

**Vegetation Characteristics Expressed Through Transformed MODIS Data: A  
MODIS Tasseled Cap**

**by  
Sarah E. Lobser**

**A THESIS**

**submitted to**

**Oregon State University**

**in partial fulfillment of  
the requirement for the  
degree of**

**Master of Science**

**Presented November 19, 2004  
Commencement June 2005**

Master of Science thesis of Sarah E. Lobser presented on November 19, 2004.

APPROVED:

---

Major Professor, representing Forest Science

---

Head of the Department of Forest Science

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

---

Sarah E. Lobser, Author

AN ABSTRACT OF THE THESIS OF

Sarah E. Lobser for the degree of Master of Science in Forest Science presented on November 19, 2004.

Title: Vegetation Characteristics Expressed Through Transformed MODIS Data: A MODIS Tasseled Cap

Abstract approved:

---

Warren B. Cohen

The MODIS NBAR (MOD43B4) data space is explored in terms of biophysical variables with the objective of formulating a MODIS transformation relevant to global vegetation studies. The basic ideas of transform formulation were borrowed from the development of the TM Tasseled Cap transformation, but with differences in sampling strategy and rotation procedures. A random sample was taken from a year's worth of the full global extent of NBAR tiles to yield a sample representative of the full range of spectral variation in the Earth. Principal Components Analysis (PCA) was used to investigate the temporal dynamics of the sample before and after the exclusion of snow and ice from the dataset. This investigation uncovered the potential for the formulation of a MODIS Tasseled Cap transformation, and multidimensional scaling methods were used to rotate the sample data space to match the orientation of the TM Tasseled Cap. The sample in

the resulting MODIS transformation space was observed to match the orientation and general biophysical characteristics of the TM Tasseled Cap. The utility of the transformation for global vegetation studies was explored by observing the characteristics of the MODIS Land Cover product (MOD12Q1) and the MODIS Vegetation Continuous Fields product (MOD44) in the transformed sample space. Observed temporal trends in vegetation phenology were observed at specific sites with known vegetation dynamics. Though the MODIS Tasseled Cap was formulated using NBAR data, the coefficients can be applied to the raw MODIS Surface Reflectance product (MOD09) as well, as determined by the one to one relationship seen in both the raw and transformed NBAR and Surface Reflectance data.

## TABLE OF CONTENTS

	<u>Page</u>
Chapter One: Introduction.....	1
Chapter Two: Vegetation Characteristics Expressed Through Transformed MODIS Data: A MODIS Tasseled Cap .....	4
Chapter Three: Conclusions.....	79
Bibliography.....	80

## LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Temporal variation in loadings of the first three principal components of the global sample before the removal of pixels containing snow.....	9
2	Schematic diagram of NBAR land tiles downloaded for sampling.....	12
3	Geographic distribution of sample pixels.....	13
4	Geographic distribution of pixels collected from two different time periods.....	14
5	Temporal variation in sample size for four MODIS tiles.....	15
6	Temporal variation in the loadings of the first three principal components of the global sample after masking out pixels that contained snow.....	16
7	Loadings from the first three principal components of the full global sample.....	17
8	Paired cumulative distribution functions demonstrating how well the global sample matches the Earth, in each of the seven MODIS spectral bands.....	18
9	Comparison of TM scene 46/29 from August 2000 and the contemporaneous MODIS tile.....	24
10	Scatterplots comparing the raw bands of MODIS and TM data for TM scene 46/29.....	26
11	Comparison of global sample PCA loadings and TM Tasseled Cap coefficients.....	29
12	Density plots of the global sample in PCA space.....	32
13	IGBP density plots in MODIS PCA space.....	34
14	Density plot of simulated TM global sample in TM Tasseled Cap space.....	37

## LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Temporal variation in loadings of the first three principal components of the global sample before the removal of pixels containing snow.....	9
2	Schematic diagram of NBAR land tiles downloaded for sampling.....	12
3	Geographic distribution of sample pixels.....	13
4	Geographic distribution of pixels collected from two different time periods.....	14
5	Temporal variation in sample size for four MODIS tiles.....	15
6	Temporal variation in the loadings of the first three principal components of the global sample after masking out pixels that contained snow.....	16
7	Loadings from the first three principal components of the full global sample.....	17
8	Paired cumulative distribution functions demonstrating how well the global sample matches the Earth, in each of the seven MODIS spectral bands.....	18
9	Comparison of TM scene 46/29 from August 2000 and the contemporaneous MODIS tile.....	24
10	Scatterplots comparing the raw bands of MODIS and TM data for TM scene 46/29.....	26
11	Comparison of global sample PCA loadings and TM Tasseled Cap coefficients.....	29
12	Density plots of the global sample in PCA space.....	32
13	IGBP density plots in MODIS PCA space.....	34
14	Density plot of simulated TM global sample in TM Tasseled Cap space.....	37

## LIST OF TABLES

<u>Table</u>		<u>Page</u>
1	Comparison of MODIS and TM spectral bands .....	8
2	Summary of QA bits used for masking the raw NBAR dataset to exclude unwanted pixels from the sample. ....	11
3	MODIS Tasseled Cap coefficients.....	41
4	IGBP land cover units. ....	45
5	Summary of the eight BigFoot sites used in analysis. ....	53
6	MODIS tiles used for comparison between NBAR and Surface Reflectance.....	63

# **Vegetation Characteristics Expressed Through Transformed MODIS Data: A MODIS Tasseled Cap**

## **Chapter One: Introduction**

There has always been an interest in global change studies and global mapping projects. Yet, the possibility of performing such large-scale studies has been limited until relatively recently. The development of new technologies in the past few decades has allowed the scientific community to explore questions of global concern, and to begin to understand such large-scale processes as climate change. One such technology of great value and expanding potential is remote sensing. Our ability now to launch sensors into space that record biophysical information about the Earth is truly remarkable, and the acquisition of data that covers the entire globe has a tremendous impact on the scale of science that can be done.

The science of remote sensing has emerged to tackle the challenges of designing sensors to collect data of interest, and extracting useful information from the acquired data. It is a rapidly expanding science, experiencing exponential growth since the mid-1960s with the number of publications doubling at regular intervals (Cracknell and Hayes, 1993). Of particular relevance to this study is the wealth of knowledge acquired about the remote sensing of vegetation. Because satellite imagery provides multitemporal coverage of the full geographic extent of the Earth, remotely sensed data possesses significant potential for monitoring vegetation dynamics at large scales (Myneni, 1997).

The particular reflectance signatures common to vegetative life forms have been well known since the first sensors were built, and sensors have since been

developed with the key attributes of these signatures in mind. The remote sensing community has also devoted a significant amount of time and effort to the development of processing techniques to extract important biophysical information from this data (e.g. Frohn, 1998; Huete, 1999). Vegetation indices, for example, have been paramount in utilizing remotely sensed data for vegetation studies. These indices are dimensionless, radiometric measures that function as indicators of relative abundance and activity of green vegetation (Jensen, 2000).

Parallel to the progress made in vegetation indices has been the interest in multidimensional data transformations. In contrast to single indices, which commonly use a subset of the available spectral bands, a multidimensional transformation may yield multiple indices and take greater advantage of the information content of the data. The first such transformation developed for vegetation studies was the Tasseled Cap transformation, formulated by Kauth and Thomas (1976). The transformation was subsequently extended to the Landsat Thematic Mapper (TM) (Crist, 1984). The TM Tasseled Cap transforms Landsat TM data into a three-dimensional space defined by axes related to biophysical variables. To date, the application of the Tasseled Cap Transformation has been limited to the regional scale. The sheer quantity of Landsat data needed to study global vegetation patterns is prohibitive to doing such large-scale studies.

The new suite of sensors aboard NASA's Earth Observing System (EOS) was designed in part to make global data more available. Of these, the Moderate Resolution Imaging Spectroradiometer (MODIS) has a configuration of seven spectral bands that are specifically designed for land applications, with spatial

resolutions ranging from 250 m to 1 km (Justice, 1997). The launch of MODIS has expanded the potential of remote sensing to be used in global vegetation studies. The remote sensing community is now confronted with the challenge of developing techniques for utilizing the vast quantities of data available.

One possible avenue of study is the formulation of transformations for MODIS data. Transformations are often sensor-specific, so while a TM Tasseled Cap transformation may not necessarily work with the MODIS sensor, a similar transformation may be built which achieves the same goals. The first is one of data reduction; all seven spectral bands of the MODIS sensor are correlated to some extent, and over ninety-nine percent of the variation can be accounted for in three principal component bands. This reduction in the data volume facilitates faster processing, and eases interpretability by allowing the user to visualize the data space in three dimensions. The second goal of transformation building is to tie the data space directly to biophysical characteristics in the landscape. The TM Tasseled Cap Transformation is a rotated PCA (Principal Components Analysis) space with axis named 'brightness', 'greenness', and 'wetness.' These axes serve as indices; brightness is highly correlated with albedo, greenness is related to amount of photosynthetically active vegetation present, and wetness responds to moisture content in both leaves and soils.

In this paper, the MODIS data space is explored to uncover opportunities for transformation development in pursuit of these two goals. A transformation is then built and explored in terms of global vegetation characteristics and dynamics.

## **Chapter Two: Vegetation Characteristics Expressed Through Transformed MODIS Data: A MODIS Tasseled Cap**

### **Introduction**

The application of the Tasseled Cap transformation to Landsat imagery has played a significant role in vegetation studies of the past few decades. Because the Tasseled Cap has been so successful in application to Landsat data, both for data reduction and determination of vegetation characteristics, there is interest in developing a transformation for MODIS data as well.

The launch of the MODIS sensor aboard the Terra platform provided extensive opportunities for vegetation analysis at the regional and global scale. The development of a globally relevant transformation specific to the MODIS sensor could not only illuminate the structure of the data space but also deepen our understanding of global vegetation dynamics as expressed through spectral characteristics.

### **Background on the Tasseled Cap**

The Tasseled Cap transformation was first formulated by Kauth and Thomas (1976) for the Landsat Multispectral Scanner (MSS). It was developed based on the principle that the four band MSS data primarily occupied a single plane, and this plane was related to orthogonal biophysical features termed “brightness” and “greenness”. By changing the viewing angle of the data space to ‘look’ directly at this plane, the Tasseled Cap transformation simultaneously reduces the

dimensionality of the data while maximizing its interpretability in terms of vegetation analysis.

A similar transformation was developed for the Landsat Thematic Mapper (TM). The information content of this sensor was increased by the addition of two new reflectance bands, and though the brightness and greenness features found in MSS data were recognizable in the TM data space, a third significant feature was also found to respond to the moisture content of both soils and vegetation. This feature was thus termed “wetness” (Crist, 1983). The three indices created by the development of the Tasseled Cap transformation have proved useful in a broad range of vegetation studies (Cohen, 1992; Oetter, 2001; Dymond, 2002; Skakun, 2003), and the Tasseled Cap space has been likened to a spectral mixing space (Small, 2004).

### **Objectives**

A MODIS Tasseled Cap transformation was not sought after at the outset. Rather, the objective in this study was to develop a MODIS data transformation that would be relevant to global vegetation studies. The process of transformation development involved a full exploration of relevant biophysical variables as expressed in the global MODIS data space, and a manipulation of that space to yield its most useful projection. What emerged was the MODIS Tasseled Cap.

A transformation should allow the data to be expressed in terms of a few axes related to fundamental biophysical features of the landscape. In the case of the TM Tasseled Cap, these features are orthogonal linear combinations of the spectral bands extracted by rotating the principal component space of the raw data. This

method of using principal components to explore the structure of the data space and extract axes related to biophysical features is accepted as a basic statistical operation, and is applied here to the MODIS sensor.

### **Need for global sampling**

In previous studies that have explored spectral space to uncover data structures related to scene components to derive spectral vegetation indices, there have been several methods used for sampling. However, to our knowledge, none of these have involved collecting a random sample across space and time. The ideal sample in this type of study represents the full range of spectral variability in the biophysical features of interest. Previously, attempts have been made to represent a wide variety of soil and vegetation conditions in the sample (e.g. Crist, 1983). However, one way to ensure that the sample represents the full range of global conditions is to collect a random sample that spans the spatial extent of the terrestrial Earth.

Principal components are based on the covariance matrix of the dataset, and are thus highly dependent on the particular sample used. The key to developing a globally relevant transformation based on principal components is to begin the process with a sample that covers the full variation of global spectral characteristics. Because global vegetation patterns change through time, it is necessary to sample through both time and space.

Developing a global sample has implications beyond the scope of this study. Understanding the characteristics of the global spectral space and its multitemporal dynamics provides a context for linking geographically and temporally disparate

vegetation studies to a common data space. A global sample provides a standard to which smaller scale samples may be compared.

### **Sampling Considerations**

The challenges of selecting a representative global sample are many, not the least of which is cloud cover that can obscure large regions of the globe almost continually for months at a time. A simple random sampling of the global population of terrestrial pixels having no cloud cover would lead to the preferential selection of cloud-free regions of the earth. Therefore, a geographically stratified approach was taken, whereby uniformly-sized regions of pixels were sampled separately to achieve a better geographic coverage of the globe.

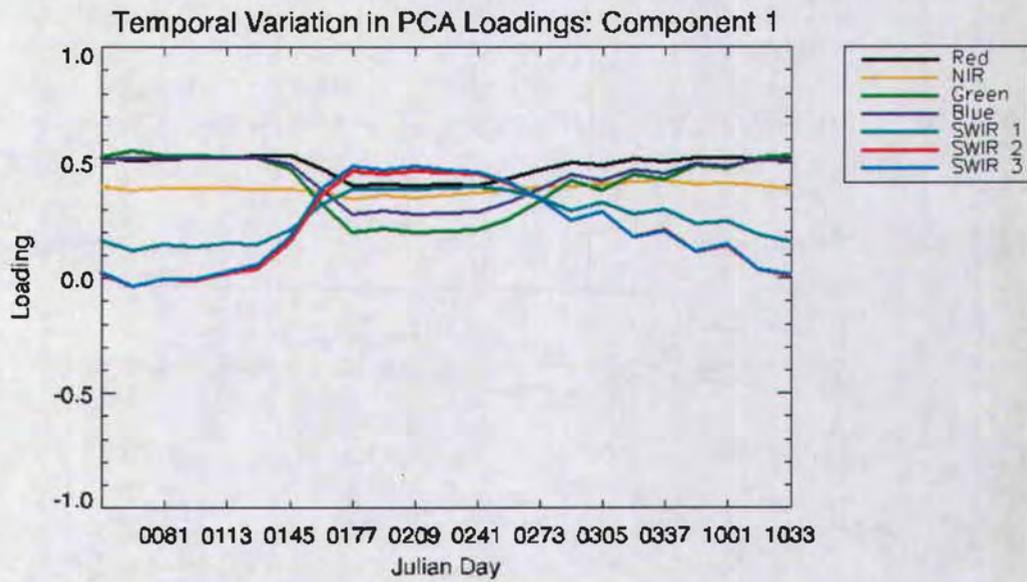
A second challenge was the unknown temporal dynamics of the principal components. The data transformation needs to be globally relevant, but temporally relevant as well, and the data structure of a full global sample that spans time and space may not be representative of the sample at one particular date bin. To more fully understand the temporal dynamics of the Earth's spectral data structure, a set of principal components was extracted from each date bin of an initial random sample taken from a cloud-free global extent of MODIS Nadir BRDF Adjusted Reflectance Data (NBAR, MOD43B4). NBAR was used as input into the formulation process, rather than raw reflectance, because it standardizes reflectance to a specific view and illumination geometry—noon sun and nadir view. This minimizes artifacts in the sample related to variable geometry (Schaaf, 2002). The NBAR spectral bands are comparable to those of Landsat TM, as

summarized in table 1. The relationship between MODIS and TM will be discussed later in greater.

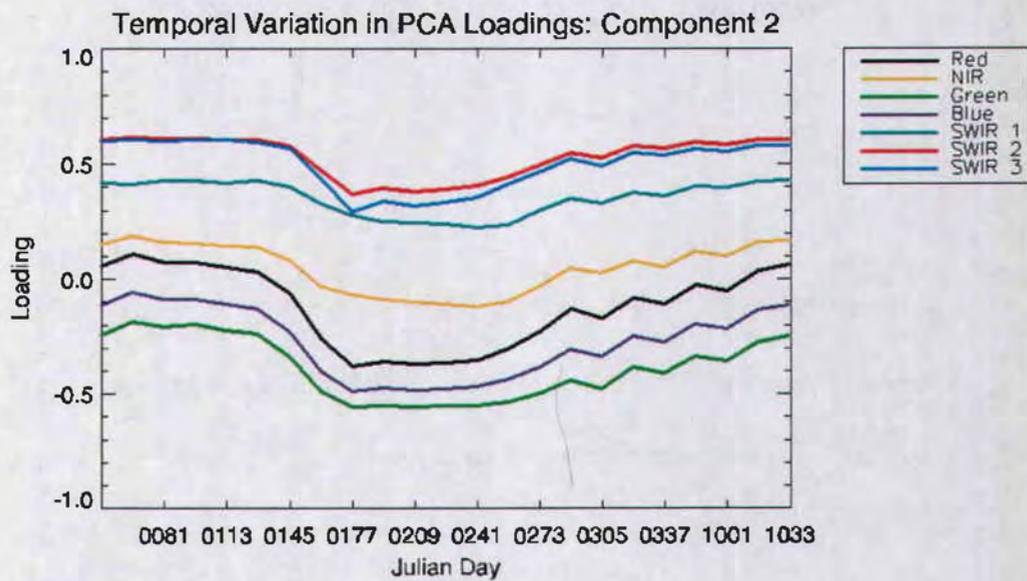
**Table 1. Comparison of MODIS and TM spectral band widths.**

	MODIS spectral band widths Wavelength in nm		TM spectral band widths Wavelength in nm	
	From	To	From	To
Red	620	670	630	690
NIR	841	876	760	900
Blue	459	479	450	520
Green	545	565	520	600
SWIR 1	1230	1250	NA	NA
SWIR 2	1628	1652	1550	1750
SWIR 3	2105	2155	2080	2350

The first principal component of an initial sample was quite temporally dynamic, and this shifting in the structure of the data through time appeared to be primarily due to snow cover. The extreme spectral difference between snow and snow-free areas dominated the variance of the sample, and the shifting geographic extent of snow coverage through time caused great variation in the loadings of the first two principal components (Figure 1).



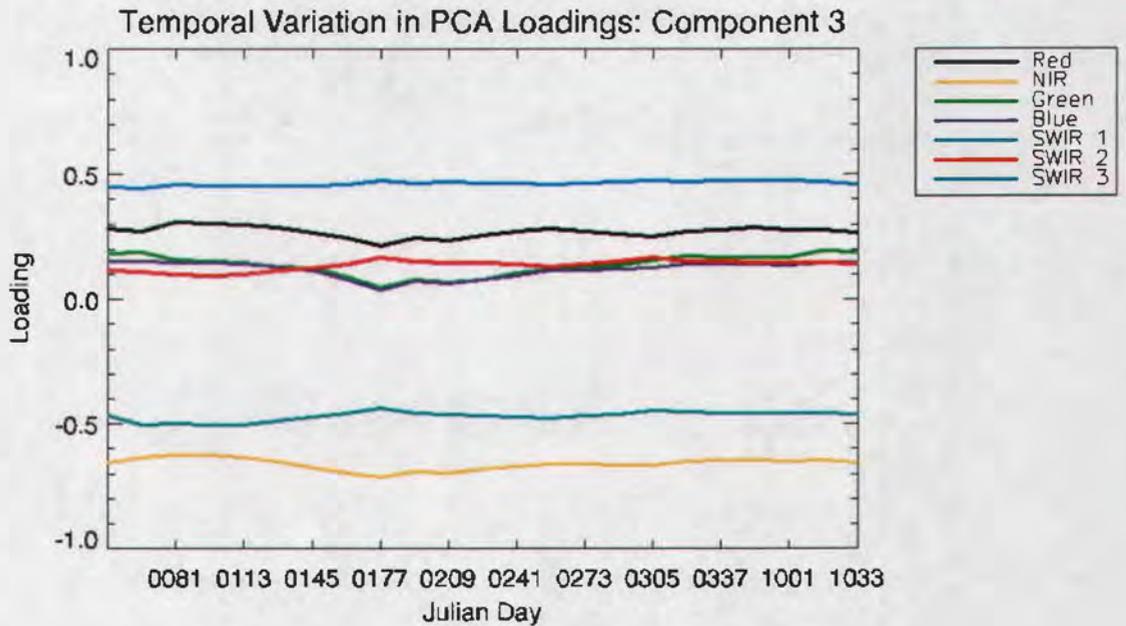
a)



b)

**Figure 1 a-c.** Temporal variation in loadings of the first three principal components of the global sample before the removal of pixels containing snow.

Figure 1 continued.



c)

It became crucial to eliminate the snow cover from the sample to stabilize the temporal dynamics of the data structure. Since the focus of the study was to develop a transformation applicable to vegetation analysis, the exclusion of snow from the sample was not disadvantageous.

The sampling was performed a second time using the methods described below.

### **Sampling Strategy**

A global extent of MODIS NBAR 16-day composites spanning the year from February 2000 to February 2001 was acquired from which to draw the sample.

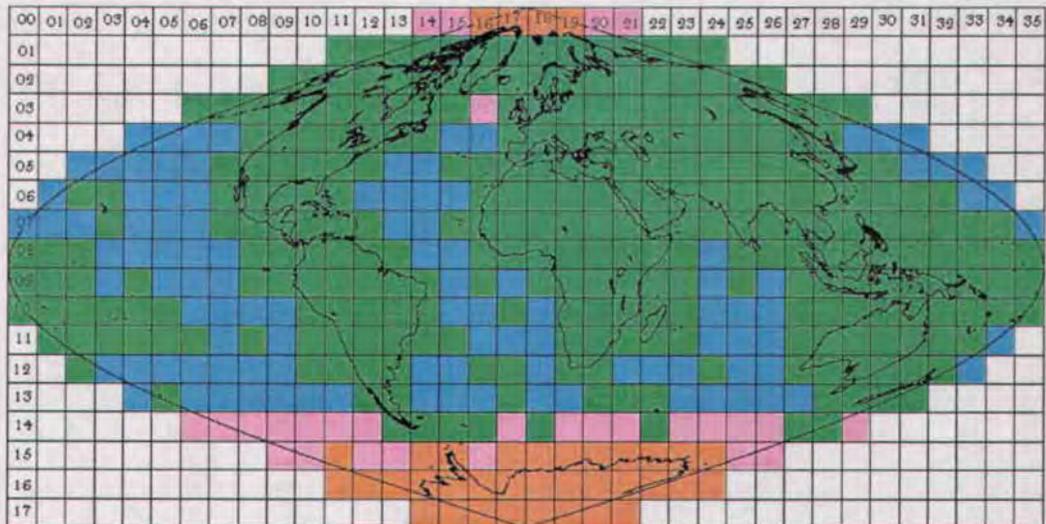
Accompanying each tile was a 32-bit QA dataset (table 2) that was used to exclude certain pixels from the sample.

**Table 2. Summary of QA bits used for masking the raw dataset to exclude unwanted pixels from the sample.**

QA Bits	Description	Usage
00-01	Mandatory QA	Processed, good quality pixels were extracted from the sample.
04-07	Surface Type	Only terrestrial pixels were sampled.
16-17	Snow Cover	Only pixels where snow cover was absent were sampled.

Only pixels with QA bits matching those above were considered in the sample, resulting in a snow-free, water-free, cloud-free sample.

Figure 2 shows the geographic distribution of imagery produced by the MODIS NBAR science team. All images, or tiles, that included some land (in green) were downloaded for sampling. Each of the 6550 tiles in the original dataset was split into 25 sections. A mask was then created using the accompanying QA dataset to remove the snow, clouds, and water. One tenth of one percent of the remaining pixels was randomly selected, and the selections from all tiles were compiled to form the global sample.



**Figure 2. Schematic diagram of NBAR land tiles (in green) downloaded for sampling. The horizontal tile number (above) and the vertical tile number (left) is used to identify the geographic location of each tile.**

The quality of data in the sample varied. When sufficient high quality MODIS observations are available to adequately sample the BRDF, the Ross-Thick/Li-Sparse semiempirical BRDF model is used to compute NBAR. Otherwise, a lower quality magnitude inversion is performed which couples *a priori* knowledge of the surface anisotropy with any high quality MODIS observations that are available (Schaaf et al., 2002). Pixels computed using the primary model are flagged in the QA dataset.

It was desirable to use only the highest quality pixels in further analysis, thus all pixels with a mandatory QA flag of '00' to denote good quality were extracted to form the global sample. It may have been more desirable to mask out pixels lacking this quality flag initially, along with the snow, water, and clouds.

However, as described in the next section, further investigation proved the sample to be acceptable.

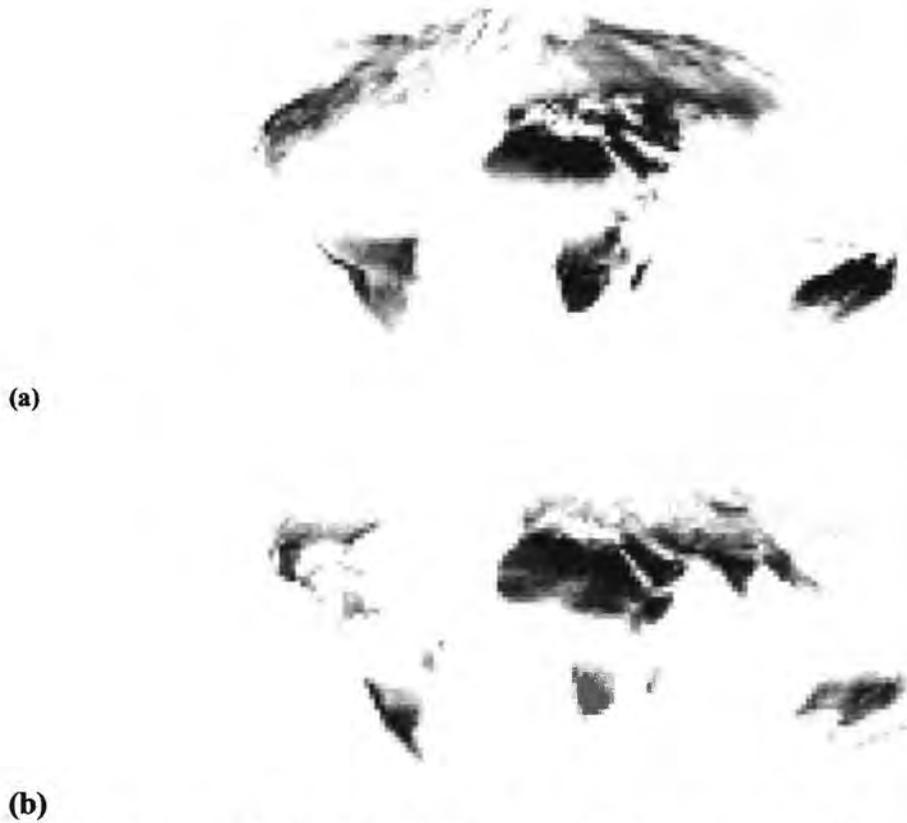
### **Sample Assessment**

Figure 3 shows where the sample was actually drawn from geographically. Darker areas signify a higher number of pixels drawn from that particular region. As expected, perpetually cloud-free areas such as the Saharan desert contribute more heavily to the sample than regions such as the tropics where cloud cover obscures large portions of the land surface for a significant portion of the year.



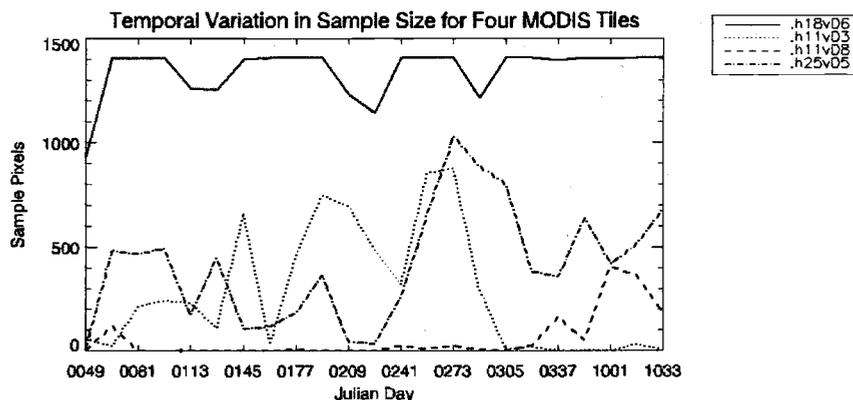
**Figure 3. Geographic distribution of sample pixels.**

The geographic range of the sample also changes over time, and figure 4 shows this change from northern hemisphere summer to winter.



**Figure 4. Geographic distribution of sample pixels collected from two different time periods:  
(a) June 25 through August 28 and (b) January 1 through March 5.**

To get an idea of the actual number of pixels drawn for the sample from different regions, four different tiles were compared on numbers of pixels chosen for the sample over time. These tiles were chosen based on their apparent differences in temporal cloud cover dynamics. Figure 5 shows the shifting number of pixels selected from each of these four tiles.

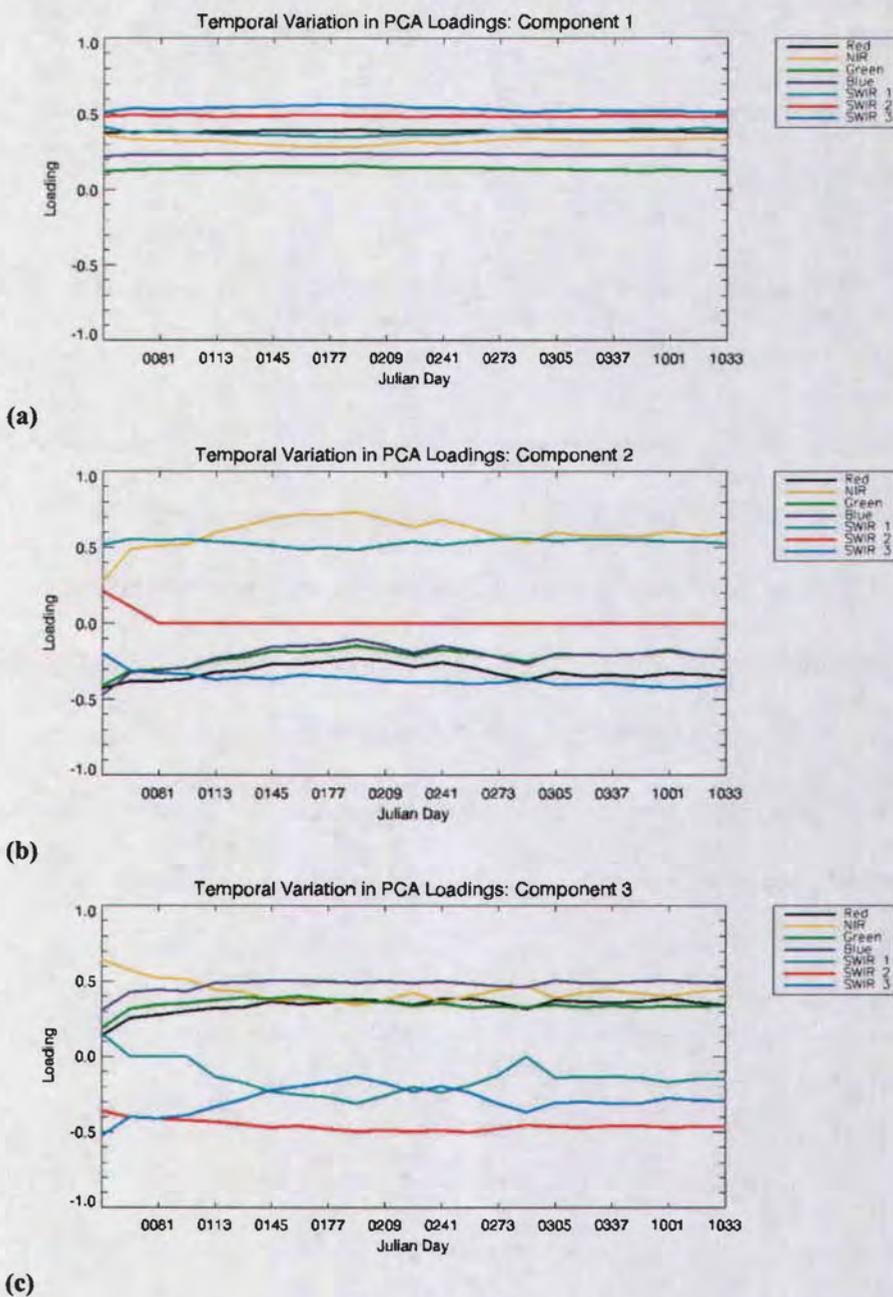


**Figure 5. Temporal variation in number of pixels randomly selected for the global sample from four MODIS NBAR tiles, labeled here by their horizontal and vertical tile numbers.**

As discussed before, snow cover was eliminated from the sample because its spectral characteristics dominated the covariance structure of the MODIS bands and its changing geographic extent caused significant temporal shifting in the first principal component loadings. It is apparent, though, that the geographic extent of the global sample itself shifts considerably over time due to temporal changes in snow cover and cloud cover. It is therefore reasonable to expect that the principal component loadings might still vary over time. This turns out not to be the case, however.

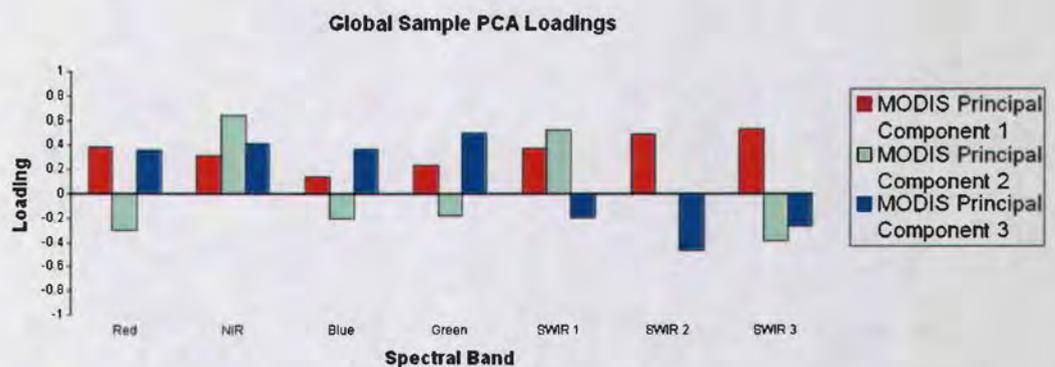
Figure 6 shows the temporal variation in principal component loadings over the course of the year. In contrast to the loadings extracted from the sample including snow, these loadings are generally more stable over time. The first principal component, accounting for over ninety percent of the variance of the sample, is particularly stable after the exclusion of snow. The stability of the sample

covariance structure suggests that a transformation formulated using the sample as a whole would represent the sample at any given point in time.



**Figure 6. Temporal variation in the loadings of the first three principal components (a-c) of the global sample after masking out pixels that contained snow.**

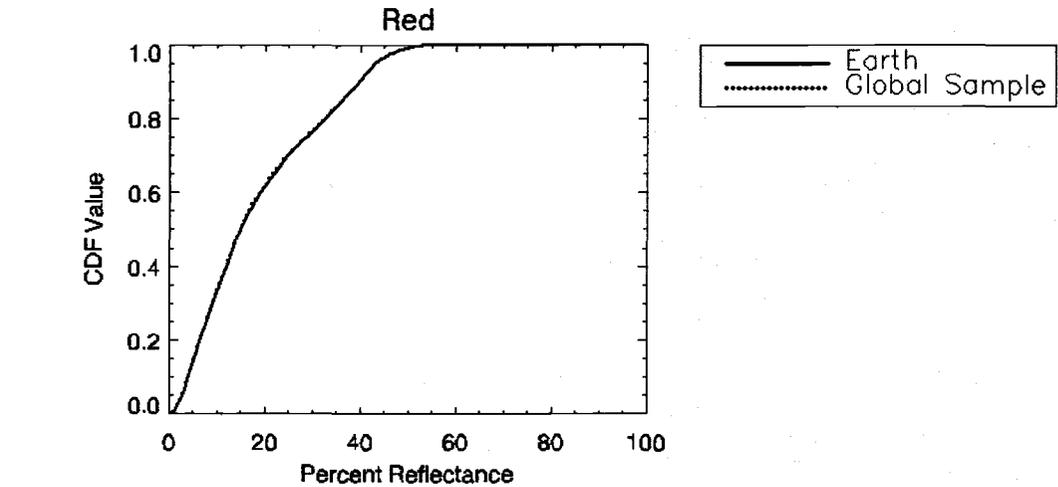
Finally, a PCA was performed using the entire global sample; the loadings are displayed in figure 7. Not surprisingly, the first principal component loadings of the full global sample were indistinguishable from any of the sets of loadings from the individual date bins. The other two components were somewhat more variable over time, and the loadings from the full sample seemed to be quite close to the temporal mean in these components. This suggests that this PCA space would represent the covariance structure of the global sample equally well at any give point in time. However, it was still unknown at this point how well the global sample represented the spectral characteristics of the globe.



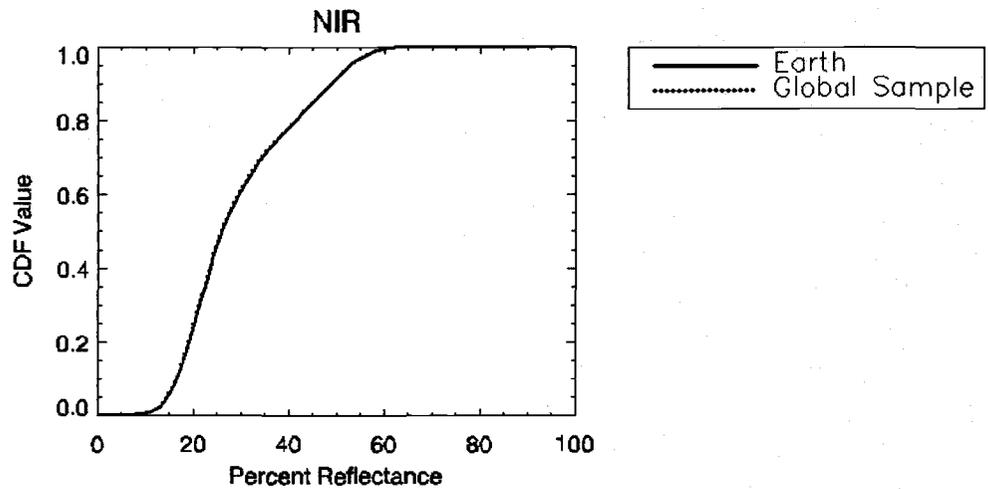
**Figure 7. Loadings from the first three principal components of the full global sample.**

To address this, it was necessary to compare the distribution of the global sample in spectral space with that of the earth. A histogram of the Earth was calculated for each MODIS spectral band, including good quality pixels that were free of snow, water, and clouds. Cumulative distribution functions built from these histograms

were compared to those calculated from the global sample as shown in figure 8, and were found to match very well.



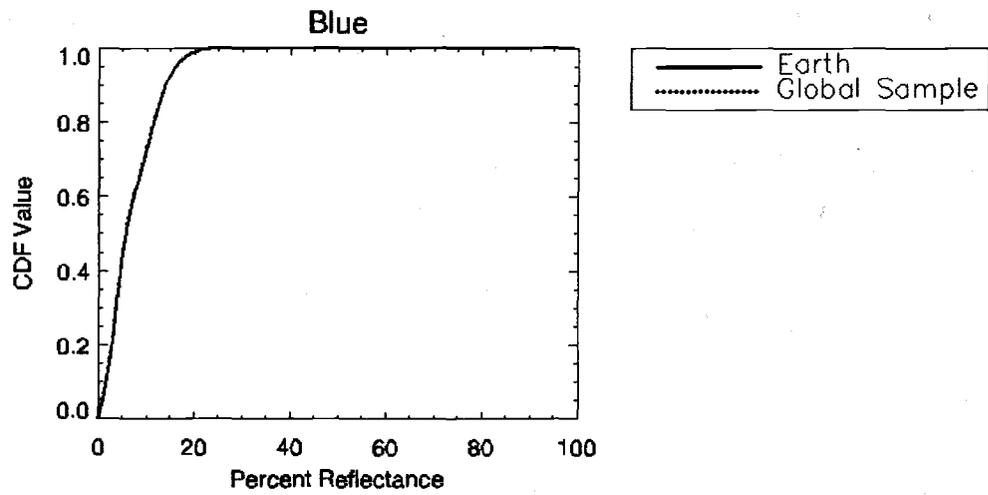
(a)



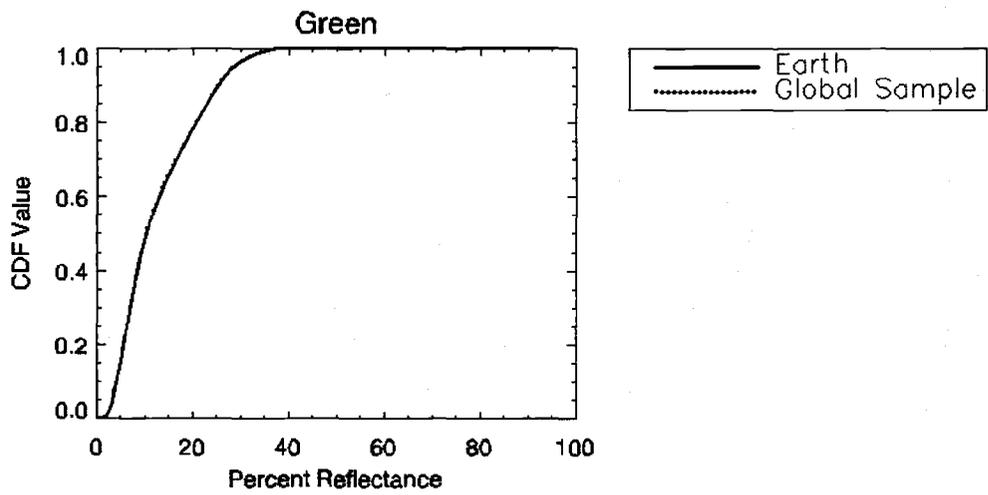
(b)

**Figure 8. (a-g) Paired cumulative distribution functions demonstrating how well the global sample matches the Earth, in each of the seven MODIS spectral bands.**

Figure 8 continued.

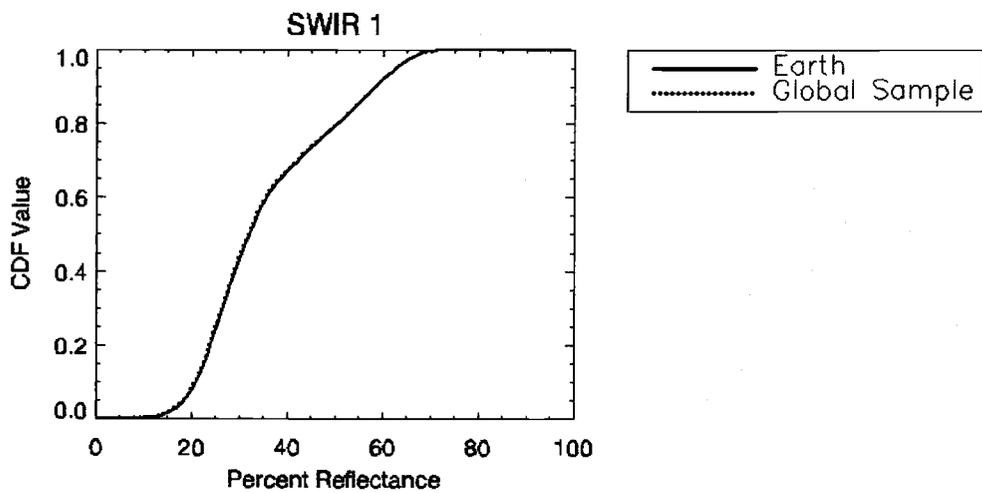


(c)

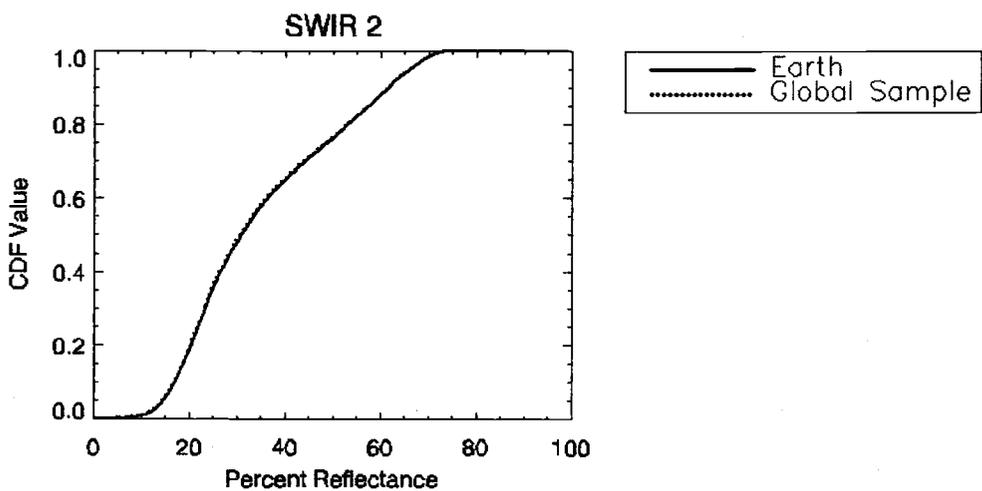


(d)

Figure 8 continued.

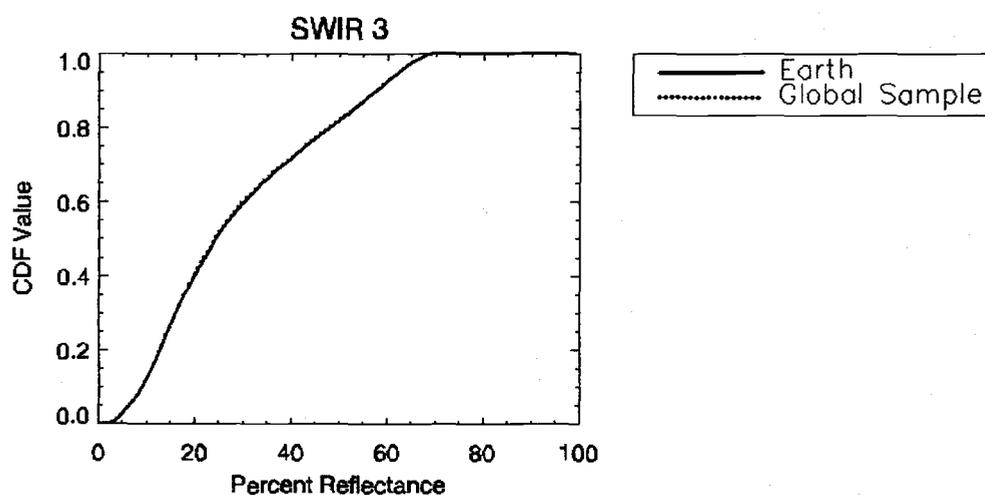


(e)



(f)

Figure 8 continued.



(g)

### Exploration of the global sample in PCA space.

The Tasseled Cap concept evolved from an initial exploration of biophysical features of agriculture and other vegetation types as expressed in terms of the principal components of MSS and TM data. The objective of this study was to broaden this type of exploration to the global scale, identifying key biophysical variables that drive the fundamental patterns in data covariance structure. The discussion has thus far dealt with the challenge of collecting a sample representative of the spectral characteristics of the globe through time. It has been confirmed that the spectral variability found in the global sample represents the

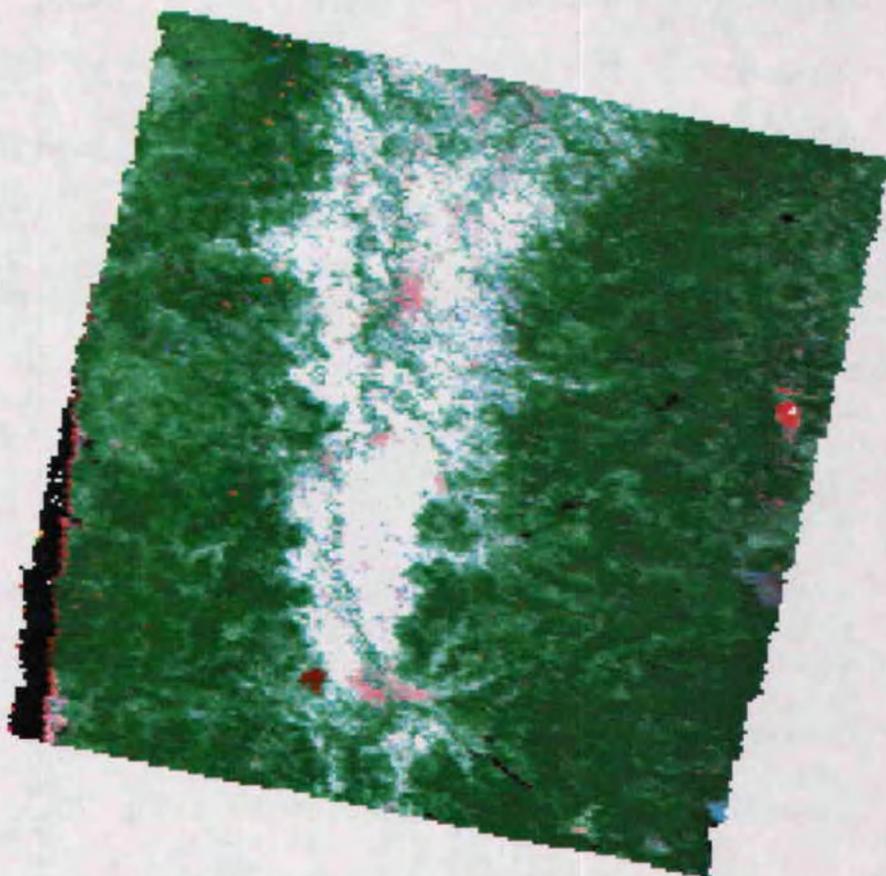
global condition well; now, exploring the expression of pertinent biophysical features in PCA space should yield a transformation relevant to the global condition.

### **Relating MODIS to TM**

It was unknown how well, if at all, the covariance structure of the NBAR product would correspond to Landsat. The MODIS sensor is a departure from Landsat, and it was of interest to us to understand how a MODIS-based transformation would correspond with the TM tasseled cap. This depends on the compatibility of the MODIS and TM covariance structures. To fully investigate the potential compatibility of the two, an initial step was taken to determine correlations between the six corresponding raw spectral bands. There are a few differences of relevance here between the MODIS and TM sensors. TM collects data in 6 spectral bands; MODIS has introduced a seventh spectral band in the SWIR region. The remaining 6 MODIS spectral bands correspond roughly to the TM spectral bands, but with some shifting in the spectral range and narrowing of band widths, as summarized in table 2. At approximately 1 km, MODIS data has a much lower spatial resolution than 30 m TM data. With a wider swath width, MODIS has significant off-nadir view angle effects. However, the NBAR product has been produced to minimize these by correcting for view angle using a BRDF model.

The impact of all these differences was examined using a method similar to Crist and Cicone (1984b). Landsat scene 46/29 from August 2000 was reprojected to SIN (MODIS projection), atmospherically corrected and translated into reflectance

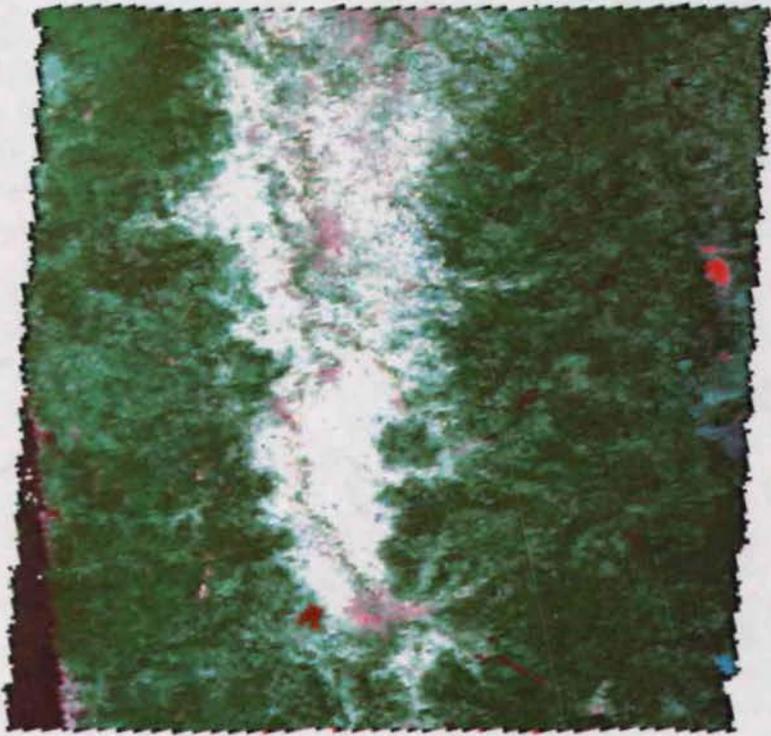
using the COST model, and spatially aggregated to 930 m to match the contemporaneous MODIS tile h09v04 in both projection and resolution. Figure 9 shows the TM and MODIS images, now in UTM for greater interpretability, with MODIS clipped to the boundaries of the TM scene. The images are presented using the spectral bands of highest correlation: Red band in red, SWIR 2 in green, and SWIR 3 in blue. Figure 10 shows scatterplots of the 6 corresponding spectral band pairs, along with their correlation coefficients. Though the spectral band ranges differ slightly and TM reflectance values are somewhat higher than MODIS values, the pairs are highly correlated. This suggests that the covariance structures of data from the two sensors are likely quite similar. In fact, the loadings of the first three principal components of the global sample, accounting for over 99% of the variance, are strikingly similar to the TM Tasseled Cap coefficients. Figure 11 compares the global sample PCA loadings with the TM Tasseled Cap coefficients. The first principal component is a weighted average of all the bands, similar to the TM Tasseled Cap brightness. The patterns in the other two components are also comparable with the TM Tasseled Cap indices: the second principal component is a contrast between visible and near-infrared bands, like greenness, and the third principal component is a contrast between the SWIR bands and everything else, similar to TM Tasseled Cap wetness.



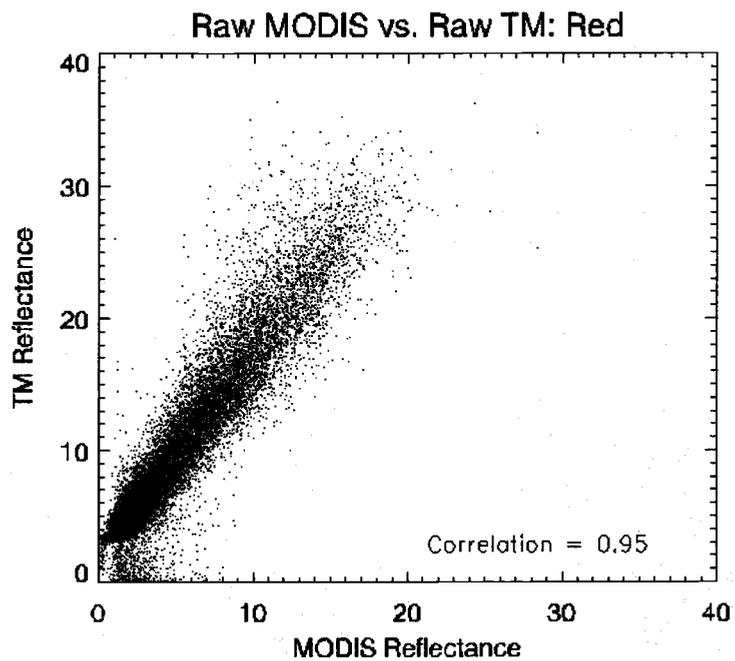
(a)

**Figure 9. Comparison of TM scene 46/29 from August 2000 and the contemporaneous MODIS tile (clipped). (a) TM scene 46/29, atmospherically corrected and translated into reflectance using the COST model, spatially aggregated to 930 m to closely match the spatial resolution of MODIS. Red: red, Green: SWIR 2, Blue: SWIR 3. (b) The corresponding clipped MODIS tile (h09 v04). Red: red, Green: SWIR 2, Blue: SWIR 3.**

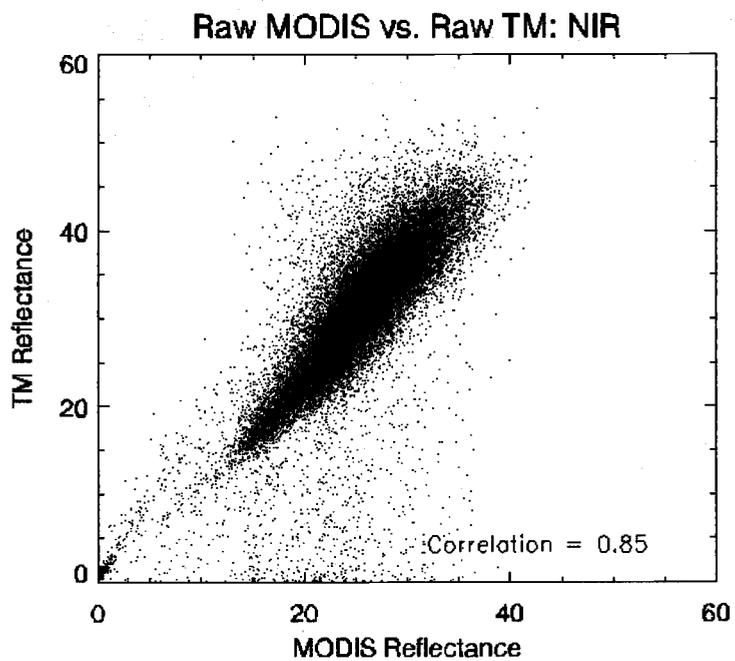
Figure 9 continued.



(b)



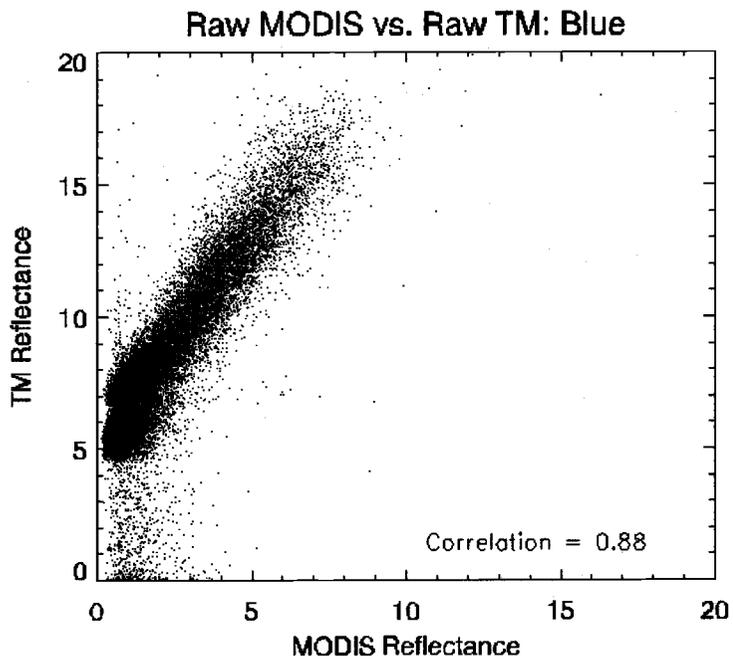
(a)



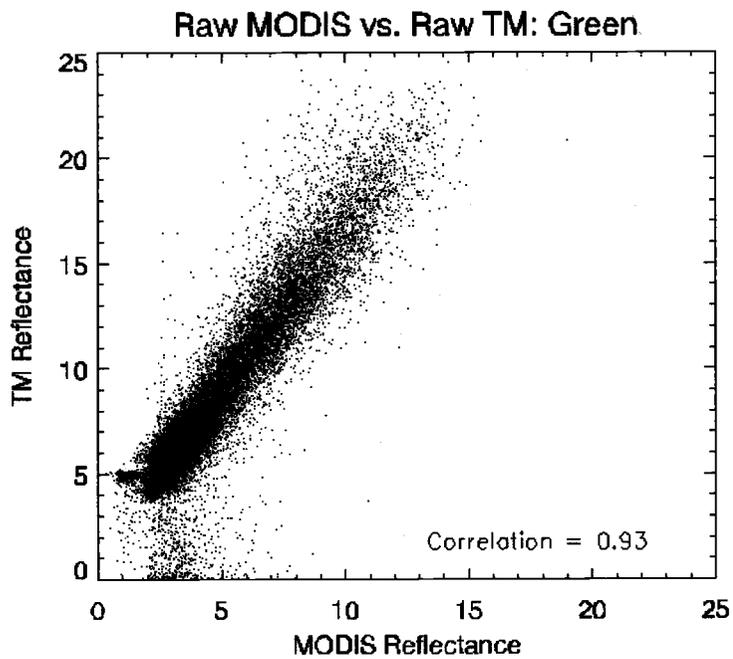
(b)

Figure 10 (a-f). Scatterplots comparing the raw bands of MODIS and TM data for TM scene 46/29.

Figure 10 continued.

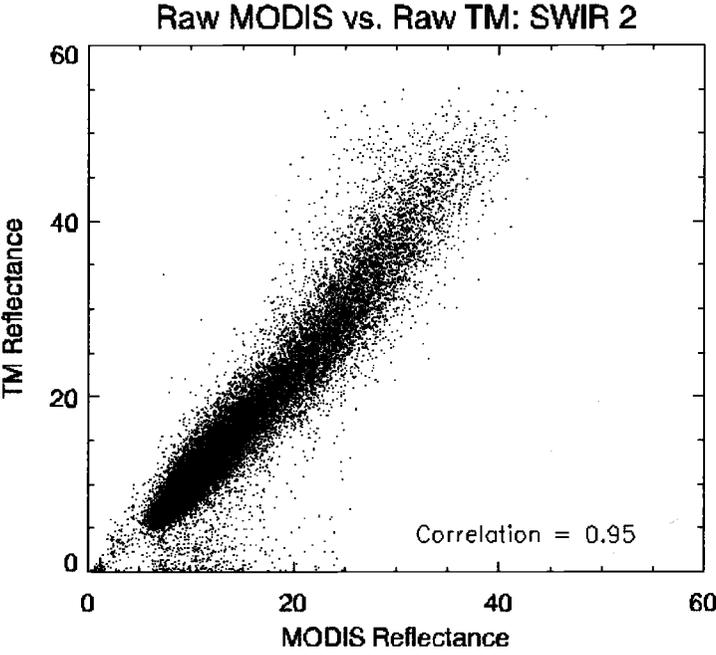


(c)

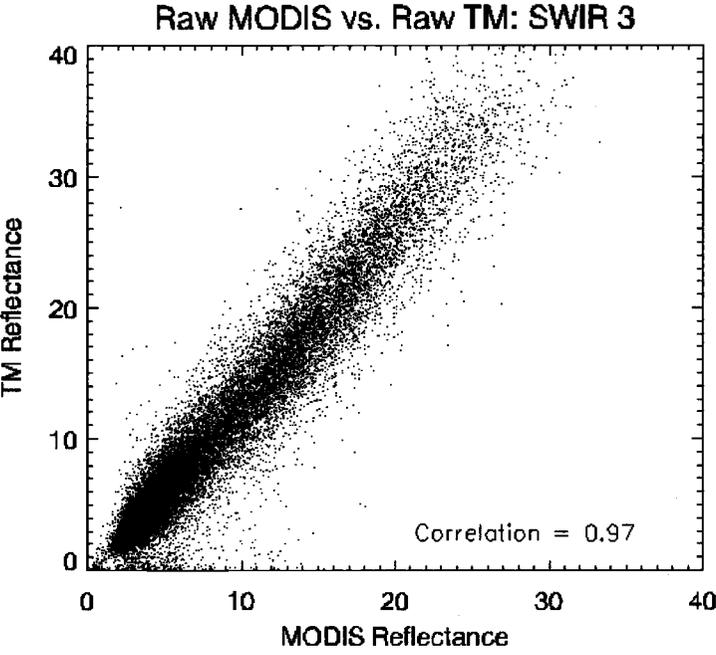


(d)

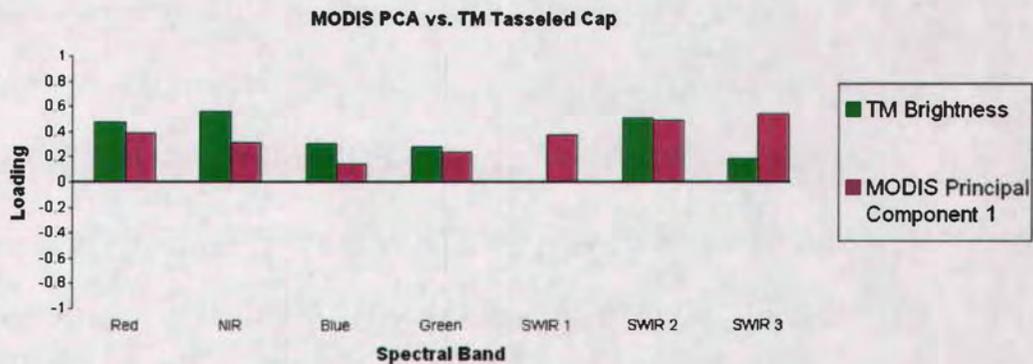
Figure 10 continued.



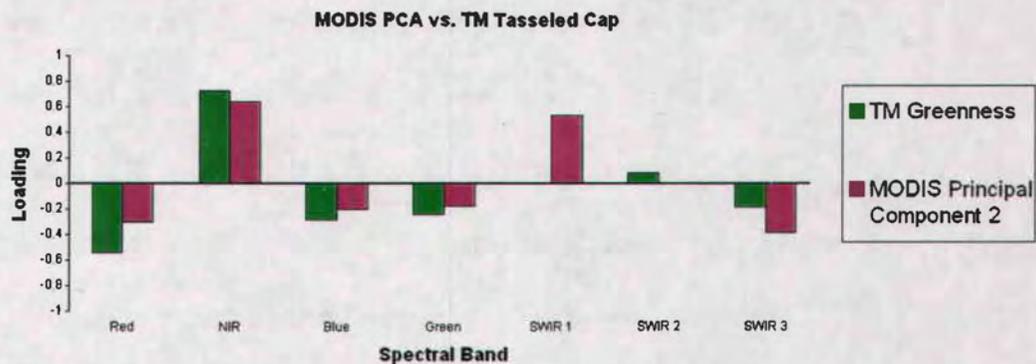
(e)



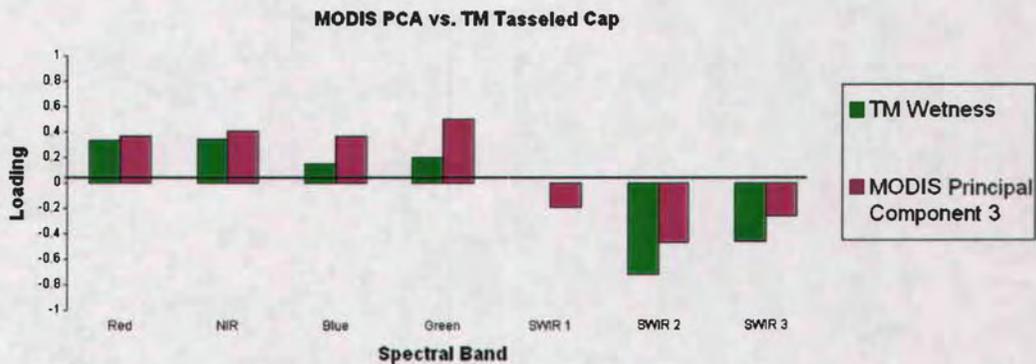
(f)



(a)



(b)

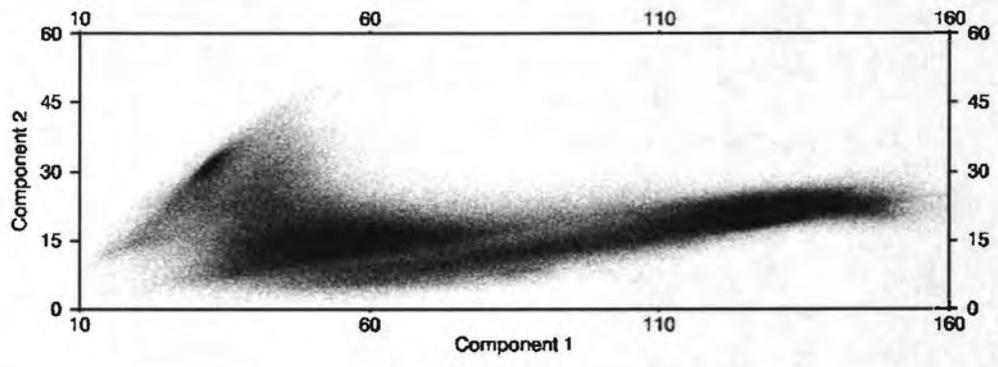


(c)

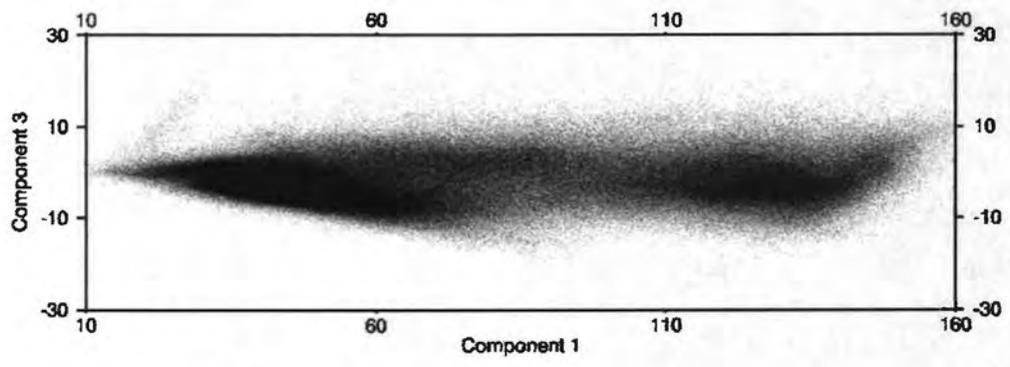
**Figure 11 (a-c). Comparison of global sample PCA loadings and TM Tasseled Cap coefficients.**

The high correlation between MODIS and TM spectral bands would lead to the expectation of a correspondence between MODIS PCA and TM PCA loadings. It was somewhat surprising, though, to find that the MODIS PCA loadings were so similar to the TM Tasseled Cap coefficients. The TM tasseled cap is a rotation of a PCA space, so the rotation may have been slight enough that the general patterns in the coefficient axes remained relatively unchanged. In parallel, a MODIS Tasseled Cap might be achieved by slightly rotating the MODIS PCA space.

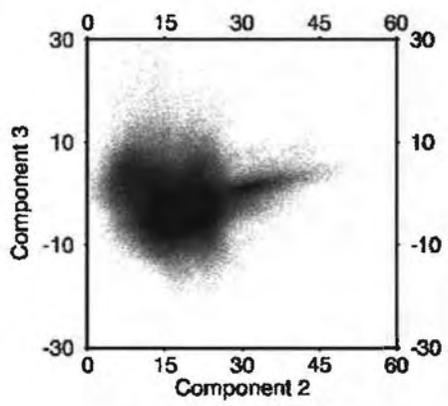
This notion is made more obvious by viewing the global sample in PCA space. Figure 12 shows a density plot of the global sample in PCA space. The distribution of the sample in space is particularly triangular, similar to what one would see in a TM Tasseled Cap space. However, the characteristic features of a Tasseled Cap would be easier to identify by adding land cover information to the sample.



(a)



(b)



(c)

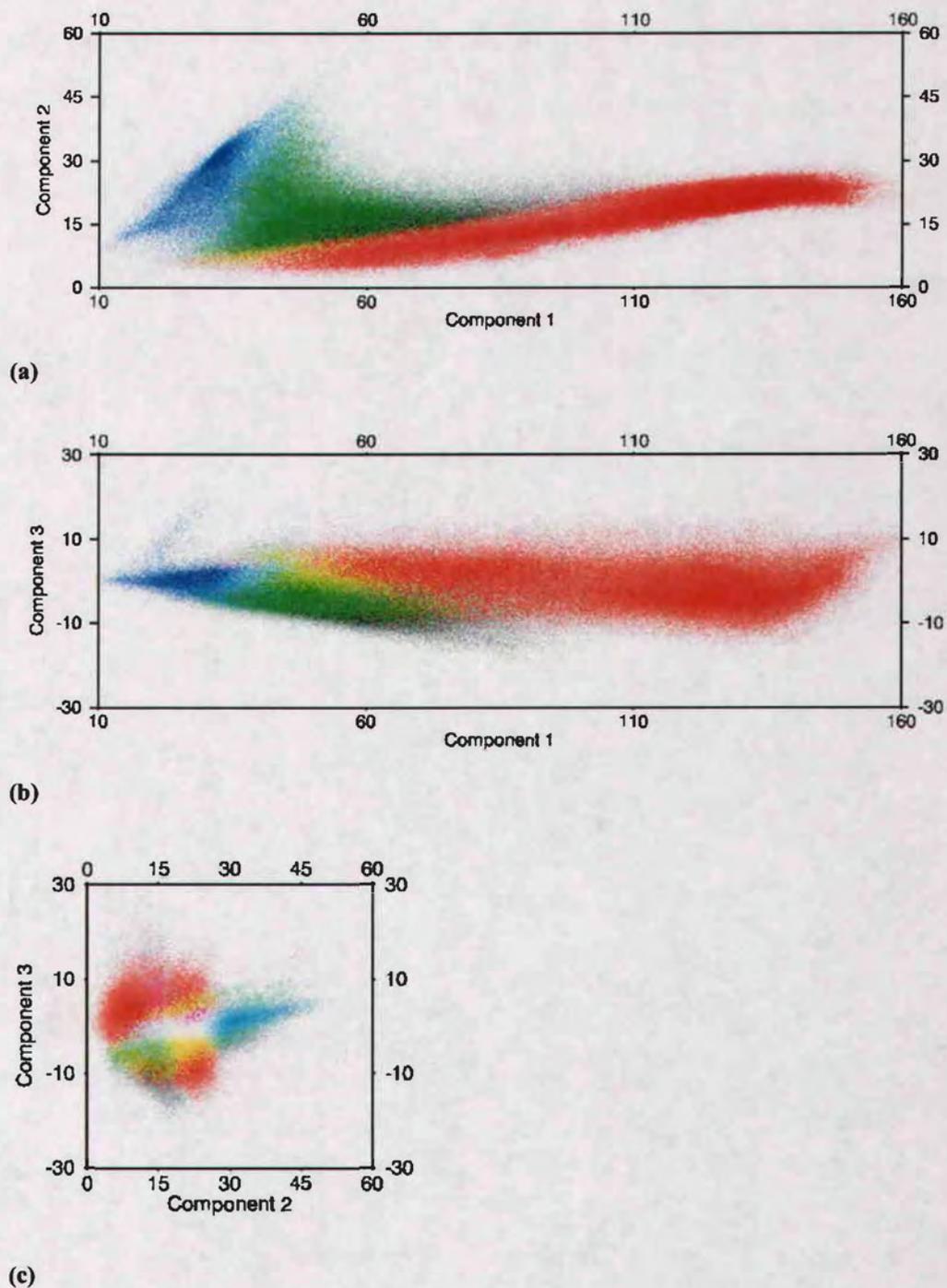
Figure 12 (a-c). Density plots of global sample in PCA space.

To accommodate this, IGBP land cover class information was acquired from the MODIS land cover product (MOD12Q1) from 2000. This product includes land cover designations using five different land cover classification schemes: International Geosphere-Biosphere Programme (IGBP), University of Maryland (UMD), Leaf Area Index/Fraction Absorbed Photosynthetically Active Radiation (LAI/FPAR), Biome BGC (BGC), and Plant Functional Types (PFT) (Friedl, 2002). Though any one of these classification schemes would have been valuable to explore in PCA space, this study focused on the IGBP system. Density plots of IGBP class groupings in PCA space further reveal a structure quite similar to that of the TM Tasseled Cap, as seen in figure 13.

This figure is a set of three density plots layered together as an RGB image. A density plot of all barren lands in the global sample, in red, is combined with density plots of croplands, in green, and forests, in blue. Cyan corresponds to areas in spectral space where the distributions of forests and croplands overlap. Yellow corresponds to areas where barren lands and croplands overlap. This method of combining three density plots in an RGB image is used in later figures as well.

The croplands group shows the characteristic Tasseled Cap shape described first by Kauth and Thomas (1976). The forestlands occupy a place named by Kauth and Thomas (1976) as the “badge of trees”. Barren and sparsely vegetated areas occupied the region common to bare soil samples. All the major components of the TM Tasseled Cap space were visible in MODIS PCA space; the only apparent difference was one of orientation. This was reasonable, as the TM Tasseled Cap is itself a rotation of a PCA space, and the data structure of TM is apparently similar

to MODIS. Note that the region occupied by water in TM Tasseled Cap space is lacking observation points in these plots. This is because we purposely excluded water from our sample.



**Figure 13 (a-c).** IGBP density plots in MODIS PCA space. Three density plots are layered to create an RGB image: barren lands in red, croplands in green, and forest lands in blue. Other colors are seen where these density plots overlap, as in cyan where forest lands (blue) overlap with croplands (green).

These results provided a unique opportunity to not only formulate a Tasseled Cap transformation for MODIS, but also to link the MODIS Tasseled Cap directly to its TM predecessor, making vegetation analysis at different scales translatable in this common data space. The challenge leading up to this is rotating the MODIS PCA space to exactly match the orientation of the TM Tasseled Cap space.

### **The rotation process**

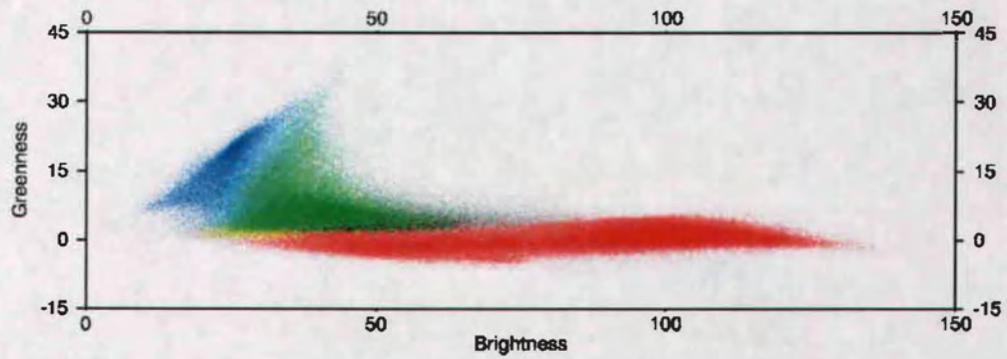
The objective in the rotation process is to change the orientation, or perspective, of the MODIS PCA space without changing the structure of the data. This could be achieved with a Procrustes rotation, a common technique in multidimensional scaling (Mardia et al., 1976). A Procrustes rotation finds a least squares solution to the problem of rotating one set of points to match a paired set of points.

Ideally, a set of MODIS points (the global sample in MODIS PCA space) would be rotated to match a paired set of TM points in TM Tasseled Cap space. To accomplish this, each point in the global sample would need a corresponding TM sample point, and this would require the acquisition of a full year of TM data spanning the globe. The costs of this undertaking, in terms of time and funding, were rather prohibitive. For the purposes of this study, it was more effective to simulate a global sample of TM data.

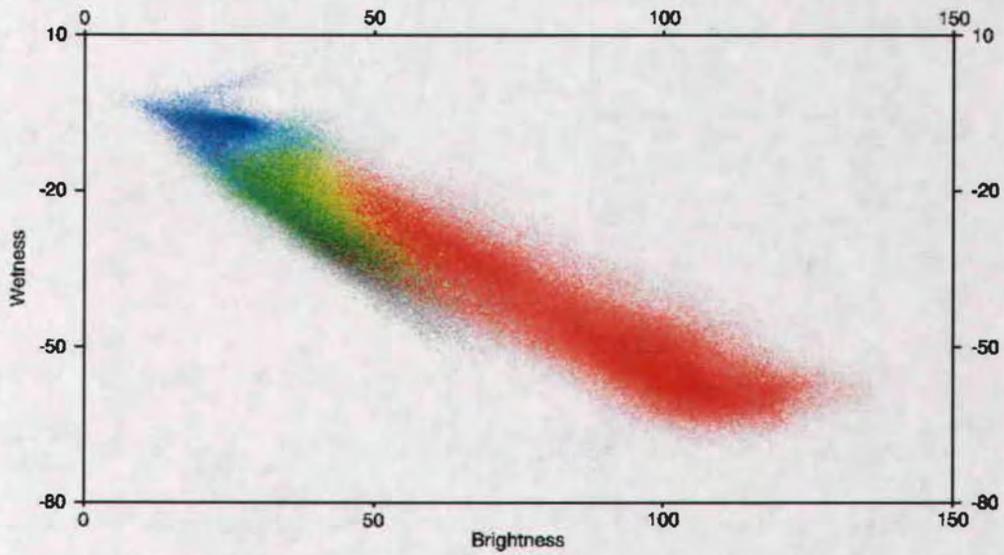
As described previously, a six-band MODIS system, composed of a subset of the spectral bands corresponding to TM, is a good proxy for the TM sensor. Using this knowledge, the global MODIS sample, minus SWIR 1, was used to simulate a TM global sample. The simulated global TM sample is simply a spectral subset of the global MODIS sample, and obviously matches the global MODIS sample point for

point. The MODIS sample in MODIS PCA space could then be rotated, using a Procrustes rotation, to match the simulated TM sample in TM Tasseled Cap space.

The simulated TM global sample was transformed into Tasseled Cap space using the coefficients provided by Crist (1985) in his TM Tasseled Cap equivalent transformation for reflectance factor data. These particular coefficients were used because, unlike MODIS, native TM data are not in reflectance units. Thus, the original Tasseled Cap coefficients developed for TM data would not work with the simulated TM sample, which was simulated using MODIS NBAR (reflectance) data. Figure 14 shows a feature space plot of the simulated TM sample in TM Tasseled Cap space, using the reflectance factor equivalent transformation coefficients. This is clearly the desired orientation, with the distribution of barren land stretched out along the greenness $\approx$ 0 line. The general shape of the simulated TM sample distribution in feature space is roughly the same as the global sample, but the increase in information content of the MODIS sensor is apparent in the greater spread in the distributions. The idea here is that the simulated TM sample is close enough in structure to the global sample that it can be used to train the orientation of the global sample data space.



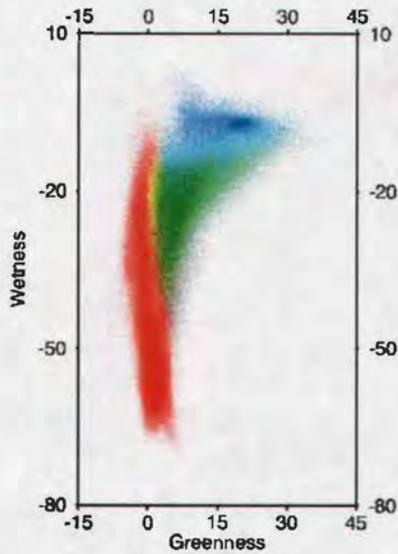
(a)



(b)

**Figure 14 (a-c).** Density plot of simulated TM global sample in TM Tasseled Cap space, using same color scheme as in figure 13.

Figure 14 continued.

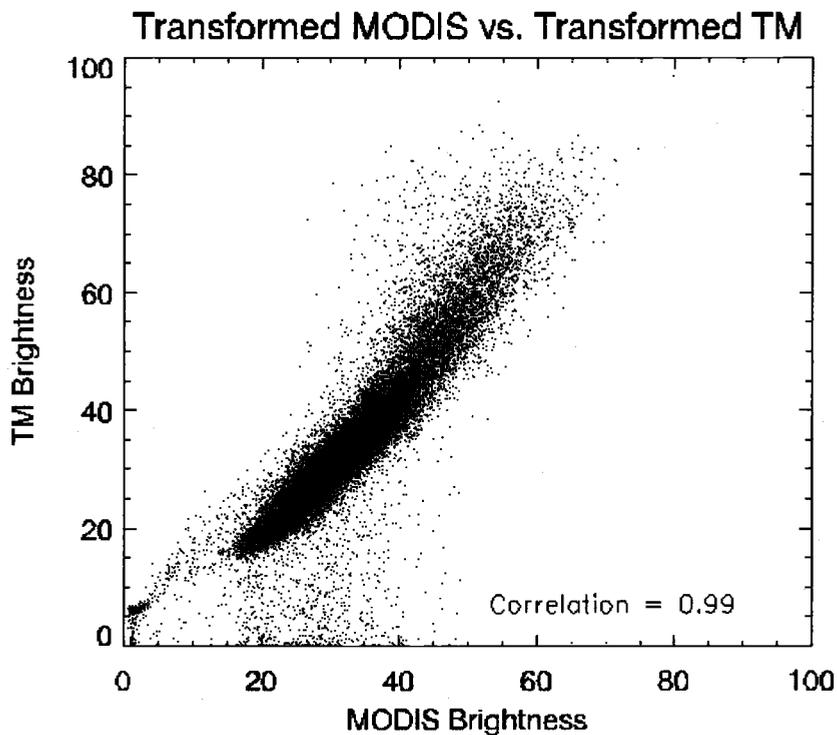


(c)

A Procrustes rotation was used to find a rotation matrix that would shift the orientation of the global MODIS sample to match that of the simulated TM sample. This least squares solution does not change the structure of the PCA space. An RMSE of 10.4 percent reflectance units reflects the added information content of the MODIS sample data.

Density plots of the global sample in the rotated MODIS PCA space reveal a structure and orientation comparable to the TM Tasseled Cap, but another step was taken to confirm the correspondence. The same Landsat scene 46/29 scene used in previous analysis, aggregated and reprojected to match its MODIS counterpart, was transformed into TM Tasseled Cap space and plotted against the

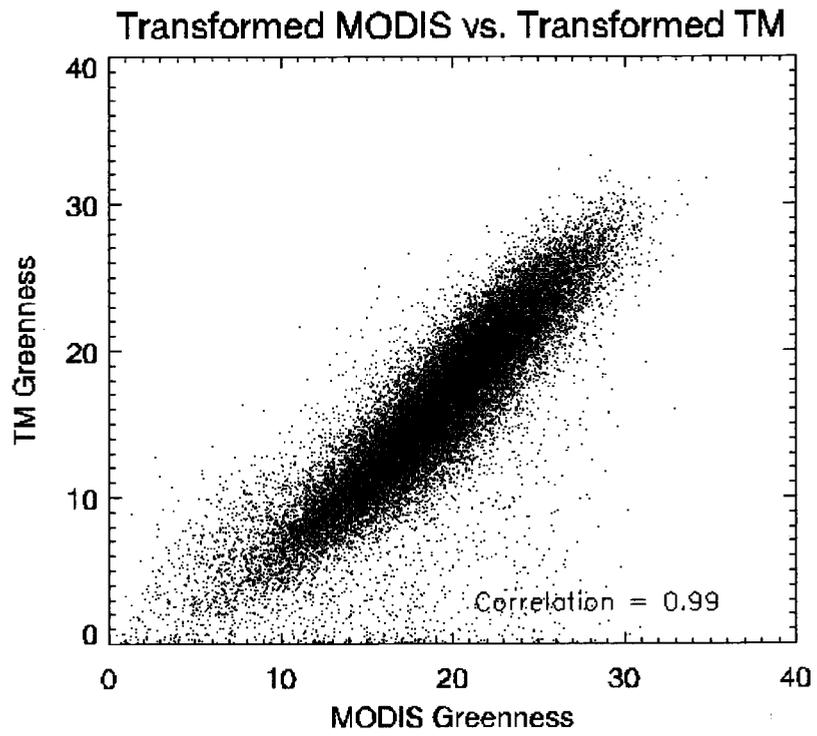
corresponding MODIS data transformed into the rotated PCA space. Results, demonstrated in figure 15, showed a very high correlation; thus the rotated MODIS PCA space became a MODIS Tasseled Cap.



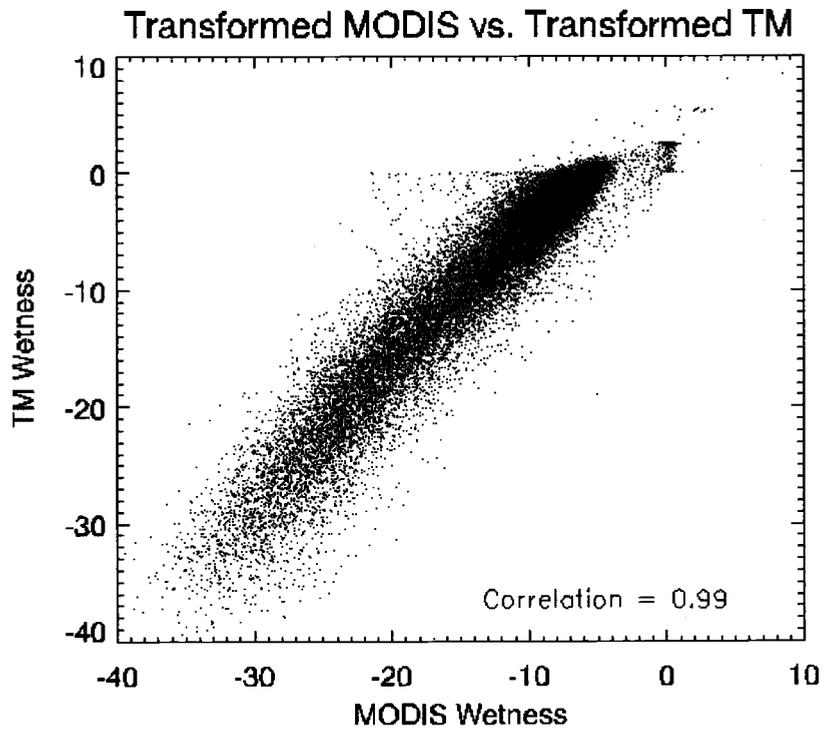
(a)

**Figure 15 (a-c). TM scene 46/29 transformed using TM Tasseled Cap coefficients compared with corresponding MODIS data transformed using MODIS Tasseled Cap coefficients.**

Figure 15 continued.



(b)

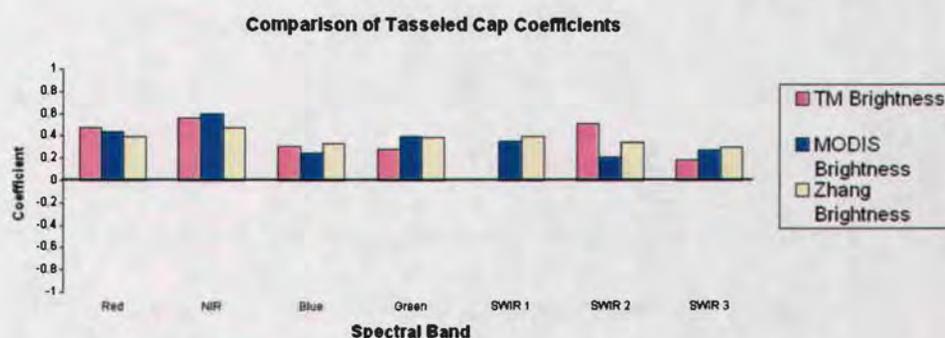


(c)

MODIS Tasseled Cap coefficients are presented in table 3, and compared in figure 16 with TM tasseled cap coefficients and a second set of MODIS Tasseled Cap coefficients developed by Zhang and others (2002). This group used a completely different sampling strategy and different methods for rotation of the MODIS PCA space. Nevertheless, the same general trends can be seen in all three sets of coefficients.

**Table 3. MODIS Tasseled Cap coefficients.**

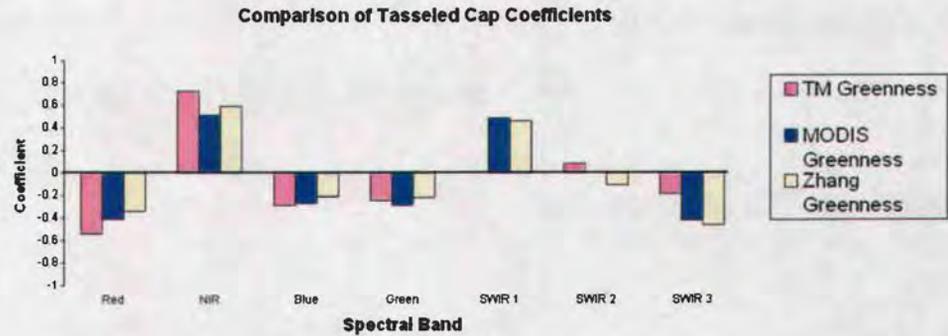
	Brightness	Greenness	Wetness
Red	0.4395	-0.4064	0.1147
NIR	0.5945	0.5129	0.2489
Blue	0.2460	-0.2744	0.2408
Green	0.3918	-0.2893	0.3132
SWIR 1	0.3506	0.4882	-0.3122
SWIR 2	0.2136	-0.0036	-0.6416
SWIR 3	0.2678	-0.4169	-0.5087



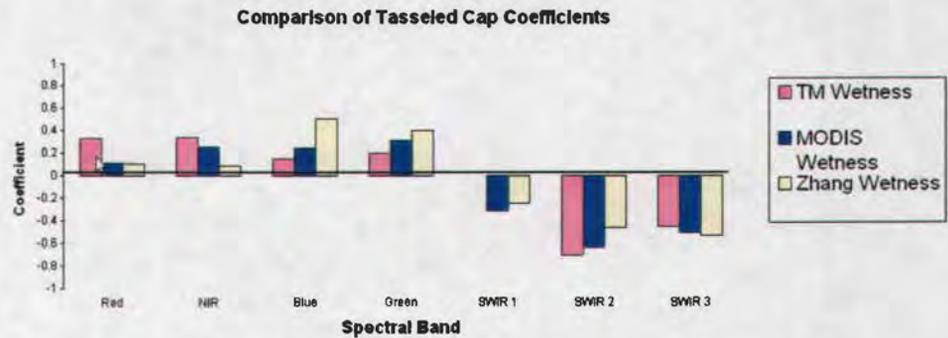
(a)

**Figure 16 (a-c). Comparison of three groups of Tasseled Cap loadings: TM Tasseled Cap, MODIS Tasseled Cap, alternative MODIS Tasseled Cap derived by Zhang et al. (2002).**

Figure 16 continued.



(b)



(c)

### Feature Space Exploration

The Tasseled Cap transformation originally got its name from the shape of crop spectral development trajectories through the growing season. Kauth and Thomas (1976) first identified this pattern in feature space; Crist and Cicone (1983) later developed the concept with the addition of the wetness index. A plot of bare soil before planting would be viewed best in the plane of soils (brightness/wetness space), then would shift up through the transition zone toward the plane of vegetation (brightness/greenness space) as the crops matured, then would “tassel

out” with senescence. A group of such plots, all starting with different soil conditions, would together form the shape of a tasseled woolen cap. The form of this distribution is identifiable in both the TM Tasseled Cap and the MODIS Tasseled Cap spaces.

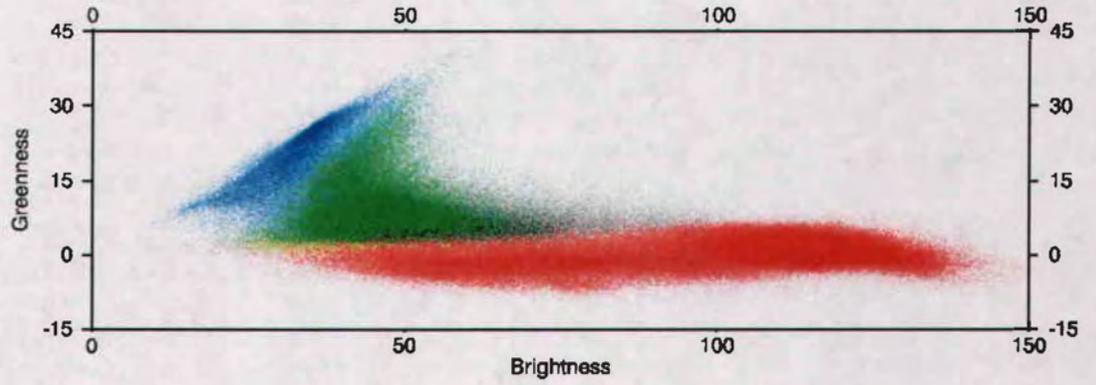
There are other land cover related features visible in the TM Tasseled Cap space. Crist (1983) identified regions in Tasseled Cap space occupied primarily by water, man-made materials, and forest. Because the Tasseled Cap maximizes spectral variation in three dimensions by means of PCA, it is expected that the Tasseled Cap would be an ideal transformation for differentiation and classification of broad global land cover classes. Also, because the Tasseled Cap rotates the data space to be expressed in terms of biophysical variables, namely brightness, greenness, and wetness, analyzing and interpreting land cover in this space is a more intuitive process. The advantage of having a global sample that includes land cover information is the ability to determine characteristics of cover type distributions empirically for the entire Earth.

Plotting the global sample in feature space in terms of IGBP land cover classes, listed in table 4, is one way to demonstrate how the characteristics of global vegetation can be interpreted using Tasseled Cap indices. Figure 17 shows some examples of these feature space plots. Three distinct clusters are apparent in Tasseled Cap feature space, and each is related to a different group of IGBP classes—one for forest, one for grasses, and one for barren regions. Barren or Sparsely vegetated areas have the largest distribution, owing to the vast amount of diversity in physical condition of these areas, which include moist and dry soil,

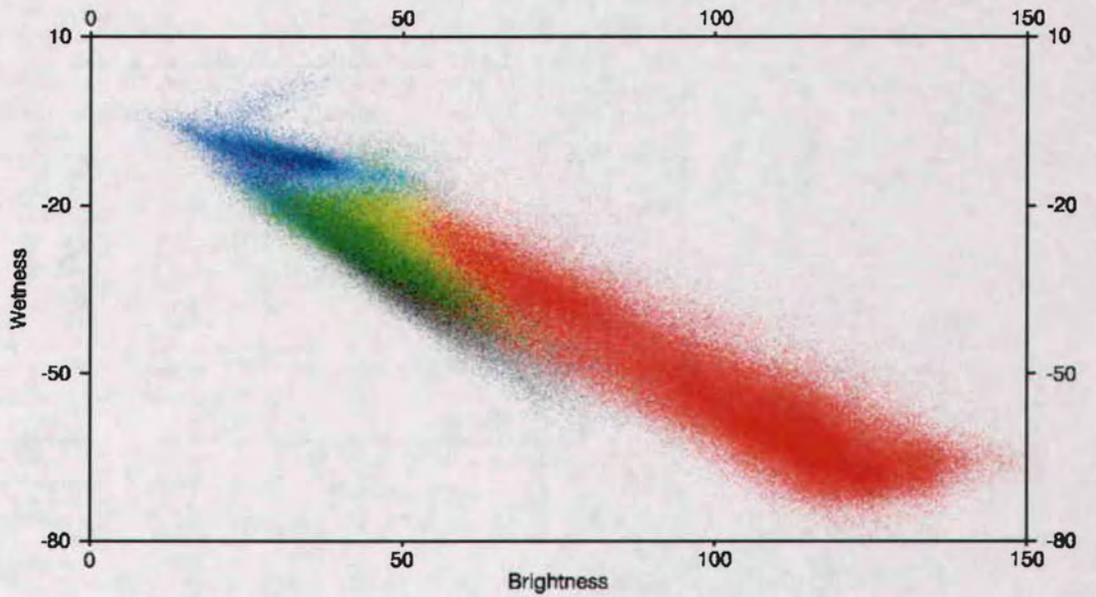
bare rock, lava, and sand. Herbaceous vegetation, including grasslands, croplands, and savannas, have the next highest amount of spectral diversity, and show a significant amount of overlap among class types. Forest classes occupy a relatively smaller region in Tasseled Cap space. Although forests are generally thought of being more structurally diverse than grasses, in spectral terms their distribution is relatively more constrained. This is due most likely to the greater temporal variation in the herbaceous vegetation types. Areas called 'grassland' or 'cropland' under the IGBP classification undergo a number of different stages of growth and senescence throughout the year, so these sample pixels represent a considerable array of phenological conditions, including bare soil, growing and mature vegetation, and senescent vegetation.

Table 4. IGBP Land Cover Units.

IGBP Land Cover Units (Strahler et al., 1999)	
Evergreen Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees remain green all year. Canopy is never without green foliage.
Evergreen Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Almost all trees and shrubs remain green year round. Canopy is never without green foliage.
Deciduous Needleleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of seasonal needleleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
Deciduous Broadleaf Forests	Lands dominated by woody vegetation with a percent cover >60% and height exceeding 2 meters. Consists of broadleaf tree communities with an annual cycle of leaf-on and leaf-off periods.
Mixed Forests	Lands dominated by trees with a percent cover >60% and height exceeding 2 meters. Consists of three communities with interspersed mixtures or mosaics of the other four forest types. None of the forest types exceeds 60% of the landscape.
Closed Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover >60%. The shrub foliage can be either evergreen or deciduous.
Open Shrublands	Lands with woody vegetation less than 2 meters tall and with shrub canopy cover between 10-60%. The shrub foliage can be either evergreen or deciduous.
Woody Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 30-60%. The forest cover height exceeds 2 meters.
Savannas	Lands with herbaceous and other understory systems, and with forest canopy cover between 10-30%. The forest cover height exceeds 2 meters.
Grasslands	Lands with herbaceous types of cover. Tree and shrub cover is less than 10%.
Permanent Wetlands	Lands with a permanent mixture of water and herbaceous or woody vegetation. The vegetation can be present in either salt, brackish, or fresh water.
Croplands	Lands covered with temporary crops followed by harvest and a bare soil period (e.g., single and multiple cropping systems). Note that perennial woody crops are classified as the appropriate forest or shrub land cover type.
Urban and Built-Up Lands	Land covered by buildings and other man-made structures.
Cropland/Natural Vegetation Mosaics	Lands with a mosaic of croplands, forests, shrubland, and grasslands in which no one component comprises more than 60% of the landscape.
Snow and Ice	Lands under snow/ice cover throughout the year.
Barren	Lands with exposed soil, sand, rocks, or snow and never has more than 10% vegetated cover during any time of the year.
Water Bodies	Oceans, seas, lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies.



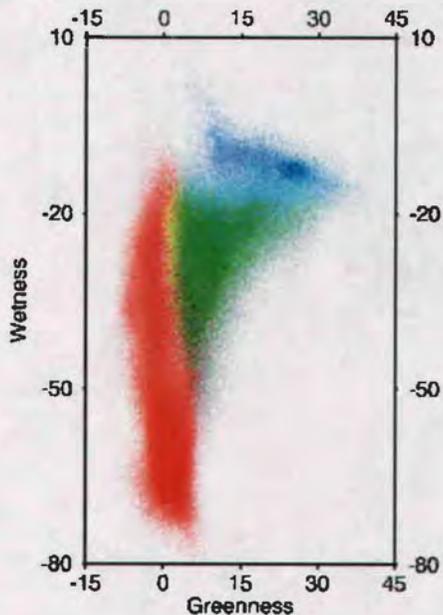
(a)



(b)

Figure 17 (a-c). IGBP density plots of the global sample in MODIS Tasseled Cap space, using same color scheme as in figure 13.

Figure 17 continued



(c)

Taking a closer look at the forest classes in figures 18-22, the phenological differences that distinguish different forest classes are apparent in their varying spectral ranges in Tasseled Cap space. Evergreen broadleaf forest occupies a slightly different region in spectral space than evergreen needleleaf forest, which is darker and less green. Deciduous broadleaf forest shows two markedly different distributions in Tasseled Cap space, relating to its two very different phenological conditions. The leaf-on condition is closer to evergreen broadleaf forest, and the leaf-off condition is closer to evergreen needleleaf forest. Mixed forests span most of the region occupied by forest classes, owing to its compositional, structural, and phenological diversity.

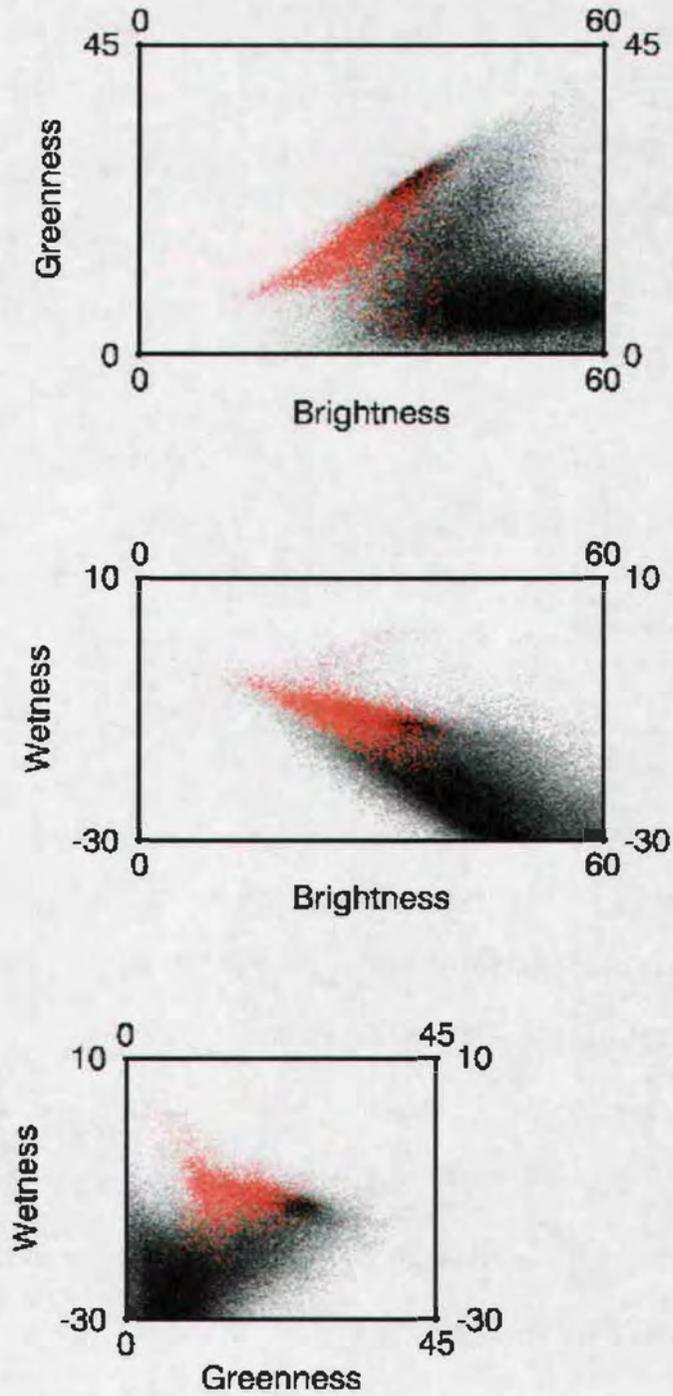


Figure 18. Evergreen needleleaf forest in MODIS Tasseled Cap space (in red), overlaid on global sample density plot (in black).

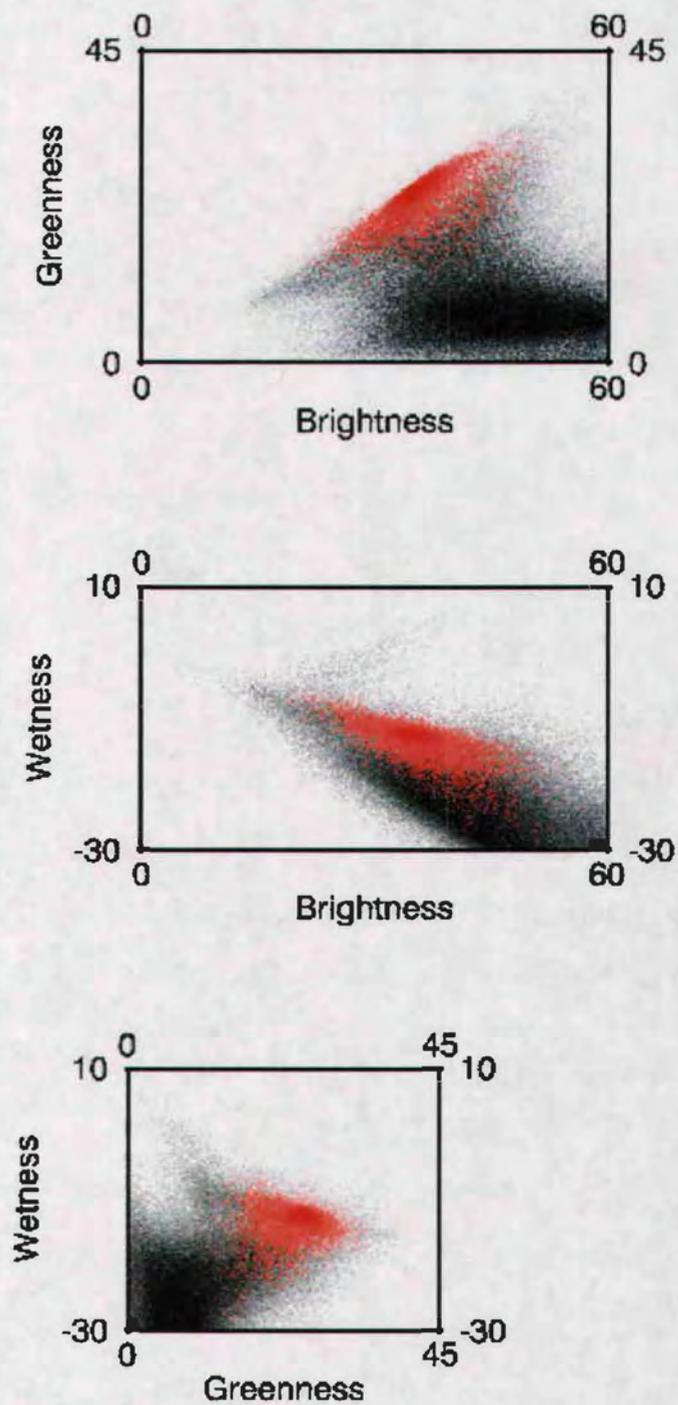
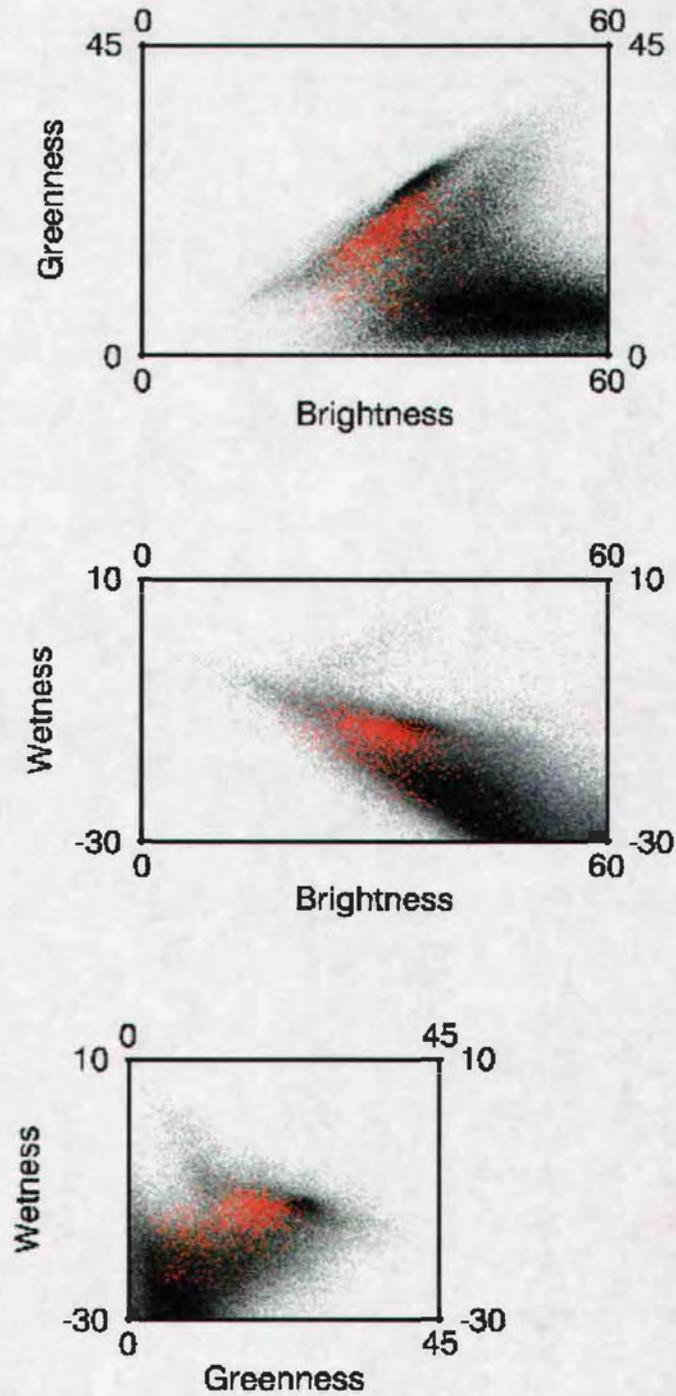
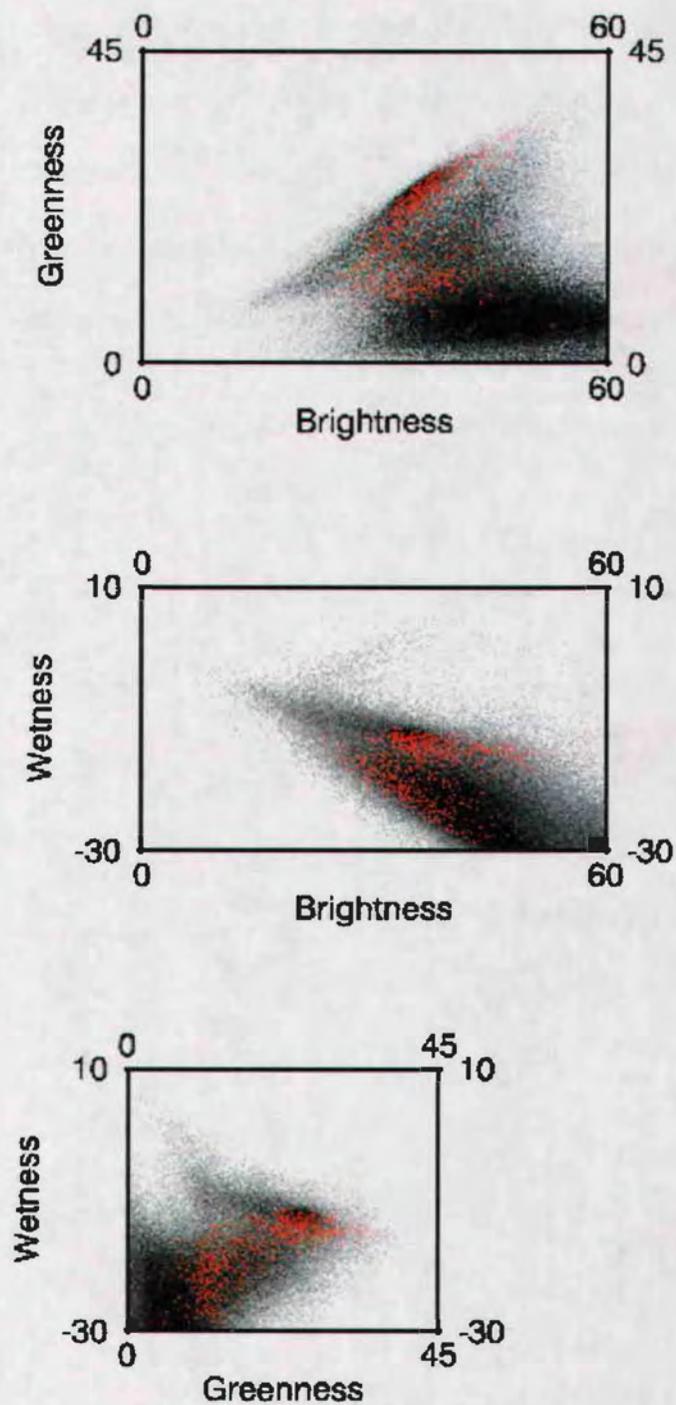


Figure 19. Evergreen broadleaf forest in MODIS Tasseled Cap space (in red), overlaid on global sample density plot (in black).



**Figure 20. Deciduous needleleaf forest in MODIS Tasseled Cap space (in red), overlaid on global sample density plot (in black).**



**Figure 21. Deciduous broadleaf forest in MODIS Tasseled Cap space (in red), overlaid on global sample density plot (in black).**

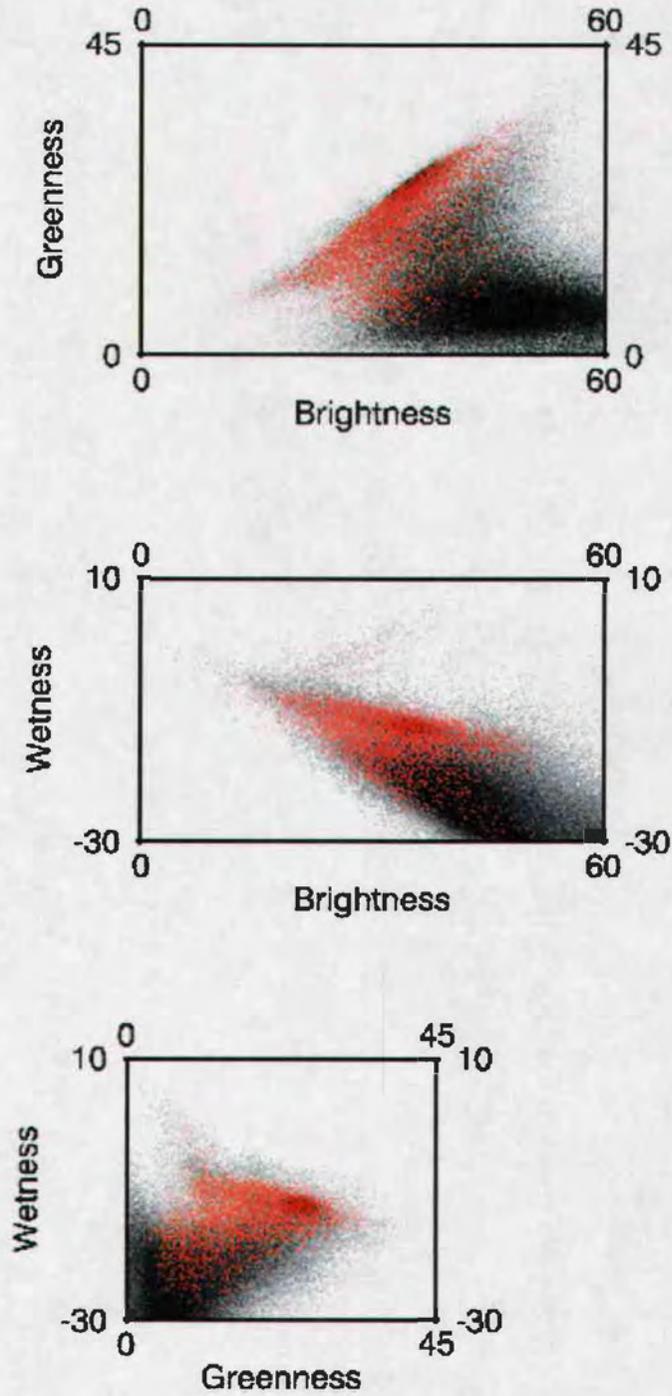


Figure 22. Mixed forest in MODIS Tasseled Cap space (in red), overlaid on global sample density plot (in black).

The distributions of these land cover classes obviously change over time, but it becomes somewhat useless to analyze temporal changes in phenology when using a global sample. For example, the phenological changes occurring in a particular land cover class in the northern hemisphere may balance out the changes in the southern hemisphere, causing the appearance of a stable spectral distribution. Thus, it is more useful to focus in on smaller scale plots where land cover is known, and the overall temporal pattern is invariable throughout the plot.

The BigFoot sites are useful toward this end (<http://www.fsl.orst.edu/larse/bigfoot/>). The BigFoot project was designed to validate certain MODIS products by linking remote sensing data with in situ measurements (Cohen and Justice, 1999). Nine sites were chosen to represent the different biomes of the Earth, and land cover is just one of the biophysical variables known in detail for the spatial extent of each site. A summary of eight sites in the Northern Hemisphere used here in analysis, including location, biome type, and land cover characteristics is given in table 5.

**Table 5. Summary of the eight BigFoot sites used in analysis.**

AGRO	Agricultural site in Illinois.
CHEQ	Mixed forest site in Wisconsin.
HARV	Harvard Forest—deciduous broadleaf forest in Massachusetts with some evergreen needleleaf forest.
KONZ	Konza Prairie—tallgrass prairie site in Kansas.
METL	Evergreen needleleaf forest site in Oregon.
NOBS	Boreal forest site (evergreen needleleaf) in Manitoba, Canada.
SEVI	Desert site in Arizona.
TUND	Arctic tundra site in Barrow, Alaska.

Figures 23 and 24 show plots of the BigFoot sites overlaid on density plots of the global sample for Northern Hemisphere summer and winter. Unfortunately, certain sites, such as TUND, have no data points for the winter months, when the land is covered in snow. The other sites express the phenological differences between the summer and winter months in the spectral differences shown in Tasseled Cap space. AGRO, an Illinois agricultural site, displays the phenological differences between a site that is planted and one that is in dormancy. This basic temporal trend ties together all the BigFoot sites with apparent seasonal changes. In the summer, most sites will migrate to brighter and greener region in spectral space, then return to their original locations in the winter.

One interesting artifact to note is the spur at the uppermost tip of the distribution, as seen in brightness-wetness space. Its placement is unusual; it is a significant departure from the regular forest spectral distribution, yet, both METL, a temperate, dry, open-canopy forest site, and HARV, a deciduous broadleaf site, occupy this region in the winter. It is likely that this spur is the spectral expression of a forest canopy with some amount of snow underneath. Though pixels containing snow were eliminated from the plot using the QA flags, the algorithm used to determine whether a pixel contains snow has an unknown threshold (Hall et al., 2002). A pixel with snow beneath a forest canopy may fall under this threshold and not get flagged.

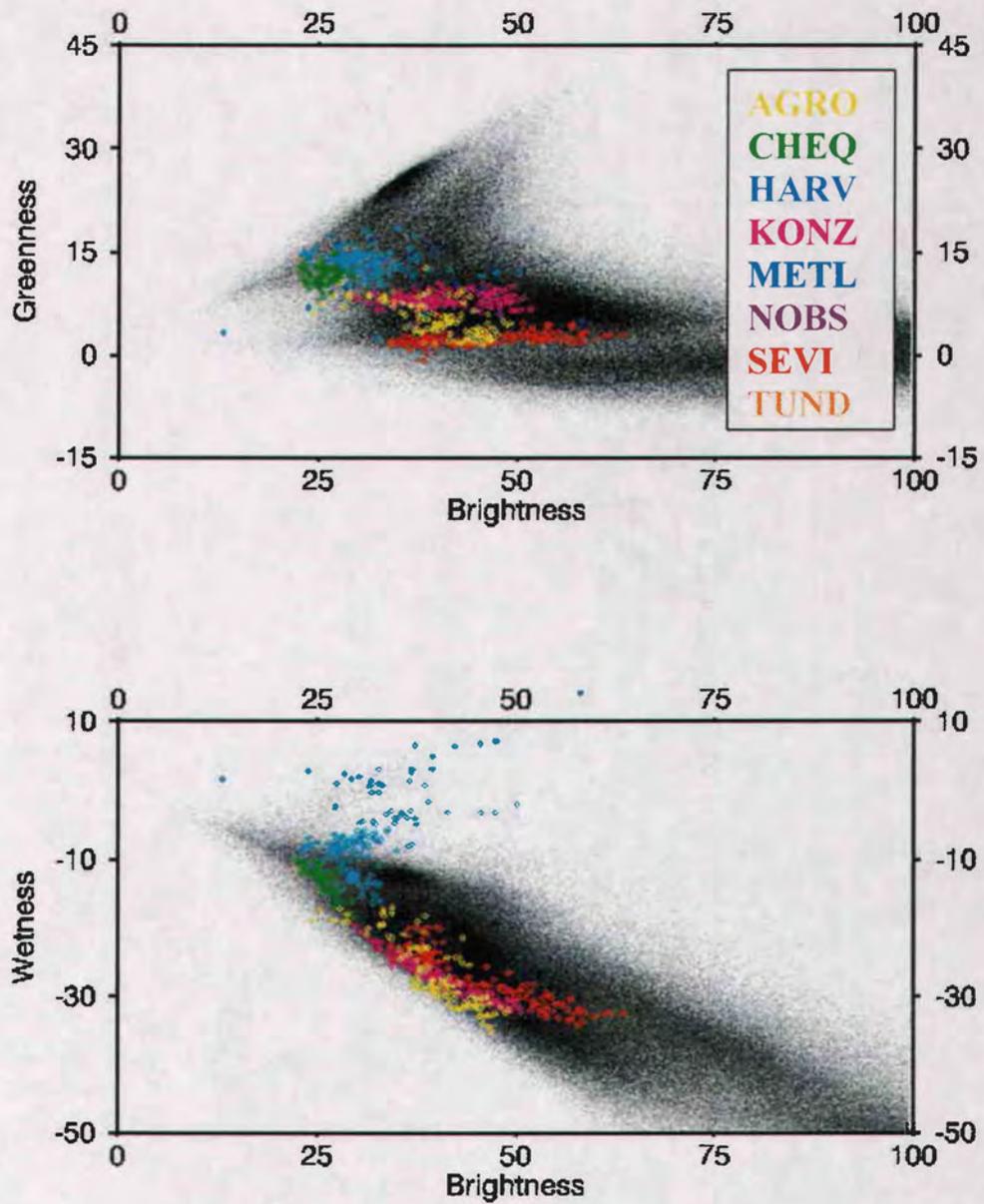
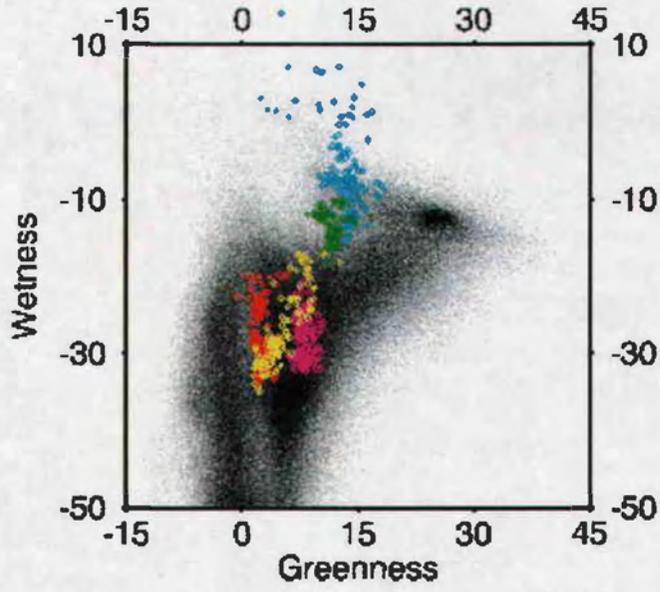


Figure 23. BigFoot NBAR data from January 1 through March 5 in MODIS Tasseled Cap space, overlaid on a global sample density plot.

Figure 23 continued.



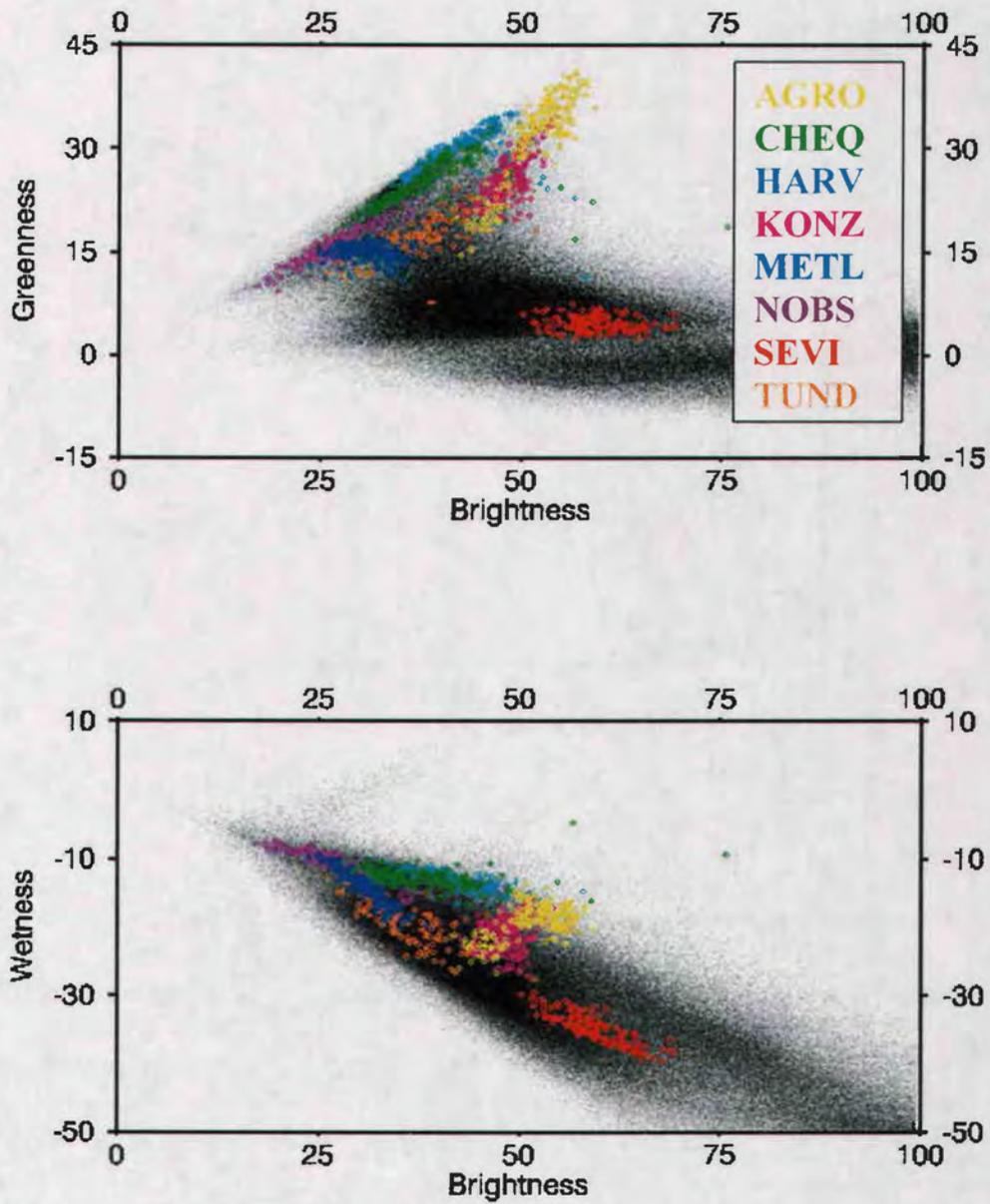
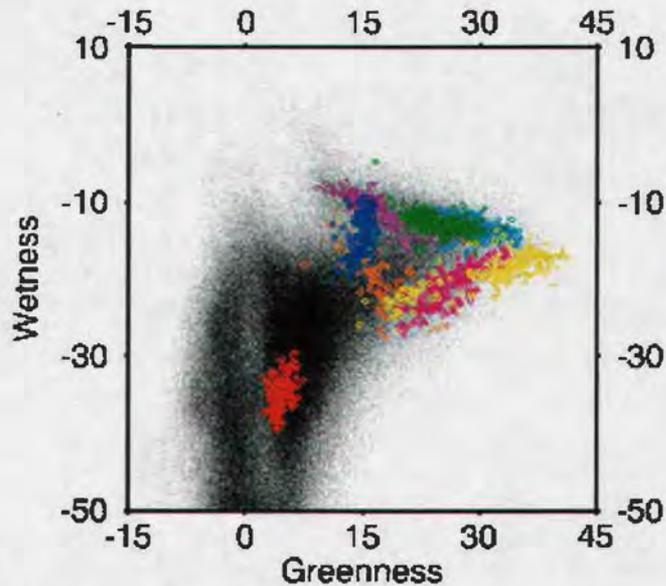


Figure 24. BigFoot NBAR data from June 25 through August 28 in MODIS Tasseled Cap space, overlaid on a global sample density plot.

Figure 24 continued.



The spectral differentiation among land cover types, observable in the global sample feature space plots, is demonstrated again by these BigFoot sites. In particular, there seems to be a very clear distinction, in all three perspectives, between forest, grass, and barren phenologies. These distinctions are even more obvious when actual phenological data is added to the global sample and viewed in feature space.

The MODIS Vegetation Continuous Fields (VCF) product (MOD44) was linked to the global sample, like the land cover product, to further explore phenological characteristics in feature space. The VCF product was designed to describe a single pixel not as a discrete class type, but as a mixture of different types

(Hansen, 2003). This product has percentage values for tree cover, herbaceous cover, and no vegetation cover.

The VCF is a 500 m product as opposed to the 1 km NBAR, so the percentage values were added to the global sample by taking the mean of the four pixels associated with each global sample pixel. Figures 25-27 show each of the three continuous fields in feature space. These plots correspond as expected with the land cover plots shown previously. Percent tree cover is highest in the same spectral region where forest classes are found, grass-type classes match up with areas of higher percentage herbaceous cover, and bare soil areas match up with areas of low or no vegetation cover.

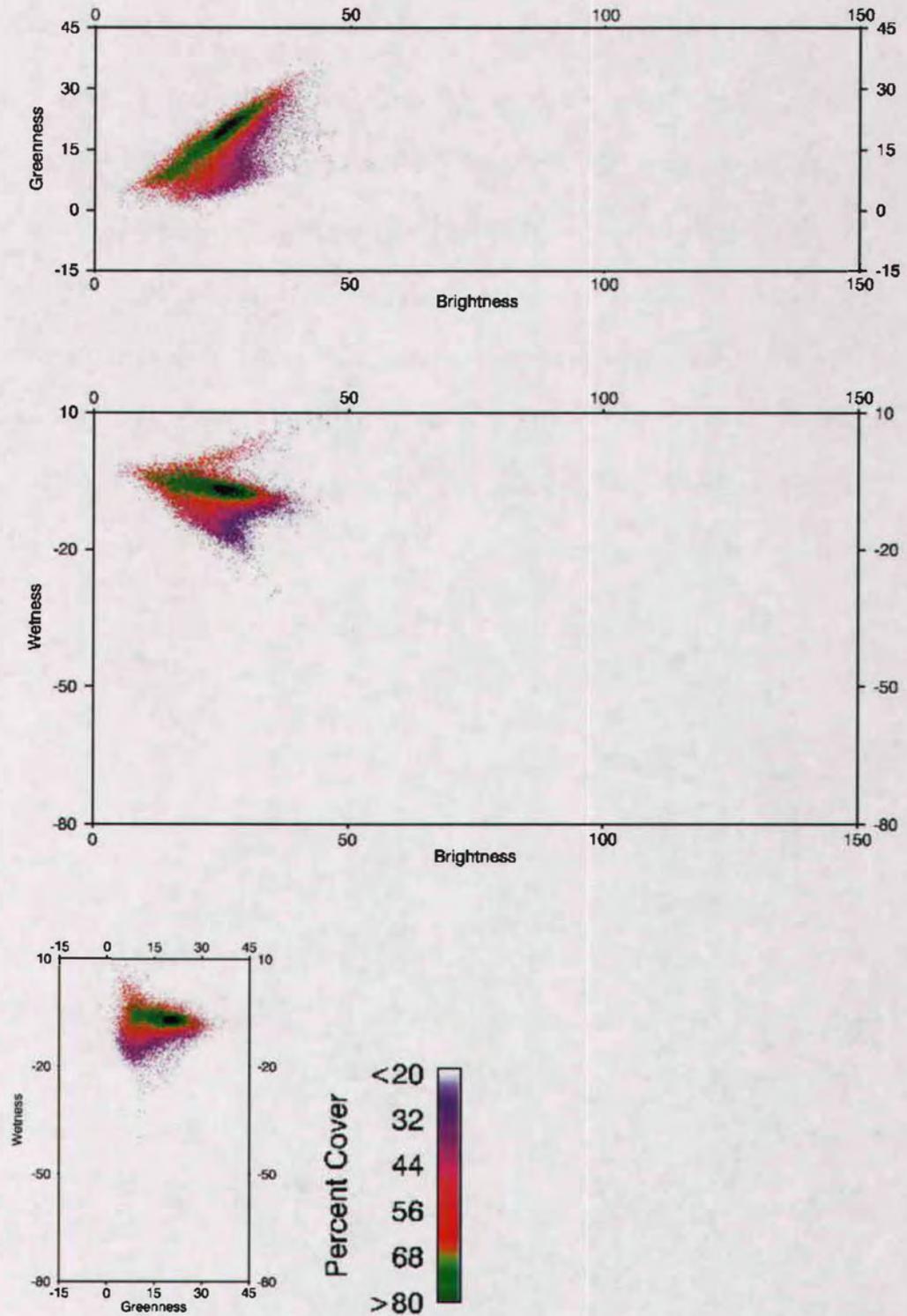


Figure 25. VCF percent tree cover in MODIS Tasseled Cap space.

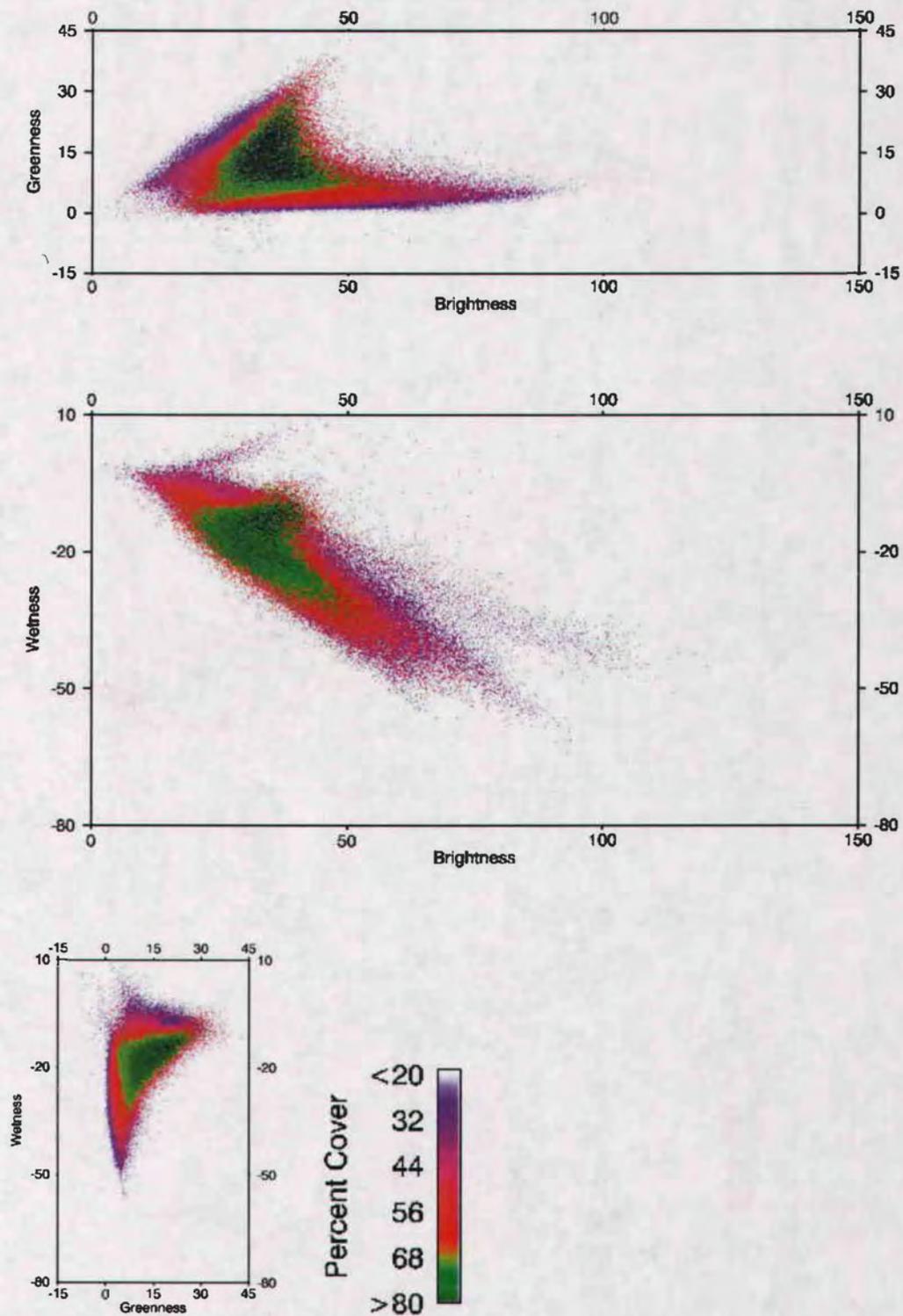


Figure 26. VCF percent herbaceous cover in MODIS Tasseled Cap space.

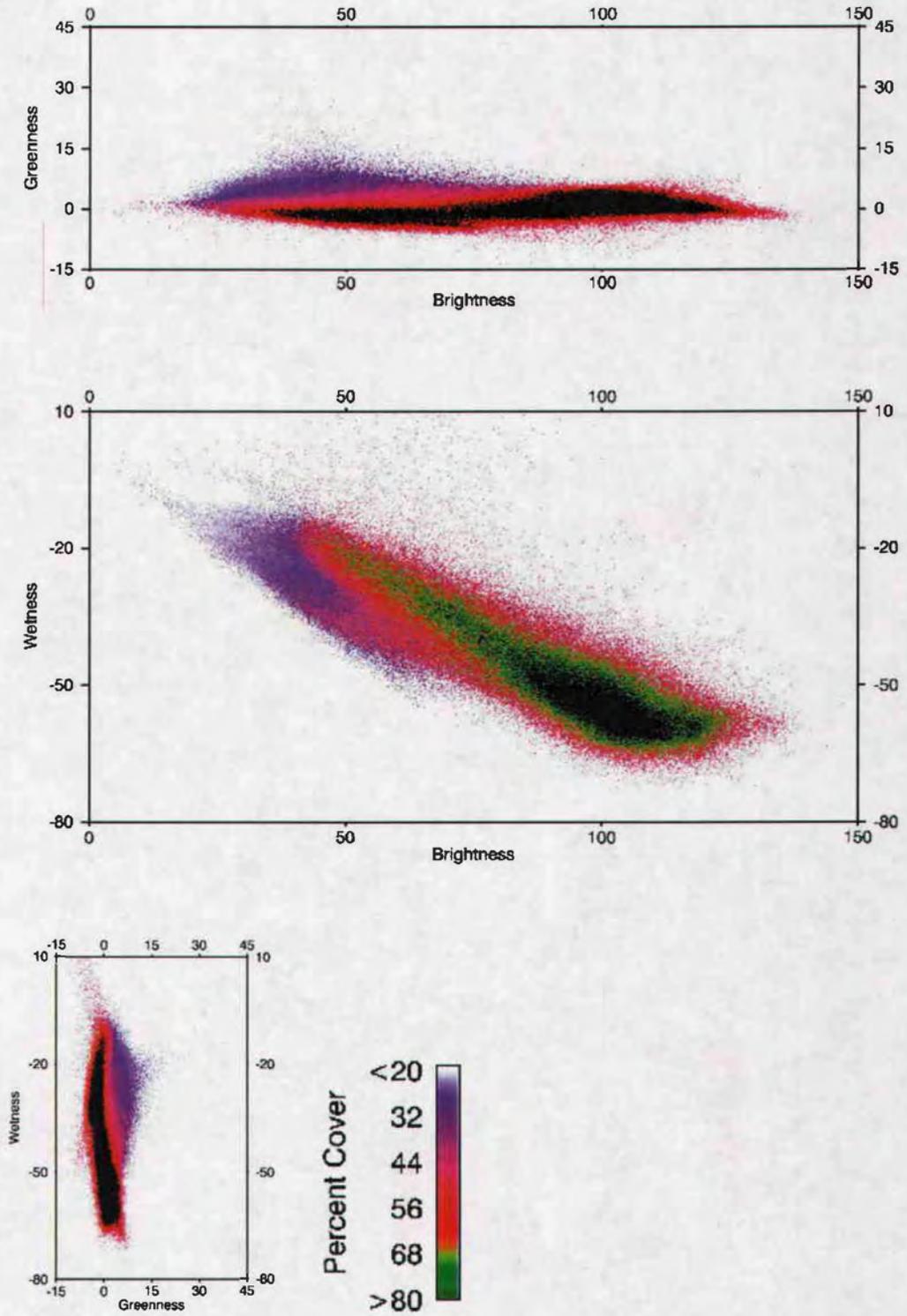


Figure 27. VCF percent bare in MODIS Tasseled Cap space.

### Application to Alternative Datasets

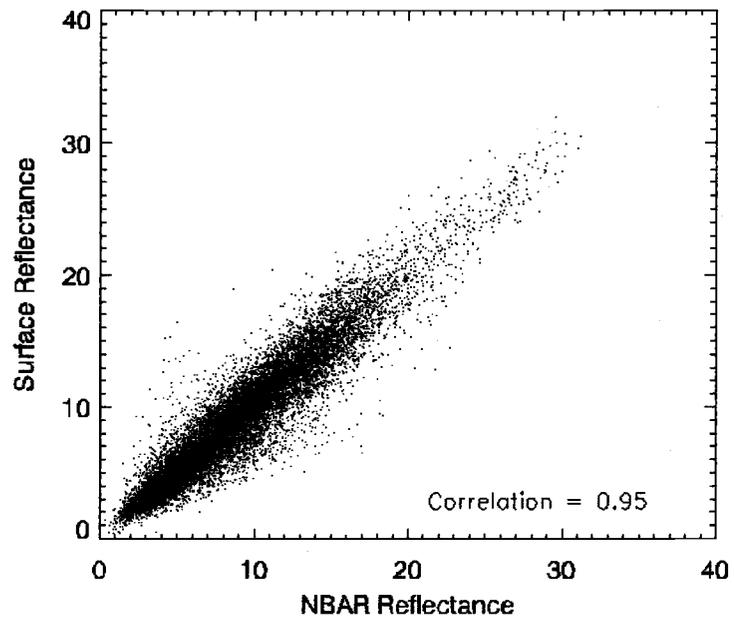
The results presented thus far in this paper have been applicable to the MODIS NBAR product. This product is derived from the original MODIS surface reflectance product (MOD09), and is designed to compensate for viewing angle geometry effects. The possibility that the MODIS Tasseled Cap coefficients could be applied to the raw surface reflectance products would expand the potential of the MODIS Tasseled Cap to be used with the original raw bands of data.

Four MODIS tiles were used to explore these possibilities, summarized in table 6. Each tile has a BigFoot site located within its boundaries, and each was selected from a different season. Thus, these four tiles should broadly represent global land cover characteristics. The accompanying QA datasets were used to eliminate snow and clouds from the tiles, and a random sample of 20,000 pairs of pixels was selected from the remaining data. There is, as expected, a strong correlation between each of the raw spectral band pairs (figure 28).

**Table 6. MODIS tiles used for comparison between NBAR (MOD43B4) and Surface Reflectance (MOD09).**

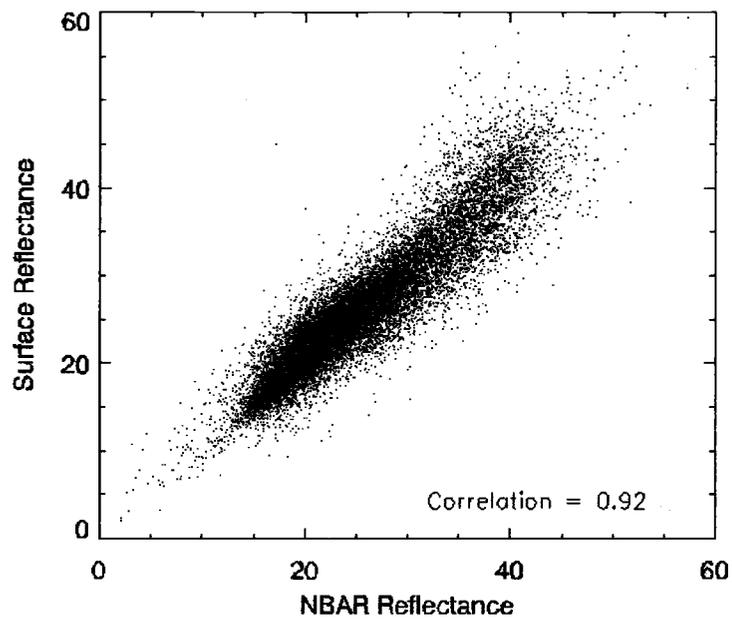
BigFoot site located in tile	Julian Day	Horizontal Tile Number	Vertical Tile Number
SEVI	1	09	05
CHEQ	273	11	04
METL	97	09	04
KONZ	177	10	05

NBAR Reflectance vs. Surface Reflectance: Red



(a)

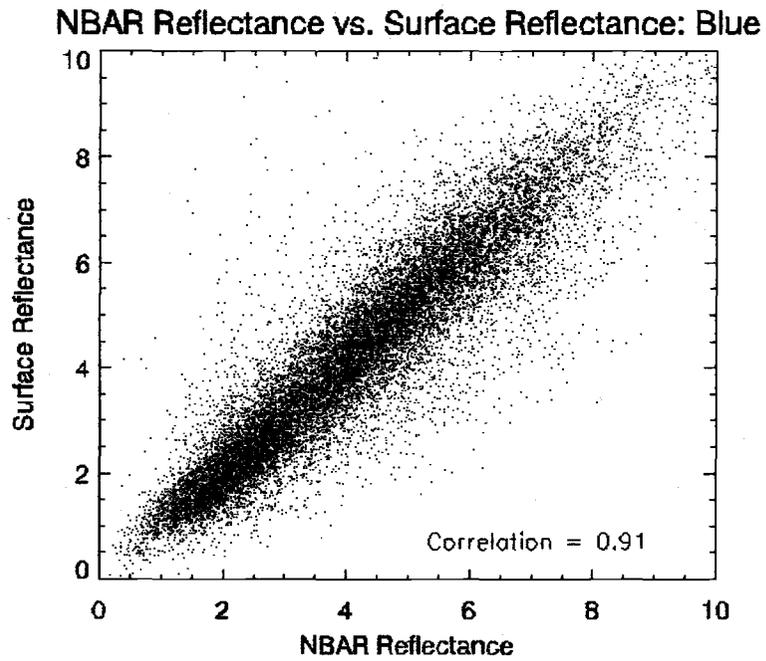
NBAR Reflectance vs. Surface Reflectance: NIR



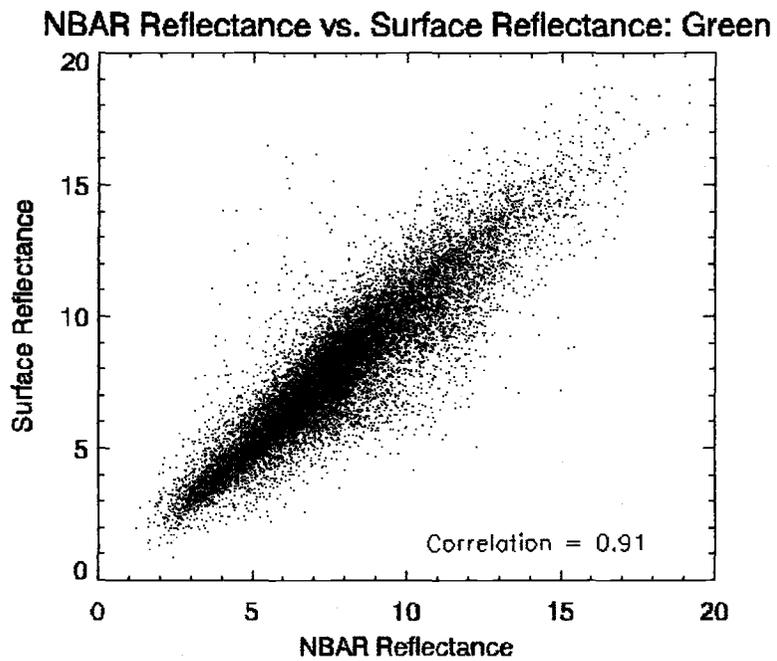
(b)

Figure 28 (a-g). Comparison of NBAR (MOD43B4) and Surface Reflectance (MOD09) raw spectral bands from four MODIS tiles.

Figure 28 continued.

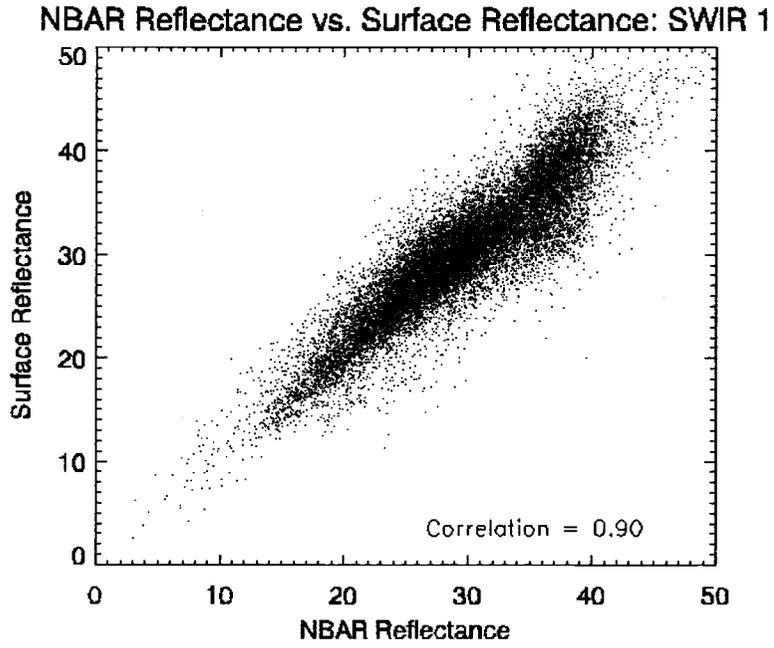


(c)

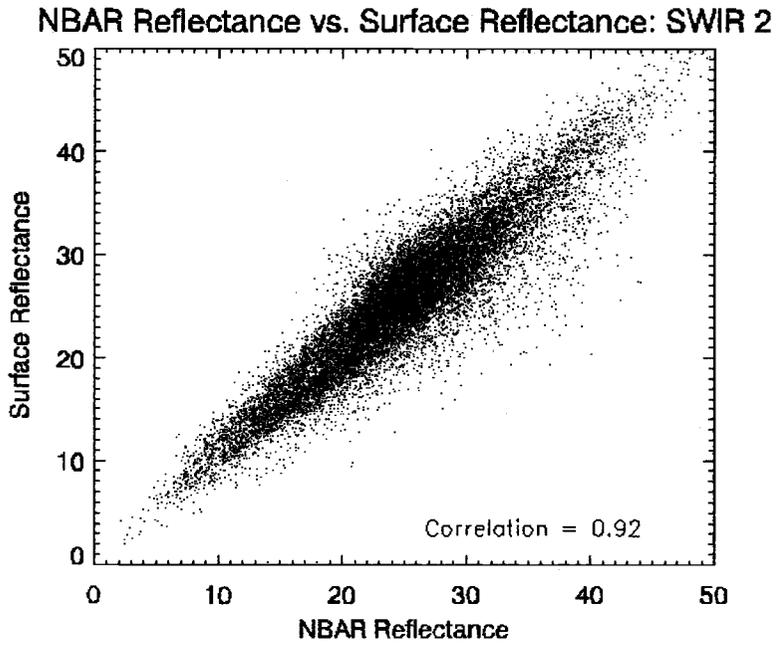


(d)

Figure 28 continued.

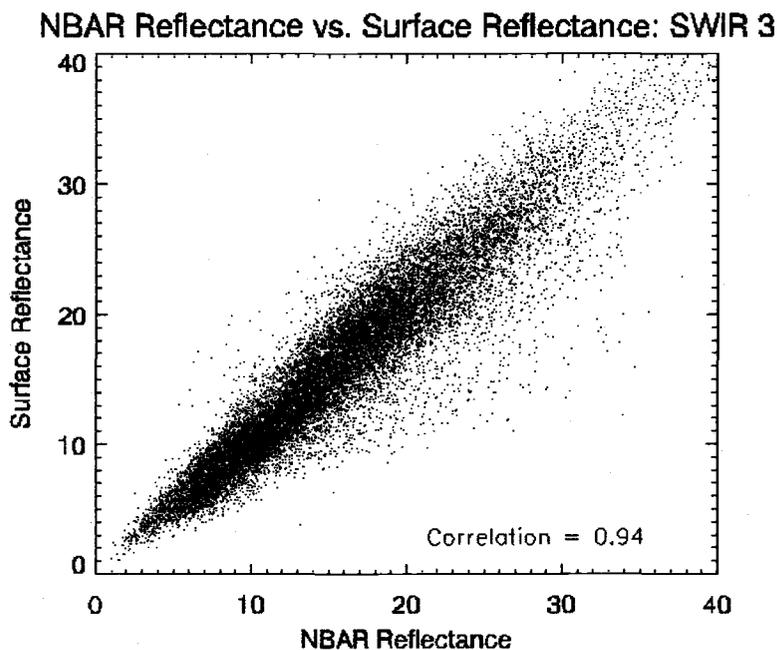


(e)



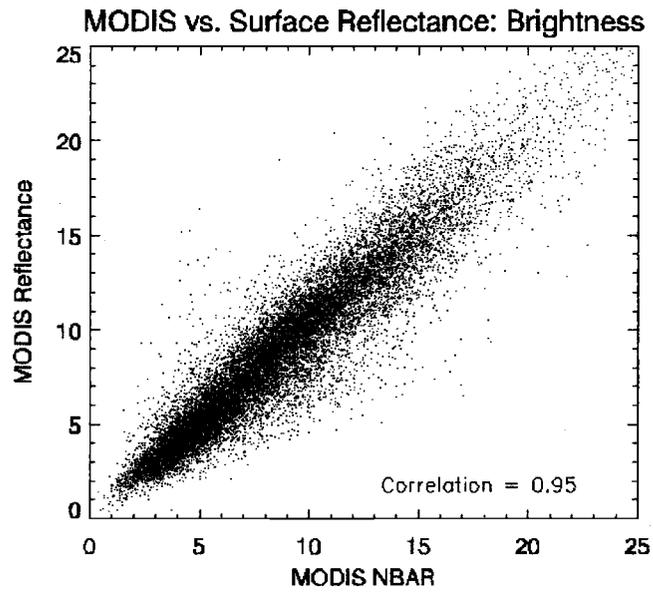
(f)

Figure 28 continued.

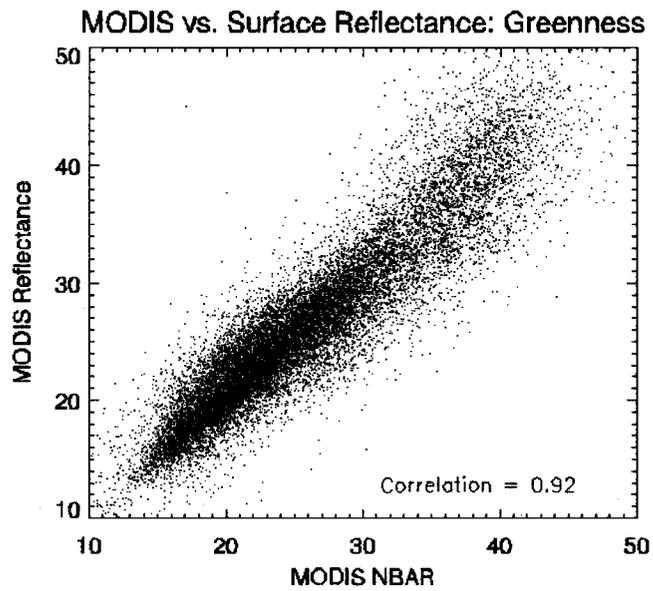


(g)

The MODIS Tasseled Cap coefficients were then applied to both the NBAR product and the raw surface reflectance of four MODIS tiles. Figure 29 shows scatterplots comparing the transformed bands of imagery, along with correlation coefficients. As expected, the raw and transformed bands are very highly correlated and show a nearly one to one relationship. Though the MODIS Tasseled Cap coefficients were developed using the MODIS NBAR product, the transformation can be applied just as effectively to the MODIS surface reflectance product.



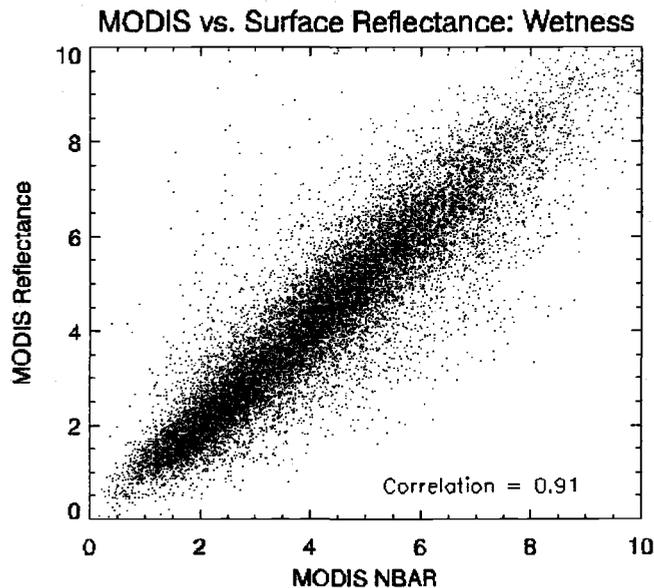
(a)



(b)

**Figure 29. Comparison of NBAR (MOD43B4) with Surface Reflectance (MOD09), both transformed using MODIS Tasseled Cap coefficients.**

Figure 29 continued.



(c)

### Summary and Discussion of Results

Much of the process of formulating the MODIS Tasseled Cap was built on the ideas and techniques used in previous studies. The concept of the Tasseled Cap is not new, and many results in the feature space exploration only serve to reinforce what is already known about vegetation spectral characteristics. There were some important differences, though, in the specific methods used in this study.

The breadth of the sampling process used here was much greater than those in previous studies. Because previous sampling efforts have been limited geographically, it was unknown how well the TM Tasseled Cap applied at the global scale. The objective here was to formulate a transformation for the MODIS

sensor that would be applicable to global vegetation studies, so it was desirable if not necessary to start with a sample representative of global vegetation conditions.

The global sample here is a random sample taken from the entire geographic extent of the globe, over the course of one year. The temporal dynamics of the sample were important to consider, as a geographically shifting sample could change the covariance structure of the spectral data enough to render a transformation meaningless at any particular date. By exploring how the covariance structure of the sample changed through time, it was found that the shifting geographic extent of the northern snow pack over time caused major shifting in the primary axis of variation. The first principal component of this preliminary sample was responding primarily to the snow/no-snow signal.

The sampling was repeated, with an extra step to mask out the snow using the NBAR QA flags. Because the objective was to build a transformation relevant to vegetation studies, it was not unfavorable to remove the snow. The covariance structure of this second sample was much more stable over time, and the principal components of the entire sample matched well with those of any individual date bin.

There are about 1.2 million pixels in this sample. This number seems rather small when considering that a single NBAR tile has 1,440,000 pixels. No single NBAR tile is fully representative of the range of vegetation condition found on the Earth. To test how representative these 1.2 million pixels were, histograms were compiled of the full global extent of NBAR tiles over the year, including good

quality pixels free of snow; cumulative distribution functions built from these and histograms of the global sample were compared and found to be nearly identical. The most comprehensive method of evaluating a sample is comparing it directly to the population it is supposed to represent. In this case, the population used for comparison was not the entire Earth over a whole year. Only good quality pixels were included, so cloud cover and snow cover eliminated a significant portion of the Earth from both population and sample. While sampling in the tropics is limited most times of the year, and in the boreal forest during the winter months, desert regions like the Sahara consistently provide ample opportunities for sampling. Thus, there is a bias in the sample toward regions of low cloud cover. This bias was acceptable, though, as the purpose of the sample was to extract principal components from the data. Differences among sample sizes of different regions were not likely to affect the primary axes of spectral variation among the entire sample.

Broad patterns in the covariance structure of the global sample seemed similar to the TM Tasseled Cap indices. This was born out by a comparison between the TM Tasseled Cap coefficients and the global sample PCA loadings. A more revealing demonstration was a density plot of the global sample in feature space. The structure of the distribution was strikingly similar to what is described by Crist and Cicone (1984a), and the particular formations are even better illuminated by coloring the sample according to IGBP class data extracted from the MODIS land cover product of 2000. The shape of the Tasseled Cap was clearly visible in the density plot of the global sample, only slightly rotated.

The results of these comparisons made sense, as the TM Tasseled Cap itself was a rotation of a PCA space. The apparent correspondence between the two data structures presented the opportunity for creating a MODIS Tasseled Cap. Rotating the MODIS PCA space to match the orientation of the TM Tasseled Cap would achieve the desired transformation for MODIS data, and would allow for easy translation between the two sensors. The rotation involved the creation of a target dataset, essentially the global sample transformed using the TM Tasseled Cap coefficients. It would have been ideal to collect a sample of real TM data that would match the global sample point for point. However, the costs associated with this undertaking were prohibitive. Rather, the six MODIS bands corresponding to the TM bands were found to be a good proxy for the TM sensor.

A Procrustes rotation was used to match the orientation of the global sample in pca space to the 6-band global sample in TM Tasseled Cap space. The mean Euclidean distance between the two samples after the rotation process was 10.4 reflectance units. This difference reflects the added information content of SWIR 1, the only MODIS band lacking a TM counterpart.

An exploration of the global sample in MODIS Tasseled Cap space yielded results that agreed with earlier findings of Crist and Cicone (1984a). Land cover types, as defined by the MODIS land cover product IGBP classes, are distributed in MODIS Tasseled Cap space as expected, with crops forming the familiar Tasseled Cap shape, surrounded by the badge of trees and a region occupied by barren land.

The expression of the global sample in terms of the MODIS VCF product agrees with the global sample IGBP distributions. In feature space, the phenology of the global sample coincides with the corresponding cover types. Investigating the agreement between two MODIS products was not an objective of this study, but feature space analysis of the global sample appears to be a simple and effective way to do so.

Temporal dynamics in spectral patterns were also explored using BigFoot sites. Again, patterns reflect findings of previous work. BigFoot sites generally shift in feature space to a brighter, greener region during the growing season. The exception seems to be with forest sites, such as NOBS and HARV, which likely have a sub-canopy snow pack during the winter. The observation of the snow related spur in the global sample distribution introduces the possibility of detecting regions with sub-canopy snow.

The MODIS Tasseled Cap loadings were formulated using NBAR data. It was expected that the coefficients could be applied just as easily to MODIS surface reflectance, and this was confirmed by comparing four NBAR tiles to the corresponding surface reflectance tiles.

There are a number of sources of uncertainty potentially affecting the formulation of the MODIS Tasseled Cap. The formulation depends first and foremost on the distribution of sample pixels in spectral space and the distribution may be slightly skewed by various artifacts in the original data. Topographic effects are not corrected in the sample, and atmospheric correction of the original data is somewhat variable across swath width. Native spatial resolution is also

variable across swath width, increasing up to five-fold in the off-nadir views (Vermote et al., 1997). These factors may be sources of error in the distribution of sample pixels in spectral space, affecting the derived principal components.

As discussed earlier, the sample is also biased toward cloud-free areas of the globe. It may be possible to correct for this by changing the sampling strategy to ensure equal sampling among various geographic regions. However, it is unknown how much this would affect the distribution of the sample in spectral space and thus the derivation of principal components. Derived principal components were relatively stable over time despite significant shifts in geographic representation, suggesting that correcting for the geographic bias in the sampling stage is unnecessary.

Once the PCA loadings were extracted, the derivation of the MODIS Tasseled Cap coefficients was dependent on the rotation process. Using the TM Tasseled Cap as a calibration tool allows for the greatest opportunity for translation between the two data spaces for vegetation analyses involving multiple data sources. However, the TM Tasseled Cap coefficients were derived using less robust methods. It would be preferable to rotate the TM Tasseled Cap space to match the orientation of an independently derived, globally relevant MODIS Tasseled Cap. Such an undertaking was beyond the scope of this study, but is an opportunity for future investigation.

## **Conclusions**

The objectives of this study were to explore the structure of MODIS data and use the knowledge gained to formulate a transformation that would be relevant to

global vegetation studies. The methods used in formulating the TM Tasseled Cap were used as a model, though it was not clear from the beginning that the end result would be a MODIS Tasseled Cap. Though much more information was introduced in the process here, with a global sample and an additional spectral band, there was no significant departure from the TM data structure. The form and characteristics of the Tasseled Cap were easily recognizable in the global sample plots, and the MODIS transformation behaves in the same way with respect to vegetation dynamics. The correspondence between the data structures of the MODIS and TM sensors made it possible to tie many of the results in this paper to concepts that originated with the formulation of the TM Tasseled Cap. There is also a more direct connection between the two transformations. Because the MODIS Tasseled Cap was designed to match the orientation of the TM Tasseled Cap, future analyses may take advantage of both types of data using transformations that are easily translatable. With the continual design and launch of new sensors, it becomes important to maintain continuity among them.

The original Tasseled Cap developed for Landsat MSS was built on a very small sample size. Successive formulations have all varied in sample size and sampling methods, but the basic patterns in vegetation spectral characteristics remain the same in every outcome. Whether or not the process of transformation development is relatively insensitive to sample characteristics, the MODIS Tasseled Cap was made robust by using a global sample; the data structure of the sample has been shown to represent every region of the globe at all times of the year. A global sample also allows for a more definitive spectral characterization of global

vegetation conditions. The development of the global sample also presented the opportunity to compare various MODIS products in a common data space. In particular, feature space plots gave a nice visual summary of how well the MODIS land cover product agrees with the VCF product for the full geographic extent of the Earth. The MODIS Tasseled Cap has proven to be a simple and powerful tool for interpreting dynamic vegetation characteristics and analyzing global remote sensing products.

## References

- Cohen, W. B. & Spies, T. A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from landsat and SPOT imagery. *Remote Sensing of Environment*, 41, 1-17.
- Cohen, W. B. & Justice, C. O. (1999). Validating MODIS terrestrial ecology products: linking in situ and satellite measurements. *Remote Sensing of Environment*, 70, 1-4.
- Crist, E. P. (1983). The Thematic Mapper Tasseled Cap—A Preliminary Formulation. *Machine Processing of Remotely Sensed Data Symposium*.
- Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17, 301-306.
- Crist, E. P. & Cicone, R. C. (1984a). A physically based transformation of Thematic Mapper data—the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22, 256-263.
- Crist, E. P., & Cicone, R. C. (1984b). Comparisons of the Dimensionality and Features of Simulated Landsat-4 MSS and TM Data. *Remote Sensing of Environment*, 14, 235-246.
- Dymond, C. C., Mladenoff, D. J., & Radeloff, V. C. (2002). Phenological differences in Tasseled Cap indices improve deciduous forest classification. *Remote Sensing of Environment*, 80, 460-472.
- Friedl, M. A., McIver, D. K., Hodges, J. F. C., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, 83, 287-302.
- Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGiromamo, N., & Bayr, K. J. (2002). MODIS Snow-Cover Products. *Remote Sensing of Environment*, 83, 181-194.
- Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., & Sohlberg, R. A. (2003). Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions*, 1, 7.
- Kauth, R. J. & Thomas, G. S. (1976). The Tasseled Cap—a graphic description of the spectral temporal development of agricultural crops as seen by Landsat.

Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, W. Lafayette, IN, 4B41-4B51.

Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press, 521 pp.

Oetter, D. R. Cohen, W. B., Berterretche, M., Maiersperger, T. K., & Kennedy, R. E. (2001). Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. *Remote Sensing of Environment*, 72, 139-155.

Schaaf et al (2002). First Operational BRDF, Albedo and Nadir Reflectance Products from MODIS. *Remote Sensing of Environment*, 83, 135-148.

Skakun, R. S., Wulder, M. A., & Franklin, S. E. (2003). Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, 86, 433-443.

Strahler, A., Muchoney, D., Borak, J., Friedl, M., Gopal, S., Lambin, E., & Moody, A. (1999). MODIS Land Cover and Land-Cover Product Algorithm Theoretical Basis Document (ATBD) Version 5.0. Boston, MA: Center for Remote Sensing, Boston University,  
[http://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod12.pdf](http://modis.gsfc.nasa.gov/data/atbd/atbd_mod12.pdf), 72 pp.

Vermote, E. F., El Saleous, N. Z., Justice, C. O., Kaufman, Y. J., Privette, J., Remer, L. C., & Tanre, D. (1997). Atmospheric correction of visible to middle infrared EOS-MODIS data over land surface, background, operational algorithm and validation. *Journal of Geophysical Research*, 102 (14).

Vikhamar, D. & Solberg, R. (2003). Snow-cover mapping in forests by constrained linear spectral unmixing of MODIS data. *Remote Sensing of Environment*, 88, 309-323.

Zhang, X. Y., Schaaf, C. B., Friedl, M. A., Strahler, A. H., Gao, F., & Hodges, J. F. C. MODIS tasseled cap transformation and its utility. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '02)*, Toronto, Canada, 24-28 June, 2002.

### Chapter Three: Conclusions

The MODIS Tasseled Cap Transformation developed here is not a significant departure from the TM Tasseled Cap. The same broad patterns in vegetation characteristics and temporal dynamics are observable in both spaces. Because the Tasseled Cap can be achieved and recognized at different scales, using different samples and different methods, it seems that the Tasseled Cap is an intrinsic property of vegetation at multiple scales.

Regardless of whether a true global random sample was required to build a globally relevant sample, the sampling process in itself yielded original results. A better understanding of temporal patterns in global spectral covariance structure was gained, and used to achieve a globally representative sample. This global sample allows for making more definitive statements about how global land cover types are distributed in spectral space; the same is true for phenological trends. The sample also provided a method for comparing multiple remote sensing products at the global scale.

It is hoped that the MODIS Tasseled Cap will be as useful as the TM Tasseled Cap in vegetation analyses, but more applicable to the global scale. Possibilities for future work include the application of the MODIS Tasseled Cap to global land cover mapping. Because the information content of the Tasseled Cap is greater than any single vegetation index, there is significant potential for its use in global vegetation monitoring.

## Bibliography

Cohen, W. B. & Spies, T. A. (1992). Estimating structural attributes of Douglas-fir/western hemlock forest stands from landsat and SPOT imagery. *Remote Sensing of Environment*, 41, 1-17.

Cohen, W. B. & Justice, C. O. (1999). Validating MODIS terrestrial ecology products: linking in situ and satellite measurements. *Remote Sensing of Environment*, 70, 1-4.

Cracknell, A. P. & Hayes, L. W. B. (1993). *Introduction to Remote Sensing*. London: Taylor & Francis, 293 pp.

Crist, E. P. (1983). The Thematic Mapper Tasseled Cap—A Preliminary Formulation. *Machine Processing of Remotely Sensed Data Symposium*.

Crist, E. P. (1985). A TM tasseled cap equivalent transformation for reflectance factor data. *Remote Sensing of Environment*, 17, 301-306.

Crist, E. P. & Cicone, R. C. (1984). Application of the Tasseled Cap concept to simulated Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, 50, 343-352.

Crist, E. P. & Cicone, R. C. (1984a). A physically based transformation of Thematic Mapper data—the TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing*, GE-22, 256-263.

Crist, E. P., & Cicone, R. C. (1984b). Comparisons of the Dimensionality and Features of Simulated Landsat-4 MSS and TM Data. *Remote Sensing of Environment*, 14, 235-246.

Dymond, C. C., Mladenoff, D. J., & Radeloff, V. C. (2002). Phenological differences in Tasseled Cap indices improve deciduous forest classification. *Remote Sensing of Environment*, 80, 460-472.

Friedl, M. A., McIver, D. K., Hodges, J. F. C., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., & Schaaf, C. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment*, 83, 287-302.

- Frohn, R. C. (1998). *Remote Sensing for Landscape Ecology*. Boca Raton, FL: Lewis Publishers, 99 pp.
- Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGiromamo, N., & Bayr, K. J. (2002). MODIS Snow-Cover Products. *Remote Sensing of Environment*, 83, 181-194.
- Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., & Sohlberg, R. A. (2003). Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS Vegetation Continuous Fields Algorithm. *Earth Interactions*, 1, 7.
- Kauth, R. J. & Thomas, G. S. (1976). The Tasseled Cap—a graphic description of the spectral temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Purdue University, W. Lafayette, IN, 4B41-4B51.
- Huete, A. & Justice, C. (1999). MODIS Vegetation Index (MOD 13) Algorithm Theoretical Basis Document. Greenbelt: NASA Goddard Space Flight Center, <http://modarch.gsfc.nasa.gov/MODIS/LAND/#vegetation-indices>, 129 pp.
- Jensen, J. R. (2000). *Remote Sensing of the Earth: An Earth Resource Perspective*. Upper Saddle River, New Jersey: Prentice Hall, 544 pp.
- Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., Strahler, A., Lucht, W., Myneni, R. B., Knyazikhin, Y., Running, S. W., Nemani, R. R., Wan, Z. M., Huete, A. R., van Leeuwen, W., Wolfe, R. E., Giglio, L., Muller, J. P., Lewis, P., & Barnsley, M. J. (1997). The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1228-1249.
- Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press, 521 pp.
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G., & Nemani, R. R. (1997). Increased plant growth in the northern high latitudes from 1981-1991. *Nature*, 386, 698-702.
- Oetter, D. R. Cohen, W. B., Berterretche, M., Maiersperger, T. K., & Kennedy, R. E. (2001). Land cover mapping in an agricultural setting using multiseasonal Thematic Mapper data. *Remote Sensing of Environment*, 72, 139-155.
- Schaaf et al (2002). First Operational BRDF, Albedo and Nadir Reflectance Products from MODIS. *Remote Sensing of Environment*, 83, 135-148.

Skakun, R. S., Wulder, M. A., & Franklin, S. E. (2003). Sensitivity of the thematic mapper enhanced wetness difference index to detect mountain pine beetle red-attack damage. *Remote Sensing of Environment*, 86, 433-443.

Strahler, A., Muchoney, D., Borak, J., Friedl, M., Gopal, S., Lambin, E., & Moody, A. (1999). MODIS Land Cover and Land-Cover Product Algorithm Theoretical Basis Document (ATBD) Version 5.0. Boston, MA: Center for Remote Sensing, Boston University,  
[http://modis.gsfc.nasa.gov/data/atbd/atbd\\_mod12.pdf](http://modis.gsfc.nasa.gov/data/atbd/atbd_mod12.pdf), 72 pp.

Vermote, E. F., El Saleous, N. Z., Justice, C. O., Kaufman, Y. J., Privette, J., Remer, L. C., & Tanre, D. (1997). Atmospheric correction of visible to middle infrared EOS-MODIS data over land surface, background, operational algorithm and validation. *Journal of Geophysical Research*, 102 (14).

Vikhamar, D. & Solberg, R. Snow-cover mapping in forests by constrained linear spectral unmixing of MODIS data. *Remote Sensing of Environment*, 88, 309-323.

Zhang, X. Y., Schaaf, C. B., Friedl, M. A., Strahler, A. H., Gao, F., & Hodges, J. F. C. MODIS tasseled cap transformation and its utility. *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '02)*, Toronto, Canada, 24-28 June, 2002.