DRAFT

International Global Burned Area Satellite Product Validation Protocol Part I – production and standardization of validation reference data (to be followed by part II – accuracy reporting)

L. Boschetti[1], D.P.Roy[2], C.O.Justice[1]
[1] Department of Geography, University of Maryland, College Park, USA;
email: luigi.boschetti@hermes.geog.umd.edu, justice@hermes.geog.umd.edu
[2] Geographic Information Science Center of Excellence,
South Dakota State University, USA; email: david.roy@sdstate.edu

Introduction

The potential research, policy and management applications of satellite products place a high priority on providing statements about their accuracy (Morisette et al. 2006). Inter-comparison of products made with different satellite data and/or algorithms provide an indication of gross differences and possibly insights into the reasons for the differences. However product comparison with independent reference data is needed to determine accuracy (Justice et al. 2000). Validation is the term used here, and more generally, to refer to the process of assessing satellite product accuracy by comparison with independent reference data. Validation is required to provide accuracy information to help users decide if and perhaps how to use a product, and, combined with product quality assessment (Roy et al. 2002), to identify needed product improvements (Morisette et al. 2002, Strahler et al., 2006).

The purpose of this document is to present an international protocol for the generation of independent reference validation ("truth") data from high resolution satellite data, needed for validation of moderate spatial resolution continental to global scale burned area products. The objective is to promote international collaboration and sharing of validation datasets among projects, and to make it possible to share archives of validation data, as a joint initiative of the Committee on Earth Observation Satellites (CEOS) Land Product Validation (LPV) Subgroup (http://lpvs.gsfc.nasa.gov/) and GOFC GOLD (Global Observation of Forest and Land Cover Dynamics) Fire (http://gofc-fire.umd.edu).

The need for a validated long term record of global burned area was initially established in the context of the international global observing system (GCOS- Global Climate Observing System /GTOS- Global Terrestrial Observing System, 1997) and was refined by CEOS and the GCOS to meet the needs of the United Nations Framework Convention on Climate Change. These needs have long been advocated by fire product producers and product users, for example: at the International Geosphere-Biosphere Programme (IGBP) Data and Information Services (DIS) working group meeting on remote sensing of fires, held in Toulouse, March 19-20 1998 (Ahern et al, 2001); the first meeting of the GOFC-GOLD Fire community in Ispra, November 3-5, 1999; and at the joint GOFC/CEOS Land Product Validation Fire Satellite Product Validation Workshop in

Lisbon, Portugal, July 9-11, 2001 (Rasmussen et al. 2001). With the increasing availability of global and regional burned area data sets produced using different satellite data and algorithms, there is an urgency to develop broadly accepted validation procedures and validation data sets (Boschetti and Roy 2008). There are several outstanding issues in the development of a burned area product validation methodology. These include the need to develop protocols for validation data sampling, validation data content and format, and statistical accuracy assessment metrics and reporting (Justice et al. 2000; Trigg and Roy 2007).

Burned area validation data should be spatially explicit maps generated with no or minimal error, that show the areas that burned (i.e. were affected by fire), the areas that were unburned, and the areas that were not, or could not be, mapped due, for example, to cloud or missing data (Roy et al. 2005). As the effects of fire may be observed for some time after fire occurrence, burned area validation data should also describe the time period that the mapped burns occurred over. None of this is trivial. Mapping burned areas using satellite data is complex because burned areas may be confused spectrally with phenomena such as flooding, cloud and relief shadow and because the spectral signature of burned vegetation varies as a function of factors including the fire behavior, the pre-fire surface properties and the time since burning (Perirea 2003, Lentile et al. 2006, Roy et al. 2005).

This document draws largely from the validation on Southern Africa undertaken by the GOFC-GOLD-Fire Southern Africa Fire Network using Landsat Thematic Mapper / Enhanced Thematic Mapper (TM/ETM+) data to validate the MODIS (Moderate Resolution Imaging Spectroradiometer) burned area product with methodology described in Roy et al. (2005) and results in Roy and Boschetti (2008). Further testing of the protocol has been undertaken for Australia in collaboration with the GOFC-GOLD Australian regional fire network. The protocol has been reviewed and refined through a number of international workshops and meetings in partnership between the GOFC GOLD Fire Implementation Team and the LPV sub-group. Additional regional initiatives are underway in Europe and India.

Validation reference data

Key to any accuracy assessment is the provision of representative, independent validation reference data that is inherently more accurate than the product to be evaluated.
Moderate resolution data are intrinsically affected by omission and commission errors, solely due to their spatial resolution (Boschetti et al., 2004). Those errors (due to the presence of mixed pixels, and not to be confused with misclassification *stricto sensu*) can bias significantly the areal estimates, affecting the results of the validation exercise. The issue is particularly relevant for fragmented land cover classes – which is frequently the case of burned areas – making moderate resolution data unsuitable for the quantitative validation of global burned area products. The use of moderate resolution data is instead acceptable for the systematic quality control of the product, where the emphasis is on the qualitative characterization of the performance of the product.

Consequently, when satellite data are used as reference data, they should have higher spatial resolution than the data used to generate the burned area product, and spectral and radiometric resolution adequate for the unambiguous identification of burned areas, as well as encompassing the same time period as the burned area product being validated.
While ground data and aerial surveys can potentially yield reference data more accurate than high resolution remote sensing data, their potential use for validating continental and global scale products is very limited. Aircraft campaigns are expensive to undertake in a regionally or globally representative manner and are difficult to coordinate with cloud-free conditions at the time of satellite overpass. Similarly, although ground based fire measurements may provide useful information, they are difficult to coordinate over large areas (Cardoso et al. 2005).

This document does not identify a preferred remote sensing system for the acquisition of the reference data. Landsat data have been widely used for validation and for simplicity this document will make reference to Landsat TM/ETM+ data. However, the considerations are general ones, and they are immediately applicable to any other high resolution (e.g. 10-60m spatial; resolution) data.

Three aspects are emphasized:

1. Temporal criteria for the selection of reference data
2. Thematic content of the reference data
3. Format of the reference data for long term archival

1. Temporal criteria for the selection of validation reference data

Given that burned areas are a non-permanent land cover change, it is necessary to define the temporal interval described by the validation reference data. For example, in areas where forests burn, fire affected areas may remain observable in satellite data for years, while in grass/shrubland systems burned areas may disappear within a single fire season. The length of time that the spectral signature of burned areas is detectable in satellite data after a fire depends on the physical evolution of the post-burn surface, (vegetation re-growth, dissipation of ash and charcoal by wind and rain) and on the spectral bands available for the analysis (Eva and Lambin, 1996; Trigg and Flasse, 2000).

It is always preferable to use *two* TM acquisitions and then map the area that burned between the acquisition dates. In this way, fires that occurred before the first acquisition date are not mistakenly mapped as having burned between the two acquisition dates. Further, using two acquisitions provides several interpretative advantages over single date data for mapping burned areas. These include a reduction in the likelihood of spectral confusion with spectrally similar static land cover types (e.g. water bodies, dark soil), and the option to interpret the data by mapping relative changes rather than using single image classification approaches (Roy et al. 2005).

For a few particular cases, such as in certain boreal forest systems, burned areas are unambiguously visible in satellite data for more than a single fire season, and the landscape is sufficiently homogeneous to map the burned area with confidence. Provided that the timing of burning can be derived from other sources, such as active fire detections, or reports from forest services, it is acceptable to use a single TM acquisition.

Examples illustrating the consequences of an incorrect choice of the time interval between the two images are presented in appendix A.

2. Thematic content of the validation reference data

The required information is:

a) **Mapped region**, i.e., the region covered by the intersection of the two TM acquisitions, or smaller region if less than the whole scene was mapped.
b) **Burned areas** interpreted as having occurred between the two TM acquisition dates.
c) **Unmapped areas** within the mapped region that could not be interpreted, e.g., because they were covered by clouds or shadows in one or both TM images, or areas whose spectral characteristics could not be unambiguously interpreted.

In this way, parts of the TM scenes that could not be interpreted, or that were not mapped, will not be mistakenly considered as unburned when the validation reference data are compared with the moderate resolution burned area product. Examples of mapping of burned, unburned and unmapped regions are presented in appendix B

While it is not possible to prescribe a fixed technique for the interpretation of the higher resolution data, it is important to highlight that the use of the interpretation results as validation reference data, requires that the data be an acceptable proxy for reality, and so should be generated with minimum error. Visual interpretation of the Landsat data by a well trained expert generally yields the most accurate results, while automatic classification algorithms provide results of unknown accuracy, which themselves would require further validation

3. Format of the validation reference data

Validation reference data must be made available, with appropriate descriptive information, that allows their use by any user. It must be possible to use the validation reference data for computing standard accuracy metrics against any generic moderate resolution product, and it must be possible to co-register them to ancillary datasets, e.g. land cover, vegetation indices or biophysical parameters. Consequently, the validation reference data should be made available:

- In a well documented format (e.g. binary, self describing HDF format, GeoTiff, ARC/INFO vector export file).
- Either as raster data at the original spatial resolution of the TM data, or as vector data that has the same geographic accuracy as the vector digitization process
- With the geographic information provided (projection, datum, pixel size, coordinates of the image corners).

- With complete identification of the satellite data used for the production of the thematic data (e.g., in the case of Landsat data, path, row, acquisition dates)
- With information on the burned area minimum mapping unit used (i.e. the minimum size of the burned areas mapped, if bigger than the pixel size of the reference data).
- With the original satellite data used for the production of the validation reference data, whenever possible (except in the case of copyright issues). This will enable quality control, and further characterization of information such as fire severity.

When the validation data are part of a CEOS stage 2[1] or stage 3[2] validation dataset (http://lpvs.gsfc.nasa.gov/) , the criteria used for the selection of the data must be also documented. In the case of a stage 2 validation dataset, how the different high spatial resolution locations cover a range of representative conditions of the product should be documented. In the case of a stage 3 dataset, the sampling design should be described in such a way that the accuracy metrics computed using the dataset can incorporate the sampling probability.

Acknowledgements: we are grateful to Kevin Tansey, Stefan Maier, Tatiana Loboda, Lilia Manzo Delgado, Isabel Cruz and Gerardo Lopez for their useful comments on the first version of the document.

Citation: This document is largely based on the following papers:

- Roy, D.P., Frost, P., Justice, C., Landmann, T., Le Roux, J., Gumbo, K., Makungwa, S., Dunham, K., Du Toit, R., Mhwandagara, K., Zacarias, A, Tacheba, B., Dube, O., Pereira, J., Mushove, P., Morisette, J., Santhana Vannan, S., Davies, D., 2005, The Southern Africa Fire Network (SAFNet) regional burned area product validation protocol, International Journal of Remote Sensing, 26:4265-4292.
- Roy, D. and Boschetti, L.,2009, "Southern Africa Validation of the MODIS, L3JRC and GlobCarbon Burned Area Products", IEEE transactions on Geoscience and Remote Sensing, *47, 4, 1032 – 1044, doi:10.1109/TGRS.2008.2009000*.

[1] Stage 2 – Product accuracy has been assessed by a number of independent measurements, at a number of locations or times representative of the range of conditions portrayed by the product.

[2] Stage 3 - Product accuracy has been assessed by independent measurements in a systematic and statistically robust way representing global conditions (and so encompass a range of the more important conditions and factors that influence product performance).

References

Ahern F. J., Belward A., Elvidge C., Goldammer J., Grégoire J.-M., Justice C.O. [2], Pereira J., Prins E. M., Stocks B., 2001. The fire component of Global Observation of Forest Cover: a plan of action. In Ahern F., Goldammer, G., Justice C.O. (Eds), **Global and Regional Vegetation Fire Monitoring From Space: Planning a Coordinated International Effort**, SPB Academic Publishing, The Hague, The Netherlands, 267-290.

Boschetti L., P.A. Brivio and S.P. Flasse, Pareto boundary: a useful tool in the accuracy assessment of low spatial resolution thematic products, IGARSS IEEE International Geoscience and Remote Sensing, session Data Correction and Calibration, Anchorage, 20-24 September 2004, volume VI: 3759-3762.

Boschetti, L., Brivio, P. A., Eva, Hugh D., Gallego, J., Baraldi, A. and Grégoire, J-M., A Sampling Method for the Retrospective Validation of Global Burned Area Products, IEEE-Transactions on Geoscience and Remote Sensing, Volume 44, Issue 7, Part 1, July 2006 Page(s):1765 - 1773 Digital Object Identifier 10.1109/TGRS.2006.874039

Cardoso, M.F., Hurtt, G.C., Moore, B., Nobre, C.A., Prins, E., 2003, Projecting future fire activity in Amazonia, Global Change Biology, 9:656-669.

H. Eva, and E. Lambin, “Remote sensing of biomass burning in tropical regions: sampling issues and multisensor approach,” Remote Sens. Environ.t, vol. 64, pp. 292-315, 1998.

Justice, C.O., Belward, A., Morisette, J., Lewis, P., Privette, J. & Baret, F., (2000), Developments in the 'validation' of satellite sensor products for the study of land surface. International Journal of Remote Sensing, 21, 3383-3390.

Lentile, L.B., Holden, Z., Smith A.M.S., Falkowski M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E.and Benson, N.C., 2006 Remote sensing techniques to assess active fire and post-fire effects, *International Journal of Wildland Fire*, 15, 3, 319-345

Morisette J.T., Privette, J.L., & Justice, C.O. (2002). A framework for the validation of MODIS land products. Remote Sensing of Environment, 83, 77-96.

Morisette, J.T., F. Baret, S. Liang, (2006). Special issue on Global Land Product Validation, IEEE TGARS, 44(7) 1695-1697.

Pereira JMC (2003) Remote sensing of burned areas in tropical savannas. International Journal of Wildland Fire 12, 259-270.

Strahler, A., Boschetti, L., Foody, G., Friedl, M., Hansen, M., Harold, M., Mayaux, P., Morisette, J., Stehman, S., Wodcock, C., 2006. Global Landcover Validation: Recommendations for Evaluation and Accuracy Assessment of Global Landcover Maps, Luxembourg, Office for Official Publication of the European Communities, EUR 22156 EN, 58p.

Rasmussen, K., Russell-Smith, J., Morisette, J.T., 2001, Establishing a validation site network for remote sensing applications to fire research: a joint activity between GOFC-Fire and the LPV subgroup, White paper available at http://modis.gsfc.nasa.gov/MODIS/LAND/VAL/CEOS_WGCV/GOFC_LPV_fire_sites.pdf

Roy, D.P, Borak, J, Devadiga, S., Wolfe, R., Zheng, M., & Descloitres, J., (2002b), The MODIS land product quality assessment approach, Remote Sensing of Environment, 83:62-76.

Roy, D.P., Frost, P., Justice, C., Landmann, T., Le Roux, J., Gumbo, K., Makungwa, S., Dunham, K., Du Toit, R., Mhwandagara, K., Zacarias, A, Tacheba, B., Dube, O., Pereira, J., Mushove, P., Morisette, J., Santhana Vannan, S., Davies, D., 2005, The Southern Africa Fire Network (SAFNet) regional burned area product validation protocol, International Journal of Remote Sensing, 26:4265-4292.

Roy, D.P., Trigg, S.N., Bhima, R., Brockett, B., Dube, O., Frost, P., Govender, N., Landmann, T., Le Roux, J., Lepono, T., Macuacua, J., Mbow, C., Mhwandangara, K., Mosepele, B., Mutanga, O., Neo-Mahupeleng, G., Norman, M., Virgilo, S., 2006, The utility of satellite fire product accuracy information – perspectives and recommendations from the southern Africa fire network, IEEE Transactions on Geoscience and Remote Sensing, Land Product Validation Special Issue, Vol. 44, No. 7, 1928-1930.

Roy, D. and Boschetti, L., "Southern Africa Validation of the MODIS, L3JRC and GlobCarbon Burned Area Products", submitted to IEEE transactions on Geoscience and Remote Sensing, June 2008.

S. Trigg, and S. Flasse, “Characterizing the spectral - temporal response of burned savannah using in situ spectroradiometry and infrared thermometry,” Int. J. Remote Sens., vol. 21, pp. 3161-3168, 2000.

Trigg, S.N and Roy D.P., 2007, A focus group study of factors that promote and constrain the use of satellite derived fire products by resource managers in southern Africa, Journal of Environmental Management, 82:95-110.

A) Time difference between the two images

Image 1: 23 Oct 2000

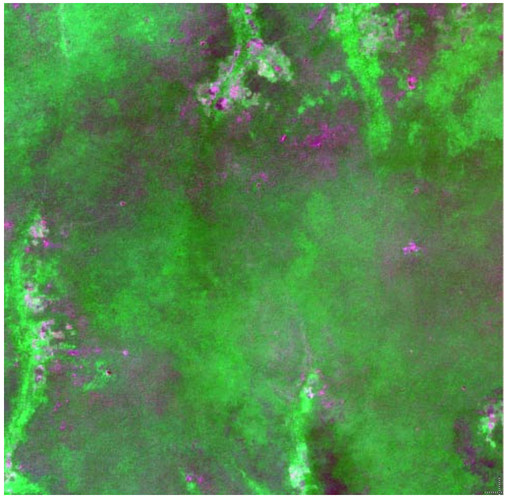
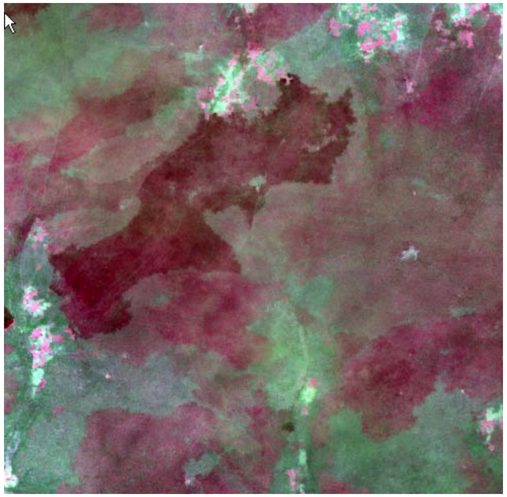


Image 2: 11 Jan 2001



INCORRECT: Excessive distance between the acquisition of the two images; the time interval is longer than the persistence time of the burned area spectral signal, and some of the older burned areas in image 2 cannot be reliably identified

Image 1: 3 Sept 2001

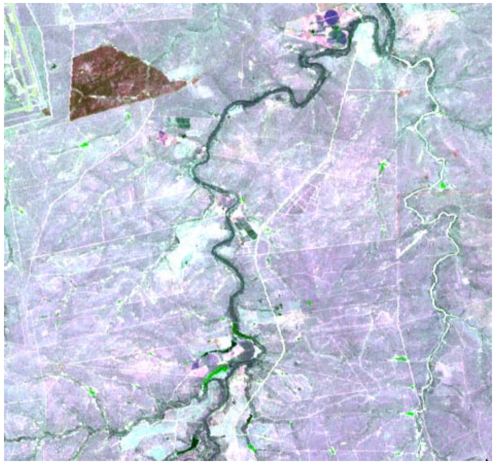
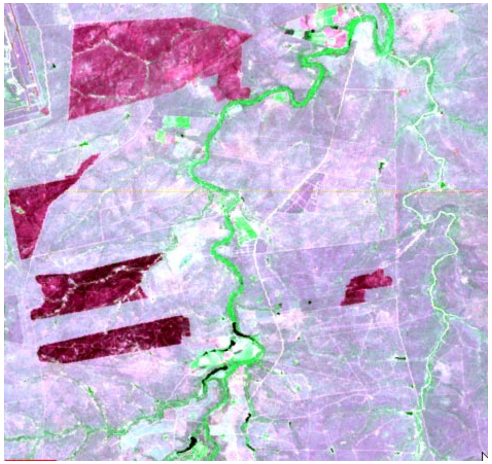


Image 2: 5 Oct 2001



CORRECT: the time interval is shorter than the persistence time of the burned area spectral signal, and all the areas burning between the acquisition of the first and the second image are clearly identifiable

B) Mapping the changes between the two dates

B.1. Burned vs. unburned

Image 1: 18 August 2001

Image 2: 3 Sept 2001

Interpretation

Image 1: 10 Sept 2001

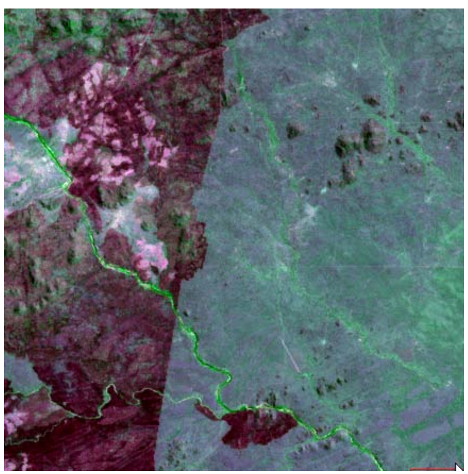
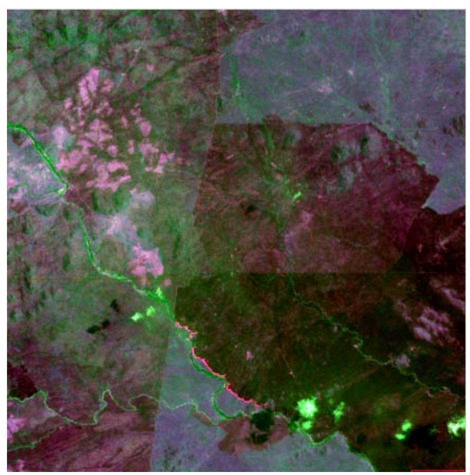


Image 2:12 Oct 2001



Interpretation

Only the portion of the burned area which burns between the two dates is digitised as burned (**red**), while the areas already burned in the first image are considered unburned (**black**)

B.2. Unmapped areas

Image 1: 23 Aug 2001

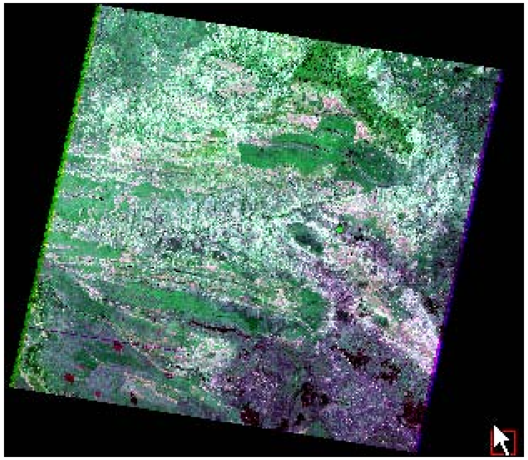
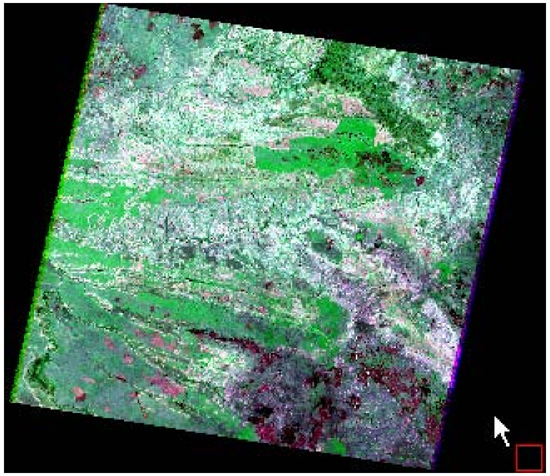
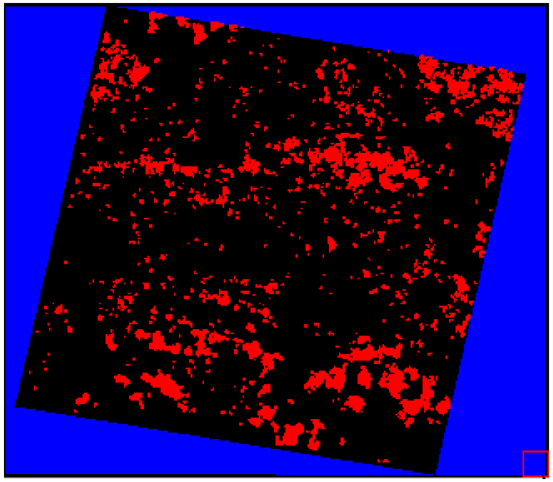


Image 2: 24 Sept 2001



Interpretation



It is important to define the footprint of the image, to differenciate between unburned (**blue**) and unburned (**black**)

Image 1: 10 Sept 2001

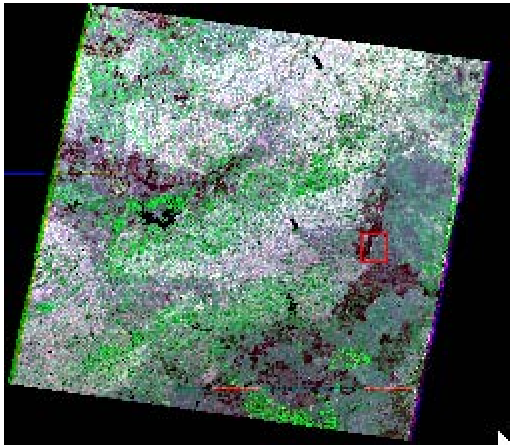
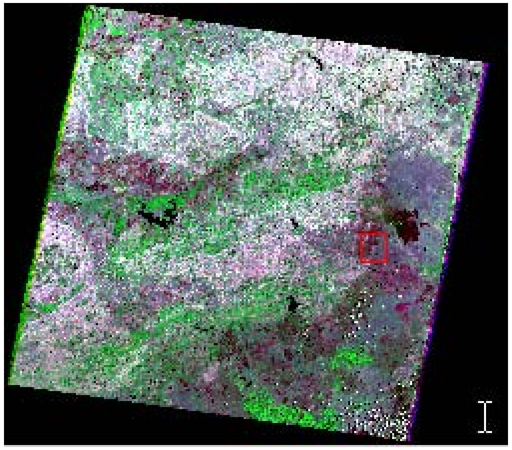
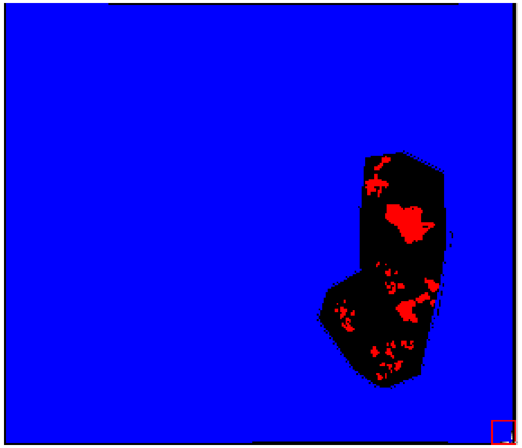


Image 2:12 Oct 2001



Interpretation



If a portion of the image cannot be interpreted because of the quality of the data, or the characterstics of the terrain, it must be labeled as unmapped (**blue**), not as unburned.

B.2. Unmapped areas

Image 1: 10 Sept 2001

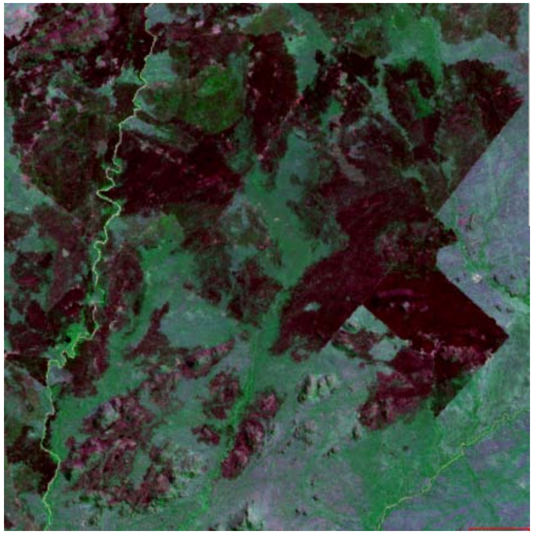
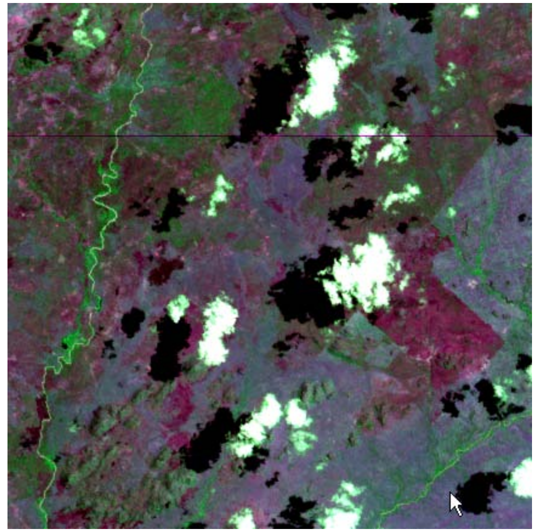
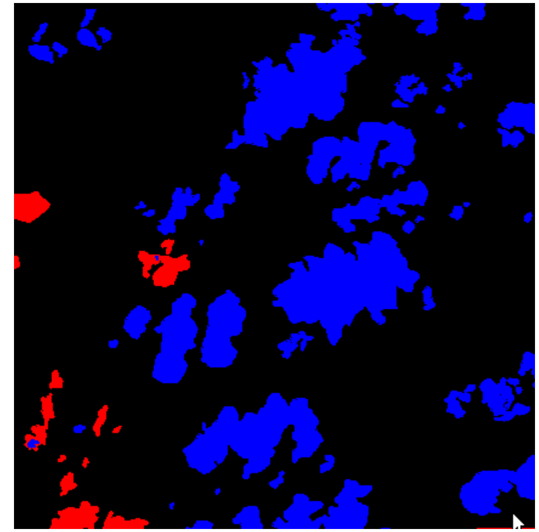


Image 2:12 Oct 2001



Interpretation



Clouds and cloud shadows that make the interpretation impossible on either image must be digitised and labeled as unmapped (**blue**)