

BRIDGING THE LANDSAT DATA GAP: EVALUATING ASTER AS AN ALTERNATIVE

By

FORREST R. STEVENS

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2009

© 2009 Forrest R. Stevens

To my father, Bob and with many loving memories of my mother, Stephenie

ACKNOWLEDGMENTS

My sincerest gratitude must be given to my peers, colleagues, friends and family. Their support and guidance, thoughts and questions, and help through many problems untold are reflected in every word you read here. To name only a few I would like to thank first my wife, Gabriela Stocks, for her prodding and understanding, and an uncompromising intellect that leaves no claim unchallenged. I would like to acknowledge the quality and constant motivation provided by my peers in the University of Florida Geography Department. Thanks also to the researchers collaborating in the MAP region of South America. Without the help of Dr. Michael Binford, Dr. Matt Marsik, Karla Rocha, Natlia Hoyos, Amy Duchelle and Andrea Chavez, much of this study's data and legwork required for its use would be incomplete.

This research was partially funded under a National Science Foundation Human and Social Dynamics grant entitled, "AOC: Infrastructure Change, Human Agency, and Resilience in Social-Ecological Systems." I would like to thank Dr. Stephen J. Perz for his generous support for travel and fieldwork related to this study's data collection.

My sincerest gratitude must also go out to my adviser, Dr. Jane Southworth. Her enthusiasm for research, the methods developed and presented here, and the rigor that she espouses and upholds for her students and their work is critical to our success. In addition, I would like to sincerely thank my other committee members, Dr. Timothy Fik and Dr. Greg Kiker, for their patience, guidance and help in crafting my study and bringing it along from thought to data and paper.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	7
LIST OF FIGURES	8
LIST OF OBJECTS	10
ABSTRACT.....	11
CHAPTER	
1 INTRODUCTION.....	12
2 METHODS.....	20
Study Region	20
Data Preparation	21
Data Analysis.....	27
3 RESULTS.....	37
Untransformed Comparisons.....	37
Theil-Sen Bisector Regressions.....	39
Transforming ASTER Data and its Outcomes	40
Transforming ASTER and its Effects on Calculated Vegetation Indices.....	42
4 DISCUSSION AND CONCLUSION	65
Study Implications.....	65
Conclusion.....	69
APPENDIX	
A ORIGINAL SOURCE CODE FOR DISCUSSED ANALYSES.....	70
ASTER Top-of-Atmosphere Reflectance Conversion from At-Sensor Radiance	70
Theil-Sen and H-M-S Intercept Regression	73
ASTER Tasseled Cap Calculations and Conversion.....	75
License for Appended Source Code	77
B SAMPLE ENVI CONFIGURATION TEMPLATES FOR ATMOSPHERIC CORRECTION.....	78
Notes on Atmospheric Correction in ENVI™.....	78

ASTER FLAASH/MODTRAN Configuration File for ENVI.....	78
Landsat FLAASH/MODTRAN Configuration File for ENVI.....	79
LIST OF REFERENCES	81
BIOGRAPHICAL SKETCH	85

LIST OF TABLES

<u>Table</u>	<u>page</u>
1-1 Key sensor and platform characteristics for the ASTER and Landsat ETM+ systems.....	18
2-1 Data acquired and subset for ASTER and ETM+ comparisons.....	34
3-1 Transformation coefficients as calculated for the specified year and band combinations.	46
3-2 Transformation results of original data after applying the derived coefficients from Table 2-1.....	47
3-3 Transformation results of data after applying the independently derived coefficients from Table 2-1 to MY01 data.....	48
3-4 Comparison of derived vegetation indices from post-transformation MY01 ASTER and original ETM+ data.....	49

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 System response curve comparison for the EOS ASTER and Landsat ETM+ sensors.	19
2-1 Study region map showing the areas subset from the various data sources used.....	35
2-2 Outline of the workflow for each data file type, grouped by year and subset.	36
3-1 Comparison by band of estimated Theil-Sen bisector slopes for ETM+ vs. ASTER L1A and L1B FLAASH surface reflectances derived for the MS01 subset.	50
3-2 The K-M-S bisector intercept estimates for ETM+ vs. ASTER L1A and L1B FLAASH surface reflectances for each band combination, corresponding to the slope estimates from Figure 3-1.	50
3-3 Illustrating the effects of ETM+ B7 MY01 being regressed on ASTER MY01.	51
3-4 Comparison of estimated, post-transformation Theil-Sen bisector slopes for ETM+ vs. ASTER L1A FLAASH MS01 surface reflectances that were spectrally resampled to match ETM+ B7.	52
3-5 Comparison of estimated, post-transformation K-M-S bisector intercepts matching the slopes from Figure 3-3.	52
3-6 Visual representation of transformation effects on estimated regression coefficients for the ETM+ MS01 regressed on ASTER MS01 pre- and post-transformation.	53
3-7 Transformation effects are shown as measured by metrics of spread and centrality for pre- and post-transformation of ASTER MS01 by its own MS01 regression coefficients.	54
3-8 Each of the five matching ETM+ MY01 bands and their diagnostic regressions are plotted against the corresponding transformed ASTER MY01 data.	55
3-9 Transformation effect as measured by the pre- and post-transformation regression coefficients for ETM+ MY01 regressed on ASTER MY01 bands after transforming with MS01 estimated coefficients.....	56
3-10 Post-transformation effects on MY01 (transformed by MS01) as measured by metrics of spread and centrality.	56
3-11 Scatterplots and their diagnostic Theil-Sen bisector regressions of ETM+ derived NDVI and EVI2 vegetation indices, against the MY01 (transformed by MS01) derived indices.	57

3-12	The spatial distribution of EVI2 values calculated for the MY01 subset after transformation using estimated Theil-Sen bisector coefficients from the MS01 sample.	58
3-13	This map shows the artifacts created by mapping the difference of visible red bands, calculated by subtracting the ETM+ band from the MY01, untransformed data.	59
3-14	The visible artifacts, though still present are clearly located near the center of the error distribution as the differencing of the transformed ASTER MY01 visible red and ETM+ band shows.	60
3-15	Scatter plot of ETM+ visible red data against the transformed ASTER MY01 band.	61
3-16	Calculated EVI2 layer with changed pixels based on the transformed MY01 (transformed by the estimated coefficients of MS01) overlayed.	62
3-17	Image difference map of MY01 (transformed by MS01 coefficients) minus ETM+ EVI2 data.	63
3-18	Illustration of pixel locations of visible red difference shifts within scatterplots of the calculated vegetation indices in post-transformation data.	64

LIST OF OBJECTS

<u>Object</u>	<u>page</u>
A-1 GNU General Public License Version 3 (.txt file 35 KB)	77

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Master of Science

BRIDGING THE LANDSAT DATA GAP: EVALUATING ASTER AS AN ALTERNATIVE

By

Forrest R. Stevens

May 2009

Chair: Jane Southworth

Major: Geography

Longitudinal studies of land use and land cover change (LULCC) increasingly utilize remotely sensed data from multiple sensors. Comparing multiple-sensor data presents challenging, technical problems that are often poorly documented and that, without solutions, severely limit many analytical tools for multi-date change detection. This study compares same-day data collected by the Landsat 7 ETM+ and ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) sensors from the South American tri-national border of Brazil, Peru, and Bolivia. Same-day imagery minimizes environmental variability between collected data and allows for an accurate assessment of sensor-specific differences in data products. A non-parametric, Theil-Sen bisector regression approach is introduced as both a comparison and transformation method, allowing for more robust cross-sensor comparisons in the presence of common sources of error in remotely-sensed data. Aspects of multi-sensor data comparisons including techniques for calibration, image registration, resampling techniques, and noise are addressed. The use of the new EVI2 vegetation index is also explored across both ASTER and Landsat data. The results of these comparisons and their relevance to land change science are used to develop guidelines for further research.

CHAPTER 1 INTRODUCTION

Earth-observing satellites are relied on for their ability to collect longitudinal land change data, allowing researchers to study land use and land cover change over time. The motivation for such study varies widely across diverse scientific fields, but as noted by researchers spanning the human and natural sciences, the methods and applications for monitoring land change have given rise to a new “land change science” (Gutman et al. 2004). Though approaches vary nearly as much as the diverse fields of its research community, one critical component for studying human-environment interactions and land change science is agreed upon: the ability to observe and monitor changes of the Earth’s surface (e.g., Rindfuss et al. 2004). Existing at the nexus of a variety of land change monitoring requirements, the jointly developed platforms and sensors of the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS) Landsat program are heavily cited as the most important land change monitoring tools over the last 35 years (e.g., Cohen and Goward 2004; Woodcock et al. 2001). Among those citations are many reasons for the Landsat program’s success, but most call on the Landsat platforms and their sensor characteristics, the spatial, radiometric, temporal, and spectral resolutions, coupled with the scales at which they operate, and how they match a broad spectrum of land monitoring applications (Jensen 2005).

Much of the NASA Landsat program’s success is founded on uninterrupted access to remotely-sensed, multi-spectral land imaging data since 1972. With the failure of the scan line corrector (SLC) of the ETM+ sensor on Landsat 7 in 2003 (Markham et al. 2004) and subsequent solar array drive problems on the aging Landsat 5 platform in 2005-2006 (United States Geological Survey (USGS) 2006), ongoing longitudinal research utilizing Landsat data seems increasingly threatened. The threat is highly aggravated by the fact that the replacement for both

aging Landsat platforms, slated to be launched with the Landsat Data Continuity Mission (LDCM), is estimated to begin collecting data at the earliest in 2012 (United States Geological Survey (USGS) 2007). Even with efforts to eke utility from Landsat 7 SLC-off data (e.g., Trigg et al. 2006) and prolong the life of Landsat 5 by correcting its spectral and radiometric drift as well as maintain data download coverages, a high likelihood exists of a multi-year “data gap” (United States Geological Survey (USGS) 2008).

In addressing the likelihood and consequences of this data gap, Wulder et al. (2008b) highlighted not only Landsat’s many applications, operating considerations and successes, but also discussed several Landsat alternatives currently available among active, Earth-observing satellites. Among these alternatives is the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), a sensor that orbits on the Terra platform. The Terra satellite was launched as part of NASA’s Earth Observing System (EOS) in 1999 and is a joint effort by the Japanese Ministry of Economy, Trade and Industry and Japan’s Earth Remote Sensing Data Analysis Center (ERSDAC). The ASTER sensor and its platform share many similarities with Landsat. These similarities extend across its spectral and radiometric ranges, temporal availability, and nearly identical orbital characteristics (Table 1-1), that in general make both good tools for land change scientists. However, despite these similarities, important differences do exist. These differences include data access and availability limitations, spatial resolution and geometry configurations, and perhaps most importantly, differences in radiometric and spectral range characteristics.

Data access and availability is a vital characteristic of the Landsat program and its utility for longitudinal studies of land change. The Landsat data archive and its importance for being able to extend land change study back through time has been well explored (e.g., Markham et al.

2004; Vicente-Serrano et al. 2008; Woodcock et al. 2001; Wulder et al. 2008a). Landsat's success in this area lies primarily in its relatively uninterrupted timeline of data that has been collected, downloaded and archived for the past 35 years. For us to consider any gap-filling alternative, its ability to continue that timeline must be closely examined, and few sensor/platform combinations are adequate (United States Geological Survey (USGS) 2007). ASTER is perhaps one of the least appropriate with regards to temporal continuity because of a limited ability to collect and download data.

The ASTER sensor is considered “on-demand,” meaning that scene collection must be tasked by researchers for any observations considered “local,” (including urgent and emergent events). Non-local data collection tasks may fall into the longer-term mission goals of regional monitoring (glacial change, volcanoes, and Long-Term Ecological Research (LTER) field sites) or the one-time coverage, global mapping project (Abrams et al. 2002). But across these priorities, due to memory and download limitations the ASTER system can only acquire about 650 scenes per day (approximately 2.34 million km²), compared to Landsat 7's estimated 300 scenes per day (approximately 9.44 million km²) (National Aeronautics and Space Administration (NASA) 2008). Of those 650, about 150 of ASTER's scenes are typically processed to the more user-friendly, Level 1B and other derived products.

Though not temporally or spatially contiguous to an extent matching Landsat, ASTER data may still be considered a useful tool for ongoing, long-term studies of land change. Two other considerations must be made: 1) the scale of the research question is limited by the logistics of a smaller footprint and spatial scale; and 2) the spectral differences between Landsat and ASTER data should be understood. The first issue of scale is first considered by the nature of the question being asked of the data and the characteristics of the study region. If ASTER is

applicable to the region and question given the data availability issues already discussed, then many areas of active research may be drawn on to inform the logistics of utilizing remotely sensed data across multiple spatial and temporal scales via techniques such as resampling, upscaling and image fusion (Hay et al. 1997; Southworth et al. 2006; Wang et al. 2004; Wu 2004; Zhukov et al. 1999).

This study is focused primarily on the second of these issues, how comparable are Landsat and ASTER data across their overlapping spectral bands? And where systematic departures exist is it possible to empirically transform ASTER data to better match the Landsat data? These issues are most commonly side-stepped if the research question employs only discrete land-cover classifications, utilizing signatures derived from data within individual scenes and within the time series or for each sensor. This approach is common in research dealing with land cover conversion, where describing transitions from one theoretically distinct class to another is the primary goal, even when utilizing multi-scalar or multi-sensor data (e.g., Asner et al. 2005; Capolsini et al. 2003). The issue can also be avoided if land change is being modeled or described using data reduction methods such as multi-date principal components analysis, decision-tree classification or logistic regression.

The issue of spectral comparability becomes paramount, however, when moving beyond simple classification approaches or when conversion from one discrete class to another is less important than degradation, for example. Often land cover classifications are convenient fictions used to simplify a more complex, continuously varying landscape with the hope that it will facilitate our understanding of the processes affecting its change. More often now, these simplistic representations of the landscape are being passed over in favor of measuring change in non-discrete measures of land cover characteristics. These studies may employ multi-spectral

indices, such as the ubiquitous Normalized Difference Vegetation Index (NDVI) or more complex manipulations of multi-spectral data. To compare these continuous measures between time-steps, researchers must rely on careful calibration of data at each time step, otherwise systematic bias is introduced. Calibration, most often occurring through radiometric normalization or by atmospheric correction, is often difficult even among time series involving a single sensor, let alone those involving sensors with different spectral characteristics (Vicente-Serrano et al. 2008).

Though the ASTER visual and near-infrared (VNIR) and shortwave-infrared (SWIR) sensors were designed to be compatible with the Landsat sensors it followed, especially in terms of applied science, bandpass modifications were applied (Figure 1-1). These modifications included changes to the bandpass widths of the VNIR and SWIR1 bands and the separation and shift of the ETM+ B7 band into five separate bands in the ASTER SWIR sensor. These sensor differences leave researchers with distinct spectral incompatibilities between ASTER and Landsat data, differences that preclude direct comparison of measured reflectances due to differing average spectral responses within the overlapping regions.

This study introduces an empirical approach with the goals of describing and correcting systematic differences in spectral measurements between ASTER and Landsat 7 ETM+ data. Though similar to other comparison and cross-validation approaches utilizing multiple sensors and modeled reflectance values (Teillet et al. 2007), this study uses multiple same-day dataset pairs, radiometrically calibrated and atmospherically corrected to give accurate estimates of surface reflectance. Similar approaches were developed to describe differences between ASTER and MODIS data (Gao and Masek 2008; Miura et al. 2008) and for comparing ASTER and Landsat thermal infrared (TIR) measurements (Chen and Zhou 2004). However, comparisons of

ASTER and Landsat ETM+ VNIR/SWIR data using robust statistical techniques have not yet been described in the literature. This study attempts to fill that gap.

The matched, overlapping surface reflectance datasets are statistically compared using a combination of bisector nonparametric, Theil-Sen and Hettmansperger-McKean-Sheather (H-M-S) regression. Within the remote sensing literature this is a novel approach, robust to issues common in studies utilizing remotely sensed data. Through their description and use here the methodology will hopefully be acknowledged for its utility in these types of applications. The regressions are used to describe the association between bands and to transform the ASTER data to more closely match Landsat ETM+ surface reflectances. Vegetation indices, including the newly formulated Enhanced Vegetation Index 2 (EVI2) (Jiang et al. 2008), are calculated from the transformed ASTER datasets and compared with those calculated from the same-day ETM+ data in order to assess whether combining ASTER and ETM+ datasets in non-classification-based, longitudinal studies of land change is viable. Specifically, the research questions to be addressed by this study are

1. How highly associated are ASTER and Landsat data under ideal, same-day conditions?
2. By using coefficients estimated from bisector, nonparametric regression can ASTER data be transformed to more closely match Landsat data?
3. Do derived, continuous datasets from post-transformation ASTER data, such as vegetation indices like NDVI and EVI2, more closely match Landsat-derived indices?
4. What are the implications for using ASTER data, with or without transformations applied, to extend longitudinal studies in conjunction with Landsat data?

Table 1-1. Key sensor and platform characteristics for the ASTER and Landsat ETM+ systems.

	ASTER (Terra platform) ^a	Landsat 7 ETM+ ^b
Swath and Pointing	VNIR and SWIR: 60 km, push-broom TIR: 60km, cross-track scanning ±24° cross-track pointing, VNIR ±8.54° cross-track pointing, SWIR & TIR	185 km, cross-track scanning
Orbit	705 km altitude Sun-synchronous Re-visit: 5 days (see off-nadir, cross-track pointing ability) Cross equator: ~10:15-10:30 AM	Identical orbit, Terra orbits ~15-16 minutes behind Landsat 7 Re-visit: 16 days (nadir only) Cross equator: ~10:00-10:15 AM
Footprint (samples x lines)	~60 km (across-track) x 60 km 4200 x 4100 (L1A VNIR) 2100 x 2048 (L1A SWIR) 700 x 700 (L1A TIR) 4200 x 4980 (L1B VNIR) 2100 x 2490 (L1B SWIR) 700 x 830 (L1B TIR)	~185 km (across-track) by 170 km 6600 x 6000 (VNSWIR) 3300 x 3000 (TIR) 13200 x 12000 (Pan)
Spatial Resolution	15 m (VNIR, bands 1-3N) 30 m (SWIR: bands 4-9) 60 m (TIR: bands 10-14)	30m (VNSWIR: bands 1-5, 7) 60 m (TIR: band 6) 15 m (Pan.: band 8)
Radiometric Resolution	8-bits (VNSWIR) 12-bits (TIR) Uncertainty: < 4% (for VNIR, terrain, gain dependent)	8-bits Uncertainty: < 5%
Geolocation Accuracy	< 0.2 pixels intra-telescope < 15 m relative ±50 m absolute	< 0.17 pixels inter-band < 8 m relative ±250 m absolute

^a(Abrams et al. 2002), ^b(National Aeronautics and Space Administration (NASA) 2008)

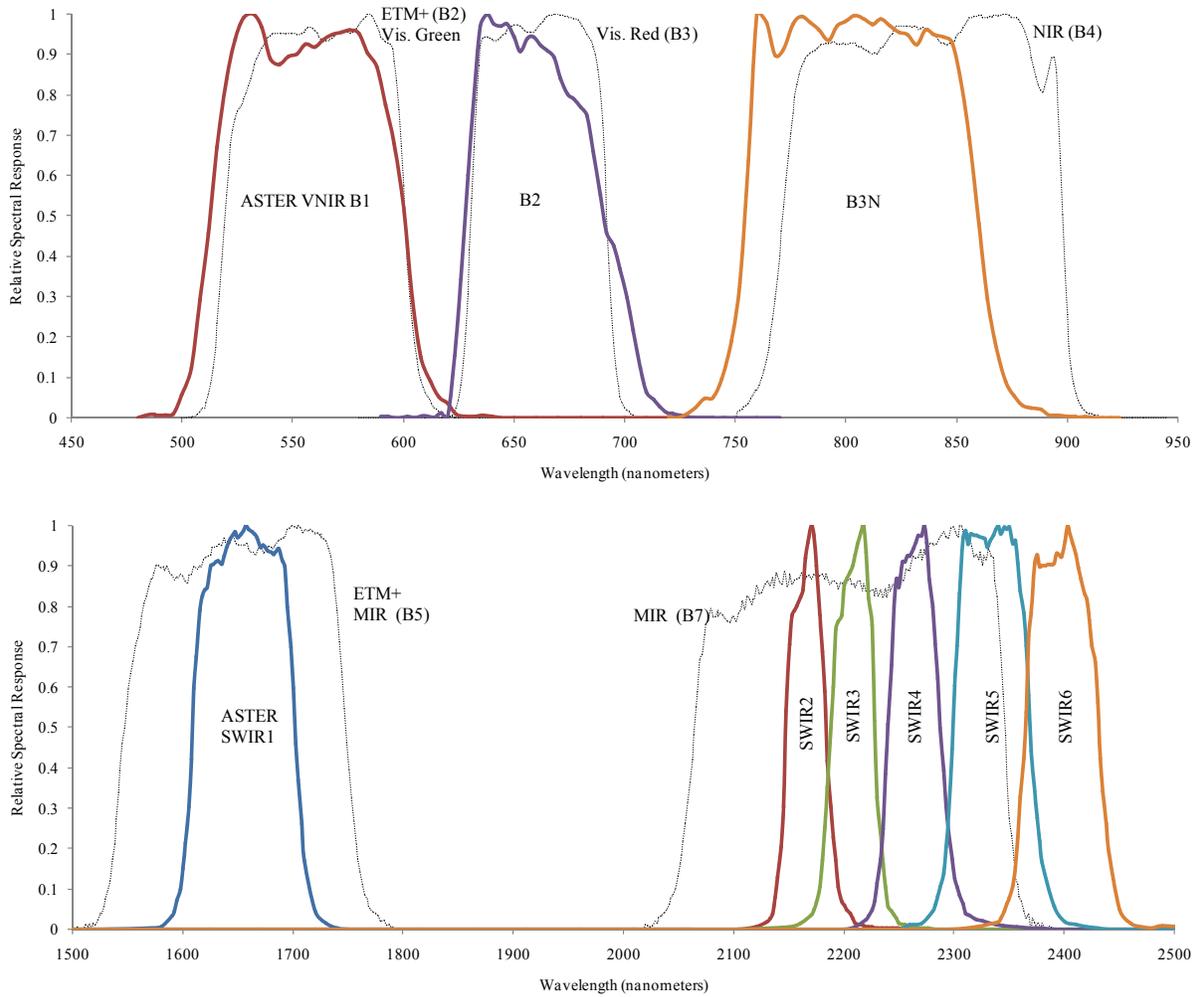


Figure 1-1. System response curve comparison for the EOS ASTER and Landsat ETM+ sensors. Subtle shifts in bandpasses were designed into the ASTER sensor to provide better contrast along the “red edge.” This is visualized by the differences in ASTER B2 and B3N compared to the ETM+ Vis. Red and NIR bands. The width of the overlapping the ASTER SWIR1 bandpass was also greatly reduced. The largest deviation is displayed by the separation of the ETM+ B7 bandpass into the ASTER SWIR2-6 bands.

CHAPTER 2 METHODS

Study Region

The study region represents an area where long-term measurement of land change is critical. The area is located in the southwestern Amazon, near the shared borders of the Peruvian state Madre de Dios, Acre in Brazil, and the Bolivian department of Pando (MAP). The MAP tri-national frontier is still largely forested and has been designated a global biodiversity hotspot (Myers et al. 2000). Though the larger scope of work covers eight Landsat scenes and is concerned primarily with transitions measured by discrete classifications, several smaller subsets are under study for more subtle shifts in forest structure and vegetation dynamics. One of these areas is to the west of the Interoceanic Highway in Madre de Dios that connects the towns of Assis, Brazil in the north, Iñapari, and Tauhamañú, Peru in the south. (Figure 2-1). The northern river in the study area is the Rio Acre, and the southern is the Tauhamañú. Development on the northern border with Brazil, observable in pink in Figure 2-1, are the sister towns of Assis, Brazil and Inapari, Peru. The Peruvian town of Tauhamañú is visible in the SE corner.

The MAP region exhibits mostly humid, warm tropical climatic conditions. Long run averages indicate 1600-2750 mm of rain per year, and mean temperatures of 24-32 degrees C. The dry season ranges from 3-5 months, with the warmest part of the summers occurring between October and November, while the coldest occurs in July (Ministério do Meio-Ambiente Cooperação Brasil-Alemanha 2000).

Rivers have historically played a very important role in the transportation of goods and people throughout the MAP region. They provide resources and fish for food, though recently due to long-term drought conditions many rivers in the area have been severely depleted of fish and other traditionally harvested aquatic resources. But the large Acre and Madre de Dios

watersheds, including the Tauhamanú River visible in the study region subset (Figure 2-1), are still regionally important for transportation and some resource uses.

The geology and soils in the MAP region are a result of the pre-cambrian formation, and then tertiary and quaternary sedimentation from the Andes. The soils within the region, like many in the Amazon, are fairly poor in nutrients. They are also poor draining material and exhibit high flood risks. Furthermore, soil nutrients are typically vegetation structure dependent, and rely on dynamics at or near the surface for nutrient cycling. So agriculture is limited primarily to swidden-type systems, although alluvial plain soils of the major rivers are typically more fertile (Ministério do Meio-Ambiente Cooperação Brasil-Alemanha 2000).

For this study a representative area that included tropical forest, water, and developed areas was needed. The area covered by the overlapping ASTER and ETM+ datasets was chosen because it provided a representative mixture of land cover types and representative reflectance common to the tropical, forested regions of Western Amazonia. This includes both rivers and ongoing development. By ensuring typical land covers are represented, the data samples provide enough dynamic range and variability to sub-samples extracted from the subsets so that statistical tests are not biased.

Data Preparation

ASTER and ETM+ image pairs from the region with multiple year overlaps were selected that met this study's requirements. Between 2000 and 2007, 130 ASTER scenes were identified within 20 km of the MAP region's Interoceanic Highway having less than 10% cloud cover. Of those, six were located that had significant overlap with a single, same-day Landsat footprint. Four of these were chosen as suitable and of these, three were selected for further processing based on scene conditions, overlap and the need to capture representative land cover for the region. For the three scenes both ASTER L1A and L1B (version 003) data were ordered and

downloaded through the USGS Global Visualization Viewer (GLOVIS) web interface (Table 2-1).

For these three ASTER scenes, the corresponding same-day ETM+ scene was acquired for the two years included in this study (Table 2-1). The relatively large overlap area between years (see MY00 and MY01 in Figure 2-1) is an area with relatively undisturbed tropical forest. The overlap areas of the 2001 scenes included both forest and developed areas (MS01).

In order to make valid comparisons between datasets, all data were pre-processed, corrected for sensor calibration and geometry effects, converted to common units and co-registered as accurately as possible. Without doing so, direct pixel-to-pixel comparisons are invalid and any empirical relationships between ETM+ and ASTER data would be inapplicable across other datasets. The general outline of the workflow for each data set is presented (Figure 2-2). ETM+ data were pre-processed to the same units and derivations by using the ENVI 4.3 software's calibration features to convert the raw digital numbers to at-sensor radiance and top-of-atmosphere reflectance (ITT Visual Information Solutions 2007). There are a lack of resources that describe ASTER data preparation in detail, therefore, the methodology for ASTER processing is more well-documented here in the hope that this study is more useful to the research community.

The ASTER HDF files were downloaded with their metadata and opened directly in ENVI. By processing the HDF file in ENVI, the individual bands are extracted from the single HDF file into logical groups (visual/near-infrared (VNIR B1-3N), backward-looking near-infrared (B3N), short-wave infrared (SWIR1-6), and thermal-infrared (TIR B1-5). After extraction, the ENVI software automatically converts each ASTER band to at-sensor radiance (ASRAD, measured in

Watts($\text{m}^{-2}\text{sr}^{-1}\mu\text{m}^{-1}$)) by reading the gain and bias coefficient for each band's wavelength and version from the HDF metadata file (if available).

ASTER L1A data are downloaded as the raw DN that were measured and captured by the sensor. The metadata included with each HDF file contain the calibration coefficients and georeferencing grid that are required to convert these DN into at-sensor radiance and orient the scene. ENVI applies the radiometric conversion automatically upon loading, however, georeferencing must be completed on the VNIR, SWIR and TIR bands separately. The ENVI ASTER georeferencing tool was run with the VNIR and SWIR bands using triangulation and 1500 gridded warp points. These settings provided the best correspondence with the ERSDAC processed L1B data. During georeferencing no resampling occurs because each pixel location is given exact geographic coordinates which are saved in separate layers of the same dimension as the raw data. The georeferenced L1A data were then layerstacked, during which the 30 m pixels of the SWIR bands were subsampled to match the 15 m pixels of the VNIR data. During the layerstacking process, the data is resampled, gridded and transformed to match the georeferenced coordinates. Resampling during this stage was completed once using the nearest-neighbor (NN) and separately using the cubic convolution (CC) method. The two separate methods were used in order to explore both resampling techniques and their effects on the empirical relationships with ETM+ data.

The ASTER L1B data, having been already georeferenced to specifications and accuracies provided by ERSDAC (Abrams et al. 2002), were then layerstacked, combining the VNIR and SWIR bands (B1-3N, SWIR1-6) into a single file. During layerstacking the bands are resampled using the nearest neighbor (NN) process, as the approximately negative eight degree, along-path scene rotation is converted to north-up, and the 30 m SWIR pixels are subsampled to match the

15 m VNIR pixels. It should be noted that L1B data are by-default already resampled during processing at ERSDAC using CC, but a special request may be made during ordering to resample the data using NN. For this study's purposes the NN processing was not requested, since the L1A data should provide a NN comparison for empirical testing, and the L1B default settings are likely to be the most common in data used by the broader research community.

The ASTER L1A and L1B data, after georeferencing and layerstacking, ideally overlay with almost perfect accuracy (confirmed by layerstacking bands from each and flickering or compositing them). Minor differences and pixel shifts occur due to resampling technique, particularly with NN resampling, but since pixel sizes are identical and the geometric grid from which the data were georeferenced are the same, the layerstacks from the two different processing levels should need no further registration. Along-track scenes from each year were mosaicked using a simple cutline across overlapping areas.

Level 1A and 1B ASTER data's georeferencing information is not orthorectified and has a ± 50 m absolute geolocation accuracy (Abrams et al. 2002). Since ETM+ L1G data has less accurate absolute positional accuracy (Table 1-1) and a coarser resolution across the VNIR bands, image-to-image registration was performed using the ASTER L1A data as a base, reference image. The ASTER L1A, L1B and ETM+ data were roughly subset to the areas represented by the MY00 and MS01 boundaries in Figure 2-1. These subsets were imported to ERDAS Imagine 9.1 (ERDAS 2006) and its AutoSync, automatic tie-point generating tool. By using the ASTER L1A data as a reference more than 200 tie points were automatically generated for each subset paired with its corresponding ETM+ data. However, automated image registration is far from error free, as large numbers of automatically generated tie points with an accumulation of small errors can still generate very low root mean square error (RMSE) values,

and yield questionable image registrations. Careful documentation and experimentation with tie point generation was conducted, yielding pixel-to-pixel RMSE values of less than 0.2 pixels and upon visual verification extremely good overlay accuracies. Also during the registration output stage, the ETM+ data were resampled using NN to 15 m pixel sizes to match the ASTER layerstacks.

The correctly georeferenced and registered data were then imported into ENVI for all subsequent processing and analyses. The ASTER at-sensor radiance (ASRAD) was converted to top-of-atmosphere reflectance (TOAR, also known as at-sensor reflectance, apparent reflectance, etc.) using a custom IDL script (Appendix A). This script implements the algorithm and coefficients for spectral solar irradiance as estimated by the World Radiation Center (Yarbrough et al. 2005) and outputs the top-of-atmosphere reflectance (TOAR), a unitless ratio measure.

Accounting for the effects of the atmosphere and solar irradiance differences on the solar radiation transmission is critical when looking at multi-date imagery (Jensen 2005). Since we are using same-date imagery to derive empirical relationships between ETM+ and ASTER data it is tempting to forego atmospheric correction. But since the goal is to employ these empirically-derived transformations across multiple dates empirical relationships must be based on reflectance measurements that are free from bias.

By using same-date imagery such bias is minimized, however the slight difference in acquisition time (14-30 minutes), sun angle and sensor geometry (ASTER pointing angles) may still introduce error. Simple methods of correcting for atmospheric effects, such as dark/light-object subtraction and scaling may be acceptable for image-to-image comparison across dates (Chavez 1996), however to derive functional relationships between datasets for application

across multiple dates these subtle differences should be corrected for. Therefore, more accurate estimates of at-surface reflectance are ideal.

Many cross-calibration and efforts are aimed at deriving such estimates using first-principles radiometric derivations and modeled sensor characteristics (Liu et al. 2004; Teillet et al. 2007). Though these approaches are vital for characterizing theoretical sensor responses and radiometric drift over time, they often focus on ideal surfaces and estimated spectral response curves for simulation purposes rather than on applications for real-world data. To create accurate estimates of surface reflectance the at-sensor-radiance subsets for ASTER L1A, L1B and ETM+ were corrected using the ENVI Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) module. It uses the MODTRAN4 radiative transfer code and is a sophisticated model derived from radiometric first principles (Berk et al. 1999; Cooley et al. 2002). The details for its use and settings are supplied in Appendix B.

During the comparison of each ASTER VNIR and SWIR band, it was apparent that no one band in the ASTER SWIR was equivalent to the ETM+ B7. From a remote sensing perspective this is highly troublesome as many optical indices rely on B7, especially those correcting for the reflectance of soil, fire scars, etc. As an alternative to choosing a single SWIR band to use as a proxy for B7, the ASTER SWIR2-6 bands were spectrally resampled using ENVI to match the B7 bandpass. The process, very common when processing hyperspectral data, assumes critical sampling and uses a Gaussian model to combine multiple bands into one estimate. The software calculates contributions to the resampled band output from the multiple inputs based on the full width at half maximum (FWHM) wavelength characteristics for each band. The resulting output file for each ASTER input was a six band dataset, with null values for the visible blue band (B1),

data identical to ASTER inputs for B2-5, and the resampled SWIR2-6 bands combined as the equivalent to B7.

Data Analysis

The layerstacks for each of the overlaying datasets, both ASTER and ETM+, were compared using a variety of statistical methods and sampling techniques. For each subset, MS01, MY00 and MY01, the ASTER L1A resampled with both NN and CC, were compared with ASTER L1B and ETM+ data. Bivariate comparisons were conducted by taking a random sample of 10,000 pixel locations from each subset and extracting the data from each pixel location across all of the layers of both TOAR and FLAASH reflectance data. These samples represented a small fraction of the total number of pixels in each subset ($2434 \times 1995 = 4,855,830$ for MS01, $2000 \times 2967 = 5,934,000$ for MY00 and MY01), however it was a large enough sample to capture most of the variability from each scene and minimize the effects of spatial autocorrelation. Spatial autocorrelation in regression errors and analyses is a concern because of the scalar mismatch between ASTER and Landsat VNIR bands, as well as the bias present, however minimal, by remaining registration errors.

The primary diagnostic and transformation test used to describe the relationship between ASTER and ETM+ data was the estimated slope and intercept coefficients from statistical regressions. Though it is common to see ordinary least squares (OLS) regression used in remote sensing literature for describing associations between remotely-sensed data, derived products and the biophysical characteristics relating to them, OLS use for this type of data regularly violates several assumptions required for the unbiased estimates of regression coefficients.

As summarized by Fernandes and Leblanc (2005) the most common violation occurs when there are measurement errors in both the “independent” and “dependent” data. They suggest that parametric estimators used in remote sensing are often biased because of inherently non-normal

and heteroscedastic data, censoring (occurs commonly in remote sensing applications where detectors are saturated beyond measurement capability), as well as common occurrence of outliers. Though many of these issues can be corrected for by truncation, scaling and sub-sampling, a more systematic approach is preferable. To counteract these biases when functional and structural regression situations apply they, like others have successfully employed the non-parametric Theil-Sen (T-S) estimator for the slope (Akritas et al. 1995; Fernandes and Leblanc 2005; Olthof et al. 2005).

Theil-Sen regression is a nonparametric regression methodology that uses the median of all slopes of unique, pair-wise observations taken from the dependent and independent datasets as the best estimate for the population slope, $\hat{\beta}$. After the median slope is estimated, the intercept is estimated for each pair of dependent and independent observations by calculating the remainder a from the standard linear relationship, $y = \hat{\beta}x + a$. The median of all calculated a 's is the Hettmansperger-McKean-Sheather (H-M-S) best-estimate for the regression intercept, or $\hat{\alpha}$ (Hollander and Wolfe 1999). T-S regression is highly robust to up to ~29% outliers, nonlinearities, and heteroscedasticity (Fernandes and Leblanc 2005). The assumptions for this non-parametric estimator state:

- A1. At each of n fixed values, x_1, \dots, x_n , of the independent (predictor) variable x , we observe the value of the response random variable Y . The values of x are assumed to be distinct and $x_1 < x_2 < \dots < x_n$. The straight line model is:

$$Y_i = \alpha + \beta x_i + e_i, i = 1 \dots n \tag{2-1}$$

Furthermore the x 's are known constants and α (intercept) and β (slope) are unknown parameters.

- A2. The random variables e_1, \dots, e_n are a random sample from a continuous population that has median 0.

Therefore, the assumptions placed on OLS are mostly relaxed with the only restriction placed on the distribution of the errors being that they have a median of zero.

Concerns regarding OLS' applicability and other parametric regression estimators apply to this study because it explores the functional relationship between two data sources where *both* have measurement error with heteroscedasticity and outliers. The issue of measurement error needs to be considered separately. The "dependent" or truth that is being measured by each set of data is in reality the true surface reflectance at each pixel's location on the ground. This is being measured, with various types of bias and error by *both* the ASTER and ETM+ data. It is not that we are predicting ETM+ data, but instead we are trying to match the error distribution across the range of ETM+ values present. This yields the expectation, which makes sense intuitively as well as statistically, that there be a 1:1 correspondence with errors distributed symmetrically (and not strictly homoscedastically).

Therefore, though a transformation for one data source is sought, neither data source is considered "independent." Errors in measurement for both ASTER and ETM+ data are present due to radiometric and sensor inefficiencies, calibration and registration errors, resampling technique employed, among others. Much research has relied on OLS or a simple Theil-Sen regression in the past with at least partial success, but for this study it is not clear whether we should be minimizing errors in the "dependent" variable (y axis) for each value along the independent (x axis) or the other way around.

So using a regression approach that treats the data symmetrically is preferred. Research in other fields, including astronomic (Feigelson and Babu 1992; Isobe et al. 1990) and ecological literature (Warton et al. 2006), suggests that the OLS methodology of minimizing RMSE or errors in the dependent variable (y axis) for each value along the independent (x axis) is flawed if

both have measurement error or the relationship between “dependent” (predicted) and “independent” (predictor) is unclear. In these instances OLS bisector regression, as described and calculated by Isobe et al. (1990) shows the highest degree of robustness and is stable under more conditions than other symmetric techniques. These techniques include reduced major axis and orthogonal regression, however both suffer from limitations when data has characteristics matching that of remotely-sensed reflectances. Therefore, Isobe et al. and others, recommend the OLS bisector estimated slope coefficient. It is essentially the angular average of the two $\hat{\beta}_1 = \text{OLS}(Y|X)$ and $\hat{\beta}_2 = \text{OLS}(X|Y)$ slopes:

$$\hat{\beta}_{bisector} = (\hat{\beta}_1 + \hat{\beta}_2)^{-1} [\hat{\beta}_1 \hat{\beta}_2 - 1 + \sqrt{(1 + \hat{\beta}_1)(1 + \hat{\beta}_2)}] \quad (2-2)$$

With the bisector intercept estimated as the following:

$$\hat{\alpha}_{bisector} = \bar{y} - \hat{\beta}_{bisector}(\bar{x}) \quad (2-3)$$

However, because it relies on OLS regression, all of the problems experienced with regards to violated assumptions inherent to the nature of remotely-sensed data still apply. This is particularly true with regards to the estimates of variance associated with the estimator. For this reason this study has used an approach unused in the remote sensing literature by combining Theil-Sen slope estimation with the OLS bisector approach. By replacing the estimates of $\text{OLS}(Y|X)$ and $\text{OLS}(X|Y)$ with Theil-Sen slope estimates, and estimating the K-M-S bisector intercept using the Theil-Sen bisector slope and using the median instead of means for the calculation of the intercept.

Processing time is a consideration when calculating Theil-Sen slopes and sub-sampling or resampling is necessary for large datasets. Since all pair-wise combinations of observations from both independent and dependent datasets must be calculated for the 10,000 pairs of observations 4,995,000 unique slopes must be calculated. Due to memory allocation limitations, it was

necessary to estimate slopes, intercepts and the standard errors associated with each by implementing a computer-intensive resampling technique. A sub-sample of 200 pixel locations were extracted without replacement from each original subset and Theil-Sen bisector regressions were run for each. This sub-sample was performed 10,000 times. The resampled “sampling distribution” of Theil-Sen and K-M-S coefficients was normally distributed (Kolmogorov-Smirnov tests were employed for normality testing). The means of each sampling distribution converged on the large-sample estimated coefficients, and the calculated standard deviations are presented as estimates of each coefficient’s variability.

For each 10,000 pixel band comparison, the median of the absolute differences, bias, and Wilcoxon signed rank statistic were calculated. The median of absolute differences (MAD) provides a relative measure of how wide the spread is around the 1:1 line for each comparison. A smaller MAD indicates less bivariate error and a “tighter” correspondence to the 1:1 relationship. Bias was calculated between two samples by taking the differences (Y_i (dependent) – x_i independent) for each pair of observations within the 10,000 pixel sample and determining what percentage of those differences fell below zero, and subtracting that from 50%. Negative bias values indicate more than 50% of the differences were below zero. This corresponds to relationships where more observations fall below the 1:1 line than above it. The closer the bias is to 0% the more evenly distributed around the 1:1 line the bivariate relationship is.

The Wilcoxon signed rank test was used as a statistical measure of location and hence is related indirectly to the bias, where standardized values (approximately normally distributed, $\sim N(\mu=0, \delta=1)$) are calculated and correspond to the departure from the large sample approximation of the estimated number of positively signed ranks (Hollander and Wolfe 1999). This is a non-parametric test for a shift in location (difference in median) between two

populations, with the assumption that the calculated differences are distributed symmetrically around a common median.

After estimating transformation coefficients, the MS01 subset data were used to calculate two vegetation indices, the normalized difference vegetation index (NDVI) and Enhanced Vegetation Index 2 (EVI2) (Jiang et al. 2008). NDVI is an extremely common vegetation index that has a highly associated with net primary productivity, leaf area index (LAI), absorbed photosynthetically active radiation and is in general used to evaluate the relative condition of vegetation across a wide spectrum of land cover types (Jensen 2005). The formula for NDVI utilizes the visible red (R) and near-infrared (NIR) measures of estimated surface reflectance.

$$\text{NDVI} = \frac{(\text{NIR}-\text{R})}{(\text{NIR}+\text{R})} \quad (2-4)$$

NDVI, however, has been shown to saturate at even moderate levels of vegetation (Huete et al. 1997) and therefore loses utility for detecting subtle changes in vegetated areas such as the tropical forests in this study's region. Therefore the Enhanced Vegetation Index (EVI) was proposed as an alternative (Liu and Huete 1995). It is calculated by including a gain correction and adjusts NDVI with the visible blue band, thereby saturating at a much higher level of vegetation productivity. However, because ASTER does not include a band in the visible blue part of the spectrum, the normal EVI cannot be calculated. A two-band alternative has been suggested and calibrated with optimal parameters that uses only the red and NIR bands, but maintains a very high degree of correlation with EVI (Jiang et al. 2008).

$$\text{EVI2} = 2.5 \frac{(\text{NIR}-\text{R})}{(\text{NIR} + 2.4\text{R}+1.0)} \quad (2-5)$$

It is critical to note that calibrated indices such as EVI2 must be calculated using estimated surface reflectance values ranging from 0.0 to 1.0. If scaled data are used then the calibration

coefficients no longer correctly account for the relative ranges in the red and NIR bands and will yield outputs outside the range of -1.0 to 1.0.

ASTER data transformed by estimated coefficients from an independent subsample were used to calculate NDVI and EVI2. The same was done for ETM+ data from the corresponding subset and these vegetation indices were compared pre- and post-transformation. The differences between the post-transformed were then used to examine the spatial variability of the performance of the empirical transformation on the ASTER data.

Table 2-1. Data acquired and subset for ASTER and ETM+ comparisons.

Subset	ASTER	Landsat 7 ETM+
MS01		
Granules	AST_L1A_00307292001151125 AST_L1B_00307292001151125 (Processed, ver. 03, 2007-09-11)	L71003068_06820010729_HDF.L1G Path 003, Row 068
Acq. Date/Time	2001-07-29 15:11:25	2001-07-29 14:59:00
Sun Elevation	52.641	45.649
Sun Azimuth	38.642	47.462
Point Angle	2.878°	0.000 (Nadir)
Resampled	NN (L1A), CC (L1A and L1B)	CC
MS01, MY01		
Granules	AST_L1A_00307292001151134 AST_L1B_00307292001151134	
Date/Time	2001-07-29 15:11:34	
Sun Elevation	52.109	
Sun Azimuth	38.286	
Point Angle	2.878°	
Resampled	NN (L1A), CC (L1A and L1B)	
MY00		
Granules	AST_L1A_00307262000151816 AST_L1B_00307262000151816 (Processed, ver. 03, 2007-09-11)	p003r068_7t20000726_z19 Path 003, Row 068
Acq. Date/Time	2000-07-26 15:18:16	2000-07-26 14:37:20
Sun Elevation	52.597	45.598
Sun Azimuth	38.823	46.316
Point Angle	0.022°	0.000 (Nadir)
Resampled	NN (L1A), CC (L1A and L1B)	NN

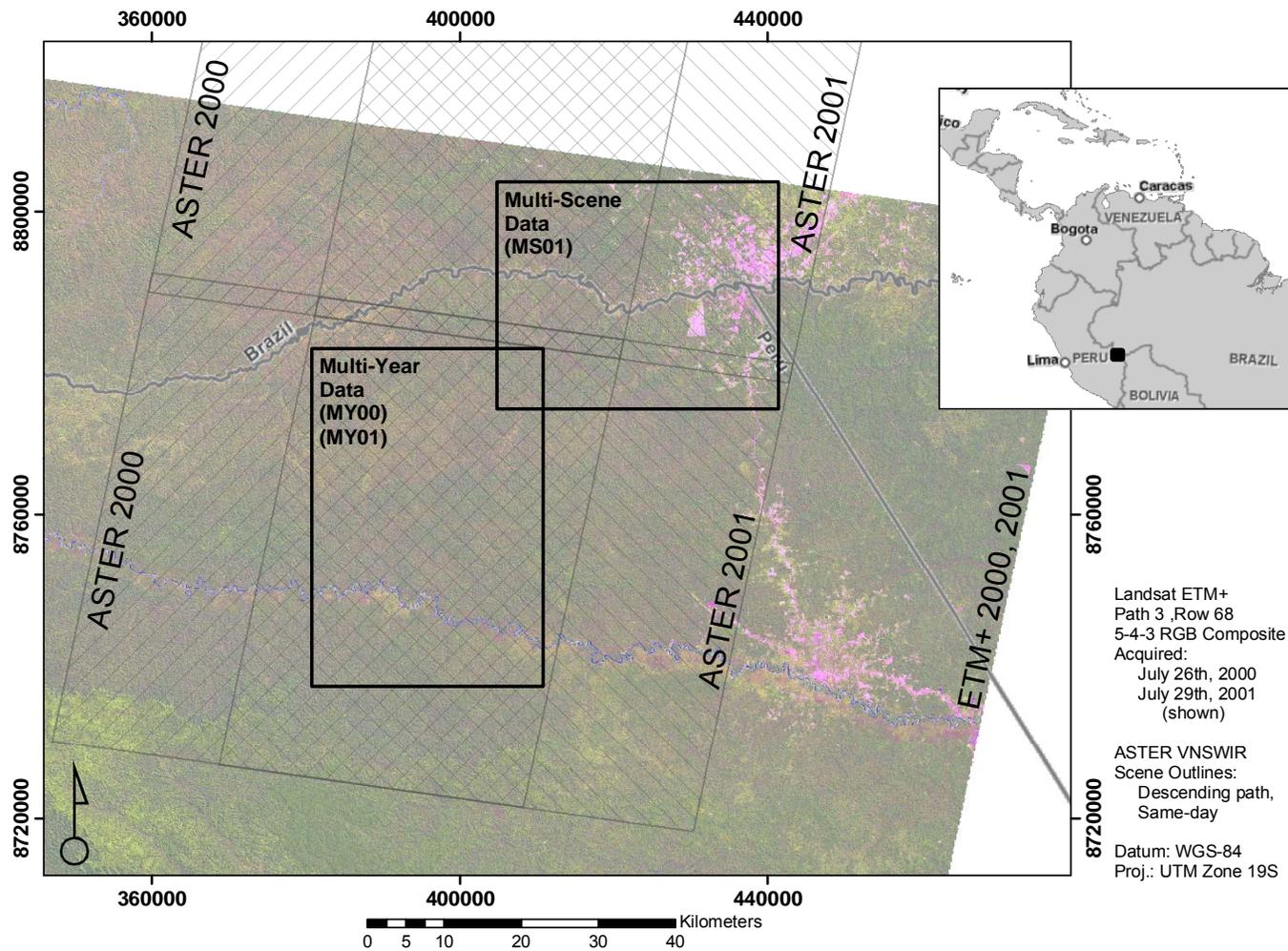


Figure 2-1. Study region map showing the areas subset from the various data sources used. Overlapping regions from two years were acquired and subset into three pairs of ASTER L1A and L1B data: Multi-Scene 2001 (MS01), Multi-Year 2000 (MY00) and Multi-Year 2001 (MY01). There were three corresponding subsets of the ETM+ data, two for the 2001 scene, corresponding to the subsets for MS01 and MY01, and one for 2000, corresponding to the MY00 subset.

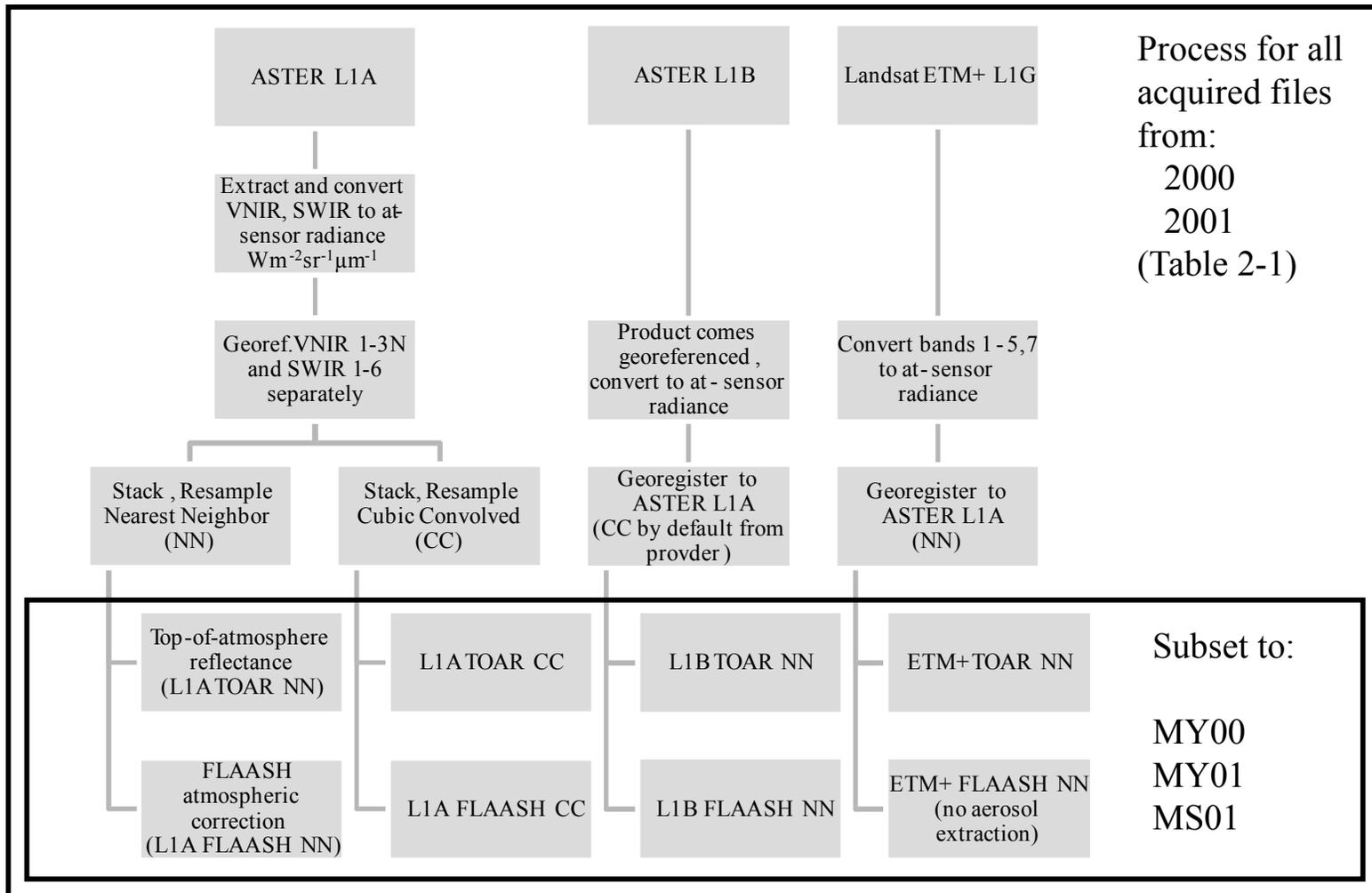


Figure 2-2. Outline of the workflow for each data file type, grouped by year and subset. All initial data extraction, calibration to at-sensor radiance (ASRAD) and top-of-atmosphere reflectance (TOAR) were performed using ENVI. Geo-registration was completed using ERDAS Imagine 9.1 and its AutoSync tool in order to maximize registration accuracy. Finally, all georegistered data were atmospherically corrected using the ENVI FLAASH module yielding estimates for surface reflectance.

CHAPTER 3 RESULTS

Untransformed Comparisons

The results from initial comparisons using Theil-Sen bisector regression across the three subsets (MS01, MY00, MY01) were focused on narrowing down the number of subset and band combinations to use in subsequent analyses. Initial comparisons of top-of-atmosphere reflectance data were highly variable and the bias introduced between years and scenes due to sensor geometry and differential atmospheric effects for varying bandpasses was considered unacceptable and therefore further analyses with TOAR data were ruled out. All subsequent analyses were completed using the FLAASH, atmospherically corrected surface reflectance estimates.

For each of the three subsets there were three sets of ASTER surface reflectance data generated: Level 1A (L1A) nearest-neighbor (NN) resampled data, L1A cubic convolved (CC) resampled data, and Level 1B (by default resampled using CC). For ETM+ there was one set of FLAASH surface reflectance data sampled for each subset from the appropriate year. The L1A CC data were compared with the L1B data with results showing virtually no difference across all of the metrics compared. This confirms that properly aligned and calibrated, L1A CC data are equivalent to L1B, as we should expect, and so L1A CC comparisons are not presented here.

Across all overlapping bandpasses, Theil-Sen bisector and K-M-S bisector slopes were calculated to compare L1A NN, L1B and ETM+ data. The results for the MS01 subset, which is representative of the other two, are displayed in Figures 3-1 and 3-2. As expected for data generated from the same sensor L1A NN has extremely high agreement with L1B data, with slope estimates centered on and confidence intervals bracketing 1.0. Intercept estimates for all of the ASTER band combinations are statistically not different from 0.0. Comparing both ASTER

datasets to ETM+ shows higher agreement among the VNIR bands (slopes closer to one, intercepts closer to zero) than among the comparisons with the SWIR bands. Recall that the ASTER SWIR1 band overlaps with ETM+ B5, while the remaining SWIR2-6 overlap with ETM+ B7, and the comparisons in Figures 3-1 and 3-2 reflect that. The regression estimates show that no one band in the SWIR2-6 range is clearly better associated with ETM+ than any other, so spectrally resampled data were used in all subsequent comparisons.

Also visible in Figures 3-1 and 3-2 is the slightly better agreement between L1B and ETM+ data than for L1A NN data. This slight difference can be attributed to the smoothing effect that cubic convolution resampling has on the ASTER data. However, further analyses were conducted using L1A NN resampled data in order to encompass the greater amount of variability in the original, unsmoothed L1A NN data.

To illustrate the regression estimates as both transformation tool and diagnostic method, a scatter plot presenting ETM+ B7 plotted against the resampled SWIR2-6 surface reflectance estimates from the MY01 subset are displayed in Figure 3-3. The top plot shows the pre-transformation relationship between SWIR2-6 and ETM+ B7. The estimated regression line is plotted as a solid line for the data as shown in each part of the plot. In the first pane untransformed data represents a slope of 2.06 and an intercept of -0.079. As displayed, the pre-transformed ASTER reflectance estimates are scaled at low reflectance to be much higher than the ETM+ estimates (falling below the 1:1 line). After the linear transformation by multiplying the ASTER resampled data by the estimated slope and adding the estimated intercept the plotted relationship falls neatly along the 1:1 line. Error in both of the measured variables produces a symmetric, representative oval shape to the scatter plot. Since the regression was estimated by the Theil-Sen bisector method the regression line is calculated to bisect that oval (as opposed to

non-bisecting regressions that minimize the vertical distance to the regression line in the estimated errors).

Theil-Sen Bisector Regressions

The transformation represented in Figure 3-3 is the result of using the Theil-Sen bisector estimate for the slope and K-M-S bisector estimate for the intercept. These coefficient estimates, calculated for each subset are presented in Table 3-1 and henceforth when transformations are named they are referring to those regression coefficients. The coefficient estimates for each band differ significantly between subsets. This may be the result of differences in land cover between the two subset footprints, as well as the change from year-to-year. However, the range is large enough, especially between slopes in the NIR and B5 vs. SWIR1 comparisons that a common transformation across all datasets may be difficult to estimate in this study.

The effects of each transformation, applied to the ASTER data from which it was derived is presented in Table 3-2. The results are largely expected, as all diagnostic Theil-Sen bisector slopes are not statistically different from 1.0 except for the SWIR2-6 coefficient in the MY00 subset. Similarly, all estimated K-M-S intercepts are not statistically significantly different from 0.0 except for the SWIR2-6 for the MY00 relationship. An important result is the extreme reduction in bias in the post-transformation data across all bands relative to the untransformed bias presented in Table 3-1. This represents the shift of a large percentage of the ASTER data across the 1:1 line, just as illustrated in Figure 3-3, which is exactly what the bi-sector transformation should accomplish.

Furthermore, the relatively small standardized Wilcoxon signed rank scores indicate non-significant differences in the median between the original ETM+ data and transformed ASTER bands. Exceptions to these generally positive transformation results occur in all three subsets' band combinations, indicating that despite large decreases in bias and median differences there

are still slight, unexplained differences in the association. It is important to consider, however, the sample size of 10,000 from which these metrics are calculated is large and even slight deviations in median are detected as significant in rank sum calculations and coefficient estimates.

To determine which transformation to carry forward as a possible candidate for subsequent analyses, each subset was transformed with the coefficients derived from the other subsets (transformations are named in Table 3-1). Figures 3-4 and 3-5 show the comparison of estimated bisector regression coefficients of the MS01 data. The MS01 data were left untransformed, transformed by its derived MS01 coefficients, and also transformed by the independently-derived MY00 transformation. The MY00 transformation is predicted the most likely to be unsatisfactory across multiple datasets, since MY00 land cover is relatively uniform and there is the possibility that bias is introduced between years despite efforts to minimize this. The results show that transforming by its own coefficients provides the most correction. However, the independent transformation by MY00 coefficients does result in significant shift toward a 1:1 relationship, even for coefficients derived from such disparate data.

Transforming ASTER Data and its Outcomes

By examining the MS01 transformation coefficients (Table 3-1) the estimated effects across all bands trend smaller (less scaling and shifting) than in the other two transformations. This is likely due to the generally greater variability in land cover present in the MS01 subset and the larger range of values to estimate. Since this study is in search of a more general functional transformation to allow direct comparison of ASTER and ETM+ in a wide variety of land cover situations, the MS01 coefficients were chosen to conduct further tests with. The transformation effect on the ASTER data was measured using the Theil-Sen bisector regression coefficients, bias, absolute median of differences, and standardized Wilcoxon signed rank scores.

To better visualize the effects of transformation on bands critical to vegetation monitoring (visible red, NIR, MIR (B7)), radar plots were constructed with each of these metrics. Figures 3-6 and 3-7 display the metrics in combination with absolute bias as the marker size (maximum absolute bias of 50% is represented in each as the size of the circle). The visualized transformation effect of MS01 on its own data is the observed change between Tables 3-1 and 3-2 for MS01, pre- and post-transformation. Movement toward the origin in this case represents a better approximation of the 1:1 ideal.

The MY01 data was chosen as an independent test for the MS01 and MY00 transformation coefficients. These two sets of coefficients represent the two ends of the spectrum, with regards to the amount of linear scaling applied during transformation. MS01, with the exception of the visual green band, has slopes closer to 1.0 and intercepts closer to 0.0 than the MY00 transformation. This again, is likely due to the more varied land cover and the increased variability in surface reflectances represented. MY00 on the other hand shows larger amounts of linear scaling and is likely more suited for transforming data representing large tracts of forest, like in the data the coefficients were derived from.

The MY01 data were transformed using both MS01 and MY00 coefficients and the results are presented in Table 3-3. The results show that the MY00 transformation tends to over-correct the data. MY00 decreases bias overall but causes it to be larger in the positive direction than the results from the MS01 transformation. Because of this, MY00 is considered to be an overly aggressive transformation, likely to decrease bias relative to untransformed data, but less likely than MS01 to increase the agreement between ETM+ and ASTER comparisons. To visualize the preferred effects of the MS01 transformation and its ability to reduce errors along the 1:1 relationship, scatterplots were constructed for each MY01 band combination (Figure 3-8). The

1:1 lines represent the ideal relationship and the grey solid line represents the linear transformation applied to the underlying data. To visualize the transformation effect of the MS01 on the MY01, radar plots were again constructed (Figures 3-9 and 3-10).

Transforming ASTER and its Effects on Calculated Vegetation Indices

The MY01 pre- and post-transformed surface reflectances were used to generate NDVI and EVI2 estimates. These were compared to estimates derived from ETM+. The bivariate comparisons of these vegetation indices are presented in Table 3-4. The results show that the estimated slopes and intercepts for NDVI are not significantly different from one and zero respectively, in the untransformed or transformed data. This is likely due to the non-linearity introduced by saturation of ETM+-derived NDVI values. ASTER NDVI saturates less quickly due to a refinement in the visible red and NIR bands that better separates the “red edge” (Miura et al. 2008). But the effects of transformation show an overall trend toward the 1:1 relationship in both sets of post-transformation derived data. But as we observe in Table 3-3, the MS01 transformation performed much better when comparing the derived vegetation indices and their metrics of association. This extends to a Wilcoxon signed-rank result that did not significantly differ from zero for the EVI2 comparison, indicating that the transformation of the visible red and NIR ASTER bands can translate to very closely matching median results over a large sample size.

Scatter plots of the vegetation indices derived from the transformed MY01 (by MS01) are presented in Figure 3-11. The saturated NDVI values are clearly displayed, showing that for this region transformation is unlikely to increase the utility of NDVI measurement in distinguishing vegetation conditions, given the clustering at the extreme high-end of NDVI values. However, the pre- and post-transformation plots for EVI2 show that transformations applied to the visible

red and NIR ASTER bands can effectively increase the correspondence and measured association between derived indices of ASTER and ETM+.

To this point departures from the expected 1:1 association have been described statistically and numerically. However in any remote sensing application error is not likely to be spatially independent. To explore how the errors from the underlying data are translated onto the landscape, post-transformation ASTER EVI2 and ETM+ EVI2 results were compared in a small subset of the MY01 dataset. The transformed ASTER EVI2 data are presented in Figure 3-12. The EVI2 data displayed shows good dynamic range across a small region with high vegetative production, in both forested, wetland and riparian areas.

The two bands used to calculate EVI2 are the visible red and NIR (see Equation 2-5). By looking at the transformation applied by the MS01 coefficients (Table 3-1 and Figure 3-8), we see that the effect should be rather subtle. Recall that the underlying data are estimated surface reflectances and are scaled, unitless values representing the percent of solar radiation reflected from the Earth's surface and take values between 0.0 and 1.0. The effect on the visible red band for the MS01 transformation is summarized by:

$$\text{ASTER}_{\text{Transformed Red}} = -0.00133 + 1.01077(\text{ASTER}_{\text{Untransformed Red}}) \quad (3-1)$$

Because reflectances are very low in the visible red part of the spectrum for this study area (as there is an abundance of highly photosynthetically active vegetation), the negative intercept coefficient shifts the reflectance values towards the origin by over one-tenth of one percent. This transformation is visualized in difference images of the visible red band (Figures 3-13 and 3-14).

What is immediately apparent in the difference image of the untransformed data are visual artifacts across the scene. By looking at the raw, unprocessed original ETM+ visible red band the artifacts can be clearly identified. These types of artifacts in ETM+ imagery (banding,

striping, noise, etc.) are common, especially in tropical imagery where very low dynamic ranges and significant haze is present. The fact that these artifacts are visible across a color-ramped range of 0.01 reflectance units indicates that they may be very important in analyses, especially given the bulk of the data in the visible red region is located between reflectances 0.00 and 0.05 (Figure 3-8). The same difference, calculated using the transformed ASTER data, indicates a trend pulling the differences between the highs and lows of the artifacted region towards zero (Figure 3-14). The artifacts in the ETM+ region clearly occur across the variation contained in its data. By transforming the ASTER visible red distribution towards the median of the ETM+ data the difference image reflects a shift where the highs and lows across the artifacted regions are centered more towards zero than they were in the untransformed data (Figure 3-13).

To observe the effect of the transformation on individual pixels a scatterplot of the ETM+ visible red data vs. transformed ASTER data for the region highlighted in Figures 3-13 and 3-14 was constructed (Figure 3-15). The pixels that changed signs from being positive in the pre-transformed difference image to negative in the post-transformed are highlighted in the scatterplot in magenta. If it were plotted, these pixels would be located on the right of the 1:1 line in the corresponding untransformed scatterplot, but after transformation they shift to the left of the 1:1 line. By identifying these pixels individually, we can plot where on the image these shifts in sign occurred (Figure 3-16). By comparing the location of the change pixels with the location of the color shift in the difference images it is apparent that the effects of transformation are not spatially independent. However, by looking at the difference image of post-transformed ASTER EVI2 image minus the ETM+ EVI2 estimates (Figure 3-17) we can see in general relatively low variability across the image, with most errors in the ± 0.05 EVI2 scale.

Last, we can identify where these pixels reside relative to the 1:1 line in both of the transformed EVI2 and NDVI scatterplots. We can observe directly how though the shifted pixels are spatially autocorrelated in the difference image, they are represented without bias in the EVI2 estimates. Similarly, the changed pixels show slightly below the 1:1 line for the transformed NDVI results. These visual representations of course do not take into consideration the relative density of these pixels, however because the visible red band showed very close to a 1:1 association before transformation, the shifted slice of pixels near the 1:1 line represents a large portion of pixels in the visible red band.

Table 3-1. Transformation coefficients as calculated for the specified year and band combinations.

Transformation Name	ETM+ FLAASH	ASTER L1A, NN FLAASH	K-M-S Bisector Intercept (α)	SD_{α} (δ_{α})	Theil-Sen Bisector Slope (β)	SD_{β} (δ_{β})	Untransformed Bias
MS01*	MS01:	MS01:					
	Vis. Green	B1	-0.00388	0.00033	0.89736	0.01204	-37.95%
	Vis. Red	B2	-0.00133	0.00030	1.01077	0.01828	-7.13%
	NIR	B3N	0.01586	0.00182	0.95870	0.00650	5.55%
	MIR (B5)	SWIR1	-0.07755	0.00171	1.38874	0.01137	-41.17%
	MIR (B7)	SWIR2-6 Resampled	-0.05692	0.00188	1.67401	0.02808	-40.23%
MY00*	MY00:	MY00:					
	Vis. Green	B1	-0.00717	0.00056	0.99106	0.02496	-39.91%
	Vis. Red	B2	-0.00483	0.00042	1.12833	0.03052	-20.44%
	NIR	B3N	-0.01212	0.00228	1.04309	0.00805	-0.63%
	MIR (B5)	SWIR1	-0.12961	0.00283	1.71166	0.01875	-45.88%
	MIR (B7)	SWIR2-6 Resampled	-0.09823	0.00331	2.23753	0.04998	-48.45%
MY01*	MY01:	MY01:					
	Vis. Green	B1	-0.00538	0.00058	0.88481	0.02174	-44.38%
	Vis. Red	B2	-0.00219	0.00042	1.00838	0.02714	-14.70%
	NIR	B3N	0.02991	0.00186	0.90067	0.00679	4.27%
	MIR (B5)	SWIR1	-0.09326	0.00242	1.50478	0.01645	-44.93%
	MIR (B7)	SWIR2-6 Resampled	-0.07878	0.00377	2.06198	0.05963	-44.32%

* The MS01 subset contains far more development and more water and wetland pixels relative to the MY datasets (Figure 2-1). This may account for relative differences in estimated coefficients.

Table 3-2. Transformation results of original data after applying the derived coefficients from Table 2-1.

Transformation Performed	ETM+ FLAASH	ASTER FLAASH	K-M-S Bisector Intercept (α)	SD_{α} (δ_{α})	Theil-Sen Bisector Slope (β)	SD_{β} (δ_{β})	Post-Transf. Bias	Median of Absolute Diff.	Std. Wilcoxon Signed Rank Score (Z)	SR p-value
MS01	MS01: V. Green	MS01: B1	0.00002	0.00028	0.996	0.0134	0.04%	0.00323	-1.2	0.120
	Vis. Red	B2	0.00003	0.00028	1.005	0.0181	0.44%	0.00289	-1.8	0.040
	NIR	B3N	0.00022	0.00192	1.000	0.0068	-0.89%	0.01681	-1.4	0.088
	MIR (B5)	SWIR1	-0.00140	0.00108	1.009	0.0082	-2.96%	0.00769	-4.6	0.000*
	MIR (B7)	SWIR2-6	-0.00137	0.00093	1.027	0.0169	-4.23%	0.00466	-3.8	0.000*
MY00	MY00: V. Green	MY00: B1	-0.00007	0.00038	1.000	0.0252	-0.39%	0.00341	-1.2	0.110
	Vis. Red	B2	-0.00014	0.00029	1.027	0.0271	-0.18%	0.00315	-0.6	0.259
	NIR	B3N	-0.00002	0.00219	0.999	0.0077	-0.51%	0.01764	-1.1	0.125
	MIR (B5)	SWIR1	-0.00225	0.00138	1.019	0.0107	-0.20%	0.00825	-0.1	0.479
	MIR (B7)	SWIR2-6	-0.00240†	0.00105	1.048‡	0.0211	-0.66%	0.00496	-0.5	0.292
MY01	MY01: V. Green	MY01: B1	0.00045	0.00045	0.977	0.0245	0.03%	0.00298	0.0	0.484
	Vis. Red	B2	0.00013	0.00036	0.995	0.0269	0.46%	0.00287	-2.4	0.008*
	NIR	B3N	0.00163	0.00208	0.994	0.0075	-0.30%	0.01593	-1.0	0.155
	MIR (B5)	SWIR1	-0.00159	0.00136	1.012	0.0107	-0.44%	0.00703	-0.8	0.203
	MIR (B7)	SWIR2-6	-0.00146	0.00149	1.029	0.0290	-0.43%	0.00411	-0.6	0.277

* Indicates a p-value for the standardized Wilcoxon signed-rank score below 0.01.

† Indicates that the estimated K-M-S intercept (α) confidence interval calculated as $\pm(1.96)(SD_{\alpha})$ does not contain the expected value of 0.0.

‡ Indicates that the estimated Theil-Sen Slope (β) confidence interval calculated as $\pm(1.96)(SD_{\beta})$ does not contain the expected slope of 1.0.

Table 3-3. Transformation results of data after applying the independently derived coefficients from Table 2-1 to MY01 data.

Transformation Performed	ETM+ FLAASH	ASTER FLAASH	K-M-S Bisector Intercept (α)	SD_{α} (δ_{α})	Theil-Sen Bisector Slope (β)	SD_{β} (δ_{β})	Bias	Median of Absolute Diff.	Std. Wilcoxon Signed Rank Score (Z)	SR p-value
None	MY01: V. Green	MY01: B1	-0.00538†	0.00058	0.885‡	0.0217	-44.38%	0.00840	-84.8	0.000*
	Vis. Red	B2	-0.00219†	0.00042	1.008	0.0271	-14.70%	0.00300	-41.9	0.000*
	NIR	B3N	0.02991†	0.00186	0.901‡	0.0068	4.27%	0.01720	-10.6	0.000*
	MIR (B5)	SWIR1	-0.09326†	0.00242	1.505‡	0.0164	-44.93%	0.01900	-84.7	0.000*
	MIR (B7)	SWIR2-6	-0.07878†	0.00377	2.062‡	0.0596	-44.32%	0.01154	-85.4	0.000*
	MS01	MY01: V. Green	MY01: B1	-0.00112†	0.00048	0.966	0.0242	-14.35%	0.00333	-37.7
Vis. Red		B2	-0.00074	0.00038	0.994	0.0269	-6.24%	0.00296	-17.9	0.000*
NIR		B3N	0.01521†	0.00197	0.938‡	0.0071	-2.81%	0.01644	-7.4	0.000*
MIR (B5)		SWIR1	-0.01070†	0.00147	1.095‡	0.0116	4.29%	0.00694	-11.9	0.000*
MIR (B7)		SWIR2-6	-0.00959†	0.00173	1.251‡	0.0354	16.00%	0.00435	-42.8	0.000*
MY00		MY01: V. Green	MY01: B1	0.00101†	0.00042	0.894‡	0.0219	8.27%	0.00325	-21.3
	Vis. Red	B2	0.00196†	0.00030	0.914‡	0.0241	5.16%	0.00314	-18.6	0.000*
	NIR	B3N	0.03936†	0.00178	0.867‡	0.0065	4.82%	0.01767	-12.0	0.000*
	MIR (B5)	SWIR1	0.01890†	0.00114	0.893‡	0.0093	16.86%	0.00896	-43.7	0.000*
	MIR (B7)	SWIR2-6	0.01042†	0.00116	0.951	0.0268	36.91%	0.00855	-78.7	0.000*

* Indicates a p-value for the standardized Wilcoxon signed-rank score below 0.01.

† Indicates that the estimated K-M-S intercept (α) confidence interval calculated as $\pm(1.96)(SD_{\alpha})$ does not contain the expected value of 0.0.

‡ Indicates that the estimated Theil-Sen Slope (β) confidence interval calculated as $\pm(1.96)(SD_{\beta})$ does not contain the expected slope of 1.0.

Table 3-4. Comparison of derived vegetation indices from post-transformation MY01 ASTER and original ETM+ data.

Transformation Performed	ASTER MY01 Index	K-M-S Bisector Intercept (α)	SD_{α} (δ_{α})	Theil-Sen Bisector Slope (β)	SD_{β} (δ_{β})	Bias	Median of Absolute Diff.	Std. Wilcoxon Signed Rank Score (Z)	SR p-value
None	NDVI	0.04848	0.02807	0.962	0.0314	15.23%	0.02104	-44.9	0.000*
	EVI2	0.06271†	0.00583	0.891‡	0.0118	9.66%	0.02722	-24.4	0.000*
MS01	NDVI	0.03289	0.02904	0.968	0.0322	5.13%	0.01892	-15.7	0.000*
	EVI2	0.03453†	0.00610	0.928‡	0.0121	-0.30%	0.02536	-1.2	0.113
MY00	NDVI	0.10536†	0.02598	0.878‡	0.0285	-4.38%	0.02084	-15.8	0.000*
	EVI2	0.07537†	0.00552	0.852‡	0.0110	3.00%	0.02705	-7.0	0.000*

* Indicates a p-value for the standardized Wilcoxon signed-rank score below 0.01.

† Indicates that the estimated K-M-S intercept (α) confidence interval calculated as $\pm(1.96)(SD_{\alpha})$ does not contain the expected value of 0.0.

‡ Indicates that the estimated Theil-Sen Slope (β) confidence interval calculated as $\pm(1.96)(SD_{\beta})$ does not contain the expected slope of 1.0.

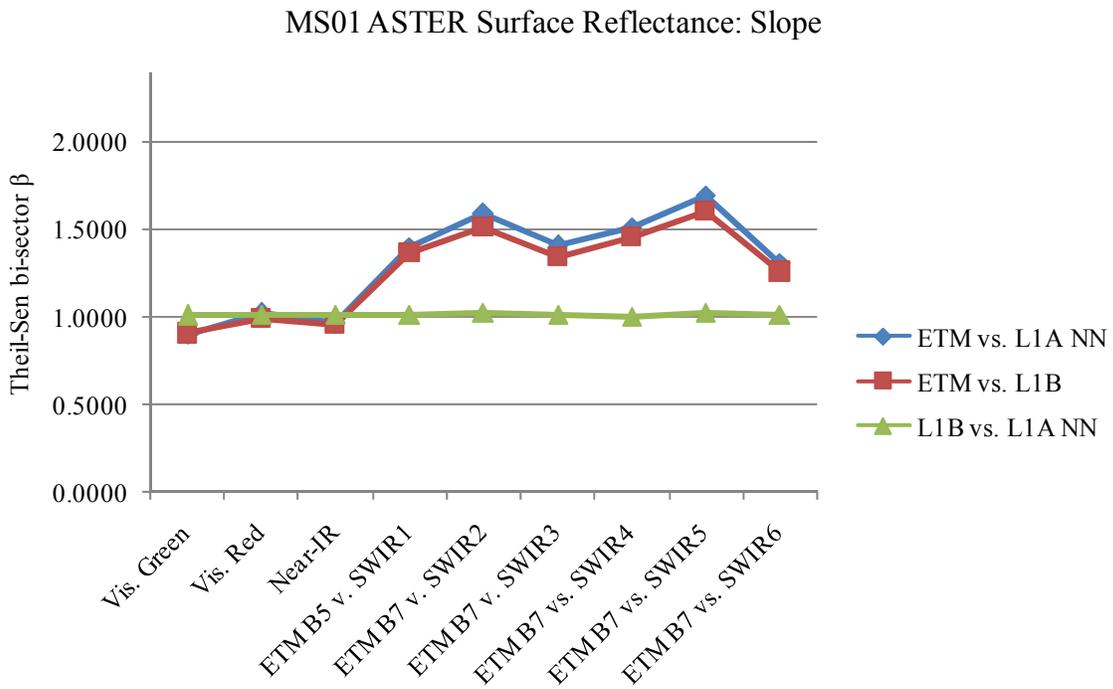


Figure 3-1. Comparison by band of estimated Theil-Sen bisector slopes for ETM+ vs. ASTER L1A and L1B FLAASH surface reflectances derived for the MS01 subset.

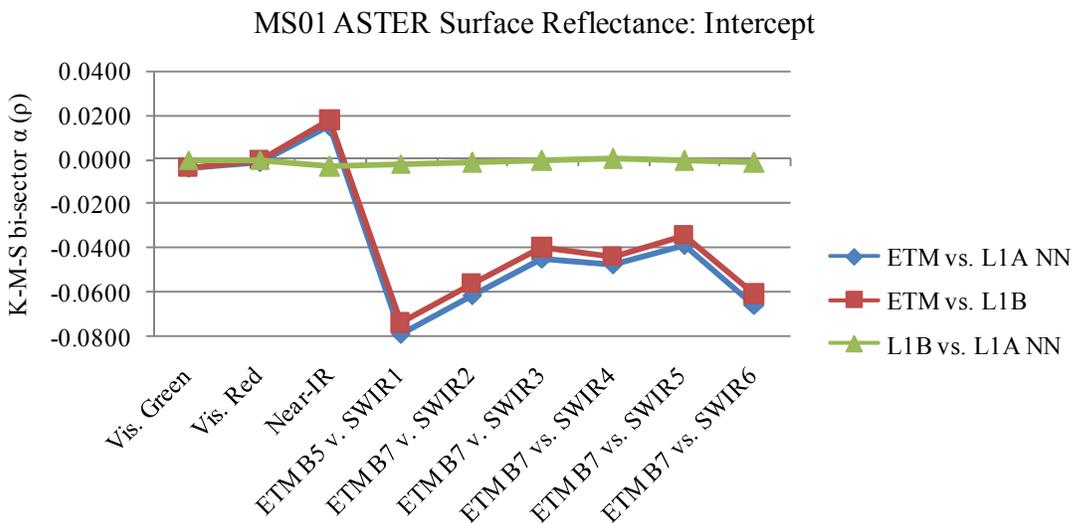
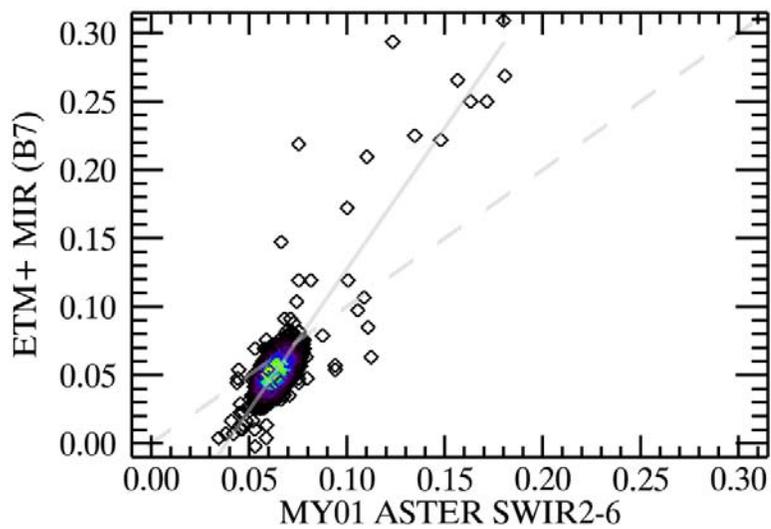


Figure 3-2. The K-M-S bisector intercept estimates for ETM+ vs. ASTER L1A and L1B FLAASH surface reflectances for each band combination, corresponding to the slope estimates from Figure 3-1.

B) Pre-transformed Relationship



A) Post-transformed Relationship

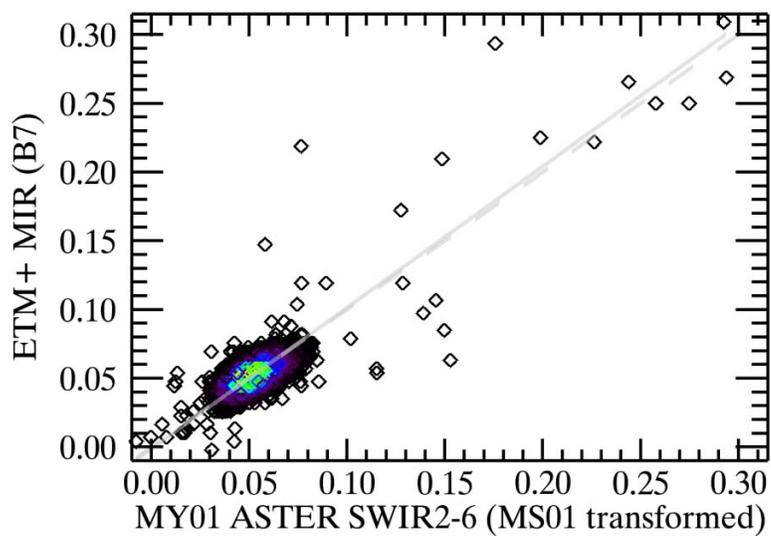


Figure 3-3. A) Illustrating the effects of ETM+ B7 MY01 being regressed on ASTER MY01. Data are presented without transformation but with plotted estimated regression line (solid) compared to the 1:1 line (dashed). B) The post-transformation results for the same data are observed to fall almost completely on the 1:1 line and as expected split the data cloud symmetrically along the medial axes.

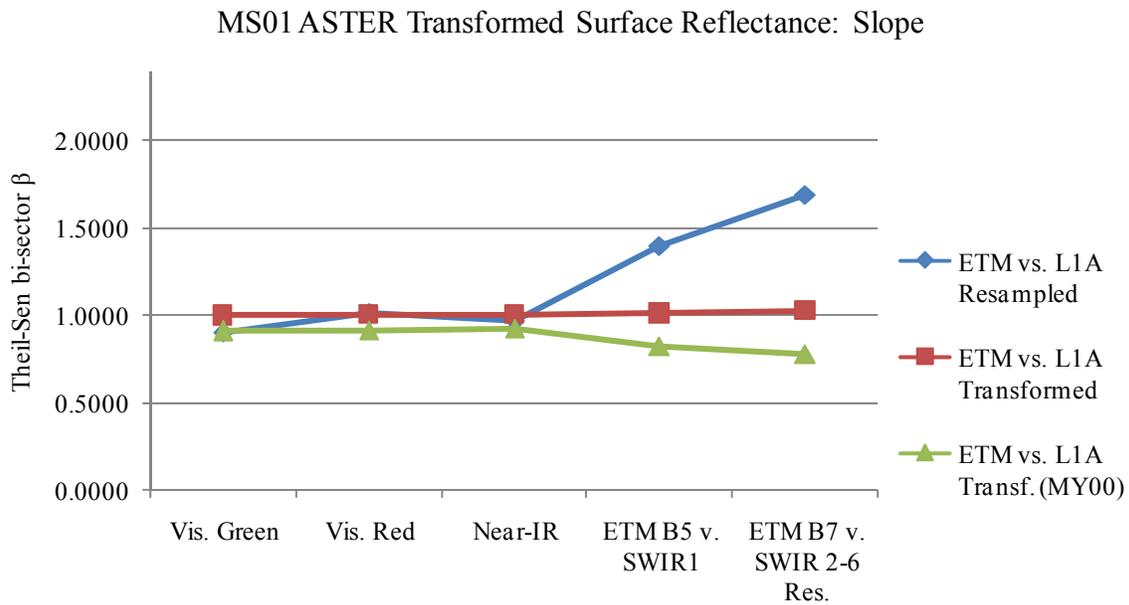


Figure 3-4. Comparison of estimated, post-transformation Theil-Sen bisector slopes for ETM+ vs. ASTER L1A FLAASH MS01 surface reflectances that were spectrally resampled to match ETM+ B7. The data shown compare the effects of the data transformed with its own estimates (MS01) to the effect of transforming by the MY01 estimated coefficients, and with both relative to the untransformed data.

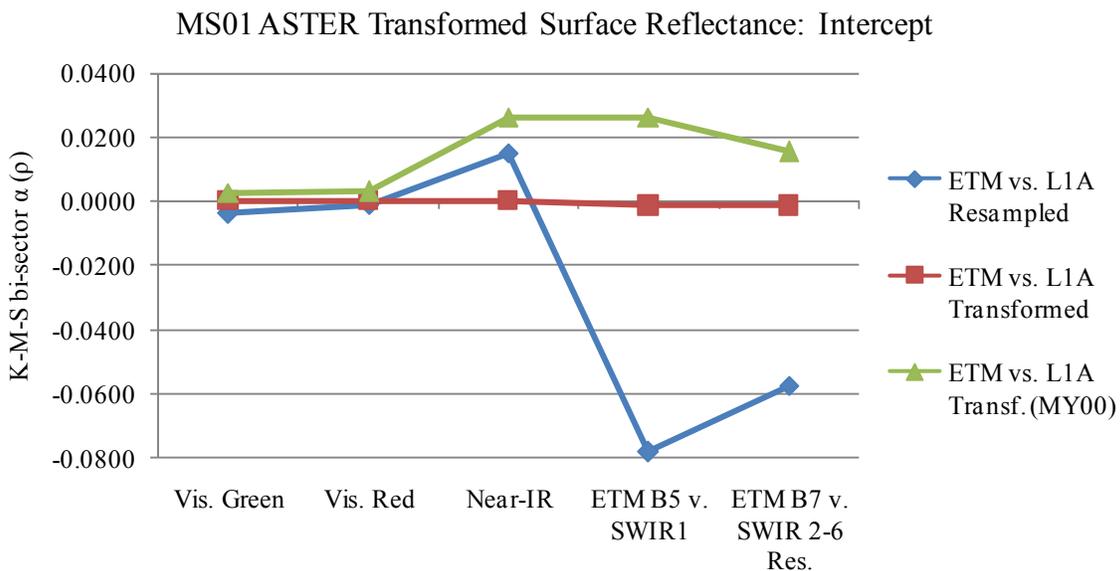


Figure 3-5. Comparison of estimated, post-transformation K-M-S bisector intercepts matching the slopes from Figure 3-3.

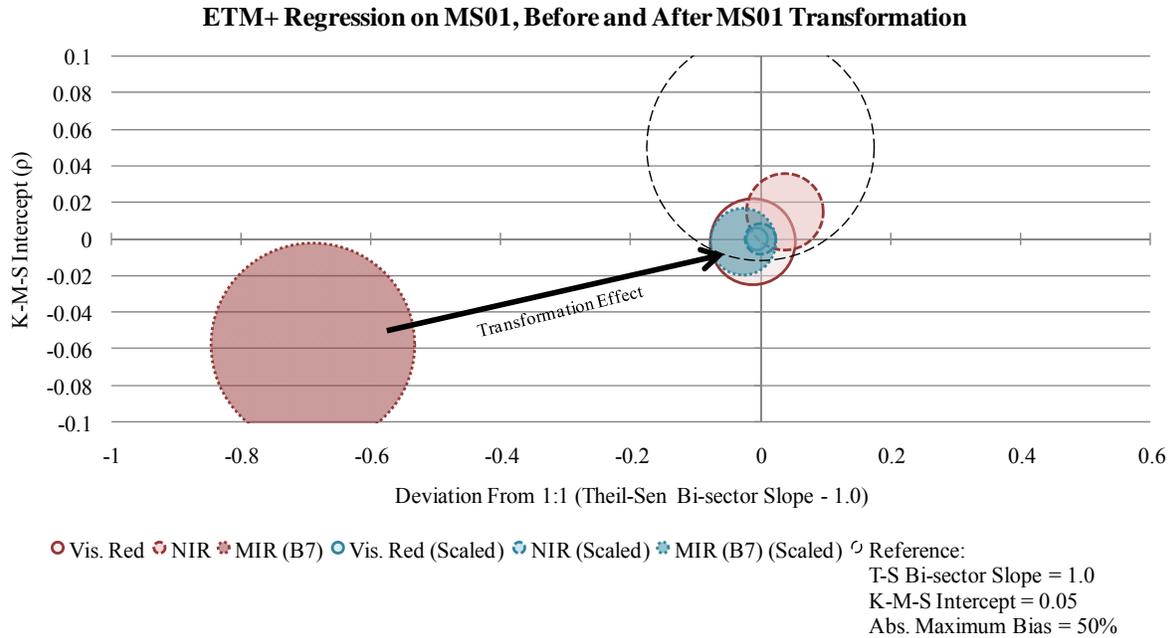


Figure 3-6. Visual representation of transformation effects on estimated regression coefficients for the ETM+ MS01 regressed on ASTER MS01 pre- and post-transformation. Note the transformation effect highlighted by the arrow on the pre-transformation coefficients for ETM+ B7 (red circles) as the slope changes to very close to 1.0 and intercept very close to 0.0 with a much smaller absolute bias (transformed, blue circles).

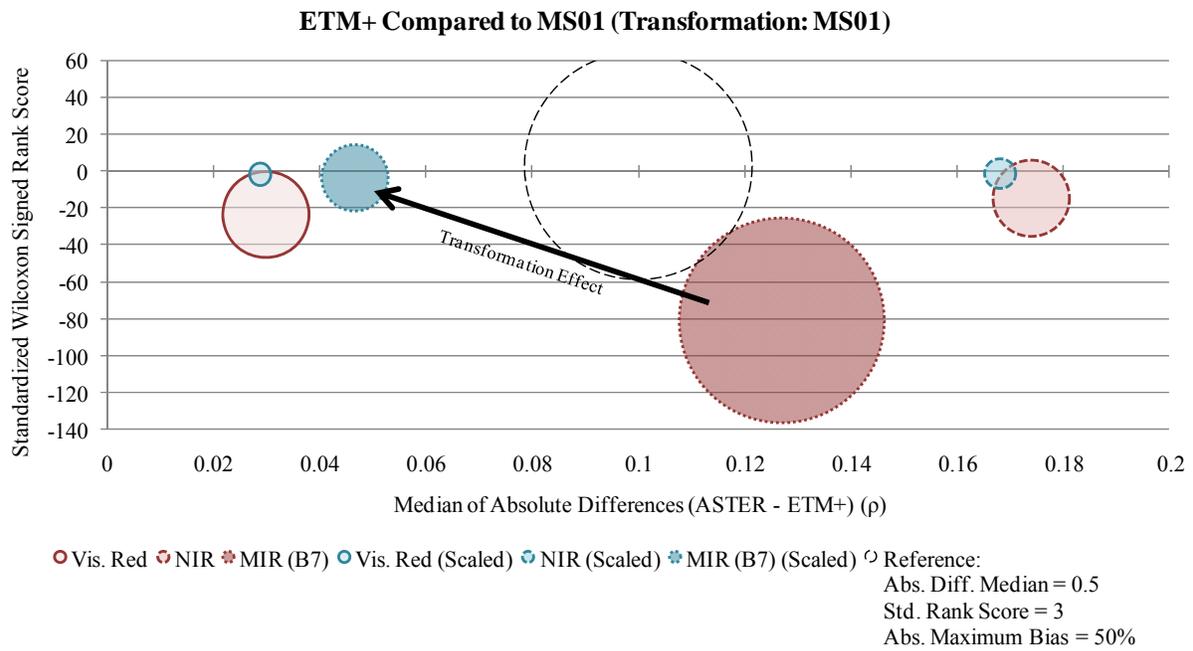


Figure 3-7. Transformation effects are shown as measured by metrics of spread and centrality for pre- and post-transformation of ASTER MS01 by its own MS01 regression coefficients. Standardized Wilcoxon signed rank scores may be interpreted as a general measure of symmetry around the median, or in the case of Theil-Sen bisector transformed data around the 1:1 line.

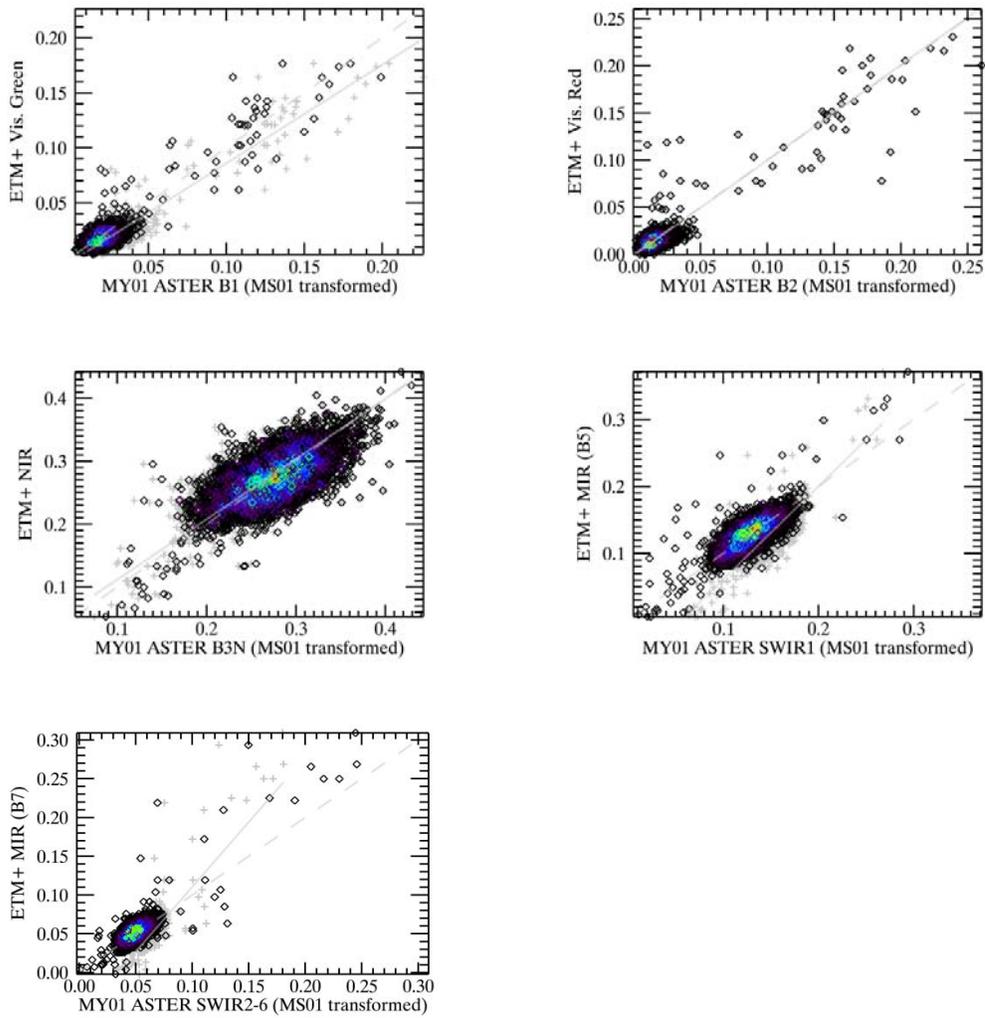


Figure 3-8. Each of the five matching ETM+ MY01 bands and their diagnostic regressions are plotted against the corresponding transformed ASTER MY01 data. The MY01 ASTER data were transformed by estimated coefficients from the MS01 data.

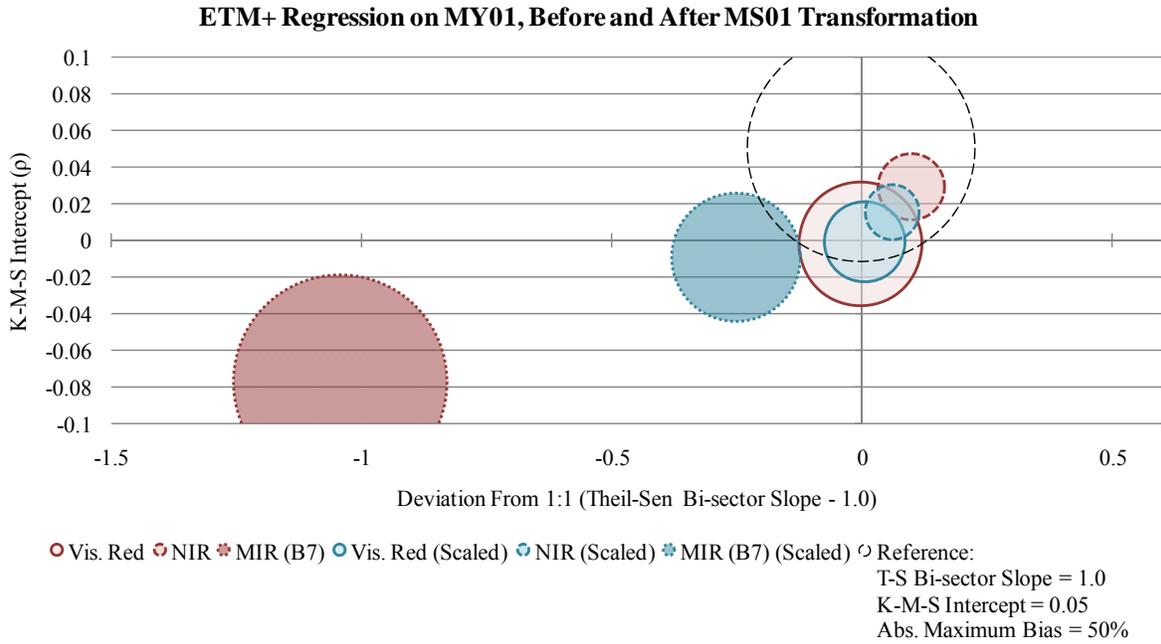


Figure 3-9. Transformation effect as measured by the pre- and post-transformation regression coefficients for ETM+ MY01 regressed on ASTER MY01 bands after transforming with MS01 estimated coefficients.

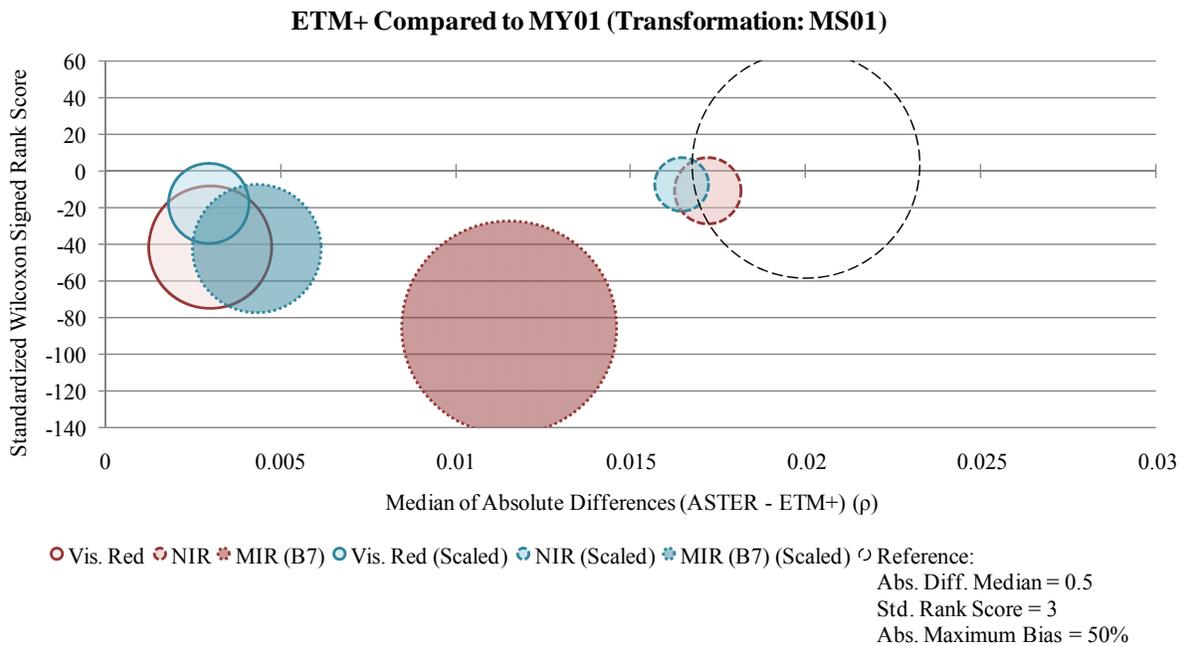


Figure 3-10. Post-transformation effects on MY01 (transformed by MS01) as measured by metrics of spread and centrality.

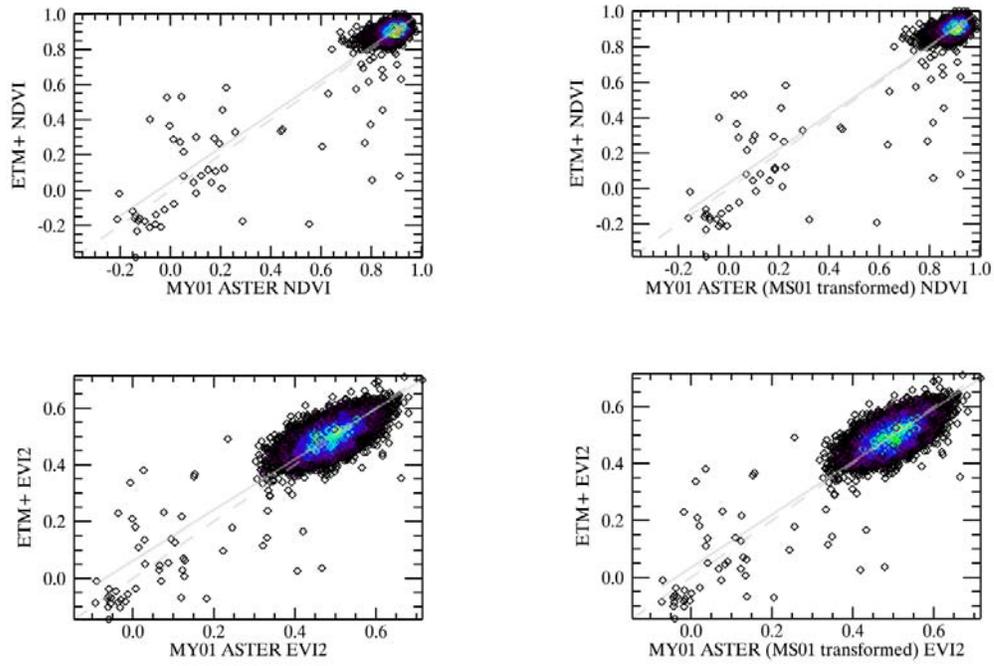


Figure 3-11. Scatterplots and their diagnostic Theil-Sen bisector regressions of ETM+ derived NDVI and EVI2 vegetation indices, against the MY01 (transformed by MS01) derived indices.

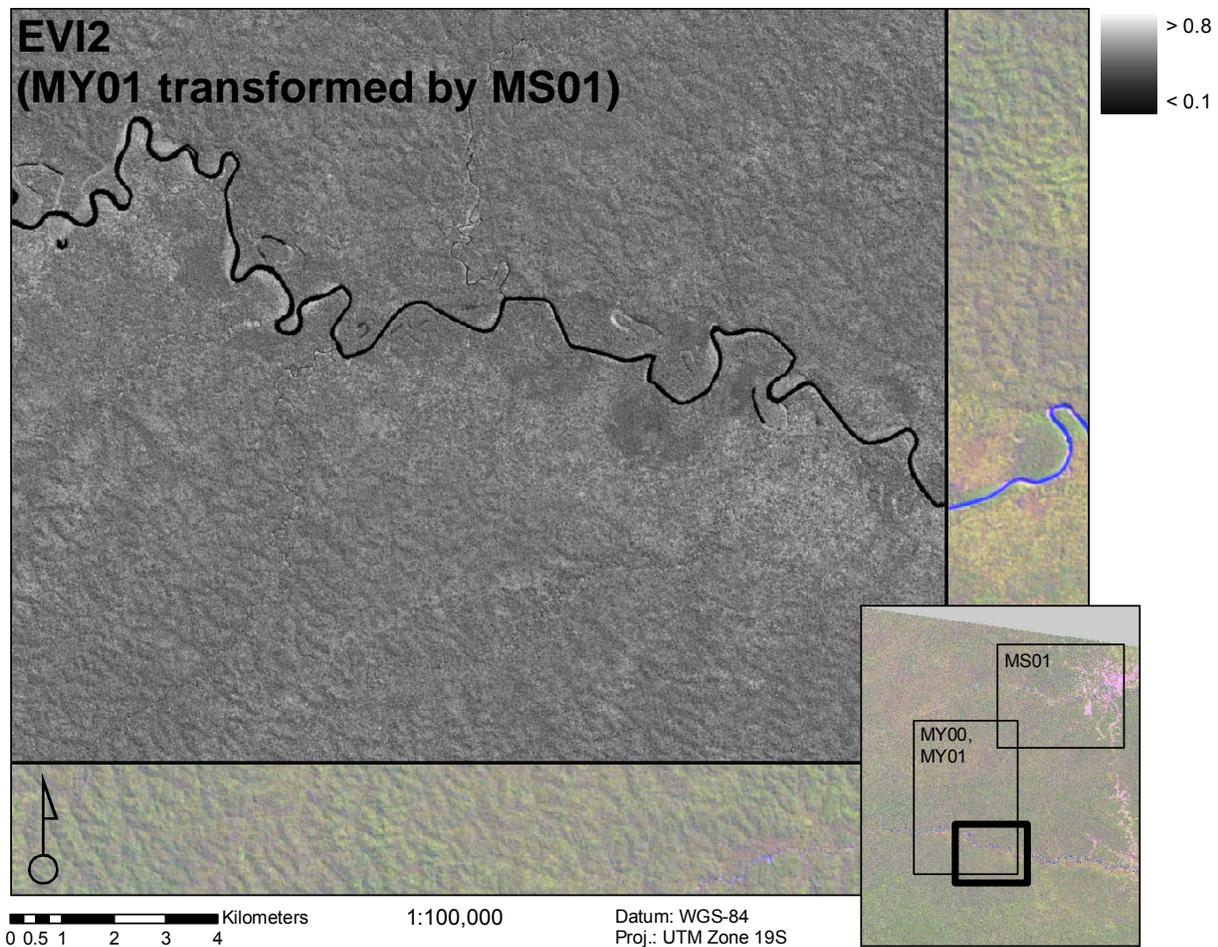


Figure 3-12. The spatial distribution of EVI2 values calculated for the MY01 subset after transformation using estimated Theil-Sen bisector coefficients from the MS01 sample. Notice the lack of saturation and no visible artifacts in the resulting data.

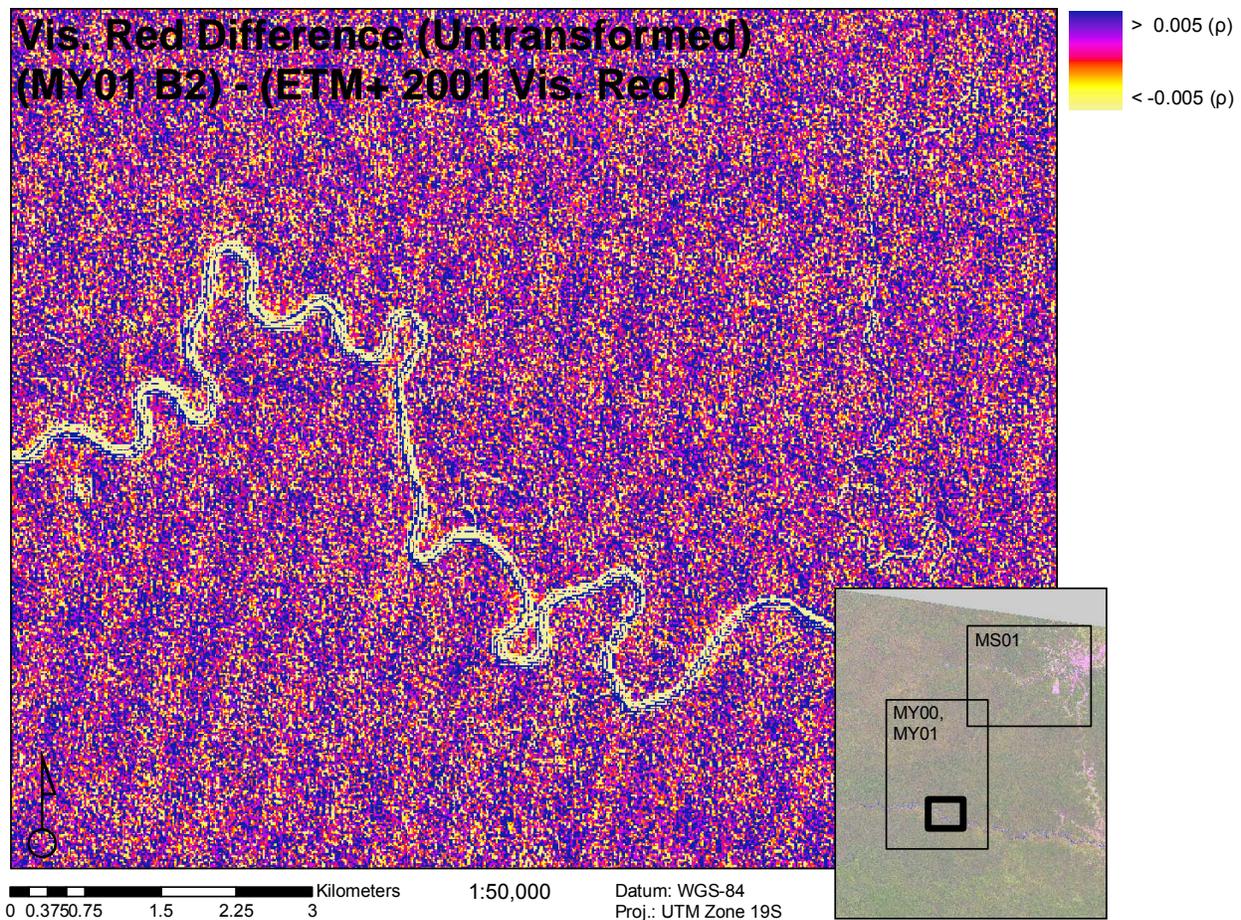


Figure 3-13. This map shows the artifacts created by mapping the difference of visible red bands, calculated by subtracting the ETM+ band from the MY01, untransformed data. Notice the regular banding in roughly a N-NE direction. This is unlikely due to registration errors or other data mis-handling, as the artifacts may be observed in the original, uncalibrated ETM+ B3. This difference image can be represented by the Vis. Red band scatterplots in in Figures 3-8 and 3-15.

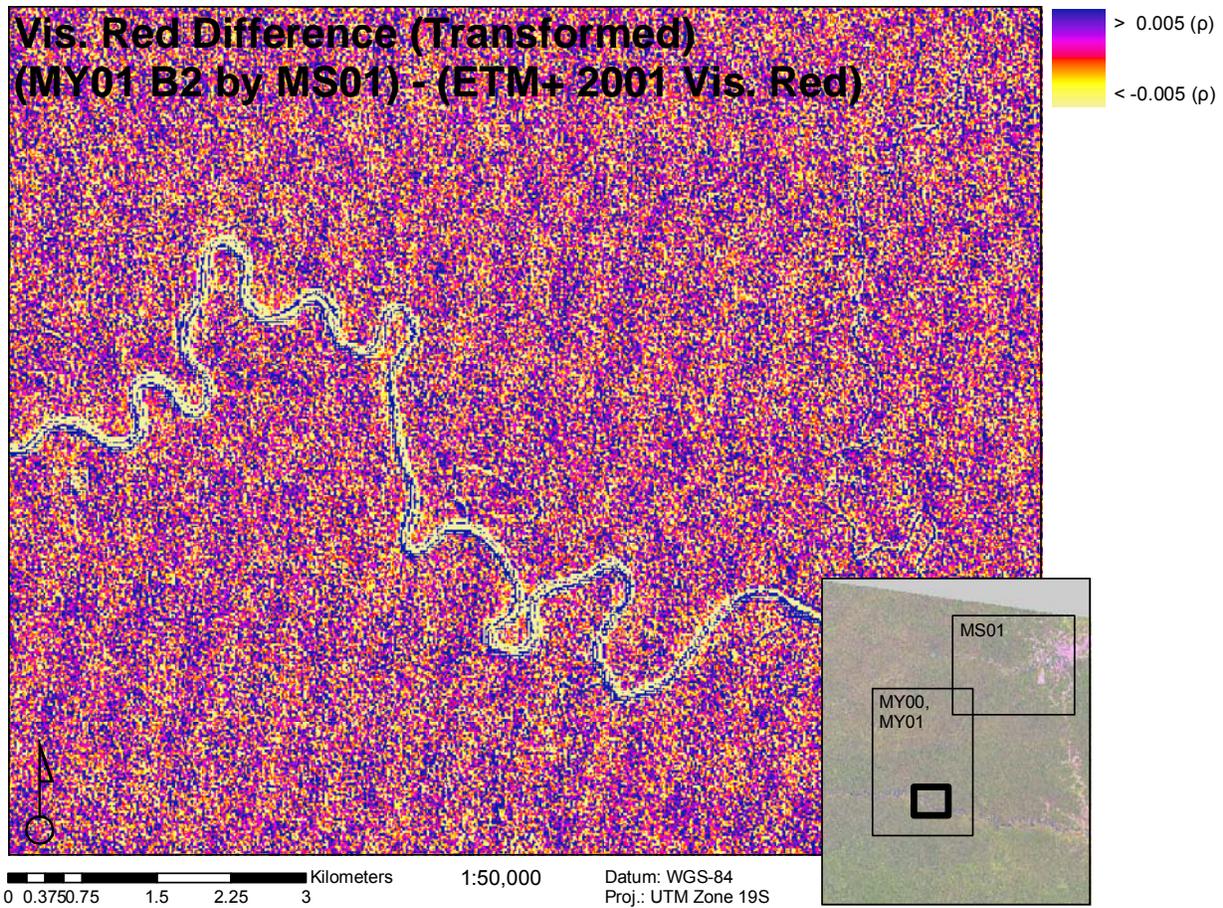


Figure 3-14. The visible artifacts, though still present are clearly located near the center of the error distribution as the differencing of the transformed ASTER MY01 visible red and ETM+ band shows. The overall distribution has shifted markedly toward the red, indicating a greater degree of symmetry around the 1:1 line in the scatterplots of 3-8 and 3-15.

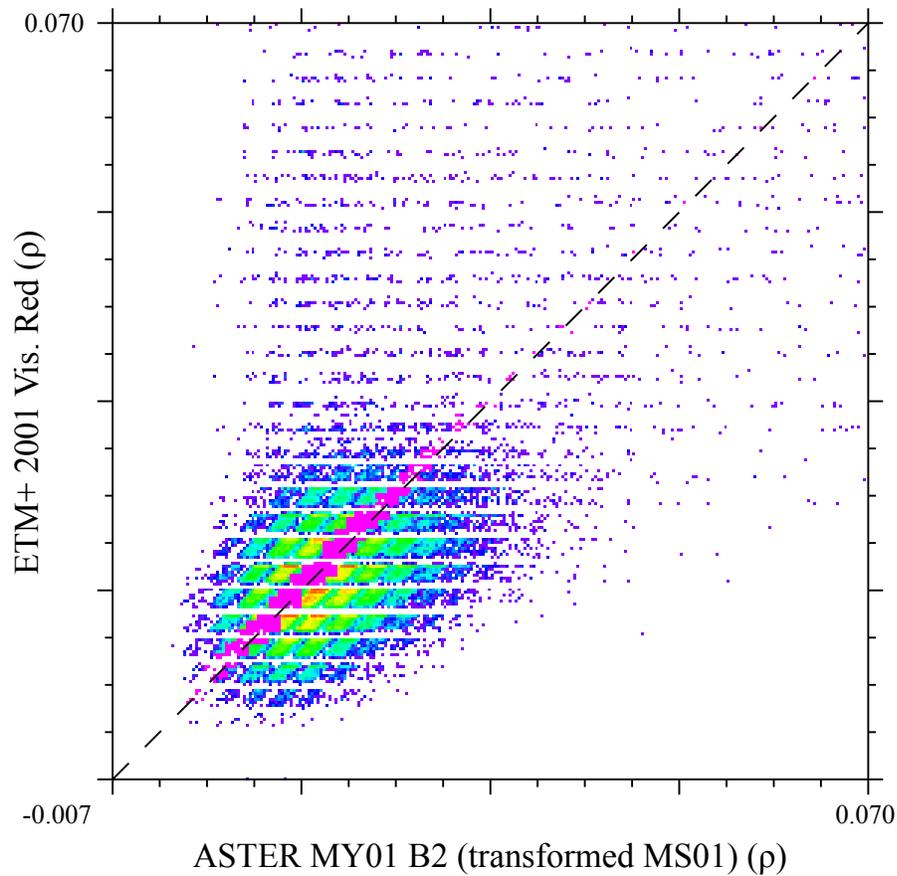


Figure 3-15. Scatter plot of ETM+ visible red data against the transformed ASTER MY01band. The highlighted area in magenta represent pixels within the region that prior to transformation fell to the right, of the 1:1 line in the figure, and hence prior to transformation representing pixels with positive differences in Figure 3-13 and negative differences in Figure 3-14.

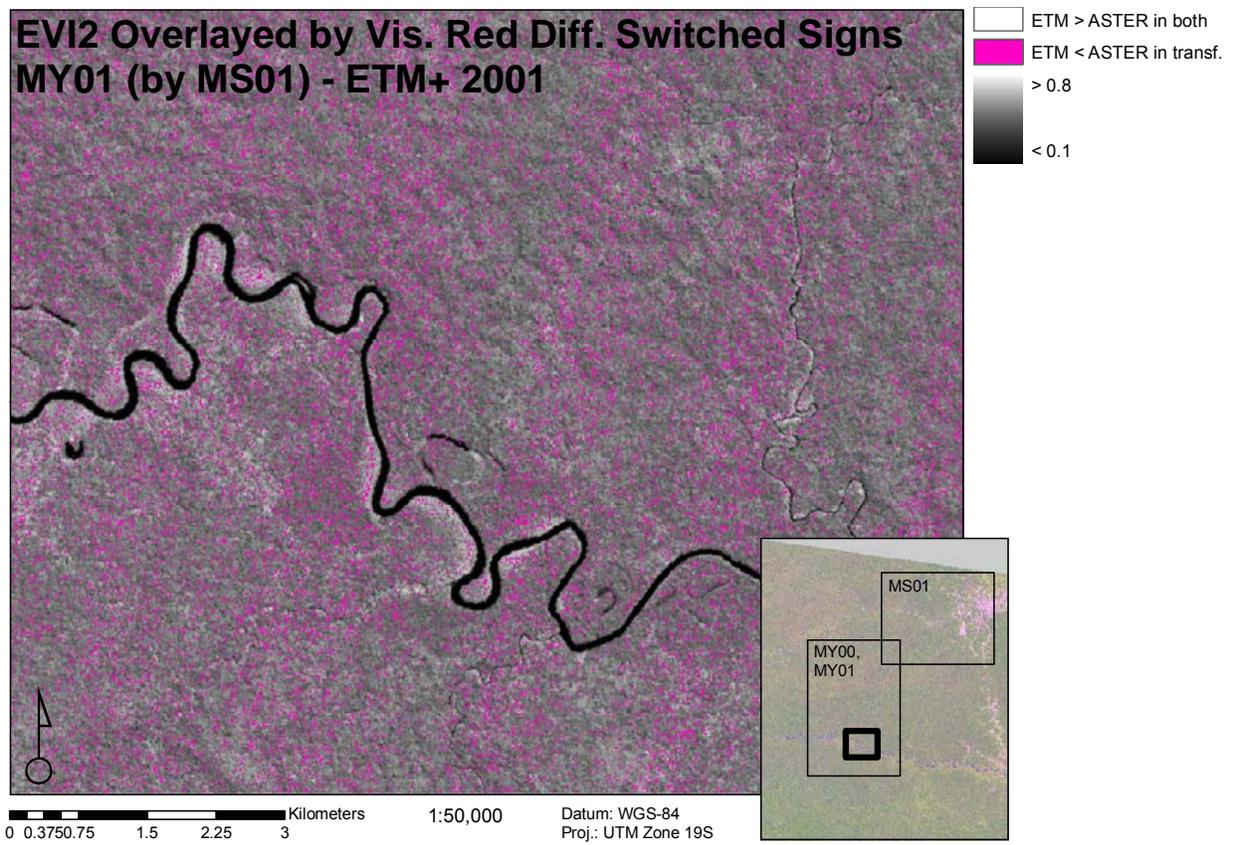


Figure 3-16. Overlaying the calculated EVI2 layer based on the transformed MY01 (transformed by the estimated coefficients of MS01) are the pixels from Figure 3-15 identified as having changed signs (i.e. moved from the right to left of the 1:1 line or changed from more purple to more yellow in Figures 3-13 and 3-14).

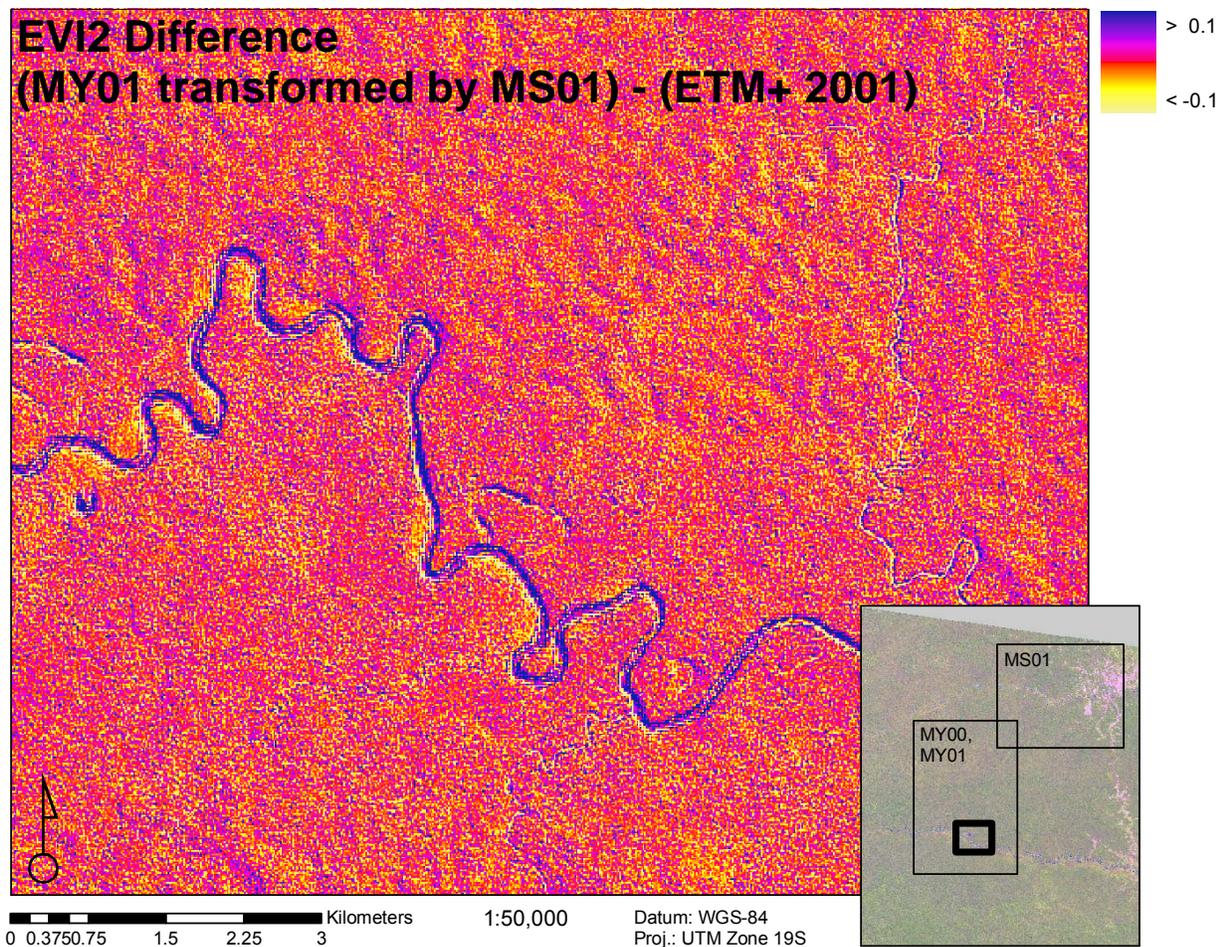


Figure 3-17. This image difference map of MY01 (transformed by MS01 coefficients) minus ETM+ EVI2 data shows that despite the spatial clustering of positive to negative differences in the visible red band, as displayed in Figure 3-16, EVI2 seems relatively robust to this segregation. EVI2 generally are observed to fall within the range of ± 0.05 , except for regions inundated with water.

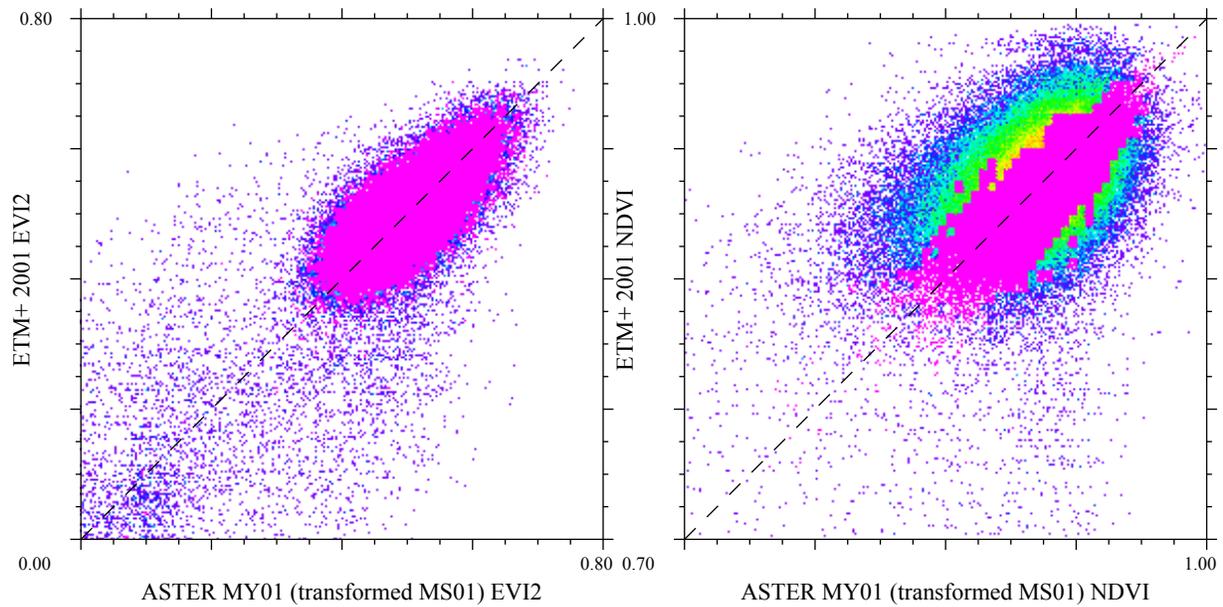


Figure 3-18. Illustration of pixel locations of visible red difference shifts within scatterplots of the calculated vegetation indices in post-transformation data. Note how EVI2 is quite symmetric about the 1:1 line relative to NDVI, and that the shifted difference pixels are symmetrically distributed as well. This indicates that EVI2 is more robust to the effects of changes in the visible red band as well as departures from a strict 1:1 relationship.

CHAPTER 4 DISCUSSION AND CONCLUSION

Study Implications

This study finds that functional associations between same-day ASTER and ETM+ data are predictable for a representative tropical forested region in the Western Amazon. Results show the utility of Theil-Sen bisector regression in not only describing but also in transforming the linear relationships in these remote sensing data. Since both ASTER and ETM+ data contain various types of measurement error, the applicability of a true “independent” and “dependent” relationship is questionable and bisector regression is necessary to fit the expected 1:1 relationship in these data.

Recent studies in the remote sensing literature have recognized the use for bisector regression (Ji et al. 2008), just as some have also incorporated the use of Theil-Sen estimators (Olthof et al. 2005). However, using Theil-Sen regression to minimize the impact of outliers and heteroscedasticity is rarely employed in tandem with the bisector approach. This study shows the Theil-Sen bisector applicability may be extremely useful when describing relationships in the presence of both bivariate error and when data do not meet the assumptions of OLS regression. This situation is extremely common in remote sensing and land change science, where we are comparing two or more calculated metrics or measures that, both being estimates of some “ground” truth, have error and outliers, multi-modality and/or non-normality. For example this is often the case when comparing vegetation indices or estimated reflectance where there is an implicit expectation of a 1:1 relationship between independent measurements (e.g., Ji et al. 2008).

Theil-Sen bisector regression should be free from these biases while offering most of the power of an OLS approach, for both small and large sample sizes, without being affected by

heteroscedasticity, outliers and parametric considerations (Isobe et al. 1990; Wilcox 1998). To date, T-S bisector regression use appears limited to astronomy and astrophysics (Bauer et al. 2002) but a full exploration of its use and implications are not in the literature. One reason for this is that estimating coefficients, standard errors and confidence intervals requires significant computing resources and custom software. Furthermore, sampled analyses likely must be bootstrapped or resampled, which adds extra burden on the researcher.

This study suggests that future research should focus on describing limitations of the Theil-Sen bisector approach. This should include a comparison with OLS bisector and other bisector regression techniques for both description and transformation of remotely sensed data since they may be easier to employ given they are more well understood and standard error estimates for coefficients can be arithmetically derived (Isobe et al. 1990).

Though the results showed very good results in same-day transformations leading to low bias and relatively high agreement between same-day ASTER and ETM+ data across all bands, a representative functional relationship was not explicitly arrived at. This study showed that significant variability in transformation coefficients may occur across each of the subsets used. The best transformation with the least aggressive, linear scaling and highest cross-date and subset applicability was the MS01 transformation (Table 3-1). Its use in transforming the MY01 dataset significantly lowered bias, absolute median of differences and standardized Wilcoxon rank scores, and in general increased the agreement greatly with ETM+ data. Further refinement of the transformations presented here may be possible and should be pursued in future research.

To facilitate such research a semi-automated system for ASTER and ETM+ comparison and transformation is available upon request. As a by-product of the custom software developed to conduct the research presented, an ENVI 4.3 and IDL 6.3 (ITT Visual Information Solutions

2007) solution for spectrally resampling ASTER data, generating Theil-Sen bisector transformation coefficients, creating scatterplot diagnostics, and providing general statistics may be a vital tool for other researchers. As a more user-friendly tool it can be modified in the future to allow for the definition of regions-of-interest that represent “unchanged” overlap between time-steps in prior ETM+ data coupled with newer ASTER data. These overlapping areas may be compared and cross-calibration based on Theil-Sen bisector transformations may be automatically applied. The software generates bivariate and statistically compares the transformation outcomes. This approach may even be extended to non Landsat data, presenting an opportunity for a more general tool for cross-date or cross-scene calibration of any two sets of data where bivariate measurement error increases the difficulty in finding functional relationships and where Theil-Sen bisector regression may be useful.

The various sources of such error were minimized to the greatest extent possible in this study. The use of same-date imagery was ideal for minimizing the effects of solar irradiation and atmospheric attenuation. The use of a sophisticated atmospheric correction technique (FLAASH) also minimized error that is introduced by varying atmospheric conditions coupled with sensor response. Registration errors were also minimized to the greatest extent possible, however, after the inherent sensor differences between ASTER and ETM+, registration error may be the largest source of bivariate error in this study.

The effects of misregistration on long-term change detection in classified studies can be extreme. Research shows that errors of only one fifth of a pixel may lead to change detection errors of 10% or more (Dai and Khorram 1998). When looking at reflectances or derived continuous indices the effects of registration error are far more subtle, since differences between images are from a continuous distribution rather than a categorical misclassification. The

impacts of resampling, either by smoothing in the case of bilinear and CC resampling, or data censoring and duplication in the NN case, also need to be better understood. This study showed higher levels of agreements with ETM+ for CC ASTER data than NN ASTER data. This is not surprising given that slight misregistration and the scalar mismatch in pixel size likely account for a large degree of the observed variability around the 1:1 lines. Therefore, this study may be expanded to include comparisons of various types and levels of smoothed ASTER data with ETM+ data. Spectral fidelity was maximized here, yet for finding a more general functional relationship this may not be the best approach. Smoothing may ultimately help in increasing inter-subset agreement on transformation coefficients.

The argument for a general set of transformation coefficients for ASTER data to match ETM+ can not be found from this study's results. However, same-day results show great promise for increasing the level of agreement between ASTER and ETM+ surface reflectance values across all bands. This may also be an argument for the previously mentioned use of invariant regions-of-interest between dates, and their use to develop study-specific transformations. Also of particular note were the results for agreements of resampled SWIR2-6 ASTER bands and ETM+ B7. The observed associations in transformed data are very encouraging, as the lack of comparable bands in this region was a significant limitation for using ASTER with ETM+ in longitudinal studies.

Finally the, successful comparison of ASTER and ETM+ derived EVI2 is extremely encouraging for the study of vegetation dynamics in tropical areas. The saturation limits of NDVI and lack of a blue band in the ASTER data make vegetation studies using continuous indices much more difficult. But high levels of agreement pre- and more so post-transformation show that EVI2 is an extremely viable solution to the vegetation index problem faced in similar

study areas. Ongoing research should compare ASTER-derived EVI2 with Landsat-derived EVI, just as in recent work with MODIS and AVHRR (Ji et al. 2008). Also research using the techniques presented here should extend comparisons of ASTER- and Landsat-derived products beyond vegetation indices to those such as “tasseled cap,” and various other analyses. By first transforming ASTER data, tasseled cap results using Landsat coefficients may be compared to those principal components coefficients presented for ASTER (Yarbrough, Eason and Kuzmaul (2005) (Appendix A).

Conclusion

Though the long-term viability of the Landsat program will no doubt continue, the near-term likelihood of data gaps exists. It should be noted that the Landsat data gap is not just a concern for individual scientists or for U.S.-based research initiatives alone. Wulder et al. (2008) identify at least six global land monitoring programs that rely on Landsat data for their ongoing research efforts. Though ASTER could not be considered a drop-in replacement for any of these larger-scale monitoring programs, researchers working at a spatial and temporal scale appropriate to ASTER, and who intend to use the hugely useful data derived from such long-term studies may utilize techniques presented here. The Theil-Sen bisector regression approach eases comparison and transformation of data and may be more widely applicable across remote-sensing applications. And by minimizing the impacts of spectral differences between ASTER and ETM+ the techniques presented here may allow for ASTER data to be used not just in classification-based studies of land change, but also in more direct comparisons. Doing so gives us one more potential bridge past data gaps by using ASTER in areas where the appropriate spatial and temporal scales allow.

APPENDIX A
ORIGINAL SOURCE CODE FOR DISCUSSED ANALYSES

ASTER Top-of-Atmosphere Reflectance Conversion from At-Sensor Radiance

```
1 ; ENVI and IDL Utilities for ASTER/Landsat Processing
2 ; Copyright (C) 2008 Forrest R. Stevens
3 ;
4 ; This program is free software: you can redistribute it and/or modify
5 ; it under the terms of the GNU General Public License as published by
6 ; the Free Software Foundation, either version 3 of the License, or
7 ; (at your option) any later version.
8 ;
9 ; This program is distributed in the hope that it will be useful,
10 ; but WITHOUT ANY WARRANTY; without even the implied warranty of
11 ; MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
12 ; GNU General Public License for more details.
13 ;
14 ; You should have received a copy of the GNU General Public License
15 ; along with this program. If not, see <http://www.gnu.org/licenses/>.
16
17
18 PRO ASTER_FILE_TOA_REFLECTANCE, hdf_file=hdf_file, append_name=append_name
19 ; The only argument here should be the hdf_filename. The at-sensor-radiance
20 ; file that is fed into the top-of-atmosphere calculation is assumed to have
21 ; '_asrad.dat' at the end of the root name (without the processing date/time).
22 ; For example, the parameter above should be the first of the following list,
23 ; but the rest of the files are assumed to follow the naming convention as
24 ; displayed:
25 ; AST_L1B_00307292001151134_20070319101905_28000.hdf (parameter above)
26 ; AST_L1B_00307292001151134_20070319101905_28000.hdf.met (metadata file from
27 ; which solar elevation
28 ; is read)
29 ; AST_L1B_00307292001151134_asrad.dat (at-sensor radiance file)
30 ; AST_L1B_00307292001151134_toar.dat (output)
31
32 if N_params() GT 3 then begin
33 MESSAGE, 'Syntax - ASTER_FILE_TOA_REFLECTANCE, hdf_filename="data_filename_root", $
34 append_name="appended name"'
35 endif
36
37 ;
38 ; Open the HDF file:
39 ;
40 infile = ''
41 if KEYWORD_SET(hdf_file) then begin
42 hdf_filename = hdf_file
43 endif else begin
44 ; Use the ENVI_SELECT function if you want to select a file on the fly rather than
45 ; use a filename passed on the command line...
46 hdf_filename = envi_pickfile(title='Input HDF Filename:')
47 endelse
48
49 ; Set the input file root name from the HDF file name:
50 data_fileroot = strmid(hdf_filename, 0, (strlen(hdf_filename) - 24))
51
52 ;
53 ; Extract solar elevation angle from the metadata file:
54 ;
55 nlines = file_lines(hdf_filename + '.met')
56 sarr = strarr(nlines)
57 openr, unit, hdf_filename + '.met', /get_lun
58 readf, unit, sarr
59 free_lun, unit
60
61 if strpos(hdf_filename, 'L1B') ne -1 then begin
62 ; Solar elevation line for L1B is hopefully always line 776:
63 text_line = sarr[775]
64 solar_elevation = strmid(text_line, 34, 9)
```

```

65 endif else begin
66 ; Solar elevation line for L1A is hopefully always line 758:
67 ; Apparently not always, it's different for the 2000 data, found on line 817:
68 if strpos(hdf_filename, '2000') ne -1 then begin
69     text_line = sarr[816]
70 endif else begin
71     text_line = sarr[757]
72 endif
73
74 solar_elevation = strmid(text_line, 30, 18)
75 endif
76
77
78 ;
79 ; Open the input file
80 ;
81 append_name_text = ''
82 if KEYWORD_SET(append_name) then begin
83     infile = data_fileroot + 'asrad' + append_name + '.dat'
84     append_name_text = '_toar' + append_name
85 endif else begin
86     infile = envi_pickfile(title='Input ASRAD Filename:')
87     append_name_text = '_toar'
88 endif
89
90 envi_open_file, infile, r_fid=fid
91 if (fid eq -1) then begin
92     MESSAGE, 'Input file ' + infile + ' not found!'
93 endif
94
95 if (fid eq -1) then return
96 envi_file_query, fid, dims=dims, ns=ns, nl=nl, nb=nb, interleave=interleave, $
97     data_type=data_type, xstart=xstart, ystart=ystart, $
98     bnames=bnames, fname=fname, fwhm=fwhm, wl=wl
99 pos = indgen(nb)
100
101 ;
102 ; Create outfile string:
103 ;
104 outfile = ''
105 if KEYWORD_SET(outfile_name) then begin
106     outfile = outfile_name
107 endif else begin
108     ; Assumes the filename of the input has the first part of the AST_ specification
109     ; of the form: "AST_L1A_00307292001151134"
110     filestart = strpos(fname, "AST_", /REVERSE_SEARCH)
111     outfile = strmid(fname, 0, filestart + 25) + append_name_text + $
112         strmid(fname, 3, /REVERSE_OFFSET)
113 endif
114
115 ; Pull the map information for the current file to save with the final output...
116 map_info = envi_get_map_info(fid=fid)
117
118
119 ; NOTE: This is SPATIAL tiling, not spectral, like as seen in the ASTER TCA calculations.
120 ; Therefore we set the interleave to BSQ (this is regardless of whether the band order
121 ; is BSQ, BIL, or BIP in the original data...
122 openw, unit, outfile, /get_lun
123 tile_id = envi_init_tile(fid, pos, num_tiles=num_tiles, $
124     interleave=0, xs=dims[1], xe=dims[2], $
125     ys=dims[3], ye=dims[4])
126
127 ; Setup the progress bar, the ENVI_REPORT_STAT procedure below updates the bar
128 ; as we loop through the tiles...
129 ENVI_REPORT_INIT, ['Reading and writing data for:', outfile], $
130     title='ASTER TOAR Conversion:', base=base
131 ENVI_REPORT_INC, base, num_tiles
132
133 for tile=0L, num_tiles-1 do begin
134     image_data = envi_get_tile(tile_id, tile, band_index=band_index)
135

```

```

136     image_toa = aster_toa_reflectance(TEMPORARY(image_data), hdf_filename, $
137         FLOAT(solar_elevation), bands=[band_index])
138
139     writeu, unit, image_toa
140     ;print, i
141
142     ENVI_REPORT_STAT, base, tile+1, num_tiles
143 endfor
144
145 ;envi_file_mng, id=fid, /remove
146
147 ENVI_REPORT_INIT, base=base, /finish
148
149 free_lun, unit
150
151 envi_setup_head, fname=outfile, ns=ns, nl=nl, nb=nb, $
152     data_type=4, offset=0, interleave=0,$
153     xstart=xstart+dims[1], ystart=ystart+dims[3], $
154     bnames='TOAR '+ bnames, map_info=map_info, $
155     descrip='ASTER TOAR conversion from: ' + fname, fwhm=fwhm, wl=wl, $
156     /write, /open
157 envi_tile_done, tile_id
158 END
159
160
161 FUNCTION ASTER_TOA_REFLECTANCE, image, hdf_filename, solar_elevation, bands=bands
162
163 if N_params() LT 3 then begin
164     print, 'Syntax - Result = ASTER_TOA_REFLECTANCE( image, hdf_filename, ', $
165         'solar_zenith, /bands=[bands] )'
166     return, -1
167 endif
168
169 if NOT KEYWORD_SET(bands) then begin
170     ; This assumes that if bands weren't specified that the bands being fed would
171     ; follow the order below, specified as 1-9 in the bands array:
172     ;     VNIR: 1, 2, 3N
173     ;     SWIR: 4, 5, 6, 7, 8, 9
174
175     dimensions = size(image, /dimensions)
176     if (n_elements(dimensions) eq 2) then bands = indgen(1) $
177     else bands = indgen(dimensions[2])
178     bands = bands + 1
179 endif
180
181 julian_date = JULDAY( $
182     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 39), 2)), $
183     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 37), 2)), $
184     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 35), 4)), $
185     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 31), 2)), $
186     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 29), 2)), $
187     LONG(STRMID(hdf_filename, (strlen(hdf_filename) - 27), 2)) $
188 )
189
190 ; Debug code to test Julian date extraction from file name:
191 ;caldat, julian_date, Month1, Day1, Year1
192 ;print, Month1, Day1, Year1
193
194 ; These represent the mean solar exoatmospheric irradiance of each band of ASTER data
195 ; from the visual, near and shortwave IR bands:
196 ;     VNIR: 1, 2, 3N
197 ;     SWIR: 4, 5, 6, 7, 8, 9
198 esun = [ 1844.6, 1556.5, 1083.8, 232.74, 79.854, 74.752, 68.898, 60.002, 57.652 ]
199
200 pi = 3.14159
201
202 distance_to_sun = (1.0 - (0.01672 * COS(0.9856 * (julian_date - 4))))
203
204 solar_zenith_rad = (90. - solar_elevation) * pi / 180.
205
206 if (n_elements(bands) gt 1) then begin

```

```

207     for band = 0, (n_elements(bands) - 1) do begin
208         image[*,*,band] = float((pi * image[*,*,band] *
209             (distance_to_sun)^2) / (esun[bands[band] - 1] *
210             COS(solar_zenith_rad)))
211     endfor
212 endif else begin
213     image = float((pi * image * (distance_to_sun)^2.) / (esun[bands[0] - 1.] *
214         COS(solar_zenith_rad)))
215 endelse
216
217 return, image
218 END

```

Theil-Sen and H-M-S Intercept Regression

```

1 ; ENVI and IDL Utilities for Nonparametric Statistics
2 ; Copyright (C) 2008 Forrest R. Stevens
3 ;
4 ; This program is free software: you can redistribute it and/or modify
5 ; it under the terms of the GNU General Public License as published by
6 ; the Free Software Foundation, either version 3 of the License, or
7 ; (at your option) any later version.
8 ;
9 ; This program is distributed in the hope that it will be useful,
10 ; but WITHOUT ANY WARRANTY; without even the implied warranty of
11 ; MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
12 ; GNU General Public License for more details.
13 ;
14 ; You should have received a copy of the GNU General Public License
15 ; along with this program. If not, see <http://www.gnu.org/licenses/>.
16
17
18 PRO KENDALL, Xs, Ys, pointN=pointN, seed=seed
19 ; For problems 8.1 and 8.22, Hollander and Wolfe, 1990
20
21 ;X = [28.9, 32.8, 12, 9.9, 15, 38, 12.5, 36.5, 8.6, 26.8]
22 ;Y = [1, 7.7, 7.3, 7.9, 1.1, 3.5, 18.9, 33.9, 28.6, 25]
23
24 ;C:      3      4      4      3      4      2      4      8      1      5
25 ;D:      6      5      5      6      5      7      5      1      8      4
26 ;T:      0      0      0      0      0      0      0      0      0      0
27 ;C-D:    -3     -1     -1     -3     -1     -5     -1      7     -7     1
28 ;Total:  -14.0000
29 ;K:      -7.00000
30 ;t:      -0.155556
31
32 if KEYWORD_SET(pointN) then begin
33     if NOT KEYWORD_SET(seed) then begin
34         newseed = SYSTIME(/seconds)
35     endif else begin
36         newseed = seed
37     endelse
38
39     ; See the SIZE function for an explanation of the "type" keyword values, we're
40     ; generating a list of random indices for the two-dimensional band arrays
41     ; to set as our scatter plot points.
42
43     random_index = fix(N_ELEMENTS(Xs) * RANDOMU(newseed, pointN), type=13)
44     X = Xs[random_index]
45     Y = Ys[random_index]
46 endif else begin
47     X = Xs
48     Y = Ys
49 endelse
50
51
52 N = n_elements(X)
53 C = indgen(N)
54 D = indgen(N)

```

```

55 T = indgen(N)
56
57 X = float(X)
58 Y = float(Y)
59
60 for i = 0, (N - 1) do begin
61     pairsC = 0
62     pairsD = 0
63     pairsT = 0
64     for j = 0, (N - 1) do begin
65         if i ne j then begin
66             concordance = (X[i] - X[j])*(Y[i] - Y[j])
67             if concordance gt 0 then pairsC = pairsC + 1
68             if concordance lt 0 then pairsD = pairsD + 1
69             if concordance eq 0 then pairsT = pairsT + 1
70         endif
71     endfor
72     C[i] = pairsC
73     D[i] = pairsD
74     T[i] = pairsT
75 endfor
76 ;print, "C:", C
77 ;print, "D:", D
78 ;print, "T:", T
79 ;print, "C-D:", C - D
80 print, "Total:", total(C-D)
81 print, "K:", total(C-D)/2.
82 print, "t:", 2*(total(C-D)/2.) / (N*(N - 1))
83 END
84
85 PRO THEIL, Xs, Ys, pointN=pointN, seed=seed
86 ; For problem 9.7, Hollander and Wolfe, 1990
87
88 ;X = [0, 5000, 10000, 15000, 20000, 25000, 30000, 100000]
89 ;Y = [0.924, 0.988, 0.992, 1.118, 1.133, 1.145, 1.157, 1.357]
90
91 ;Bhat: 5.54500e-006
92
93 if KEYWORD_SET(pointN) then begin
94     if NOT KEYWORD_SET(seed) then begin
95         newseed = SYSTIME(/seconds)
96     endif else begin
97         newseed = seed
98     endelse
99
100     ; See the SIZE function for an explanation of the "type" keyword values, we're
101     ; generating a list of random indices for the two-dimensional band arrays
102     ; to set as our scatter plot points.
103
104     random_index = fix(N_ELEMENTS(Xs) * RANDOMU(newseed, pointN), type=13)
105     X = Xs[random_index]
106     Y = Ys[random_index]
107 endif else begin
108     X = Xs
109     Y = Ys
110 endelse
111
112
113 N = n_elements(X)
114 S = findgen(N*(N - 1)/2)
115 index = 0L
116
117 X = float(X)
118 Y = float(Y)
119
120 for i = 0, (N - 2) do begin
121     for j = i + 1, (N - 1) do begin
122         S[index] = (Y[j] - Y[i]) / (X[j] - X[i])
123         index = index + 1
124     endfor
125 endfor

```

```

126 Bhat = median(S, /EVEN)
127 ;print, "S:", S
128 print, "Theil Bhat:", Bhat
129 print, "H-M-S Int.:", median((Y - X*Bhat), /EVEN)
130 END

```

ASTER Tasseled Cap Calculations and Conversion

```

1 ; ENVI and IDL Utilities for ASTER/Landsat Processing
2 ; Copyright (C) 2008 Forrest R. Stevens
3 ;
4 ; This program is free software: you can redistribute it and/or modify
5 ; it under the terms of the GNU General Public License as published by
6 ; the Free Software Foundation, either version 3 of the License, or
7 ; (at your option) any later version.
8 ;
9 ; This program is distributed in the hope that it will be useful,
10 ; but WITHOUT ANY WARRANTY; without even the implied warranty of
11 ; MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the
12 ; GNU General Public License for more details.
13 ;
14 ; You should have received a copy of the GNU General Public License
15 ; along with this program. If not, see <http://www.gnu.org/licenses/>.
16
17
18
19 PRO ASTER_TCA, infile_name=infile_name, outfile_name=outfile_name, append_name=append_name
20 ; I wrote this to implement the ASTER tasseled cap linear recombination as
21 ; outlined in Yarbrough, Easson and Kuzmaul (2005), Using at-sensor radiance... FRS, 2008
22
23 ; aster_tca, infile_name='AST_L1A_00307262000151816_asrad_nn.dat', $
24 ;   outfile_name='AST_L1A_00307262000151816_tca_nn.dat'
25
26
27 if N_params() GT 3 then begin
28   MESSAGE, 'Syntax - ASTER_TASSELED_CAP, infile_name="input file", ' + $
29   'outfile_name="output file", append_name="text to append to output name"'
30 endif
31
32 append_name_text = ''
33 if KEYWORD_SET(append_name) then begin
34   append_name_text = append_name
35 endif else begin
36   if NOT KEYWORD_SET(outfile_name) then begin
37     append_name_text = '_tca'
38   endif
39 endelse
40
41 ;
42 ; Open the input file
43 ;
44 infile = ''
45 if KEYWORD_SET(infile_name) then begin
46   infile = infile_name
47
48   envi_open_file, infile, r_fid=fid
49   if (fid eq -1) then begin
50     MESSAGE, 'Input file ' + envi_filename + ' not found!'
51   endif
52 endif else begin
53   ; Use the ENVI_SELECT function if you want to select a file on the fly rather than
54   ; use a filename passed on the command line...
55   envi_select, title='Input Filename', fid=fid, $
56   pos=pos, dims=dims
57 endelse
58
59 if (fid eq -1) then return
60 envi_file_query, fid, dims=dims, ns=ns, nl=nl, nb=nb, interleave=interleave, $
61 data_type=data_type, xstart=xstart, ystart=ystart, $

```

```

62     bnames=bnames, fname=fname, fwhm=fwhm, wl=wl
63     pos = indgen(nb)
64
65     ;
66     ; Create outfile string:
67     ;
68     outfile = ''
69     if KEYWORD_SET(outfile_name) then begin
70         outfile = outfile_name
71     endif else begin
72         ; Assumes the filename of the input has the first part of the AST_ specification
73         ; of the form: "AST_L1A_00307292001151134"
74         filestart = strpos(fname, "AST_", /REVERSE_SEARCH)
75         outfile = strmid(fname, 0, filestart + 25) + append_name_text + '$'
76         strmid(fname, 3, /REVERSE_OFFSET)
77     endwhile
78
79     ; Pull the map information for the current file to save with the final output...
80     map_info = envi_get_map_info(fid=fid)
81
82     transform_matrix = [ '$
83         [0.63400000,0.62500000,0.44600000,0.09300000,-0.01500000,$
84           0.00600000,0.00000000,-0.00100000,-0.00100000], '$
85         [0.04700000,-0.57600000,0.63200000,0.51100000,-0.06500000,$
86           0.00800000,0.00300000,-0.00400000,-0.00300000], '$
87         [0.76800000,-0.49800000,-0.35100000,-0.19800000,-0.00700000,$
88           -0.00200000,-0.00200000,-0.00100000,0.00300000], '$
89         [-0.00900000,-0.14100000,0.46400000,-0.77400000,-0.39200000,$
90           -0.02700000,-0.01600000,-0.06200000,0.03800000], '$
91         [-0.01000000,-0.05200000,0.13200000,-0.17500000,0.48200000,$
92           0.53400000,0.55100000,-0.26700000,-0.23600000], '$
93         [-0.00900000,-0.05270000,0.14300000,-0.17600000,0.40600000,$
94           -0.30900000,0.12400000,0.78200000,-0.24200000], '$
95         [-0.00900000,-0.04800000,0.12400000,-0.14400000,0.48300000,$
96           0.36200000,-0.63000000,0.07000000,0.44200000], '$
97         [-0.00300000,-0.03200000,0.08300000,-0.08800000,0.36600000,$
98           -0.48800000,-0.35100000,-0.51200000,-0.47600000], '$
99         [0.00000000,-0.01900000,0.05400000,-0.04500000,0.27700000,$
100        -0.50000000,0.40000000,-0.21500000,0.68000000] '$
101     ]
102
103     openw, unit, outfile, /get_lun
104     tile_id = envi_init_tile(fid, pos, num_tiles=num_tiles, '$
105       interleave=(interleave > 1), xs=dims[1], xe=dims[2], '$
106       ys=dims[3], ye=dims[4])
107
108     ; Setup the progress bar, the ENVI_REPORT_STAT procedure below updates the bar
109     ; as we loop through the tiles...
110     ENVI_REPORT_INIT, ['Reading and writing data for:', outfile], '$
111       title='ASTER Tasseled Cap Transform:', base=base
112     ENVI_REPORT_INC, base, num_tiles
113
114     for tile=0L, num_tiles-1 do begin
115         image_data = envi_get_tile(tile_id, tile)
116         ; We read in the spectral slice (an array of NS x NB) and take a median
117         ; slice of it. The DIMENSION variable tells the median function to return
118         ; an array of one dimension smaller than the original with the median operating
119         ; over the specified dimension number, in this case, the NB dimension. So we
120         ; get an array of NS long as our output data...
121         size_vec = size(image_data)
122         n_cols = size_vec[1]
123         n_data = size_vec[2]
124
125         ; The number of linear combinations is equal to the number of rows in
126         ; the transformation matrix, but the transform needs a column for every
127         ; band in the original image data (i.e. to generate combinations for only
128         ; the first three TCA channels you could remove rows 4-9).
129         n_combines = (size(transform_matrix))[2]
130         image_tca = fltarr(n_cols, n_combines)
131         for i=0, nb-1 do begin
132             for j=0, n_data-1 do begin

```

```

133         image_tca[* , i] = image_tca[* , i] + transform_matrix[j, i] * (image_data)[* , j]
134     endfor
135 endfor
136
137 writeu, unit, image_tca
138 ;print, i
139
140 ENVI_REPORT_STAT, base, tile+1, num_tiles
141 endfor
142
143 ;envi_file_mng, id=fid, /remove
144
145 ENVI_REPORT_INIT, base=base, /finish
146
147 free_lun, unit
148
149 envi_setup_head, fname=outfile, ns=ns, nl=nl, nb=nb, $
150     data_type=data_type, offset=0, interleave=(interleave > 1),$
151     xstart=xstart+dims[1], ystart=ystart+dims[3], bnames=bnames, map_info=map_info, $
152     descrip='ASTER Tasseled Cap Transformation for: ' + fname, fwhm=fwhm, wl=wl, $
153     /write, /open
154     envi_tile_done, tile_id
155 END
156
157
158 FUNCTION TEST_TRANSFORM, data, matrix
159     image_data = data
160     transform_matrix = matrix
161
162     size_vec = size(image_data)
163     n_cols = size_vec[1]
164     n_rows = size_vec[2]
165     n_data = size_vec[3]
166
167     ; The number of linear combinations is equal to the number of columns in
168     ; the transformation matrix:
169     n_combines = (size(transform_matrix))[2]
170     image_tca = fltarr(n_cols, n_rows, n_combines)
171     for i=0, n_combines-1 do begin
172         for j=0, n_data-1 do begin
173             image_tca[* , *, i] = image_tca[* , *, i] + transform_matrix[j, i] * (image_data)[* , *, j]
174         endfor
175     endfor
176
177     return, image_tca
178
179 END

```

License for Appended Source Code

All original and appended source code is licensed and distributed under the GNU General Public License Version 3, Copyright © 2007 Free Software Foundation, Inc. A copy of the license is available upon request from the author and appended as an attached object.

[Object A-1. GNU General Public License Version 3 \(.txt file 35 KB\)](#)

APPENDIX B
SAMPLE ENVI CONFIGURATION TEMPLATES FOR ATMOSPHERIC CORRECTION

Notes on Atmospheric Correction in ENVI™

Both radiometric and atmospheric corrections are critical when comparing data across dates, sensors or space. An advanced atmospheric correction add-on module is available in ENVI™, a remote sensing and general image analysis software package available from ITT/VIS (ITT Visual Information Solutions 2007). The module, named Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) uses the MODTRAN4 radiative transfer code and is a sophisticated model derived from radiometric first principles (Berk et al. 1999; Cooley et al. 2002). Results in this study derived from FLAASH surface reflectance estimates were found to be much more reliable than those derived from simpler cross-image atmospheric correction techniques such as dark/light-object subtraction (Chavez 1996). Since the documentation of FLAASH/MODTRAN4 implementations is scarce sample configuration files for the FLAASH module are appended here for researchers looking for a model or for those looking to duplicate atmospheric correction results for data observed under similar conditions.

ASTER FLAASH/MODTRAN Configuration File for ENVI

```
;
;ENVI FLAASH PARAMETERS TEMPLATE (4.3)
;Written Wed Jul 11 23:54:18 2007
;
; Project Parameters
enviacc.prj.radiance_file = D:\Documents\Graduate
School\Research\ASTER\Data\ASTER\AST_L1A_00307292001151134_20070914104929_16374\AST_L1A_003072920
01151134_asrad_nn.dat
enviacc.prj.reflect_file = D:\Documents\Graduate
School\Research\ASTER\Data\ASTER\AST_L1A_00307292001151134_20070914104929_16374\AST_L1A_003072920
01151134_flaash_nn.dat
enviacc.prj.filter_func_file = C:\Program Files\ITT\IDL64\products\envi44\filt_func\aster.sli
enviacc.prj.filter_func_file_index = 0
enviacc.prj.water_band_choice = 1.13
enviacc.prj.red_channel = 2
enviacc.prj.green_channel = 1
enviacc.prj.blue_channel = 0
enviacc.prj.water_abs_channel = 0
enviacc.prj.water_ref_channel = 0
enviacc.prj.kt_upper_channel = 0
enviacc.prj.kt_lower_channel = 2
enviacc.prj.kt_cutoff = 0.1000
enviacc.prj.kt_ratio = 0.4500
enviacc.prj.cirrus_channel = 0
```

```

enviacc.prj.water_retrieval = 0
enviacc.prj.user_stem_name = flaash_
enviacc.prj.modtran_directory = D:\Documents\Graduate
School\Research\ASTER\Data\ASTER\AST_L1A_00307292001151134_20070914104929_16374\FLAASH\
;
; MODTRAN Parameters
enviacc.modtran.visvalue = 40.0000
enviacc.modtran.f_resolution = 15.0000
enviacc.modtran.day = 29
enviacc.modtran.month = 7
enviacc.modtran.year = 2001
enviacc.modtran.gmt = 15.1925
enviacc.modtran.latitude = -11.2621
enviacc.modtran.longitude = -69.8538
enviacc.modtran.sensor_altitude = 705.0000
enviacc.modtran.ground_elevation = 0.2750
enviacc.modtran.view_zenith_angle = 177.1220
enviacc.modtran.view_azimuth = -90.0000
enviacc.modtran.atmosphere_model = 1
enviacc.modtran.aerosol_model = 1
enviacc.modtran.multiscatter_model = 2
enviacc.modtran.disort_streams = 8
enviacc.modtran.co2mix = 390.0000
enviacc.modtran.water_column_multiplier = 1.0000
;
; Image Parameters
enviacc.img.nspatial = 4999
enviacc.img.nlines = 4712
enviacc.img.data_type = 4
enviacc.img.margin1 = 0
enviacc.img.margin2 = 0
enviacc.img.nskip = 0
enviacc.img.pixel_size = 15.0000
enviacc.img.sensor_name = ASTER
;
; Analysis Parameters
enviacc.ana.aerosol_scaleht = 2.0000
enviacc.ana.use_adjacency = 1
enviacc.ana.output_scale = 10000.0000
enviacc.ana.polishing_res = 0
enviacc.ana.aerosol_retrieval = 0
enviacc.ana.calc_wl_correction = 0
enviacc.ana.reuse_modtran_calcs = 0
enviacc.ana.use_square_slit_function = 0
enviacc.ana.convolution_method = fft
enviacc.ana.use_tiling = 1
enviacc.ana.tile_size = 250.0000
;
; Spectral Parameters
enviacc.spc.wavelength_units = micron
enviacc.spc.lambda = [
    0.5560,    0.6610,    0.8070,    1.6560,    2.1670,    2.2090,    2.2620,
    2.3360,    2.4000]
enviacc.spc.fwhm = [
    -1.000000, -1.000000, -1.000000, -1.000000, -1.000000, -1.000000, -1.000000,
    -1.000000, -1.000000]
enviacc.img.p_input_scale = [
    10.0000,   10.0000,   10.0000,   10.0000,   10.0000,   10.0000,   10.0000,
    10.0000,   10.0000]

```

Landsat FLAASH/MODTRAN Configuration File for ENVI

```

;
;ENVI FLAASH PARAMETERS TEMPLATE (4.3)
;Written Tue Jul 17 19:49:58 2007
;
; Project Parameters
enviacc.prj.radiance_file = D:\Documents\Graduate
School\Research\ASTER\Data\Landsat\etm_3_68_2001_07_29\etm_3_68_2001_07_29_asrad.dat

```

```

enviacc.prj.reflect_file = D:\Documents\Graduate
School\Research\ASTER\Data\Landsat\etm_3_68_2001_07_29\etm_3_68_2001_07_29_flaash_noaerosol.dat
enviacc.prj.filter_func_file = C:\RSI\IDL63\products\envi43\filt_func\tm.sli
enviacc.prj.filter_func_file_index = 12
enviacc.prj.water_band_choice = 1.13
enviacc.prj.red_channel = 3
enviacc.prj.green_channel = 2
enviacc.prj.blue_channel = 1
enviacc.prj.water_abs_channel = 0
enviacc.prj.water_ref_channel = 0
enviacc.prj.kt_upper_channel = 6
enviacc.prj.kt_lower_channel = 3
enviacc.prj.kt_cutoff = 0.1000
enviacc.prj.kt_ratio = 0.4500
enviacc.prj.cirrus_channel = 0
enviacc.prj.water_retrieval = 0
enviacc.prj.user_stem_name = flaash_no_aerosol_
enviacc.prj.modtran_directory = D:\Documents\Graduate
School\Research\ASTER\Data\Landsat\etm_3_68_2001_07_29\FLAASH_No_Aerosol\
;
; MODTRAN Parameters
enviacc.modtran.visvalue = 40.0000
enviacc.modtran.f_resolution = 15.0000
enviacc.modtran.day = 29
enviacc.modtran.month = 7
enviacc.modtran.year = 2001
enviacc.modtran.gmt = 14.5900
enviacc.modtran.latitude = -11.5708
enviacc.modtran.longitude = -70.2069
enviacc.modtran.sensor_altitude = 705.0000
enviacc.modtran.ground_elevation = 0.2750
enviacc.modtran.view_zenith_angle = 180.0000
enviacc.modtran.view_azimuth = 0.0000
enviacc.modtran.atmosphere_model = 1
enviacc.modtran.aerosol_model = 1
enviacc.modtran.multiscatter_model = 2
enviacc.modtran.disort_streams = 8
enviacc.modtran.co2mix = 390.0000
enviacc.modtran.water_column_multiplier = 1.0000
;
; Image Parameters
enviacc.img.nspatial = 7891
enviacc.img.nlines = 6971
enviacc.img.data_type = 4
enviacc.img.margin1 = 0
enviacc.img.margin2 = 0
enviacc.img.nskip = 0
enviacc.img.pixel_size = 30.0000
enviacc.img.sensor_name = Landsat TM7
;
; Analysis Parameters
enviacc.ana.aerosol_scaleht = 2.0000
enviacc.ana.use_adjacency = 1
enviacc.ana.output_scale = 10000.0000
enviacc.ana.polishing_res = 0
enviacc.ana.aerosol_retrieval = 0
enviacc.ana.calc_wl_correction = 0
enviacc.ana.reuse_modtran_calcs = 0
enviacc.ana.use_square_slit_function = 0
enviacc.ana.convolution_method = fft
enviacc.ana.use_tiling = 1
enviacc.ana.tile_size = 250.0000
;
; Spectral Parameters
enviacc.spc.wavelength_units = micron
enviacc.spc.lambda = [
    0.4787,    0.5610,    0.6614,    0.8346,    1.6500,    2.2080]
enviacc.spc.fwhm = [
    -1.000000, -1.000000, -1.000000, -1.000000, -1.000000, -1.000000]
enviacc.img.p_input_scale = [
    10.0000,   10.0000,   10.0000,   10.0000,   10.0000,   10.0000]

```

LIST OF REFERENCES

- Abrams, M., Hook, S., & Ramachandran, B. (2002). *ASTER User Handbook, Version 2*. Pasadena, CA: Jet Propulsion Laboratory
- Akritas, M.G., Murphy, S.A., & Lavalley, M.P. (1995). The Theil-Sen Estimator with Doubly Censored-Data and Applications to Astronomy. *Journal of the American Statistical Association*, *90*, 170-177
- Asner, G.P., Knapp, D.E., Broadbent, E.N., Oliveira, P.J.C., Keller, M., & Silva, J.N. (2005). Selective logging in the Brazilian Amazon. *Science*, *310*, 480-482
- Bauer, F.E., Alexander, D.M., Brandt, W.N., Hornschemeier, A.E., Vignali, C., Garmire, G.P., & Schneider, D.P. (2002). The Chandra Deep Field North Survey. XII. The link between faint X-ray and radio source populations. *Astronomical Journal*, *124*, 2351-2363
- Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden, S.M., Chetwynd, J.H., Richtsmeier, S.C., Pukall, B., Allred, C., Jeong, L.S., & Hoke, M.L. (1999). MODTRAN4 radiative transfer modeling for atmospheric correction. In, *Proceedings of SPIE Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research*. Denver, CO, USA: SPIE
- Capolsini, P., Andrefouet, S., Rion, C., & Payri, C. (2003). A comparison of Landsat ETM+, SPOT HRV, Ikonos, ASTER, and airborne MASTER data for coral reef habitat mapping in south pacific islands. *Canadian Journal of Remote Sensing*, *29*, 187-200
- Chavez, P.S. (1996). Image-based atmospheric corrections revisited and improved. *Photogrammetric Engineering and Remote Sensing*, *62*, 1025-1036
- Chen, W., & Zhou, G. (2004). Comparison of satellite measured temperatures using Terra ASTER and Landsat ETM+ data. In, *Geoscience and Remote Sensing Symposium, 2004. IGARSS '04* (pp. 1723-1726): IEEE
- Cohen, W.B., & Goward, S.N. (2004). Landsat's role in ecological applications of remote sensing. *BioScience*, *54*, 535-545
- Cooley, T., Anderson, G.P., Felde, G.W., Hoke, M.L., Ratkowski, A.J., Chetwynd, J.H., Gardner, J.A., Adler-Golden, S.M., Matthew, M.W., Berk, A., Bernstein, L.S., Acharya, P.K., Miller, D., & Lewis, P. (2002). FLAASH, a MODTRAN4-based atmospheric correction algorithm, its application and validation. In, *Geoscience and Remote Sensing Symposium, 2002. IGARSS '02*. (pp. 1414-1418). Air Force Res. Lab, Arizona University, Tucson, AZ, USA: IEEE International
- Dai, X.L., & Khorram, S. (1998). The effects of image misregistration on the accuracy of remotely sensed change detection. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1566-1577

- ERDAS (2006). ERDAS Imagine 9.1. In, *Products and Services*. Norcross, GA: ERDAS, Incorporated
- Feigelson, E.D., & Babu, G.J. (1992). Linear-Regression in Astronomy. II. *Astrophysical Journal*, 397, 55-67
- Fernandes, R., & Leblanc, S.G. (2005). Parametric (modified least squares) and non-parametric (Theil-Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment*, 95, 303-316
- Gao, F., & Masek, J.G. (2008). Normalizing ASTER data using MODIS products for land cover classification. In, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXI ISPRS Congress*. (pp. 1209-1214). Beijing, China
- Gutman, G., Janetos, A.C., Justice, C.O., Moran, E.F., Mustard, J.F., Rindfuss, R.R., Skole, D., Turner, B.L.I., & Cochrane, M.A. (Eds.) (2004). *Land Change Science: Observing, Monitoring and Understanding Trajectories of Change on the Earth's Surface*. Dordrecht, The Netherlands: Kluwer Academic Publishers
- Hay, G.J., Niemann, K.O., & Goodenough, D.G. (1997). Spatial thresholds, image-objects, and upscaling: A multiscale evaluation. *Remote Sensing of Environment*, 62, 1-19
- Hollander, M., & Wolfe, D.A. (1999). *Nonparametric Statistical Methods*. New York: Wiley
- Huete, A.R., Liu, H.Q., Batchily, K., & vanLeeuwen, W. (1997). A comparison of vegetation indices global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59, 440-451
- Isobe, T., Feigelson, E.D., Akritas, M.G., & Babu, G.J. (1990). Linear-Regression in Astronomy. I. *Astrophysical Journal*, 364, 104-113
- ITT Visual Information Solutions (2007). ENVI 4.4 and IDL 6.4.1. In, *Products and Services*. Boulder, CO: ITT Visual Information Solutions
- Jensen, J.R. (2005). *Introductory Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall
- Ji, L., Gallo, K., Eidenshink, J.C., & Dwyer, J. (2008). Agreement evaluation of AVHRR and MODIS 16-day composite NDVI data sets. *International Journal of Remote Sensing*, 29, 4839-4861
- Jiang, Z.Y., Huete, A.R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, 112, 3833-3845
- Liu, H.Q., & Huete, A. (1995). A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 457-465

- Liu, J.J., Li, Z., Qiao, Y.L., Liu, Y.J., & Zhang, Y.X. (2004). A new method for cross-calibration of two satellite sensors. *International Journal of Remote Sensing*, 25, 5267-5281
- Markham, B.L., Storey, J.C., Williams, D.L., & Irons, J.R. (2004). Landsat sensor performance: History and current status. *IEEE Transactions on Geoscience and Remote Sensing*, 42, 2691-2694
- Ministério do Meio-Ambiente Cooperação Brasil-Alemanha (2000). *Zoneamento Ecológico-Econômico do Acre*. Rio Branco, Acre
- Miura, T., Yoshioka, H., Fujiwara, K., & Yamamoto, H. (2008). Inter-comparison of ASTER and MODIS surface reflectance and vegetation index products for synergistic applications to natural resource monitoring. *Sensors*, 8, 2480-2499
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403, 853-858
- National Aeronautics and Space Administration (NASA) (2008). Landsat 7 Science Data Users Handbook. In. Greenbelt, Maryland: NASA's Goddard Space Flight Center
- Olthof, I., Pouliot, D., Fernandes, R., & Latifovic, R. (2005). Landsat-7 ETM+ radiometric normalization comparison for northern mapping applications. *Remote Sensing of Environment*, 95, 388-398
- Rindfuss, R.R., Walsh, S.J., Turner, B.L.I., Fox, J., & Mishra, V. (2004). Developing a science of land change: Challenges and methodological issues. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 13976-13981
- Southworth, J., Cumming, G.S., Marsik, M., & Binford, M.W. (2006). Linking spatial and temporal variation at multiple scales in a heterogeneous landscape. *Professional Geographer*, 58, 406-420
- Teillet, P.M., Fedosejevs, G., Thome, K.J., & Barker, J.L. (2007). Impacts of spectral band difference effects on radiometric cross-calibration between satellite sensors in the solar-reflective spectral domain. *Remote Sensing of Environment*, 110, 393-409
- Trigg, S.N., Curran, L.M., & McDonald, A.K. (2006). Utility of Landsat 7 satellite data for continued monitoring of forest cover change in protected areas in Southeast Asia. *Singapore Journal of Tropical Geography*, 27, 49-66
- United States Geological Survey (USGS) (2006). Landsat 5 Resumes Operations. In, *News Releases*. Reston, VA: U.S. Geological Survey, U.S. Department of the Interior
- United States Geological Survey (USGS) (2007). Landsat Data Continuity Mission. In, *Fact Sheet 2007-3093*: U.S. Department of the Interior
- United States Geological Survey (USGS) (2008). Landsat Data Gap Studies. In, *The USGS Remote Sensing Technologies Project*. U.S. Department of the Interior

- Vicente-Serrano, S.M., Pérez-Cabello, F., & Lasanta, T. (2008). Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. *Remote Sensing of Environment*, *112*, 3916-3934
- Wang, G., Gertner, G., & Anderson, A.B. (2004). Up-scaling methods based on variability-weighting and simulation for inferring spatial information across scales. *International Journal of Remote Sensing*, *25*, 4961-4979
- Warton, D.I., Wright, I.J., Falster, D.S., & Westoby, M. (2006). Bivariate line-fitting methods for allometry. *Biological Reviews*, *81*, 259-291
- Wilcox, R.R. (1998). A note on the Theil-Sen regression estimator when the regressor is random and the error term is heteroscedastic. *Biometrical Journal*, *40*, 261-268
- Woodcock, C.E., Macomber, S.A., Pax-Lenney, M., & Cohen, W.B. (2001). Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sensing of Environment*, *78*, 194-203
- Wu, J. (2004). Effects of changing scale on landscape pattern analysis: scaling relations. *Landscape Ecology*, *19*, 125-138
- Wulder, M.A., Butson, C.R., & White, J.C. (2008a). Cross-sensor change detection over a forested landscape: Options to enable continuity of medium spatial resolution measures. *Remote Sensing of Environment*, *112*, 796-809
- Wulder, M.A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R., & Woodcock, C.E. (2008b). Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment*, *112*, 955-969
- Yarbrough, L.D., Easson, G., & Kuzmaul, J.S. (2005). Using at-sensor radiance and reflectance tasseled cap transforms applied to change detection for the ASTER sensor. In, *2005 International Workshop on the Analysis of Multi-Temporal Remote Sensing Images: IEEE*
- Zhukov, B., Oertel, D., Lanzl, F., & Reinhäkel, G. (1999). Unmixing-based multi-sensor multiresolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, *37*, 1212-1226

BIOGRAPHICAL SKETCH

Forrest Robert Stevens was born in 1976 in Pocatello, Idaho. He grew up among the dry-land wheat farms, cattle ranches and rolling foothills in Arbon Valley, and attended its two-room schoolhouse. After turning eight years old, his family moved back into town, living in Pocatello, where he completed the rest of his elementary, junior and senior high school education. In 1994, he became the third generation of his family on his mother's side, Stephenie Kahm Stevens, to attend and graduate from Pocatello High School. He completed his Bachelor of Arts degree at the University of Chicago in 1998, finishing with high honors and majoring in biology, while specializing in ecology and evolution. In 2004, he started his graduate studies at the University of Florida, where he chose to pursue degrees in Geography and plans to earn a Ph.D. He then received his M.S. from the University of Florida in the Spring of 2009.