathrm{d}L(\theta_{\mathrm{s}},\phi_{\mathrm{s}};\theta_{\mathrm{v}},\phi_{\mathrm{v}};\lambda)}{\mathrm{d}E(\theta_{\mathrm{s}},\phi_{\mathrm{s}};\lambda)} \tag{2}$$

where f_{r} is the BRDF (sr^{-1}) and dL is the reflected radiation for an incident beam of intensity dE at wavelength λ. θ and φ are the zenith and azimuth angles respectively. The subscripts s and v denote angles in the Sun and view directions respectively. The BRDF is commonly expressed as the bidirectional reflectance factor $\rho=f_{\mathrm{r}}\pi$.

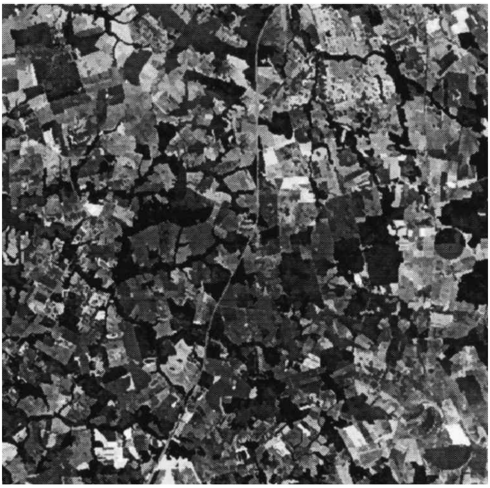
Since most current remote sensing instruments such as AVHRR and Landsat TM record measurements at fixed illumination and view geometries at any given time, surface anisotropy could introduce errors in multi-temporal analysis of these measurements unless they are corrected and normalized for consistent geometry (Holben 1986, Cihlar *et al.* 1994, Burgess and Pairman 1997). Researchers have shown that BRDF models can be used to infer surface properties such as the green leaf area index, and the amount of photosynthetically active radiation absorbed by plant canopies (Asrar *et al.* 1989, Myneni *et al.* 1995). The BRDF information is also critical for deriving broad band albedo for climate and energy balance modelling (Pinker and Laszlo 1990). Therefore, understanding surface BRDF is important in studies of land cover dynamics.

Reflectance measurements from future Earth Observing System (EOS) sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Multi-angle Imaging Spectroradiometer (MISR) are expected to provide the capability to retrieve surface BRDF operationally through a combination of well tested algorithms (Diner *et al.* 1996, Strahler and Muller 1997). However, these algorithms

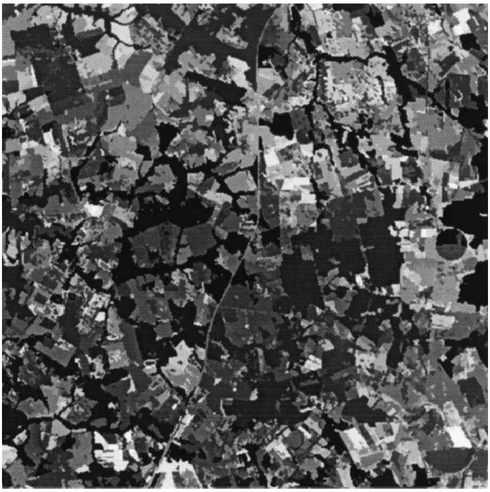
Figure 4. (*a*) First principal component image derived from the six reflective bands of TM. (*b*) Output image after SNF filtering. (*c*) Segmentation results.



(*a*)



(*b*)



(*c*)

have not been previously applied at a 'global scale', and there are several questions that need to be addressed:

- How do these algorithms perform at a global scale?
- What are the computational requirements? and how do we optimize the computational performance?

Our objective is to answer some of these questions by applying some of the BRDF algorithms chosen for MODIS and MISR to global AVHRR data.

Several studies aimed at deriving BRDF data from AVHRR data have been reported in the literature (see, e.g., Cihlar *et al.* 1994, Braswell *et al.* 1996, Privette *et al.* 1996). However, unlike these studies which used samples of pixels from imagery at varying spatial and temporal resolutions, we implemented two algorithms to retrieve BRDF data from global AVHRR images which were generated by the National Oceanic and Atmospheric Administration/National Aeronautic Space Administration (NOAA/NASA) Pathfinder program (James and Kalluri 1994).

From the suite of BRDF algorithms proposed to be implemented with data from the MODIS and MISR instruments, we have chosen the modified Walthall model (Walthall *et al.* 1985, Nilson and Kuusk 1989) and the Coupled Surface–Atmosphere Reflectance Model (CSAR) (Rahman *et al.* 1993a) for deriving BRDF from AVHRR data. We believe that these two algorithms are good candidates for addressing the previously mentioned issues, since they have been shown to work well over different cover types (e.g. Rahman *et al.* 1993a, Lewis *et al.* 1995, Russell *et al.* 1995, Diner *et al.* 1996, O'Neill *et al.* 1997, Privette *et al.* 1997).

The modified Walthall model is an empirical model, which describes the surface BRDF as a quadratic function of view, solar and relative azimuth angles:

$$\rho(\theta_v, \theta_s, \phi_v, \lambda) = a_0(\theta_v^2 + \theta_s^2) + a_1\theta_v^2\theta_s^2 + a_2\theta_v\theta_s\cos\phi + a_3 \tag{3}$$

This model has four linear coefficients which are derived by a least-squares method using the Gauss elimination technique (Stoer and Burlirsch 1993). φ is the relative azimuth angle ($\varphi = \varphi_v - \varphi_s$). The only parameter in this model that has a physical meaning is a_3, and this parameter represents the nadir reflectances for an overhead Sun.

A modified Minnaert function (Minnaert 1941), a one-term Henyey and Greenstein function (Henyey and Greenstein 1941), and a hot spot function (Pinty *et al.* 1990) are used to describe the BRDF in the CSAR model with three unknown parameters ρ_0, k, and Θ:

$$\rho(\theta_v, \phi_v, \theta_s, \phi_s, \lambda) = \rho_0 \frac{\cos^{k-1}\theta_v \cos^{k-1}\theta_s}{(\cos\theta_v + \cos\theta_s)^{1-k}} F(g)[1 + R(G)] \tag{4}$$

where

$$F(g) = \frac{1 - \Theta^2}{[1 + \Theta^2 - 2\Theta\cos(\pi - g)]^{1.5}} \tag{4a}$$

$$\cos g = \cos\theta_s \cos\theta_v + \sin\theta_s \sin\theta_v \cos\phi \tag{4b}$$

$$1 + R(G) = 1 + \frac{1 - \rho_0}{1 + G} \tag{4c}$$

$$G = (\tan^2\theta_v + \tan^2\theta_s - 2\tan\theta_s \tan\theta_v \cos\phi)^{1/2} \tag{4d}$$

ρ_0 represents nadir reflectances and the parameter k indicates the level of surface anisotropy with a range of 0 to 1. A surface with lower values of k is more anisotropic than a surface with higher values of k. Θ is a parameter that determines the relative amount of forward and backward scattering. The three unknown coefficients of the CSAR model (ρ_0, k and Θ) are determined by model inversion and iteration.

The coefficients for both models are derived statistically for a set of observations measured at different view and illumination geometries. Since the AVHRR instrument measures reflectance over a given target only once a day, data from several days have to be used to get a good sampling of the BRDF at different view and solar angles. In our study, we used only the 10-day maximum value Normalized Difference Vegetation Index (NDVI) composite data acquired during the time period 1983 to 1986, since maximum value compositing has been shown to reduce the effects of clouds and atmosphere by choosing the clearest picture during the 10-day period (Holben 1986). Each 10-day composite image from the Pathfinder AVHRR Land (PAL) data set has calibrated reflectances and brightness temperatures from all the five bands of the AVHRR, along with view and solar angles for each pixel at a spatial resolution of 8 km. Cloud condition flags (CLAVR) are also present for each pixel, which were generated using a variety of tests (Stowe *et al.* 1991). The visible and near IR reflectances from AVHRR have been corrected for Rayleigh scattering and ozone absorption. For a detailed description of the PAL data set refer to the work by James and Kalluri (1994).

In order to isolate the influence of surface phenology in the BRDF signal, the input data were divided into four quarters: January–March, April–June, July–September, October–December. Our analysis indicates that increasing the temporal resolution by more than three months introduced phenological effects in the BRDF signal, while at the same time decreasing the time period by less than three months did not provide enough information to determine the BRDF accurately. Only clear pixels identified by the PAL cloud mask were used in the BRDF analysis, and the coefficients for both the modified Walthall model and the CSAR model were determined by inversion for the four quarters.

Each global composite image is an array of 5004×2168 pixels. There are nine files for each composite, and we have processed data from 144 compositing periods. Thus 1296 files were ingested, having a total volume of 27 Gbytes.

For the modified Walthall model, 3600 Floating Point Operations (FLOPs) are required per pixel to derive the coefficients. In comparison, the CSAR model is computationally more expensive (because of its nonlinear nature) and requires 3.9 MFLOPs per pixel. Thus, for the 2.5 million land pixels in the PAL data set, 9 GFLOPs are required to solve the modified Walthall model, and 9750 GFLOPs are needed for the CSAR model. Given the peak performance of 266 MFLOPs per second of a single RS6000 processor, our estimates indicate that 3 h of processing time would be required to determine the coefficients of the modified Walthall model using a single CPU. Using a single CPU requires 42 h to solve the BRDF for the CSAR model.

To cut down the processing time, we implemented the two BRDF algorithms on the IBM SP2 machine previously described in §2. The input land data were evenly distributed among all the 16 nodes, and communication between the processors was minimized. We achieved a performance rate of 0.9 GFLOPs per second for the modified Walthall model, and 1.2 GFLOPs per second for the CSAR model on the SP2, which significantly reduced the runtime to 10 min for the modified Walthall model, and 2.5 h for the CSAR. It is clear from these results that our parallel implementation achieves an almost perfect linear speed-up.

Although the CSAR model is more computationally intensive than the modified Walthall model, the results from both models are very similar for the data set we analysed. Figure 5 shows the standard errors in channel 1 and 2 reflectances from the CSAR model and the modified Walthall model for the third quarter (July–September). The histograms of the standard errors show that both the models perform equally well at a global scale. However, since the model coefficients of the CSAR model are based on physical principles, these coefficients are closely related to the land cover type and provide more information about land cover dynamics compared to the model coefficients of the modified Walthall model (Zhang *et al.* 1998).

Figures 6(*a*) and 7(*a*) show the directional reflectances in channels 1 and 2 from the first 10-day composite in July 1983, and figures 6(*b*) and 7(*b*) show the hemispherical albedo derived from the modified Walthall model for the same bands. The hemispherical albedo $[\rho(\theta_s, \lambda)]$ is derived by integrating the BRDF over the exitance

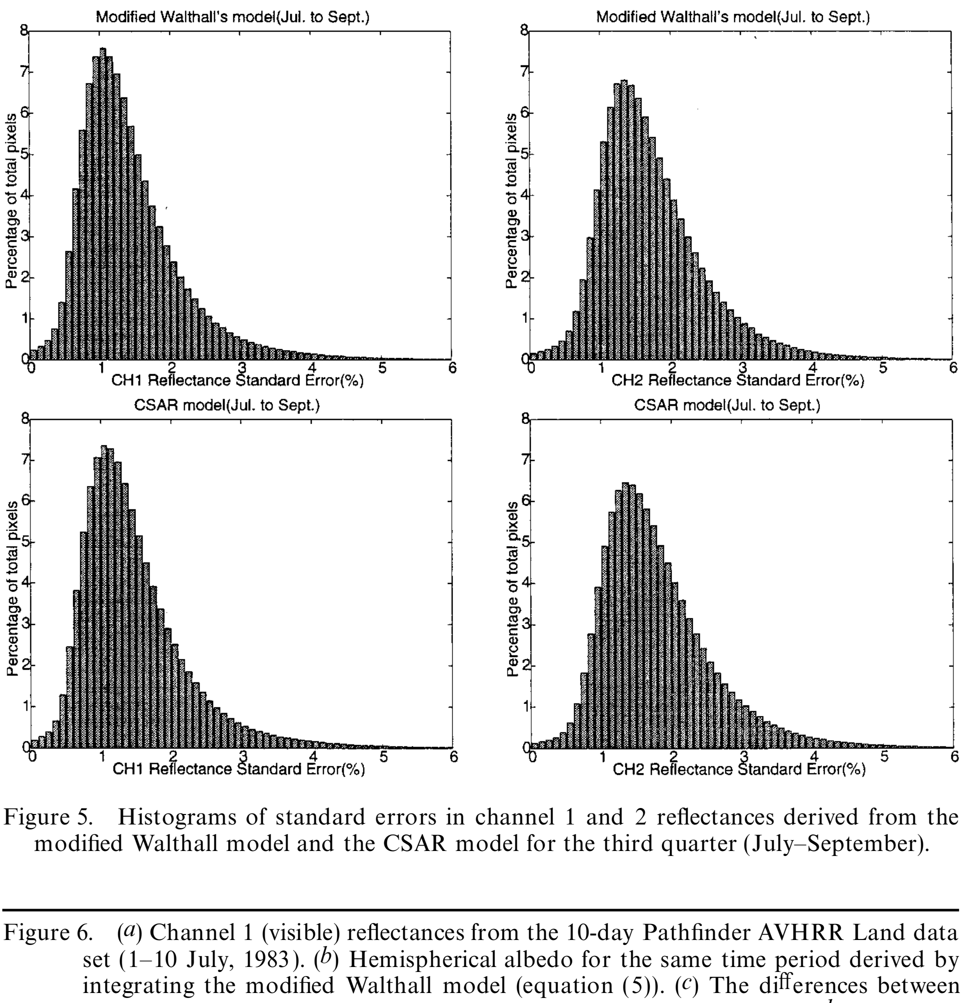
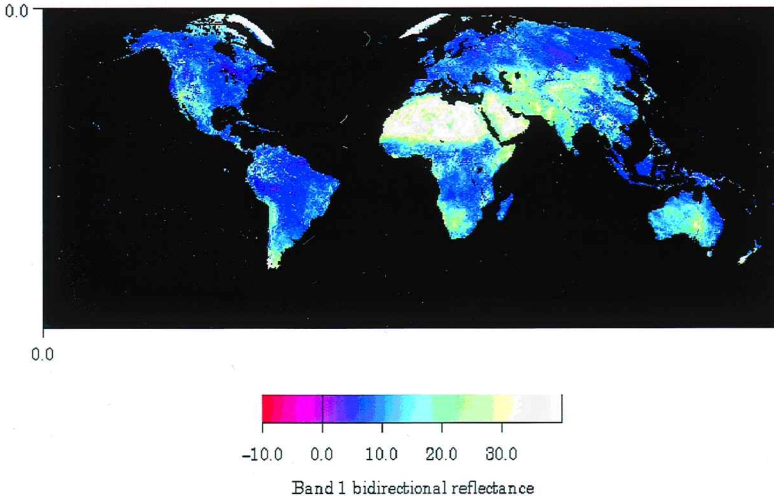
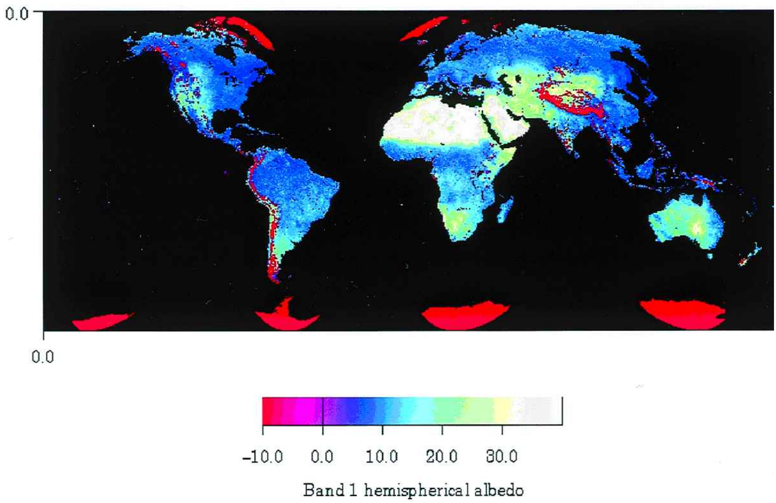
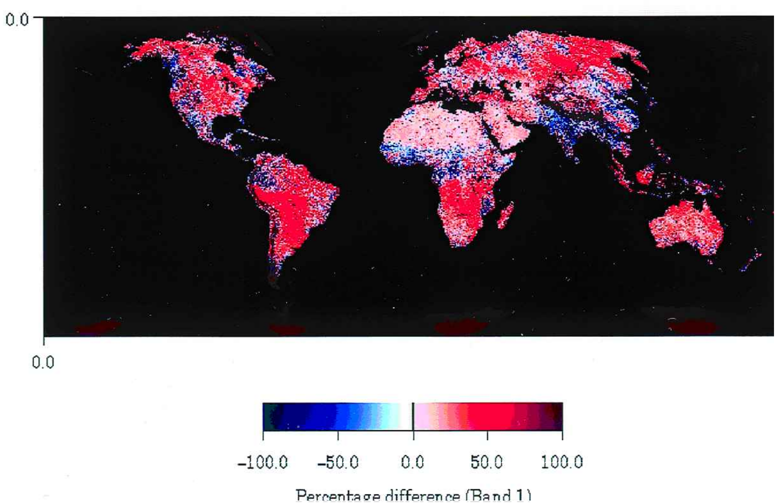


Figure 5. Histograms of standard errors in channel 1 and 2 reflectances derived from the modified Walthall model and the CSAR model for the third quarter (July–September).

Figure 6. (*a*) Channel 1 (visible) reflectances from the 10-day Pathfinder AVHRR Land data set (1–10 July, 1983). (*b*) Hemispherical albedo for the same time period derived by integrating the modified Walthall model (equation (5)). (*c*) The differences between directional reflectances (figure 6(*a*)) and hemispherical albedo (figure 6(*b*)) expressed as a percentage of the hemispherical albedo.


0.0
0.0
-10.0 0.0 10.0 20.0 30.0
Band 1 bidirectional reflectance
(*a*)


0.0
0.0
-10.0 0.0 10.0 20.0 30.0
Band 1 hemispherical albedo
(*b*)


0.0
0.0
-100.0 -50.0 0.0 50.0 100.0
Percentage difference (Band 1)
(*c*)

hemisphere for a single irradiance direction, and is a key parameter used in energy balance models:

$$\rho(\theta_s, \lambda) = \frac{1}{\pi}\int_0^{2\pi}\int_0^{\pi/2} \rho(\theta_s, \theta_v, \phi, \lambda)\cos\theta_v \sin\theta_v \, d\theta_v \, d\phi \tag{5}$$

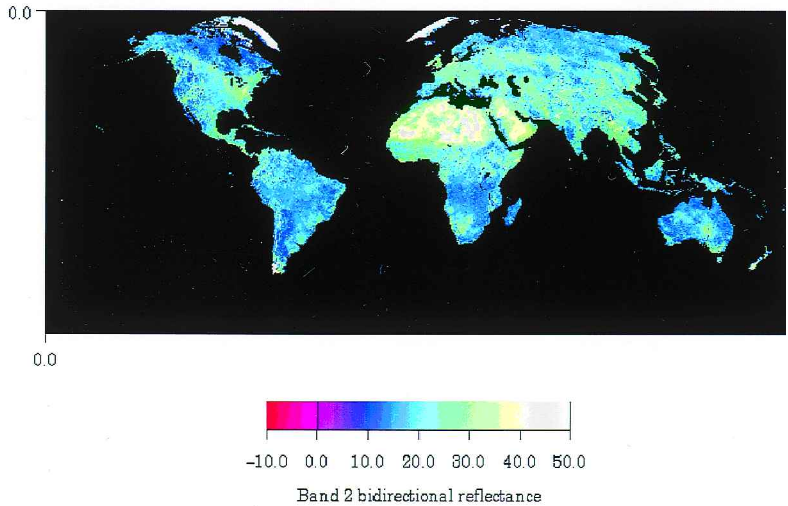
The differences between directional reflectance and hemispherical albedo, expressed as a percentage of the hemispherical albedo, are shown in figures 6(*c*) and 7(*c*) for bands 1 and 2 respectively. It can be seen from figures 6(*c*) and 7(*c*) that the differences between directional reflectances and hemispherical albedo are large, especially in densely vegetated areas such as the tropical rain forests of South America and Central Africa, and in the temperate regions of Asia and Europe. Also, since visible reflectances are more anisotropic than near-IR reflectances, these differences are larger in channel 1. Differences in channel 1 directional reflectance and hemispherical albedo are significant, having typical values of differences greater than 25% in vegetated areas. This indicates that the use of reflectances without BRDF correction could result in large errors in modelling land cover dynamics. Deserts on the other hand show smaller difference between directional reflectance and hemispherical albedo. Areas coloured in red (Himalayas, Andes, etc.) have null values since these areas were flagged as 'cloudy' by CLAVR.

For the first time, to our knowledge, we have demonstrated the feasibility of applying simple as well as complex BRDF algorithms at a global scale using high performance computing techniques. The results of this analysis are significant since they not only provide us with valuable information about the performance of these algorithms chosen for MODIS and MISR at a global scale, but also help us in gauging the computational requirements of processing large volumes of data operationally.

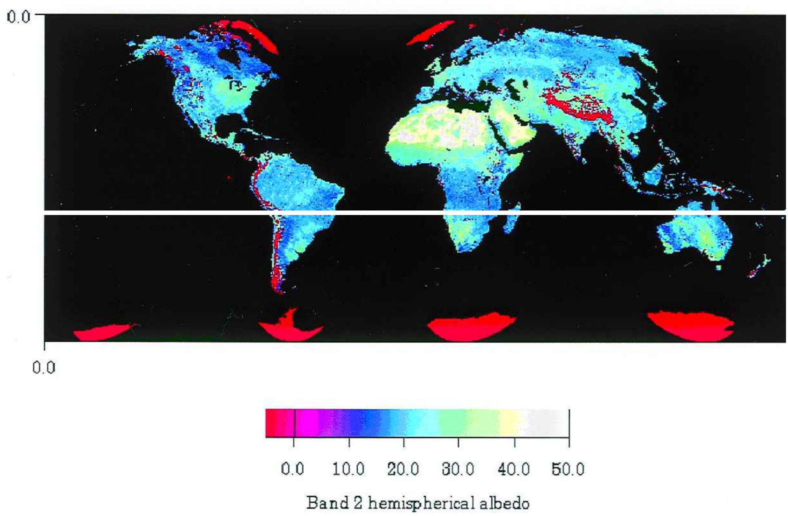
5. Designing a high performance system for processing, storing and retrieving AVHRR data

Several global data sets derived from the AVHRR instrument have been produced to study land cover dynamics since 1981. These include several versions of the Global Vegetation Index (GVI) products (Kidwell 1990, Goward *et al.* 1993, 1994, Gutman *et al.* 1994), the continental NDVI data set produced by the Global Inventory Monitoring and Modeling Studies (GIMMS) group at NASA's Goddard Space Flight Center (GSFC) (Holben 1986, Los *et al.* 1994), the Pathfinder AVHRR Land (PAL) data set (James and Kalluri 1994) and the 1 km global land data produced by the Earth Resources Observation Systems (EROS) Data Center (EDC) (Eidenshink and Faundeen 1994). A comprehensive review of these data can be found in the work of Townshend (1994). Although these data sets have found widespread use among the Earth system science community, they have several inherent limitations. Some of the limitations include availability of data in a fixed geographic projection, spatial and temporal resolution, with little or no capability

Figure 7. (*a*) Channel 2 (near-IR) reflectances from the 10-day Pathfinder AVHRR Land data set (1–10 July, 1983). (*b*) Hemispherical albedo for the same time period derived by integrating the modified Walthall model (equation (5)). (*c*) The differences between directional reflectances (figure 7(*a*)) and hemispherical albedo (figure 7(*b*)) expressed as a percentage of the hemispherical albedo.


0.0
0.0
-10.0 0.0 10.0 20.0 30.0 40.0 50.0
Band 2 bidirectional reflectance

(*a*)


0.0
0.0
0.0 10.0 20.0 30.0 40.0 50.0
Band 2 hemispherical albedo

(*b*)

for generating subsets based on user requirements. The compositing method and time interval are also static for all the previously mentioned AVHRR data sets. Some of these AVHRR data sets also have atmospheric correction applied to them. However, users who are interested in retrieving atmospheric properties from uncorrected satellite measurements and those users who want to experiment with new atmospheric correction algorithms may prefer to obtain uncorrected satellite data.

Because of the limitations of these data sets, the full potential capabilities of the AVHRR instrument cannot be exploited for special applications such as monitoring global Net Primary Production (NPP) (Prince and Goward 1995). Moreover, the requirements of users differ depending upon their specific application and usage of the data sets, and these requirements are expected to change as they gain experience in the use of these data sets and as their needs evolve (Townshend 1994). Thus, there is an important need to design a processing system that can generate AVHRR data sets following the specifications of individual users. Using the algorithms developed by the Pathfinder II group (El Saleous *et al.* 2000) we have designed an integrated processing system that can process, archive and distribute AVHRR data to the scientific community as per individual data requirements.

An end-to-end design of a system to process AVHRR data tailored to different user requirements is a complex task, especially since the volume of the raw level 1B data is about 220 Gbytes per year. Moreover, to generate a data set suitable for the study of land cover dynamics, several complex algorithms have to be applied to the raw data for navigation and geolocation of pixels, radiometric calibration, atmospheric correction and compositing (Townshend *et al.* 1994). We have designed and built a comprehensive prototype system that allows researchers to request and generate data according their needs by specifying:

- region of interest;
- map projection;
- spatial resolution;
- temporal resolution;
- land/ocean data;
- atmospheric correction;
- cloud screening;
- compositing function.

Thus, the user would specify the spatial and temporal resolution, along with a compositing criterion (e.g. maximum NDVI) and the data would be generated in the geographic projection of his/her choice. Also, the user would have the option of either obtaining a data set corrected for atmospheric affects or just calibrated data. Either land or ocean data alone can be retrieved as well.

To achieve optimum performance, the system consists of two major components. The first involves processing modules that are common to all user queries, and the second involves general indexing, search and retrieval procedures in addition to a library of processing functions that will be unique to specific queries. The first stage can be considered as a pre-processing stage, and involves ingesting satellite orbits, precise navigation using satellite ephemeris, calibration of all the five AVHRR channels, determination of cloud condition information and data quality. Algorithms developed by El Saleous *et al.* (2000) have been used for pre-processing level 1B data. It should be noted that no gridding or re-sampling of the data is done during the pre-processing stage. Once the pre-processing is complete, the data are indexed and placed on a disk array for fast retrieval and generation of user-specified products. For each pixel the following parameters are stored: latitude and longitude, calibrated reflectances and brightness temperatures from the five bands of AVHRR, view and solar geometry, cloud and quality flags, and date and time of observation.

Our data indexing scheme is novel, and can be summarized as follows. We first create a two-dimensional grid with a spatial resolution of $1° \times 1°$. Each $1° \times 1°$ cell can be considered as a 'bucket' containing all the Instantaneous Fields of View (IFOV) that have been navigated into that region. Within each bucket all the pixels are indexed using an optimal *k–d* tree spatial data structure. The *k–d* tree data structure is a hierarchical spatial data structure in which the *k*-dimensional space is recursively divided (Samet 1990). In our case, *k* represents a two-dimensional space denoted by latitudes and longitudes. Thus, the $1° \times 1°$ space is recursively divided in to finer regions hierarchically, based on latitudes and longitudes, and a binary search tree is constructed. Note that this indexing ensures that all the navigated IFOVs are preserved absolutely with no loss in information.

The *k–d* tree for our case can be described as follows. The first branch within the $1° \times 1°$ tree represents the median latitude of all the pixels within the cell. All the pixels within the $1° \times 1°$ cell are divided into two regions by the median latitude. Within each region further division is made by the median longitude of the pixels. Thus, the latitude and the longitude are used alternatively to decompose the entire space. Finally, a global index is created which can be used to retrieve the pixels for a given Cartesian coordinate. A detailed explanation of the *k–d* data structure can be found in the work of Bentley (1975). The *k–d* tree architecture is better for representing point data compared to other methods such as quadtrees (Samet 1989).

Once a user specifies a spatio-temporal query, the global index is searched and all the IFOVs that fall within the spatial bounding box specified by the user are retrieved. The next step is atmospheric correction and compositing based on user-defined criteria. Our current version of implementation supports corrections for Rayleigh scattering, ozone and water vapour absorption, and stratospheric aerosols. After these corrections have been applied, the data are binned into one of the thirty one geographic projections supported by our processing system, using a nearest neighbour inverse binning algorithm (Emery *et al.* 1989, Cracknell 1997).

Once the data are ingested into the system, it takes 22 min on an IBM RS6000 to generate an AVHRR image with all the calibrated bands along with locational and geometry information for each pixel at a spatial resolution of 8 km for data from a single day. The image generated for our benchmark includes both land as well as ocean data. Although the current system design is prototyped with AVHRR GAC data only, it can be further expanded to support geospatial raster data from other satellites as well. Storing the satellite data in a hierarchical data structure has significant advantages since it allows us to use the spatial and temporal metadata in a relational data base environment to perform Boolean searches and logical operations on multiple data sets using queries defined by the user.

6. Conclusions

In this paper we presented an overview of some of the high performance methodologies that we have implemented to study land cover dynamics. Algorithms we have developed are scalable and portable, and run on both serial as well as parallel architectures. Some of the applications are computationally intensive (e.g.

atmospheric correction, image segmentation and BRDF retrieval), while others are more input/output (I/O) bound (e.g. AVHRR data processing).

High performance computing technology has matured to the point where most new high-end servers and workstations contain multiple processors that can be programmed using standard systems software. The technical challenge in handling our applications has been in improving the overall complexity and in mapping the computation into multiple processors in such a way as to ensure load balancing across the nodes and to minimize the overhead incurred by communication and synchronization. All of our computationally intensive algorithms achieve a linear speed-up in terms of the number of processors available and, in addition, their sequential complexity is superior to any of the previously known implementations.

Our data-intensive applications, such as the on-demand generation of user-specified AVHRR data products, have required the development of novel indexing schemes and particular data placement methods across the available disks so as to achieve the maximum possible parallel I/O throughput. These techniques are currently being generalized to handle a wide variety of spatial and temporal data sets, including the fusion of multiple data sets with different spatial and temporal resolutions. The current prototype system built especially for AVHRR GAC Level 1B data provides an unprecedented flexibility in generating user-specified data products, which we hope to extend to handle data fusion and correlation of data sets from multiple sensors, especially data collected by future EOS satellites.

Although our examples show that significant improvements in processing speeds can be achieved by using high performance computing for processing satellite data, in practice, high performance computing comes at a certain cost. The learning curve for developing scalable parallel programs is steep, and parallel computers' runtime environment is more unpredictable than that of a single processor system (Pancake 1996). It is thus worthwhile investigating resources in developing parallel implementation of algorithms that are well understood and robust compared to those that are still in experimental stages of development—otherwise the initial costly investment could be largely wasted.

Data sets from future EOS satellites are expected to have routine operational corrections for atmospheric effects and BRDF at a global scale. These corrections are crucial for studying land cover dynamics and are computationally intense. We have tried to address some of the issues at both the algorithm level as well as at the computational level. Our results show that unless innovative approaches are used, processing high resolution global data sets operationally is not practical. For the first time we are able to ingest and correct complete TM scenes for atmospheric effects within a few minutes. Classification of regions in remotely sensed imagery requires the computationally intensive tasks of image enhancement and segmentation. Our parallel version of the hierarchical connected component algorithm can segment a TM scene in minutes compared to a few hours on a high performance computer. The results from our BRDF study are unique since they provide valuable information about the applicability, accuracy and processing time of different algorithms at a global scale which was not previously available. Efficient access to massive spatial data bases (such as global AVHRR data) and processing them in a real time mode requires the engineering of a processing system using a hierarchical data structure. The sequential hierarchical data structure that we used in designing the AVHRR database provides very efficient random search methods by capitalizing on the physical organization of the data. The results from our analysis are expected to provide new insights into the implementation of remote sensing algorithms at a global scale for studying land cover dynamics, and in the design of new processing methodologies using smart and innovative computational techniques.

Acknowledgments

This research was made possible by a grant from the National Science Foundation (BIR9318183). David Bader was supported by the NSF CISE postdoctoral research associateship in experimental computer science (No 96-25668). We wish to thank Jeff Masek for providing us with the Landsat computational requirements. Jeff Privette's comments on the manuscript are also acknowledged.

References

Acharya, A., Uysal, M., Bennett, R., Mendelson, A., Beynon, M., Hollingsworth, J., Saltz, J., and Sussman, A., 1996, Tuning the performance of I/O-intensive parallel applications. *Proceedings of the Fourth ACM Workshop on I/O in Parallel and Distributed Systems, Philadelphia, 27 May 1996* (Association for Computing Machinery), pp. 15–27.

Asrar, G., Myneni, R. B., and Kanemasu, E. T., 1989, In *Estimation of plant-canopy attributes from spectral reflectance measurements*, edited by G. Asrar (New York: Wiley), pp. 252–296.

Bader, D. A., and JàJà, J., 1996a, Parallel algorithms for image histogramming and connected components with an experimental study. *Journal of Parallel and Distributed Computing*, **35**, 173–190.

Bader, D. A., and JàJà, J., 1996b, Practical parallel algorithms for dynamic data redistribution, median finding, and selection. *Proceedings of the 10th International Parallel Processing Symposium, Honolulu, HI* (IEEE), pp. 292–301.

Bader, D. A., JàJà, J., Harwood, D., and Davis, L. S., 1996, Parallel algorithms for image enhancement and segmentation by region growing with an experimental study. *The Journal of Supercomputing*, **10**, 141–168.

Bentley, J. L., 1975, Multidimensional binary search trees for associative searching. *Communications of the Association for Computing Machinery*, **18**, 509–517.

Braswell, B. H., Schimel, D. S., Privette, J. L., Moore III, B., Emery, W. J., Sulzman, E. W., and Hudak, A. T., 1996, Extracting ecological and biophysical information from AVHRR optical data: an integrated algorithm based on inverse modeling. *Journal of Geophysical Research*, **101**, 23 335–23 348.

Burgess, D. W., and Pairman, D., 1997, Bidirectional reflectance effects in NOAA AVHRR data. *International Journal of Remote Sensing*, **18**, 2815–2825.

Castleman, K. R., 1996, *Digital Image Processing* (Englewood Cliffs, NJ: Prentice Hall).

Chandrasekhar, S., 1960, *Radiative Transfer* (New York: Dover).

Chang, Y. L., and Li, X., 1994, Adaptive image region-growing. *IEEE Transactions on Image Processing*, **3**, 868–872.

Cihlar, J., Manak, D., and Voisin, N., 1994, AVHRR bidirectional reflectance effects and compositing. *Remote Sensing of Environment*, **48**, 77–88.

Cracknell, A. P., 1997, *The Advanced Very High Resolution Radiometer (AVHRR)* (London: Taylor & Francis).

Diner, D. J., Davies, R., Varnai, T., Borel, C., and Gerstl, S. A. W., 1996, *MISR-Level 2 Top-of-Atmosphere albedo algorithm theoretical basis* (JPL). *http://eospso.gsfc.nasa.gov/atbd/misrtables.html*

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

El Saleous, N. Z., Vermote, E. F., Justice, C. O., Townshend, J. R. G., Tucker, C. J., and Goward, S. N., 2000, Improvements in the global biospheric record from the Advanced Very High Resolution Radiometer (AVHRR). *International Journal of Remote Sensing*, **21**, 1251–1277.

Emery, W. J., Brown, J., and Nowak, Z. P., 1989, AVHRR image navigation: summary and review. *Photogrammetric Engineering and Remote Sensing*, **8**, 1175–1183.

Fallah-Adl, H., JàJà, J., and Liang, S., 1996a, Fast algorithms for estimating aerosol optical depth and correcting thematic mapper imagery. *The Journal of Supercomputing*, **10**, 315–329.

Fallah-Adl, H., JàJà, J., Liang, S., Townshend, J. R. G., and Kaufman, Y. J., 1996b, Fast algorithms for removing atmospheric effects from satellite images. *IEEE Computational Science and Engineering*, **3**, 66–77.

Forster, B. C., and Best, P., 1994, Estimation of SPOT P-mode point spread function and derivation of a deconvolution filter. *ISPRS Journal of Photogrammetry and Remote Sensing*, **49**, 32–42.

Fraser, R. S., Ferrare, R. A., Kaufman, Y. J., Markham, B. L., and Mattoo, S., 1992, Algorithm for atmospheric correction of aircraft and satellite imagery. *International Journal of Remote Sensing*, **13**, 541–557.

Fraser, R. S., and Kaufman, Y. J., 1985, The relative importance of scattering and absorption in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, **23**, 625–633.

Goward, S. N., Dye, D. G., Turner, S., and Yang, J., 1993, Objective assessment of the NOAA Global Vegetation Index data product. *International Journal of Remote Sensing*, **14**, 3365–3394.

Goward, S. N., Turner, S., Dye, D. G., and Liang, S., 1994, The University of Maryland improved Global Vegetation Index data set for terrestrial monitoring. *International Journal of Remote Sensing*, **15**, 3365–3395.

Goward, S. N., and Williams, D. L., 1997, Monitoring the terrestrial biosphere: Landsat and Earth system science. *Photogrammetric Engineering and Remote Sensing*, **63**, 887–900.

Gutman, G. G., Ignatov, A. M., and Olson, S., 1994, Towards better quality of AVHRR composite images over land: reduction of cloud contamination. *Remote Sensing of Environment*, **50**, 134–148.

Haralick, R. M., and Shapiro, L. G., 1985, Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, **29**, 100–132.

Harwood, D., Subbarao, M., Hakalahti, H., and Davis, L. S., 1987, A new class of edge-preserving smoothing filters. *Pattern Recognition Letters*, **6**, 155–162.

Henyey, L. G., and Greenstein, T. L., 1941, Diffuse radiation in the galaxy. *Astrophysical Journal*, **93**, 70–83.

Holben, B., 1986, Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, **7**, 1417–1434.

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

Kaufman, Y. J., 1984, The atmospheric effect on seperatability of field classes measured from satellite. *Remote Sensing of Environment*, **18**, 21–34.

Kaufman, Y. J., and Holben, B. N., 1996, Hemispherical backscattering by biomass burning and sulfate particles derived from sky measurements. *Journal of Geophysical Research*, **101**, 19 433–19 445.

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

Kaufman, Y. J., and Tanre, D., 1996, Strategy for direct and indirect methods for correcting the aerosol effect on remote sensing: from AVHRR to EOS-MODIS. *Remote Sensing of Environment*, **55**, 65–79.

Kaufman, Y. J., Tanre, D., Remer, L. A., Vermote, E. F., Chu, A., and Holben, B. N., 1997, Operational remote sensing of tropospheric aerosols over land from EOS moderate resolution imaging spectroradiometer. *Journal of Geophysical Research*, **102**, 17 051–17 067.

Kidwell, K. B., 1990, *Global Vegetation Index User's Guide* (Washington DC: NOAA/NESDIS, National Climatic Data Center).

Kneizys, F. X., Shettle, E. P., Abreau, L. W., Chetwynd, J. H., Anderson, G. P., Gallery, W. O., Selby, J. E. A., and Clough, S. A., 1988, *User's guide to LOWTRAN7* (Hanscom AFB, MA: Airforce Geophysical Laboratory).

Lewis, P., Barnsley, M. J., Sutherland, M., and Muller, J.-P., 1995, Estimating land surface albedo in the HAPEX-Sahel Experiment: model-based inversion using ASAS.

International Geoscience and Remote Sensing Symposium, Florence, Italy, 10–14 July 1995 (IEEE), pp. 2221–2223.

Liang, S., Fallah-Adl, H., Kalluri, S., Jàjà, J., Kaufman, Y. J., and Townshend, J. R. G., 1997, An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research*, **102**, 17 173–17 186.

Los, S. O., Justice, C. O., and Tucker, C. J., 1994, A global 1° by 1° NDVI data set for climate studies derived from the GIMMS continental NDVI data. *International Journal of Remote Sensing*, **15**, 3493–3518.

MPI (Message Passing Interface) Forum, 1995, *MPI: A Message-Passing Interface Standard* (Knoxville, TN: University of Tennessee Press).

MPI (Message Passing Interface) Forum, 1997, *MPI-2: Extensions to the Message-Passing Interface* (Knoxville, TN: University of Tennessee Press).

Minnaert, M., 1941, The reciprocity principle in lunar photometry. *Astrophysical Journal*, **93**, 403–410.

Myneni, R. B., Maggion, S., Iaquinta, J., Privette, J. L., Gobron, N., Pinty, B., Kimes, D. S., Verstraete, M. M., and Williams, D. L., 1995, Optical remote sensing of vegetation: modeling, caveats, and algorithms. *Remote Sensing of Environment*, **51**, 169–188.

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., and Limperis, T., 1977, *Geometrical considerations and nomenclature for reflectance* (Washington DC: National Bureau of Standards).

Nilson, T., and Kuusk, A., 1989, A reflectance model for homogeneous plant canopies and its inversion. *Remote Sensing of Environment*, **27**, 157–167.

O'Neill, N. T., Zagolski, F., Bergeron, M., Royer, A., Miller, J. R., and Freemantle, J., 1997, Atmospheric correction validation of casi images acquired over the Boreas southern study area. *Canadian Journal of Remote Sensing*, **23**, 143–162.

Pancake, C. M., 1996, Is parallelism for you? *IEEE Computer Science and Engineering*, **3**, 18–36.

Pinker, R. T., and Laszlo, I., 1990, Improved prospects for estimating insolation for calculating regional evapotranspiration from remotely sensed data. *Agricultural and Forest Meteorology*, **52**, 227–251.

Pinty, B., Verstraete, M. M., and Dickinson, R. E., 1990, A physical model of the bidirectional reflectance of vegetation canopies, 2. Inversion and validation. *Journal of Geophysical Research*, **95**, 11 767–11 775.

Poropat, G. V., 1993, Effect of system point spread function, apparent size, and detector instantaneous field of view on the infrared image contrast of small objects. *Optical Engineering*, **32**, 2598–2607.

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

Privette, J. L., Eck, T. F., and Deering, D. W., 1997, Estimating spectral albedo and nadir reflectance through inversion of simple BRDF models with AVHRR/MODIS-like data. *Journal of Geophysical Research*, **102**, 29 529–29 542.

Privette, J. L., Emery, W. J., and Schimel, D. S., 1996, Inversion of a vegetation reflectance model with NOAA AVHRR data. *Remote Sensing of Environment*, **58**, 187–200.

Rahman, H., Pinty, B., and Verstraete, M. M., 1993a, Coupled surface–atmospheric reflectance (CSAR) model. 2. Semiempirical surface model usable with NOAA Advanced Very High Resolution Radiometer. *Journal of Geophysical Research*, **98**, 20 791–20 801.

Rahman, H., Verstraete, M. M., and Pinty, B., 1993b, Coupled surface–atmosphere reflectance model (CSAR). 1. Model description and inversion against synthetic data. *Journal of Geophysical Research*, **98**, 20 779–20 789.

Russell, C. A., Walthall, C. L., Irons, J. R., and Brown de Colstoun, E. C., 1995, Comparison of airborne and surface spectral bidirectional reflectance factors, spectral hemispherical reflectance and spectral vegetation indices. *Journal of Geophysical Research*, **100**, 25 509–25 522.

Samet, H., 1989, *The Design and Analysis of Spatial Data Structure* (Reading, MA: Addison Wesley).

Samet, H., 1990, *Applications of Spatial Data Structures Structure* (Reading, MA: Addison Wesley).

Sellers, P. J., Hall, F. G., Asrar, G., Strebel, D. E., and Murphy, R. E., 1992, An overview of the First International Land Surface Climatology Project (ISLSCP) field experiment (FIFE). *Journal of Geophysical Research*, **97**, 18 345–18 371.

Singh, A., and Harrison, A., 1985, Standardized principal components. *International Journal of Remote Sensing*, **6**, 883–896.

Singh, S. M., and Saull, R. L., 1988, The effect of atmospheric correction on the interpretation of multitemporal AVHRR-derived vegetation index dynamics. *Remote Sensing of Environment*, **25**, 37–51.

Stoer, J., and Burlirsch, R., 1993, *Introduction to Numerical Analysis* (New York: Springer-Verlag).

Stowe, L. L., McClain. E. P., Carey, R., Pellegrino, P., Gutman, G. G., Davis, P., Long, C., and Hart, S., 1991, Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Advances in Space Research*, **3**, 51–54.

Strahler, A. H., and Muller, J., 1997, MODIS BRDF/Albedo Product: Algorithm Theoretical Basis Document (Version 3.2). http://eospso.gsfc.nasa.gov/atbd/modistables.html

Tanre, D., Holben, B. N., and Kaufman, Y. J., 1992, Atmospheric correction algorithm for NOAA-AVHRR products: theory and applications. *IEEE Transactions on Geoscience and Remote Sensing*, **30**, 231–248.

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**, 3319–3332.

Townshend, J. R. G., Cushine, J., Hardy, J. R., and Wilson, A., 1988, *Thematic Mapper Data—Characteristics and Use* (Swindon: National Environment Research Council).

Townshend, J. R. G., Justice, C. O., Skole, D., Malingreau, J. P., Cihlar, J., Teillet, P., Sadowski, F., and Ruttenburg, S., 1994, The 1 km resolution global data set: needs of the International Geosphere Biosphere Programme. *International Journal of Remote Sensing*, **15**, 3417–3441.

Vermote, E. F., Tanre, D., Deuze, J. L., Herman, M., and Morcrette, J., 1997, Second simulation of the satellite signal in the solar spectrum, 6S: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 675–686.

Walthall, C. L., Norman, J. M., Welles, J. W., Campbell, G., and Blad, B. L., 1985, Simple equation to approximate the bidirectional reflectance from vegetated canopies and bare soil surfaces. *Applied Optics*, **24**, 383–387.

Westman, T., Harwood, D., Laitinen, T., and Pietikaeinen, M., 1990, Color segmentation by hierarchial connected components analysis with image enhancement by symmetric neighborhood filters. *Proceedings of the 10th International Conference on Pattern Recognition, Atlantic City, NJ*, 796–802.

Zhang, Z., Kalluri, S., JàJà, J., Liang, S., and Townshend, J. R. G., 1998, High performance algorithms for global BRDF retrieval. *IEEE Computational Science and Engineering*, **5**(4), 16–29.