Remote Sensing Image Information Mining applied to Burnt Forest Detection in the Brazilian Amazon

Mikhaela A. J. S. Pletsch[1], Thales Sehn Körting[1]

[1]Instituto Nacional de Pesquisas Espaciais (INPE)
Caixa Postal 15.064 – 91.501-970 – São José dos Campos – SP – Brazil

{mikhaela.pletsch; thales.korting}@inpe.br

Abstract. *Fire processes contribute to carbon dioxide emissions, main gas responsible for the Greenhouse Effect. Considering the importance of fire processes management for the detection of burnt areas in the Brazilian Amazon, the Linear Spectral Mixture Model is one of the main methods available. Nonetheless, some manual processes are required before its application, such as identifying adequate images in databases. In this manner, we have developed an approach for Remote Sensing Image Information Mining (ReSIIM), which was tested for burnt areas studies. ReSIIM stores information about well-known targets found in Remote Sensing imagery, such as cloud, cloud shadow, clear land, water, vegetation and bare soil.*

1. Introduction

Several rainforests worldwide are located in underdeveloped countries. In such a way, resources to protect and preserve the intact environment are typically scarce [FAO et al. 2011]. Although rainforests play an important role in climate regulation, they face countless threats. Among them, in the Brazilian Amazon, deforestation is often the first one to be pointed, yet during a fire, the main gas emitted is carbon dioxide, which is also the primary Greenhouse effect gas [Anderson et al. 2005a, Lashof 1991]. According to [Aragão and Shimabukuro 2010], drought years followed by fires release as much carbon (C) as deforestation processes. The negative effects of fires thus extend beyond damage to a single swath of trees. They influence global climate changes once surface radiative changes have occurred [Shimabukuro et al. 2009, Shimabukuro et al. 2015, Anderson et al. 2015, Padilla et al. 2017]. Therefore, comprehending, managing and avoiding fires in the Amazon region could meet the international demand for C emission reductions. These fires, normally induced by humans, are often applied to facilitate land use and land cover (LULC) changes. Where before there was a natural habitat supporting a wide variety of life forms, agricultural areas and pastures arise [Shimabukuro et al. 2015, Aragão et al. 2016].

Analyzing large areas of the Earth's surface in a short time is possible using Remote Sensing (RS) tools. However, describing and finding information, as well as improving data management, analysis and cataloging such a great amount of data are some of the main challenges [Li et al. 2016]. Produced in a high velocity, Earth Observation (EO) data come from several remote sensors in different data types, resolutions and scales [Datcu and Seidel 2002, Datcu and Seidel 2003, Li et al. 2016]. Considering the data deluge, production velocity and high diversity [Laney 2001], EO data is also coined as Big Data [Körting et al. 2016]. In this way, RS image catalogs were developed by institutions in order to manage and distribute EO data. Many efforts currently aim to develop efficient, easily-accessible and well-sourced catalogs. Some examples are the United States Geological Survey (USGS) [1], the European Space Agency (ESA) [2] and the Brazilian National Institute for Space Research (INPE) [3], whose goals are the management and distribution of EO data. Through these catalogs, it is possible to search for images according to user's specifications, such as location and date [Datcu et al. 2000, Li and Narayanan 2006, Stepinski et al. 2014]. On the other hand, more accurate images search tools are not available even in the most modern satellite image catalogs, and smart tools are scarce [Stepinski et al. 2014]. Such mining strategies would strive to improve geospatial data handling tools, taking into account the volume of data [Quartulli and Olaizola 2013].

For the detection of burnt areas in the Brazilian Amazon using satellite images, the Linear Spectral Mixture Model (LSMM) [Shimabukuro and Smith 1991] is one of the main approaches available. Nonetheless, some manual processes are still required before its application, such as identifying adequate images in databases [Andere et al. 2015]. If more accurate methods for searching RS images in catalogs existed, detection of burnt forest areas would be boosted along with other related research. Image searching is, hence, a time consuming step and fast results are important to guide public policies. In this context, this paper introduces a methodology of Remote Sensing Image Information Mining (ReSIIM) applied to supporting the detection of burnt forest areas in the Brazilian Amazon, providing users with the possibility of searching according to its basic target criteria, such the presence of *cloud*, *cloud shadow*, *clear land*, *water*, *vegetation* and *bare soil*. The aforementioned methodology is, however, versatile enough to be applied on any related RS applications.

2. Literature Review

2.1. Remote Sensing Image Metadata

Generally, spatial data is related to any information with absolute or relative location. In this context, the collection of a geographic phenomena information forms a spatial database [Guptill 2015]. Metadata provides detailed attributes and characteristics description of a given element [Guptill 1999], being *the data about data*. Based on RS image metadata, different search criteria can be developed in catalogs. Roughly, a catalog refers to a description list of items found in a collection [Frank 1994]. A data catalog is thus, a collection of metadata records, which is associated with search tools and data management [Guptill 1999]. Therefore, an image catalog facilitates the operations of searching, sharing and processing data to users.

RS image metadata can be identified through some approaches such as Fmask algorithm, first developed by [Zhu and Woodcock 2012] and improved by [Zhu et al. 2015]. Based on these works, [Flood and Gillingham 2017] developed a set of command line utilities in a Python module. The output of the algorithm is a single thematic raster with up to 6 different values [0, 5], representing: *null value*, *cloud*, *cloud*

[1] `https://data.usgs.gov/datacatalog/`
[2] `https://earth.esa.int/web/guest/data-access/catalogue-access`
[3] `http://www.dgi.inpe.br/catalogo/`

shadow, *clear land*, *snow*, and *water*. Clear land refers to data that are none of the aforementioned targets even though it carries information.

Regarding the information of presence of water in RS images, [Namikawa et al. 2016] extracted 5-meter spatial resolution masks from Brazilian water bodies using RapidEye images with an automated methodology. Although water bodies are usually identified due to its low reflectance, in the real world several parameters may interfere in their detection, such as suspended solids and water depth. With this in mind, the methodology was based on the color transformation from Red-Green-Blue (RGB) to Hue-Saturation-Value (HSV) and the minimum radiance from all the bands. For that, some factors were considered, as the differences in illumination and scattering throughout more than 15,000 RapidEye scenes. As a result, it was possible to classify seven classes of water in agreement with the confidence of the classified pixels ranging from 1 to 7, according to the presence of water, where 1 is more reliable and 7, less reliable. Furthermore, the degree of persistence is also available, if needed. According to the authors, although this methodology is considered simple, it is accurate to detect water bodies. [Namikawa and Castejon 2017] identified some issues which may interfere the results nonetheless, such as cloud noise, shadow in urban areas and specular reflection of sunlight.

Furthermore, regarding the information of presence of live green vegetation in RS images, [Rouse Jr et al. 1974] developed the Normalized Difference Vegetation Index (NDVI), which ranges from [-1.0, +1.0]. The definition of thresholds in NDVI for mapping vegetation is controversial, and it also varies from one region to another. However, NDVI values between 0 and 0.1 normally refer to rocks and bare soil, regardless of said controversy. Values greater than 0.1 indicate a gradual increase in *greeness* and intensity of vegetation [NOAA 2017]. On the other hand, some authors use NDVI limit of 0.2 for bare soil and vegetation [Jin et al. 2014, Liang et al. 2014b, Liang et al. 2014a], and the transition zone is considered from 0.1 to 0.2 [Liang et al. 2014a].

2.2. Burnt Forest Detection

Although some processes were developed for identifying burnt areas, outstanding questions remain in the literature, such as associated uncertainties and its causes [Anderson et al. 2005a, Aragão et al. 2016]. In [Aragão et al. 2016], studies for estimating burnt forest areas are briefly described, ranging from the first tests in the 80's to algorithms to separate burnt forests from other phenomena such as selective logging. A short compilation of burnt forest identification is available in [Lima 2013].

Generally, two main approaches are used to identify burnt forests: alterations of biophysical properties of the carbonized matter and heat release. Different channel combinations of multispectral RS images can be used to enhance complex phenomena identification [Bannari et al. 1995, Key and Benson 2006], yet they are not accurate in every biome.

Currently, burnt forest research in the Brazilian Amazon is commonly performed based on the Linear Spectral Mixture Model (LSMM) [Anderson et al. 2005b, Shimabukuro et al. 2009, Cardozo et al. 2013, Andere et al. 2015, Anderson et al. 2015]. Even for more accurate spatial resolutions, satellite data presents a *mixture problem* [Anderson et al. 2005b, Shimabukuro et al. 2009], since a pixel represents the average spectral response from all the elements located in that pixel. In this perspective, LSMM was developed aiming to depict subpixel heterogeneity [Shimabukuro and Smith 1991]. In this model, some pure pixels called *end-members* are selected by a domain specialist, deriving shade fraction images for burnt area detection. Similar spectral responses may interfere with the results though [Chuvieco and Congalton 1988, Bastarrika et al. 2011, Andere et al. 2015]. For instance, burnt areas exhibit low reflectance, as well as water and cloud shadows. Moreover, cloud coverage and fire smoke may also omit pixels with spectral response affected by fire [Aragão et al. 2016], whilst clouds and their shadows influence negatively many uses of EO data, such as inaccurate atmospheric correction and land cover classification [Zhu and Woodcock 2012]. In this context, applying LSMM methodology requires a prior step: identifying appropriate images for analysis.

3. Methodology

In this section, we present the developed methodology (Figure 1). The Remote Sensing image database is composed of Landsat 5 and 8 imagery. ReSIIM is organized in two main steps, feature extraction algorithm and metadata generation algorithm. In a decision process, the generated metadata is analyzed according to the reference data. If the result is unsatisfactory, it goes back to the ReSIIM phase in an attempt to generate more accurate metadata. Finally, some searching criteria for burnt forest detection studies takes place in the metadata database.

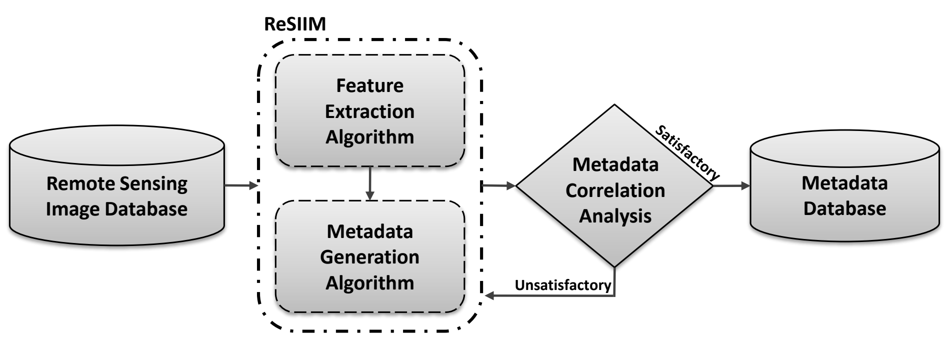

Figure 1. Methodology Flowchart.

3.1. Remote Sensing Image Information Mining (ReSIIM)

Our Remote Sensing Image Information Mining methodology (hereafter ReSIIM) aims to extract and generate metadata from satellite images, and is based on open source softwares, scripts and libraries. For that, different indices and methods consolidated in the literature were evaluated according to the correlation between these approaches and real targets along a set of images. Two main tools were used: Fmask algorithm [Zhu and Woodcock 2012, Zhu et al. 2015] for *cloud*, *cloud shadow*, *clear land* and *water*; and NDVI [Rouse Jr et al. 1974] for *vegetation* and *bare soil*. More details about them are available in Section 2. Summing up, ReSIIM is already able to identify 6 basic targets.

We describe in this work the use of ReSIIM to support burnt forest research, yet its application to other analysis is not ruled out, since ReSIIM was idealized to be continuously improved according to the scientific demands.

3.1.1. ReSIIM Metadata Examples

Examples of RS images are available in Figure 2, which shows the contrast between two scenes and the generated metadata. Figure 2-A (path/row 220/065 - 24.09.2013) is not cloudy (0.05%) in contrast with Figure 2-B (path/row 225/071 - 10.04.2015) (52.30%), which means that for burnt forest detection, A would be more suitable than B. Moreover, in A, Clear Land percentage is noticeably higher. On the other hand, it is also remarkable that Figure 2-B presents *vegetation* percentage higher than Figure 2-A. Other basics targets in both scenes do not show outstanding percentage differences.

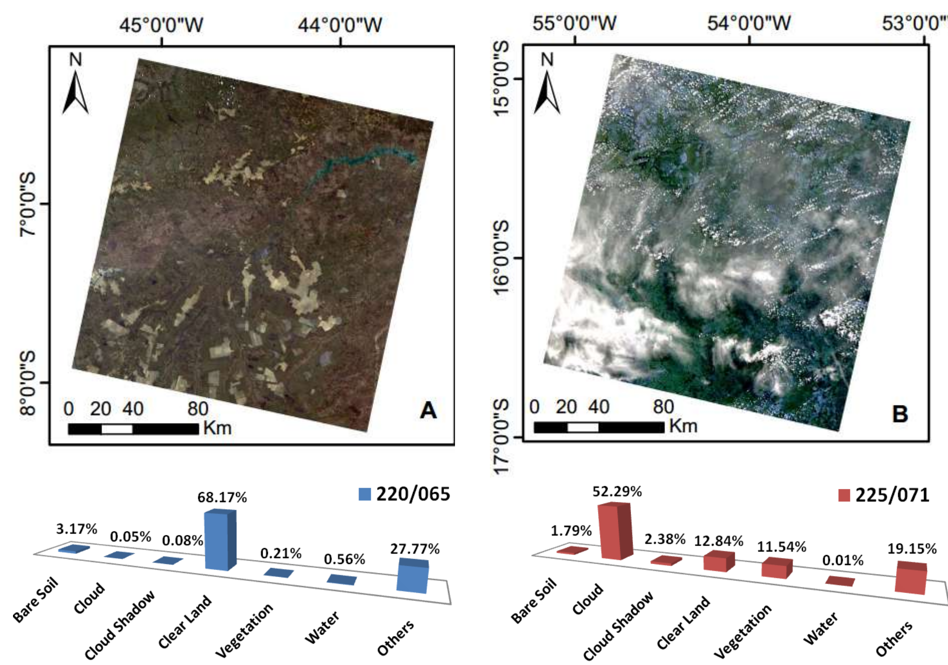

Figure 2. Examples of Remote Sensing images and the generated metadata. A – scene 220/065; B - scene 225/071 (color combination in R4G3B2 using L8 bands).

3.2. Metadata Correlation Analysis

In this section, we explain the methods used to analyze the correlation between the generated metadata and real targets along a set of images. For that, random L8 scenes along Legal Amazon were used (freely available at `https://earthexplorer.usgs.gov/`).

3.2.1. Water

The correlation analysis of *water* was performed based on the reference dataset developed by [Namikawa et al. 2016] (detailed in Subsection 2.1). This approach was taken into account, once it is more flexible when compared with a common threshold for a large amount of scenes [Namikawa et al. 2016]. The *water* reference data was filtered according to the high persistence of *water* along 4 years.

The correlation analysis was based on two main approaches: water correlation accuracy (WCA) and commission error (CE). WCA refers to the overlapped data between the target *water* generated in ReSIIM and the reference data, whilst CE refers to the data wrongly classified. For that, 10 random scenes were selected across the Brazilian Amazon. WCA ranged from 76% to 99% of correct classification, and average WCA was 90%. CE is also considered low once no cases of more than 24% were misclassified, and the average CE obtained about 10%. Omission error was not taken into account, since the reference data's spatial resolution is 5m and the classified data is 30m (L8).

3.2.2. Vegetation

Considering the several meanings that land cover *vegetation* carries, such as for agriculture and several kinds of forest, we limited a NDVI threshold in this work, according to previous analysis. We used forest mask classified by the Brazilian National Institute for Space Research (INPE) from the PRODES project, which monitors the Brazilian Amazon through satellites [INPE 2017]. The forest mask was used, thus, to derive the NDVI interval for our analysis (Figure 3). In most cases, the amount of pixels labeled as forest by PRODES was higher in NDVI values starting from 0.8. In such a manner, we empirically determined the threshold of 0.8 due to its high correlation to dense vegetation, and defined NDVI >= 0.8 for *vegetation* target detection.

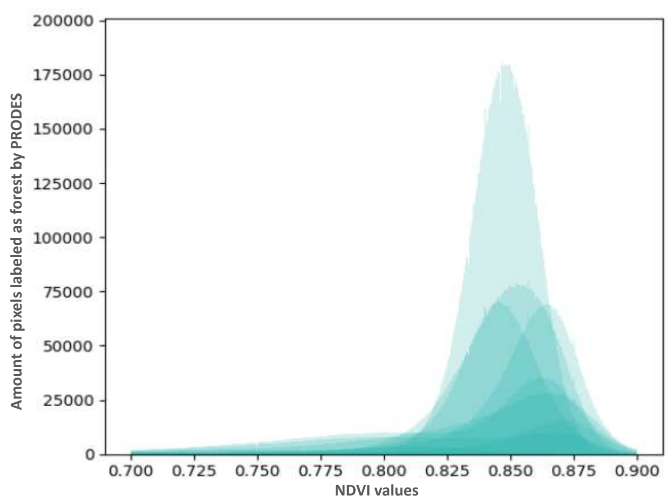

Figure 3. NDVI variance in areas identified as forest by PRODES.

Our *vegetation* correlation analysis was based on two approaches: vegetation correlation accuracy (VCA) and commission error (CE). VCA refers to the overlapped data between the target *vegetation* generated in ReSIIM and the forest reference data from PRODES, whilst CE refers to the misclassified data. For this analysis 10 random scenes were selected from the Brazilian Amazon.

The results were satisfactory, considering that average VCA was 93% and average CE was 7%. Nonetheless, it is remarkable that although VCA was higher than 92% for each scene, just in a riparian vegetation scene, VCA was 58% and CE, 42%. This outline suggests that the used threshold is not suitable for this kind of vegetation. Figure 4 presents this special case in four steps: A - original data; B - forest reference (PRODES); C - detected forest (NDVI >= 0.8); and D - final composition of the three aforementioned layers. PRODES data considers that both sides of the water body are characterized as forest. Even so, through NDVI threshold approach, this reality was not detectable. Along the bank of the river, known as riparian vegetation, NDVI values range mainly from 0.4 to 0.7.

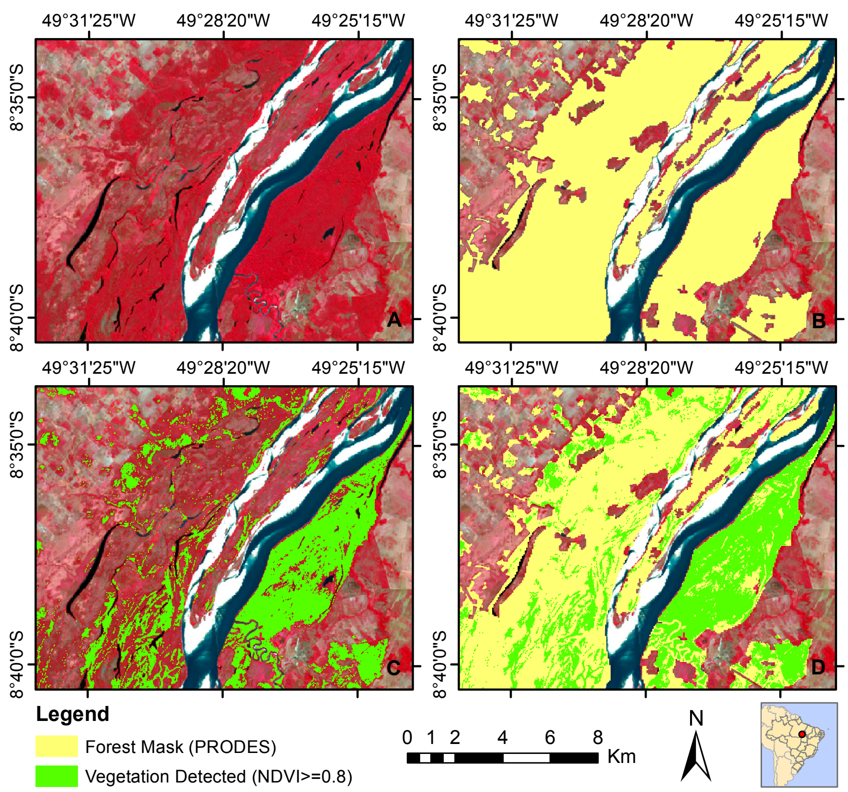

Figure 4. Outlier identified in NDVI threshold for *vegetation* target. A - Original Scene (color combination in R5G4B2 using L8 bands); B - Forest Mask (PRODES); C - Vegetation Detected (NDVI>=0.8); D - Input Image, Forest Mask, and Vegetation Detected.

3.2.3. Other Targets

Due to the lack of source data neither *cloud* nor *cloud shadow* correlation analysis were possible, as well as *clear land*.

No validation data for *bare soil* targets was also identified, thus we analyzed instead the main thresholds of NDVI found in the literature, aforementioned in Section 2. As highlighted by [Liang et al. 2014b], validation processes in this area of research are restricted due to limited data sources and methodologies.

Low NDVI values do not necessarily indicate a lack of vegetation, since the differences along vegetated areas through the year and seasons may interfere in this value [NOAA 2017]. However, in this work we employed the NDVI interval based on the literature. Therefore, we defined 0.0 <= NDVI <= 0.2 for *bare soil* target detection.

4. ReSIIM Results for Burnt Forest Detection

According to the required criteria to search for RS images in databases in burnt areas studies, some tests were performed. The tests were based only on the basic attributes available in the ReSIIM, focusing on phenomena in lieu of date or location. Summing up 60 Landsat scenes, representing almost 200 millions hectares and about 120 GB of data, two main datasets were selected to compose the database, reference data (RD) and noisy data (ND). Half of the database was composed of RD and the another half of ND. The aim of the tests is to understand what combination of metadata information is needed to maximize the RD generation and minimized ND.

Through the tests, we were able to comprehend better not only the available data, but also the role of ReSIIM in burnt forest detection. Firstly we assessed individual targets in order to support image retrieval. After that, the targets with best performance were combined and evaluated as well. An overview of the main tests is present in Table 1.

In test 1, the searching criterion was the presence of *bare soil* target in scenes not superior to 10%, yet, after analyzing the results of image retrieval from the target *cloud*, we identified that this target was misclassified as other targets such as *cloud* and *water*. For this reason, this target was not considered in the following steps.

Although clouds are crucial for image retrieval in burnt forest detection studies, tests 2 (percentage of *clouds* in scenes is not superior to 10%) and 3 (percentage of *clouds* is not superior to 20%) showed that it is not the only important searching criterion. That probably happens because in some areas of the Brazilian Amazon, it is not possible to access RS images without the presence of *clouds* along the year, due to local humidity. As well as *cloud*, in tests 4-6, it is possible to notice that *clear land* also plays an important role in image retrieval tests. Images with more than 30% of *clear land* (test 5) retrieved all the RD and none of the ND.

Shadow targets were not relevant in the process, since almost 100% of the dataset is composed of images with less than 10% *shadow* (test 7). We previously thought that *vegetation* targets would also be essential in the analysis. However, images with more than 10% of *vegetation* (test 8) are present in both RD and ND, and this is not a good searching criterion for this kind of work. The same was also identified in the target *water* (test 9). Some areas present water along the whole year, thus this kind of filter would be

Table 1. ReSIIM tests to support burnt forest detection (RD: Reference Data; ND: Noisy Data). Where *X* represents the percentage range criterion used to retrieve scenes.

Test Number	Searching Criteria	Percentage										Image Retrieval	
		10	20	30	40	50	60	70	80	90	100	RD	ND
1	*Bare Soil*	X										30	1
2	*Cloud*	X										27	0
3	*Cloud*	X	X									30	1
4	*Clear Land*	X	X	X								0	30
5	*Clear Land*				X	X	X	X	X	X	X	30	0
6	*Clear Land*					X	X	X	X	X	X	26	0
7	*Shadow*	X										30	29
8	*Vegetation*		X	X	X	X	X	X	X	X	X	12	17
9	*Water*	X										30	29
10	*Cloud*	X	X									30	0
	Clear Land				X	X	X	X	X	X	X		

more suitable if applied for more detailed and specific studies, such as along coastlines.

Finally, we analyzed the integration of *cloud* and *clear land* searching criteria (test 10), once those were the main targets identified in the aforementioned steps. The intersection between both was satisfactory, once all the images from RD and none of the ND were retrieved.

5. Conclusions

The Remote Sensing data deluge is overwhelming the capacity of institutions to manage and retrieve its content. In this context, ReSIIM is a fast and easy alternative tool for Remote Sensing image information mining. It is based on the application of well-known methods for information extraction from RS scenes, storing this information and allowing users to access land use and land cover metadata through open source software, scripts and libraries. The developed tool will support many other avenues of sustainable research, considering that it enables phenomena searching criteria in lieu of just location or date parameters, as available in current official catalogs.

ReSIIM applied to support burnt forest detection was satisfactory, retrieving all the reference data from the database. The crucial targets for our test application in burnt forest detection were *cloud* and *clear land*. More analyses are consequently required in order to combine different targets and Remote Sensing image retrieval results for further understanding of Earth phenomena, such as land use and land cover changes.

ReSIIM methodology accuracy is not absolute, since it is an indication of target correlations using mathematical models. However, ReSIIM can be continuously improved according to its user's demands. In this context, future research is also required in order to comprehend user's requirements.

Acknowledgement

We would like to thank National Council for Scientific and Technological Development (CNPq) for the research financial support.

References

Andere, L., Anderson, L., Duarte, V., Arai, E., Aragão, J., and Aragão, L. (2015). Dados multitemporais do sensor modis para o mapeamento de queimadas na amazônia. *XVII Simpósio Brasileiro de Sensoriamento Remoto. Anais... João Pessoa: INPE*.

Anderson, L. O., Aragão, L. E., Gloor, M., Arai, E., Adami, M., Saatchi, S. S., Malhi, Y., Shimabukuro, Y. E., Barlow, J., Berenguer, E., et al. (2015). Disentangling the contribution of multiple land covers to fire-mediated carbon emissions in amazonia during the 2010 drought. *Global Biogeochemical Cycles*, 29(10):1739–1753.

Anderson, L. O., Aragão, L. E. O. E. C. D., Lima, A. D., and Shimabukuro, Y. E. (2005a). Detecção de cicatrizes de áreas queimadas baseada no modelo linear de mistura espectral e imagens índice de vegetação utilizando dados multitemporais do sensor MODIS/TERRA no estado do Mato Grosso, Amazônia brasileira. *Acta Amazonica*, 35(4):445–456.

Anderson, L. O., Shimabukuro, Y. E., Defries, R. S., and Morton, D. (2005b). Assessment of deforestation in near real time over the brazilian amazon using multitemporal fraction images derived from terra modis. *IEEE Geoscience and Remote Sensing Letters*, 2(3):315–318.

Aragão, L. E., Anderson, L. O., Lima, A., and Arai, E. (2016). Fires in amazonia. In *Interactions Between Biosphere, Atmosphere and Human Land Use in the Amazon Basin*, pages 301–329. Springer.

Aragão, L. E. and Shimabukuro, Y. E. (2010). The incidence of fire in amazonian forests with implications for redd. *Science*, 328(5983):1275–1278.

Bannari, A., Morin, D., Bonn, F., and Huete, A. (1995). A review of vegetation indices. *Remote sensing reviews*, 13(1-2):95–120.

Bastarrika, A., Chuvieco, E., and Martín, M. P. (2011). Mapping burned areas from landsat tm/etm+ data with a two-phase algorithm: Balancing omission and commission errors. *Remote Sensing of Environment*, 115(4):1003–1012.

Cardozo, F. d. S., Pereira, G., Shimabukuro, Y. E., and Moraes, E. C. (2013). Análise do uso do Modelo Linear de Mistura Espectral (MLME) para o mapeamento das áreas queimadas no Estado de Rondônia no ano de 2010. In *XVI Simpósio Brasileiro de Sensoriamento Remoto*, pages 7265–7272, Foz do Iguaçu, PR.

Chuvieco, E. and Congalton, R. G. (1988). Mapping and inventory of forest fires from digital processing of tm data. *Geocarto International*, 3(4):41–53.

Datcu, M. and Seidel, K. (2002). An innovative concept for image information mining. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 84–99. Springer.

Datcu, M. and Seidel, K. (2003). Image information mining-exploration of earth observation archives. *Geographica Helvetica*, 58(2):154–168.

Datcu, M., Seidel, K., Pelizarri, A., Schroeder, M., Rehrauer, H., Palubinskas, G., and Walessa, M. (2000). Image information mining and remote sensing data interpretation. *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, 7(July):3057–3059.

FAO, F., of the United Nations, A. O., and ITTO, I. T. T. O. (2011). The State of Forests in the Amazon Basin, Congo Basin and Southeast Asia. Technical report, Rome.

Flood, N. and Gillingham, S. (2017). Pythonfmask documentation - release 0.4.4. `http://pythonfmask.org/en/latest/`. Accessed: 04 June 2017.

Frank, S. (1994). Cataloging digital geographic data in the information infrastructure: A literature and technology review. *Information Processing and Management*, 30(5):587–606.

Guptill, S. C. (1999). Metadata and data catalogues. In Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W., editors, *Geographical Information Systems: Principles and Technical Issues*, volume 1, chapter 49, pages 677–692. John Wiley & Sons, INC.

Guptill, S. C. (2015). Spatial data. In Wright, J. D., editor, *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, pages 126 – 129. Elsevier, Oxford, second edition edition.

INPE (2017). Projeto prodes: Monitoramento da floresta amazônica brasileira por satélite. `http://www.obt.inpe.br/prodes/index.php`. Accessed: 29 June 2017.

Jin, X., Zhang, Y.-K., Tang, Y., Hu, G., and Guo, R. (2014). Quantifying bare soil evaporation and its relationship with groundwater depth. *International journal of remote sensing*, 35(21):7567–7582.

Key, C. and Benson, N. (2006). General technical report - landscape assessment (la): Sampling and analysis methods. Technical Report RMRS-GTR-164-CD, USDA Forest Service.

Körting, T. S., Namikawa, L., Fonseca, L., and Felgueiras, C. (2016). How to effectively obtain metadata from remote sensing big data?

Laney, D. (2001). 3d data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6:70.

Lashof, D. (1991). The contribution of biomass burning to global warming: an integrated assessment. In *Global biomass burning. Atmospheric, climatic, and biospheric implications*.

Li, J. and Narayanan, R. M. (2006). Integrated Information Mining and Image Retrieval in Remote Sensing. In Chang, C. I., editor, *Recent Advances in Hyperspectral Signal and Image Processing*, chapter 16, pages 449—-478. Transworld Research Network, Trivandrum, India, 1 edition.

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A., et al. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:119–133.

Liang, S., Zhang, X., Xiao, Z., Cheng, J., Liu, Q., and Zhao, X. (2014a). *Longwave Emissivity*, pages 73–121. Springer International Publishing, Cham.

Liang, S., Zhang, X., Xiao, Z., Cheng, J., Liu, Q., and Zhao, X. (2014b). *Shortwave Albedo*, pages 33–72. Springer International Publishing, Cham.

Lima, A. (2013). *Influência da Cobertura da Terra na Extensão e Configuração Espacial de Áreas Queimadas em Anos de Seca Extrema na Amazônia Oriental*. PhD thesis, Tese de Doutorado. Instituto Nacional de Pesquisas Espaciais.

Namikawa, L. M. and Castejon, E. F. (2017). Mapas de lamina de Água para todo o brasil extraídos do rapideye. `http://wiki.dpi.inpe.br/doku.php?id=mapas:waterbodies`. Accessed: 02 June 2017.

Namikawa, L. M., Körting, T. S., and Castejon, E. F. (2016). Water body extraction from rapideye images: An automated methodology based on hue component of color transformation from rgb to hsv model. *Revista Brasileira de Cartografia*, 68(6).

NOAA (2017). Mgvi - normalized difference vegetation index. `http://www.ospo.noaa.gov/Products/land/mgvi/NDVI.html`. Accessed: 30 June 2017.

Padilla, M., Olofsson, P., Stehman, S. V., Tansey, K., and Chuvieco, E. (2017). Stratification and sample allocation for reference burned area data. *Remote Sensing of Environment*.

Quartulli, M. and Olaizola, I. G. (2013). A review of eo image information mining. *ISPRS Journal of Photogrammetry and Remote Sensing*, 75:11–28.

Rouse Jr, J., Haas, R., Schell, J., and Deering, D. (1974). Monitoring vegetation systems in the great plains with erts.

Shimabukuro, Y. E., Duarte, V., Arai, E., Freitas, R., Lima, A., Valeriano, D., Brown, I., and Maldonado, M. (2009). Fraction images derived from terra modis data for mapping burnt areas in brazilian amazonia. *International Journal of Remote Sensing*, 30(6):1537–1546.

Shimabukuro, Y. E., Miettinen, J., Beuchle, R., Grecchi, R. C., Simonetti, D., and Achard, F. (2015). Estimating burned area in mato grosso, brazil, using an object-based classification method on a systematic sample of medium resolution satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(9):4502–4508.

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(1):16–20.

Stepinski, T. F., Netzel, P., and Jasiewicz, J. (2014). Landex—a geoweb tool for query and retrieval of spatial patterns in land cover datasets. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(1):257–266.

Zhu, Z., Wang, S., and Woodcock, C. E. (2015). Improvement and expansion of the fmask algorithm: cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images. *Remote Sensing of Environment*, 159:269–277.

Zhu, Z. and Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in landsat imagery. *Remote Sensing of Environment*, 118:83–94.