COMBINING LANDSAT ETM+ AND TERRAIN DATA FOR SCALING UP LEAF AREA INDEX (LAI) IN EASTERN AMAZON: AN INTERCOMPARISON WITH MODIS PRODUCT

Y. E. Shimabukuro
yosio@ltid.inpe.br
L. E. O. C. Aragão
aragao@ltid.inpe.br
F. D. B. Espírito-Santo
fernando@ltid.inpe.br
Instituto Nacional de Pesquisas Espaciais - INPE
Av. dos Astronautas 1758, São José dos Campos, SP, Brazil, 12227-010

M. Williams
mat.williams@ed.ac.uk
School of GeoSciences - University of Edinburgh
Darwin Building, Mayfield Rd
Edinburgh, EH9 3JU

Abstract - The general aim of this study was to produce a continuous field of LAI to evaluate the LAI surface (MOD15 product) derived from the moderate resolution imaging spectroradiometer (MODIS), for the Tapajós region, eastern Amazonia. Our method consisted in generating regression models combining spectral data derived from Enhanced Thematic Mapper Plus (ETM+) sensor (07/30/2001) and terrain slope and altimetry information extracted from a digital terrain model. The spectral variables considered for this study are reflectance, vegetation indices (NDVI and SR) and fraction images. Using a multiple comparison test we compared the mean LAI estimated by the three models generated in this study (270m spatial resolution) with the 8 days LAI composition (08/13/2001) derived from MODIS sensor (1km spatial resolution) and also with field data. The MODIS LAI surface for the Tapajós region showed a more homogenous surface and little information about land cover and land use when we visually compared with our estimations. This fact occurs due to the 1km resolution from de MODIS LAI against the 270m resolution used in our approach. The statistics indicated that the mean LAI derived from MODIS sensor is significantly overestimated ($P<0.05$) in relation to both field and modeled data. We conclude that the approach employed here is promising for generating LAI surfaces, based on field data, for MODIS LAI validation purposes.

I. INTRODUCTION

Leaf area index (LAI) is considered a key variable in scaling up forest productivity from leaf to canopy level in various stand- and regional-scale models (e.g. Running and Coughlan, 1988; Williams et al., 1996; Sellers et al*.,* 1996). The most widely used methodology to generate continuous surfaces of LAI has been correlation of canopy biophysical properties with simple spectral indices, such as normalized difference vegetation index (NDVI), the simple ratio (SR), and the Kauth-Thomas (KT) greenness index (Hall et al., 1995).

However, the simple regression of LAI and vegetation index (VI) is linear at low LAI values, but has an asymptotical behaviour at LAI values between 3 and 5 (Turner et al., 1999). These assumptions are critical to retrieve biophysical information at high-density canopies of Tropical Rain Forests such as Amazon forest.

The MODIS LAI product (MOD15) has been operationally produced since June 2000 (Myneni et al., 2002) and is available free of charge to the users from the Earth Resources Observation System (EROS) Data Center. The use of this product by the scientific community will permit a better understanding of ecophysiological process, from the regional to the global levels. However, current and planned validation activities are needed to test and then increase product accuracy.

This study aimed specifically (1) to generate a continuous LAI surface for the Tapajós region, by combining spectral information, derived from the Enhanced Thematic Mapper Plus (ETM+) sensor images, with terrain slope and altimetry data extracted from a digital terrain model (DTM); (2) to evaluate the fit of the models proposed in this study and (3) to compare our results with a LAI surface derived from the moderate resolution imaging spectroradiometer (MODIS) in order to verify its accuracy.

II. MATERIAL AND METHODS

A. Study Area

The study was carried out at the Tapajós region in Pará State, eastern Amazonia. This is one of the experimental sites of the Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA). We selected a 13,164 km^2 area, lying between latitudes 02^o 33’ S and 04^o 00’ S and longitudes 55^o 30’ W and 54^o 30’ W.

About a half of this surface area is occupied by the Tapajós National Forest (TNF) (figure 1).

This area has a mosaic structure of secondary vegetation, pastures and bare soil inserted in a matrix of primary forest distributed in Dense Forest, which is the main vegetation type in the region, with a high number of emergent tree species, Dense Forest with uniform canopy, as well as Open Canopy forest without palms, characterized by presence of lianas (RADAMBRASIL, 1976).

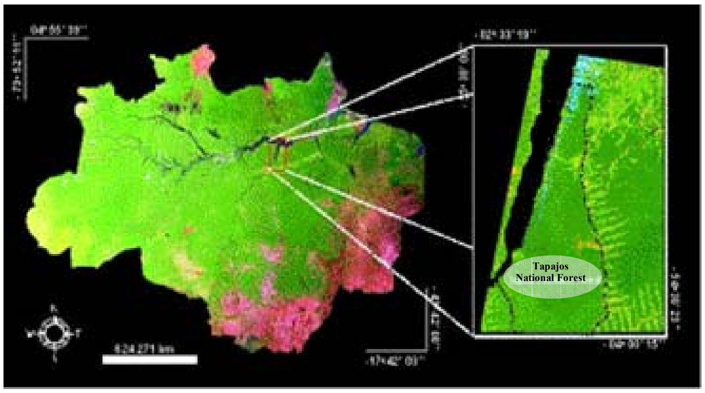

Figure 1. Location of the study area: Landsat TM mosaic of Brazilian Amazon highligthing the Tapajós region. The black dashed line on the right side image is the Tapajós National Forest border.

B. Remote Sensing Data and Processing

The study area was covered by ETM+ image (WRS 227/62) acquired on 30 July 2001. The scene was registered with a Root Mean Square (RMS) error < 1 pixel (30 m), using topographic maps at 1:100,000 scale (DSG, 1984).

The ETM+ bands (B1, B2, B3, B4, B5 and B7) were previously atmospheric corrected, and reflectance values were derived using 6S software (Vermote et al., 1997). These bands were used to create the linear spectral mixing model (Shimabukuro and Smith, 1991), and generated the fraction images corresponding to each subpixel component (shade, sunlit background and sunlit canopy). Two vegetation indices (NDVI and SR) were also calculated. The image processing methodologies of this remote sensing database are detailed in Espirito-Santo (2003).

C. Generation of the Digital Terrain Model

The DTM for the study area was built using the digitalized terrain features (contours and drainage lines, and spot elevations) extracted from the topographic maps. We created a triangular irregular network (TIN) using the drainage as constraint lines. Afterwards, we converted the TIN into two regular grids with 30 m spatial resolution, the altimetry and the terrain slope grid.

D. Field Survey

In this study we adopted a stratified LAI sampling. We divided the region in homogeneous areas according to vegetation, land use and topographic patterns based on the Landscape Unit (LU) concept (Zonneveld, 1989). The fieldwork was undertaken in October 2002. We surveyed a total of 13 LUs (from A1 to A13) (Table 1).

Table 1. Characterization of the 13 sites showing the landscape unit identification, the location along the BR-163 route, the geographic location (Lat/Long), the vegetation characteristics, the mean LAI, standard deviation (SD), and coefficient of variation (CV).

Unit	Location	Latitude	Longitude	Caracterization	Mean	SD	CV
A1	km 67	-02.85	-54.96	Primary Forest	**5.1**	0.97	19.02
A2	km 83	-03.02	-54.97	Primary Forest	**4.61**	0.71	15.40
A3	km 117	-03.35	-54.93	Pasture	**1.58**	1.01	63.92
A4	km 117	-03.36	-54.95	Primary Forest	**4.83**	0.60	12.42
A5	km 60	-02.83	-54.90	Primary Forest	**4.39**	0.34	7.74
A6	km 88	-03.08	-54.93	Fire damaged area	**3.86**	0.45	11.66
A7	km 113	-03.30	-54.94	Primary Forest	**3.49**	0.53	15.19
A8	km 83	-03.05	-54.98	Primary Forest	**4.15**	0.69	16.63
A9	km 84	-03.05	-54.93	Secondary Forest	**3.46**	0.50	14.45
A10	km 211	-04.05	-54.94	Primary Forest	**3.73**	0.37	9.92
A11	km 200	-04.01	-54.89	Primary Forest	**3.84**	0.44	11.46
A12	km 184	-03.89	-54.81	Primary Forest	**3.25**	0.28	8.62
A13	km 150	-03.64	-54.85	Primary Forest	**3.37**	0.59	17.51
Mean					**3.82**		
SD					**0.89**		
CV					**23.30**		

A pair of LAI-2000 plant canopy analysers (PCA) (LI-COR Inc., Lincoln, NE, USA) was used to measure LAI. In each of the 13 selected units we marked out three plots, each with a 0.25 ha area (50m x 50m) giving a total number of 39 plots. Measurements within the plots were carried out in a regular grid with 10 m spatial resolution. We recorded a total of 25 LAI measurements in each plot.

E. Extraction of the Spectral Information

We sampled a window of 9 per 9 pixels (7.3 ha) over the three plots in each LU. The means of the 9 x 9 windows were associated to the mean LAI value of the three plots in each site for the regression analysis. To integrate the layers into the model, we converted the grids from 30 m to 270 m spatial resolution.

F. Regression Model

We used a forward stepwise (Model 1 and 2) and a standard (Model 3) multiple regression technique to estimate LAI as the dependent variable of a multivariate set of independent variables such as spectral reflectances, endmember fractions, altimetry (Alt) and terrain slope (Sl). Using this procedure we constructed a set of 3 linear equations. For Model 1, we inputted only the variables derived from the ETM+ sensor:

$$LAI = 29.71108 - (0.26118*B5) - (0.43843*SR) - (0.08991*Shade)$$

Model 2 assimilated all satellite and terrain data:

$$LAI = 2.25184 + (0.156097*B5) - (0.03774*Sl) + (0.014629*Alt) - (0.45165*B7)$$

Finally, Model 3 was based purely on terrain information:

$$LAI = 2.072286 - (0.03339*Sl) + (0.013912*Alt)$$

G. *Intercomparison with MODIS Product (MOD15A2)*

The three resulting LAI surfaces and the LAI surface derived from measurements made by the MODIS were compared using the mean of 38 polygons covering the whole region. We applied a multiple comparison analysis of variance to investigate the differences among the mean of our LAI estimations, MODIS and field LAI.

III. Results and Discussion

The three models fit the field LAI data within the confidence level of 95% with the standard error (SE) ranging from 0.36 to 0.45. Model 2 had the higher coefficient of determination (R^2 = 0.91), followed by Model 1 (R^2 = 0.81), and Model 3 (R^2 = 0.76) (Table 2).

Table 2. Regression models statistics.

	R^2	Standard Error (SE)	*p* -level	n
Model 1	0.81	0.45	0.001	13
Model 2	0.91	0.37	0.002	11
Model 3	0.76	0.36	0.013	9

Model 3 was built using only primary forest LAI, and demonstrates that terrain variables are important in determining LAI in undisturbed areas.

The residual of estimations for each one of the 13 sites varied between 0.7 and -0.6. We could observe a systematic underestimation by the three models for the primary forest sites that had a LAI between 4.84 and 5.10 (A1, A2 and A4) and a overestimation for sites A5, A6, A7, A13 that had the LAI ranging from 3.49 to 4.38. The spectral model (Model 1) had the highest positive residual values in A1, A2 and A4 sites. Model 2 could reduce the error associated with the estimations in the high LAI domain (figure 2).

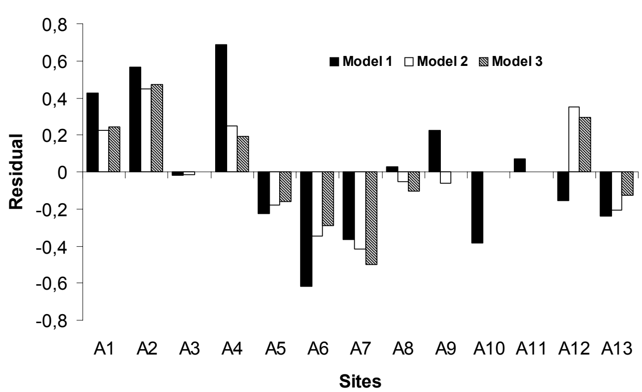

Figure 2. Comparison of the residual of LAI estimations among the three models used in the 13 sites sampled.

After the implementation of equations into the GIS we generated three LAI surfaces for the Tapajós region (figure 3).

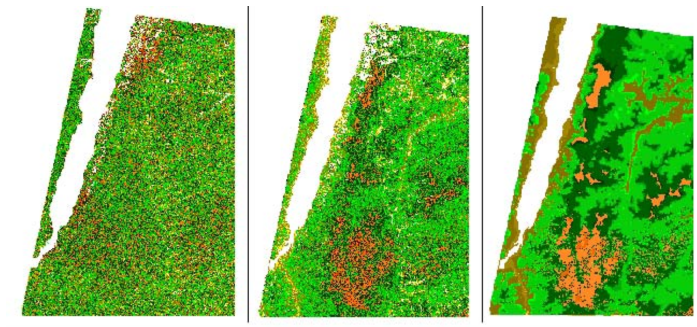

Figure 3. Comparison among the three LAI surfaces obtained by the application of the regression models. The Model 1 is purely spectral, Model 2 is a combination of spectral and terrain data, and model 3 contain just terrain information.

Despite the possibility of retrieving LAI between 4.8 and 6.0, this pattern was not so clear in the surface generated by Model 1, confirming the likely underestimation of values in sites with high LAI. On the other hand, Model 2 generated a more defined pattern of LAI distribution over the region. This surface seems to be more realistic than that of Model 1, but a cross-validation procedure must be carried out to prove the robustness of our method. Model 3 could reproduce LAI patterns in the region before any disturbance (deforestation and regeneration process).

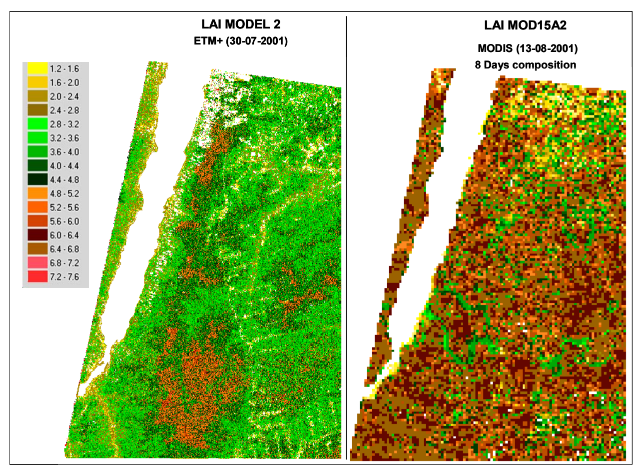

Figure 4. Comparison between LAI surfaces generated by Model 2 and by the Look-Up Table method applied for MODIS (MOD15A2 product).

The comparison of Model 2 with the MODIS LAI surface (figure 4) generated three observations: first, that MODIS surface had a more homogeneous LAI pattern; second, that MODIS LAI values are systematically higher than our estimations; and third, that the 1 km spatial resolution used to estimate MODIS LAI missed the impact of land use on LAI along the BR-163 route.

The multiple comparison applied to the four LAI surfaces and field data showed that the three models generated in this study had a similar mean for the whole region, and which was also similar to the mean of the 13 sites surveyed in the field. However, mean LAI calculated for the MODIS product was significantly higher ($p<0.05$) than our three estimations and field values (figure 5). Model 2 reproduced a mean LAI with a standard deviation and minimum and maximum values close to those found in the field (figure 5).

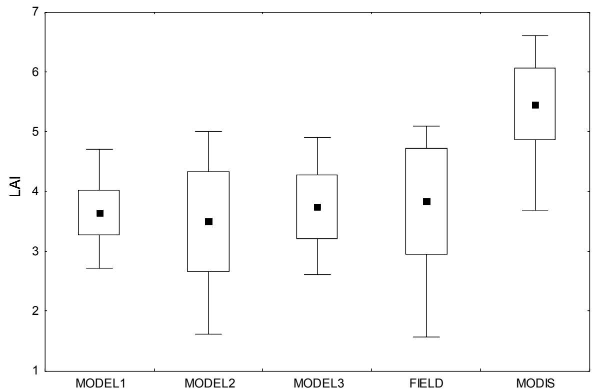

Figure 5. Comparison of the mean (black box), standard deviation (white box) and minimum and maximum (bars) values of LAI, calculated using 38 polygons over the whole Tapajós region, among the three resulting LAI surfaces and the MODIS LAI surface and also with the mean LAI of the 13 surveyed in the field. The mean LAI of MODIS surface showed a significant difference ($p<0.05$) compared with the mean estimated by Models 1, 2 and 3 and the field LAI mean.

These results suggest that MODIS Look-Up Table method used for LAI estimations (Knyazikhin et al. 1998) should be adjusted for application in this region.

IV. Conclusions

We conclude that the approach employed here is promising for generating LAI surfaces, based on field data, for MODIS LAI validation purposes. The spectral-terrain model captured more realistically the LAI hetereogeneity in the Tapajós region than did the spectral and MODIS estimations. In addition, this model showed a satisfactory fit in relation to field data, reproducing the mean, standard deviation and minimum and maximum LAI values found in the field data. MODIS LAI values are overestimating LAI values measured in the field. This fact can lead to the misinterpretation of canopy productivity and water fluxes model results in the Amazon.

Acknowledgments

This research was supported by FAPESP grant (Proc. 02/00985-2). We thank LBA-NASA Office staff at Santarém by the logistical support and CNPq and CAPES for the first author studentship.

References

[1] Diretoria de Serviço Geográfico (DSG), 1984, **Folhas SA.21-Z-B-IV, SA.21-Z-B-V, SA.21-Z-D-I, SA.21-Z-D-II, SA.21-Z-D-IV e SA.21-Z-D-V**. (Brasília: Instituto Brasileiro de Geografia e Estatística).

[2] Espírito-Santo, F. D. B., 2003, **Caracterização e mapeamento da vegetação da região da Floresta Nacional do Tapajós através de dados óticos, radar e de inventários florestais**. São José dos Campos. (INPE-10133-TDI/898). Master Thesis on Remote Sensing, 249 p.

[3] Hall, F. G., Shimabukuro, Y. E., Huemmrich, K. F. 1995, Remote sensing of forest biophysical structure using mixture decomposition and geometric reflectance models. *Ecological Applications*, **5(4)**, 993-1013.

[4] Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Dinner, D. J., Running, S. W., 1998, Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research,* **103(D24),** 32257-32277.

[5] Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, S., Zhang, Y., Smith, G.R., Lotsch, A., Frield, M., Morrisette, J. T., Votava, P., Nemani, R. R., Running, S. W., 2002, Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, **83**, 214-231.

[6] Radambrasil, 1976, Levantamento de Recursos Naturais: **Folha SA.-21 Santarém, geologia, geomorfologia, solos, vegetação e uso potencial da terra**. Projeto RADAMBRASIL. Vol. 10. Ministerio das Minas e Energia, Departamento Nacional de Produção Mineral (DNPM). Rio de Janeiro.

[7] Running, S.W., Coughlan, J.C., 1988, A general model of forest ecosystem processes for regional applications I: Hydrologic balance, canopy gas exchange and primary production process. *Ecological Modelling*, **42**, 125-154.

[8] Sellers, P.J., Randall, D.A., Collatz, G.J., Berry, J.A., Field, C.B., Dazlich, D.A., Zhang, C., Collelo, G.D., Bounoua, L., 1996, A revised land surface parametrization (SIB2) for atmospheric GCMs: Part I. Model formulation. *Journal of Climate*, **9**, 676-705.

[9] Shimabukuro, Y. E. and Smith, J. A., 1991, The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing*, **29**, 16-20.

[10] Turner, D.P., Cohen, W. B., Kennedy, R. E., Fassnacht, K.S., Briggs, J.M., 1999, Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites. *Remote Sensing of Environment*, **70**, 52-68.

[11] Vermote, E.F., Tanré, D., Deuzé, J.L., Herman, M., Morcrette, J., 1997, Second Simulation of the Satellite Signal in the Solar Spectrum 6S: an overview. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 675-686.

[12] Williams, M., Rastetter, E. B., Fernandes, D. N., Goulden, M. L., Wofsy, S. C., Shaver, G. R., Melillo, J. M., Munger, J. W., Fan, S. M., Nadelhoffer, K. J., 1996, Modelling the soil-plant-atmosphere continuum in a Quercus-Acer stand at Harvard forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. *Plant, Cell and Environment,* **19**, 911-927.

[13] Zonneveld, I.S., 1989, A fundamental concept in landscape ecology, and its applications. *Landscape Ecology,* **3(2)**, 67-86.