

Effect of Revisit Frequency and Availability of Cloud-Free Observation on
Landsat Global Burned Area Mapping

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

with a

Major in Natural Resources

in the

College of Graduate Studies

University of Idaho

by

Andrea Melchiorre

Major Professor: Luigi Boschetti, Ph.D.

Committee Members: Andrew Hudak, Ph.D.; David Roy, Ph.D; Lee Vierling, Ph.D.

Department Administrator: Charles Goebel, Ph.D.

August 2020

Authorization to Submit Dissertation

This dissertation of Andrea Melchiorre, submitted for the degree of Doctor of Philosophy with a Major in Natural Resources and titled "Effect of Revisit Frequency and Availability of Cloud-Free Observation on Landsat Global Burned Area Mapping," has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: _____ Date: _____
Luigi Boschetti, Ph.D.

Committee Members: _____ Date: _____
Andrew Hudak, Ph.D.

_____ Date: _____
David Roy, Ph.D.

_____ Date: _____
Lee Vierling, Ph.D.

Department

Administrator: _____ Date: _____
Charles Goebel, Ph.D.

Abstract

Fire is one of the most relevant disturbances affecting terrestrial ecosystems globally, altering vegetation, soil, water, and atmospheric composition. Fire causes a non-permanent land cover change, through the removal of vegetation, the deposition of charcoal and ashes, and the exposure of soil; the temporal persistence of these changes is highly variable, ranging from a few weeks in tropical savannas to years in boreal forests. Global burned area products have been systematically generated in the past 20 years from several coarse spatial resolution (250 m - 1 km) Earth Observation (EO) systems. These products are the main input in global biomass burning atmospheric emission inventories, and in the most recent studies on the role of fire in the global carbon cycle and vegetation dynamics.

Because of the non-permanent nature of burned areas, the algorithms employed for the generation of global burned area products rely on the availability of daily or near-daily observations from coarse resolution EO systems. The high revisit frequency ensures that a sufficient number of cloud-free observations are generally available globally before burned areas disappear, with few exceptions in known locations of persistent cloud cover.

The systematic generation of moderate spatial resolution (10 m - 30 m) burned area products could potentially meet the needs of a variety of fire science and applications communities, and at different scales from global (e.g., pyrogenic carbon emissions estimation) to regional scale (e.g., environmental post-fire assessment and remediation decision support). Algorithms for the generation of moderate resolution burned area maps have been recently prototyped regionally and continentally, and have the potential for global implementation. However, moderate resolution sensors have reduced temporal resolution (e.g., 16 days for Landsat) compared to coarse resolution sensors (e.g., ~1 day for MODIS), which could potentially lead to omission errors in ecosystems where the spectral signal associated with burning events disappears quickly, and cloud cover limits the number of valid observations.

My dissertation focuses on estimating the combined effect of the impermanent nature of land cover change typical of burning events and the cloud cover, which reduces the number of valid observations available to detect burns, on global burned area mapping using Landsat data. The dissertation has three objectives. The first objective (Chapter 2) is to estimate the temporal persistence time of the signal associated with burned areas, stratified by ecosystem and land cover type, making use of the global, multiyear MODIS data record. The second objective (Chapter 3) is to evaluate the suitability of the MODIS-derived cloud mask as a proxy for Landsat 7 cloud observations. Finally, the

third objective (Chapter 4) is to estimate the potential omission errors in a hypothetical global Landsat burned area product, due to the combined effect of reduced revisit frequency and cloud contamination.

Chapter 2 presents a global analysis of the burned area persistence time defined as the duration of the spectral separability of the burned / unburned areas mapped by the MODIS Global Burned Area Product (MCD64). The separability was computed by analyzing time series of normalized burn ratio (NBR) from nadir BRDF-adjusted MODIS reflectances (MCD43). Results showed that, globally, the median burned area persistence time was estimated as 29 days and 86.6% of the global area, as detected by MODIS, can be detected accurately only for up to 48 days. Furthermore, the results indicated that early and late fires had a shorter persistence time compared to fires burning in the central portion of the fire season. The results, therefore, indicate that the persistence time can be a limiting factor for mapping burned areas using moderate resolution satellite sensors, which have a low temporal resolution (e.g. Landsat 16 days, Sentinel 2A and 2B 10 days each, 5 days when used in combination).

Chapter 3 presents a comparison of Landsat and MODIS cloud data. Landsat 7 Enhanced Thematic Mapper Plus (ETM+) image cloud fractions over land were compared with collocated MODIS cloud fractions, generated by combining the MODIS-Terra global daily cloud mask product (MOD35) with the Landsat 7 ETM+ image footprints and acquisition calendar. The results showed high correlation between the MODIS and Landsat 7 ETM+ cloud fractions ($R^2 = 0.83$), negligible bias (median difference: < 0.01) and low dispersion around the median (inter-quartile range: $[-0.02, 0.06]$). These results indicated that, globally, the cloud cover detected by MODIS Terra data can be used as a proxy for Landsat 7 ETM+ cloud cover at the Landsat World Reference System (WRS) scale.

Chapter 4 builds on the previous chapters and presents the potential omission error of a hypothetical Landsat global burned area product compared to the MODIS global burned area product. The Landsat omission error was estimated as the amount of burned area detected by MODIS that would not be detected by Landsat 7 because of the combined effect of the impermanent spectral signal associated with burned areas and missing observations due to cloud cover. The simulation was informed by the MODIS global burned area product (MCD64A1), used as fire mask to define the location and timing of burning, and the MODIS-Terra cloud product (MOD35), used to determine the number of post-fire cloud-free observations available following the Landsat 7 acquisition calendar and ground swath footprints. Globally, the resulting omission error was estimated as 19% of the average annual burned area detected by MODIS, with a maximum error over forest land cover (33%) and minimum over shrubland land cover (5%). The results were derived using the acquisition calendar of Landsat 7 only, however, thanks to the aggregation of data from over 15 years of acquisitions, the results can be extended to the other existing Landsat sensors, which are positioned on the same orbit shifted by an 8-

days lag, and also to Landsat 9, which is planned to be launched in the same orbit of Landsat 7 by Spring 2021.

The findings of this research have implications for the future development of a global burned area product generated using moderate resolution EO data such as Landsat. The burned area persistence times provide an estimation of the period after the burning date in which burned areas can be mapped reliably and have implications on the length of the rolling periods, used in change detection algorithms to map burned areas. The potential omission error of a Landsat burned area product identified locations and times of the year in which the low revisit frequency of Landsat combined with the occurrence of clouds can have degrading effects on Landsat burned area maps accuracy.

Acknowledgements

I wish to thank my Major Professor Dr. Luigi Boschetti for the support and guidance I received in these years and all the members of my Ph.D. committee, Dr. Andrew Hudak, Dr. David Roy and Dr. Lee Vierling, their advice was vital for the development of my degree, my articles and this dissertation. I also wish to thank all the fellow graduate students and friends of the College of Natural Resources I had the opportunity to meet for the constructive discussions, cooperation and reciprocal help. Finally, I wish to thank the NASA Earth and Space Science Fellowship that supported financially my work.

Dedication

I dedicate this manuscript to my family, to my old friends, and all the new friends I had the fortune to meet in this long path that connects the chaotic narrow streets of Napoli to the isolated and vast buttes and valleys of the Palouse.

Table of Contents

Authorization to Submit Dissertation.....	ii
Abstract.....	iii
Acknowledgements.....	vi
Dedication.....	vii
Table of Contents.....	viii
List of Figures.....	xi
List of Tables.....	xvi
Chapter 1: Introduction.....	1
1.1. Fire in the earth system.....	1
1.2. Earth observation burned area maps.....	2
1.3. Objective and goals.....	5
1.4. References.....	7
Chapter 2: Global Analysis of the Burned Area Persistence Time With MODIS Data	14
2.1. Abstract.....	14
2.2. Introduction.....	14
2.3. Data.....	17
2.3.1. <i>MODIS global burned area product</i>	17
2.3.2. <i>MODIS global nadir BRDF-adjusted reflectance product</i>	18
2.3.3. <i>MODIS land cover product</i>	18
2.3.4. <i>Terrestrial ecoregions of the worlds</i>	18
2.4. Methods.....	18
2.4.1. <i>Theoretical basis: definition of the burned area persistence time</i>	18
2.4.2. <i>Global implementation</i>	23
2.5. Results.....	32
2.5.1. <i>Fire activity distribution</i>	32
2.5.2. <i>Ecoregion level burned area persistence time</i>	34
2.5.3. <i>Biome/Realm aggregation of the burned area persistence time</i>	36
2.6. Discussion.....	42
2.6.1. <i>Ecoregion level burned area persistence time</i>	42
2.6.2. <i>Biome/Realm level burned area persistence time</i>	43
2.7. Conclusions.....	47
2.8. References.....	48

Chapter 3: Global Evaluation of the Suitability of MODIS-Terra Detected Cloud Cover as a Proxy for Landsat 7 Cloud Conditions.....	61
3.1. Abstract	61
3.2. Introduction	61
3.3. Data	63
3.3.1. <i>Landsat cloud data</i>	63
3.3.2. <i>MODIS cloud mask and geolocation product</i>	64
3.3.3. <i>Spatial and temporal extent of the analysis</i>	64
3.4. Methods.....	65
3.4.1. <i>Computation of MODIS cloud fractions for each Landsat image</i>	65
3.4.2. <i>MODIS-Landsat cloud fraction comparison methodology</i>	65
3.5. Results	67
3.5.1. <i>Landsat 7 ETM+ image availability and global cloud histograms</i>	67
3.5.2. <i>MODIS and Landsat cloud fraction differences</i>	68
3.5.3. <i>Regression analysis</i>	69
3.6. Discussion	76
3.7. Conclusions	78
3.8. References	78
Chapter 4: Potential Landsat Global Burned Area Omission Error due to Reduced Revisit Frequency and Cloudy Observations	83
4.1. Abstract	83
4.2. Introduction	84
4.3. Data	86
4.3.1. <i>MODIS global burned area product</i>	86
4.3.2. <i>MODIS land cover product</i>	86
4.3.3. <i>MODIS cloud mask</i>	86
4.3.4. <i>Terrestrial ecoregions of the world map</i>	88
4.3.5. <i>Burned area persistence time</i>	89
4.4. Methods	89
4.4.1. <i>Spatial stratification</i>	90
4.4.2. <i>Burned area omission error estimation</i>	91
4.4.3. <i>Biome/Realm level aggregation of the results</i>	93
4.5. Results	93
4.5.1. <i>Burned area omission error at the ecoregion scale</i>	93

4.5.2. <i>Burned area omission error at the Realm/Biome scale</i>	95
4.6. <i>Discussion</i>	99
4.6.1. <i>Realm/Biome scale</i>	99
4.7. <i>Conclusions</i>	100
4.8. <i>References</i>	101
Chapter 5: Conclusion and Future Work	109
5.1. Summary and conclusion	109
5.2. Limitations and future research	110
5.3. References	114
Appendix A: Publishing Rights & Permissions	119

List of Figures

- Figure 2-1: Illustration of the method for the estimation of the burned area persistence time. The plot is generated using observations of a set of burned pixels (ba) and unburned pixels (ub) in the Savanna land cover of the “Victoria Plains tropical savanna” ecoregion in northern Australia, starting on the day of detection of the burned pixels ($jday=329$, November 25th). The red dotted lines show the distribution of dNBR of the burned pixels ($PB(dNBR)\Delta t, jday$), observed at $\Delta t = 10, 20, 30$ and 40 days after burning; the blue dotted lines show the distribution of dNBR of the unburned pixels ($PUB(dNBR)\Delta t, jday$) observed at the same time. The red and blue solid lines represent respectively the 10th percentile of the dNBR distribution of burned pixels (i.e., $FB\Delta t jday$) and the 90th percentile of the dNBR distribution of the unburned pixels (i.e., $FUB\Delta t jday$). The persistence time is calculated as the time $\Delta t * jday$ when the two distributions overlap by more than 20%, i.e., when $FB\Delta t jday$ is equal to $FUB\Delta t jday$ ($\Delta t * jday = 31$ days in this example)..... 21
- Figure 2-2: Illustration of the procedure for the estimation of the burned area persistence time. The input data for the analysis are the MODIS burned area and NBAR product. The MODIS land cover product and the TEOW ecoregions are used to define each spatial stratum defined by ecoregion and land cover. The estimated burned date extracted from the MODIS burned area product are used to allocate the pixels in the different 16-days periods (Section 3.2.1). The median persistence time for each ecoregion and land cover is computed as the median value of the average annual burned area as a function of the persistence time (Section 3.2.3). 23
- Figure 2-3: Illustration of the procedure for the definition of the fire season, showing the burned area histogram, binned in 16-days intervals t , of three representative examples of fire season. Left, unimodal fire season fully contained in a single calendar year (“Angolan woodland” ecoregion in Southern Africa); center, bimodal fire season (“Victoria Plains tropical savanna” ecoregion in northern Australia); right, unimodal fire season spanning across calendar years (“Northern Congolian tropical forest-savanna mosaic” ecoregion in northern Africa). The fire season is identified by ranking the 16-days bins in decreasing order by burned area and including in the fire season the bins necessary to reach 90% of the burned area (shown in red). The 16-days periods outside the fire season (shown in blue) were not considered in the analysis. 26
- Figure 2-4: Burned area persistence time as a function of the day of burning $\Delta t Eco, LC, t *$ (*top*) and average annual burned area for each 16-days period of the year (*bottom*), in the “Victoria Plains tropical savanna” ecoregion in Northern Australia, stratified by aggregated land cover class. The persistence time was defined only for the 16-days periods of the fire season (hatched gray area). The Forest land

cover class was not considered in this ecoregion because of negligible fire activity (i.e., less than 250 km ² over the entire study period).	28
Figure 2-5: Summary persistence time metrics at the ecoregion level, in the “Victoria Plains tropical savanna” ecoregion in Northern Australia, stratified by aggregated land cover class. <i>Top</i> : Histogram of $BAEco, LC(\Delta t^*)$, i.e., the average annual burned area as a function of the burned area persistence time. The histogram is computed using 16 day bins. <i>Bottom</i> boxplot of $BAEco, LC(\Delta t^*)$, showing the median (i.e., $\Delta tEco, LC^*$), interquartile range, minimum and maximum of the distribution.	29
Figure 2-6: Realms and aggregated biomes used as spatial units to present the global summary metrics persistence time. Oceania and Antarctic realms (in gray) were not considered, because of negligible fire activity.	31
Figure 2-7: Global burned area distribution, as detected by the MODIS global burned area product (MCD64A1) in the 2003-2016 period, stratified by aggregated land cover class. <i>Left</i> : histogram of the partition of the average annual burned area by biome and realm. <i>Right</i> : average annual burned area detected in each ecoregion. The ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.	33
Figure 2-8: Histogram (<i>Left</i>) and map (<i>Right</i>) of the fire season duration, computed for the ecoregions considered in the analysis. The ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.	34
Figure 2-9: Burned area persistence time calculated for each ecoregion, stratified by aggregated land cover class: Forest (<i>top row</i>), Shrubland (<i>middle row</i>), and Grassland & Savanna (<i>bottom row</i>). The median persistence time $\Delta tEco, LC^*$ (<i>left column</i>) and interquartile range (<i>right column</i>) are shown using a rainbow color scale; the ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.	35
Figure 2-10: Histogram bars and cumulative lines of the estimated persistence time $\Delta tEco, LC^*$ calculated for all the ecoregions and stratified by aggregated land cover class. The histograms were computed using 16-days bins, up to 96 days after the date of burn.	36
Figure 2-11: Histograms of $BAB, R, LC(\Delta t^*)$ in the Forest aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.	38
Figure 2-12: Boxplots of $BAB, R, LC(\Delta t^*)$ in the Forest aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\Delta tB, R, LC^*$), interquartile range, minimum and maximum of the distribution.	38

Figure 2-13: Histograms of $BAB, R, LC(\Delta t^*)$ in the Shrubland aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.....	39
Figure 2-14: Boxplots of $BAB, R, LC(\Delta t^*)$ in the Shrubland aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\Delta tB, R, LC^*$), interquartile range, minimum and maximum of the distribution.	40
Figure 2-15: Histograms of $BAB, R, LC(\Delta t^*)$ in the Grassland & savanna aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.	41
Figure 2-16: Boxplots of $BAB, R, LC(\Delta t^*)$ in the Grassland & savanna aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\Delta tB, R, LC^*$), interquartile range, minimum and maximum of the distribution.	41
Figure 3-1: Spatial distribution and number of Landsat 7 ETM+ images acquired from 1 January to 31 December 2002 between 70° S and 70° N. The number of images for each Landsat WRS path/row is depicted with a rainbow color scale.	67
Figure 3-2: Histograms of Landsat 7 (top) and MODIS (bottom row) cloud fractions, observed in all the 2002 study data (Figure 3-1). The histograms are quantized using 100 bins of equal width. The three MODIS cloud fraction histograms are for the three MODIS cloud definitions (Table 3-1).	68
Figure 3-3: Box and whisker plots summarizing the distribution of the difference between the MODIS and Landsat 7 cloud fractions for all the global 2002 data (Figure 3-1) considering the three MODIS cloud fraction definitions (Table 3-1). The thick vertical lines show the median of the distribution, the boxes show the interquartile range, and the whiskers show the 5 th to 95 th quantile range.	68
Figure 3-4: Scatterplot of the Landsat 7 cloud fraction plotted against the MODIS cloud fraction, considering all the 2002 study data, plotted for each MODIS cloud fraction definition (Table 3-1). The logistic regression (solid) linear regression (dotted) and the identity line (dashed) are shown superimposed on point density distributions that are generated using a 100 × 100 quantization of the axes, and are displayed with a rainbow logarithmic color scale.....	69
Figure 3-5: Scatterplot of the regression model residuals, defined as the observed minus the predicted Landsat 7 cloud fraction, plotted against the MODIS cloud fraction for the linear (top row) and the logistic regression model (bottom row) considering the 3 MODIS cloud fraction definitions (Table 3-1). The median (+) and interquartile range (x) of the residuals are shown for four MODIS cloud fraction bins: [0, 0.25], [0.25, 0.50], [0.5, 0.75], and [0.75, 1].	71

- Figure 3-6: Spatial distribution of the minimum (*top row*), median (*middle row*) and maximum (*bottom row*) residuals observed at each WRS path/row location for the linear (*left column*) and logistic (*right column*) regression using the most conservative MODIS cloud definition (#1). The residuals are depicted using the same rainbow color scale for easier comparison. 73
- Figure 3-7: Box and whisker plots summarizing the distribution of the median residual of the linear and logistic models at each WRS path/row location (i.e., Figure 3-6, middle row). The thick vertical lines show the median of the distribution, the boxes show the interquartile range, and the whiskers show the 5th- 95th quantile range. 73
- Figure 3-8: Scatterplot of the linear (left) and logistic (right) regression residuals plotted against the land fraction in each Landsat path/row, considering all the 2002 study data, for the most conservative MODIS cloud fraction definition (#1). The median (+), interquartile range (x) and 5th- 95th quantile range (°) of the residuals are shown for 0.05 land fraction bins. The point density distributions are generated using a 100 × 100 quantization of the axes, and are displayed with a rainbow logarithmic color scale. 74
- Figure 3-9: Scatterplot of the linear (top row) and logistic (bottom row) regression residuals divided between the northern (left column) and southern hemisphere (right column) plotted against the Landsat acquisition day of year, considering all the 2002 study data, plotted for the most conservative MODIS cloud fraction definition (#1) (Table 3-1). The point density distributions are generated using a 25 x 25 quantization of the axes, and are displayed with a rainbow logarithmic color scale. It should be noted that the artifact visible in all plots for days 75-90 is due to a gap from March 19th to 28th in the MODIS Terra data acquisition. The median (+), interquartile range (x) and the 5th- 95th quantile range (°) of the residuals are shown for quarterly acquisition intervals, defined as {[Jan - Mar], [Apr - Jun], [Jul - Sep], [Oct - Dec]}..... 75
- Figure 4-1: MODIS-Terra ground swaths for 03/21/2016 (red) and the daily overlapping Landsat 7 ETM+ ground swaths within 30 minutes (green). The ground swaths were generated using the Committee on Earth Observation System (CEOS) COVE Tool (Kessler et al. 2013). The figure illustrates that the nadir (< 8°) portion of the MODIS-Terra swaths always coincides with the Landsat 7 ETM+ swaths and there is a 30 minutes lag between Terra and Landsat 7 overpasses (Mercury et al. 2012)..... 87
- Figure 4-2: Map delineating the borders of the realm (red lines) and the aggregated biomes (color-coded). Oceania and Antarctic realms were not considered for this study. From Melchiorre and Boschetti (2018). 88
- Figure 4-3: TEOW map used to stratify the analysis. Ecoregions considered in the analysis are random color-coded, ecoregions with negligible fire activity were discarded and are shown in light grey..... 90

Figure 4-4: MOD35-L7 cloud mask time series extraction over each burned pixel $X_{B,j}$. The time series started on the day of burning j and lasted $\Delta t^*_{Eco,LC,j}$ days; i.e., the burned area persistence time. MOD35-L7 cloud data observed on Landsat overpass days are extracted for the whole duration of the period considered..... 91

Figure 4-5: Schematic representation of the cloud-free observations count. At each location $X_{B,j}$ classified as burned in the MCD64A1 product on Day Of Year (DOY) j , cloud mask observations were extracted following the Landsat 7 ETM+ geometry and acquisition calendar within the time series defined by the day of burning and the persistence time. The burned pixel is classified as detected from Landsat if at least one cloud-free acquisition is retrieved (green pixels) or omitted if all the acquisitions extracted are flagged as cloudy (red pixel)..... 92

Figure 4-6: Example of Landsat omission error estimate from a set of burned pixels extracted from the TEOW Kimberly Tropical Savanna, Northwest Australia, for the Grassland & Savanna land cover. The graph shows the average annual burned area calculated as a function of the number of cloud-free (NCF) observations. The estimated omission error is the average annual burned area with no cloud-free observations (NCF = 0). 93

Figure 4-7: Potential average annual burned area omission error of Landsat 7 with respect to the MCD64A1 global burned area product (left column) and relative percentage (right column) for each aggregated land cover: Forest (top row), Shrubland (middle row) and Grassland & Savanna (bottom row). Ecoregions excluded from the analysis are depicted in gray for geographic reference. 94

Figure 4-8: Average annual burned area detected in Forest aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm..... 96

Figure 4-9: Average annual burned area detected in Shrubland aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm..... 97

Figure 4-10: Average annual burned area detected in Grassland & Savanna aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm. 98

List of Tables

Table 2-1: Land cover aggregation scheme of the MODIS MCD12Q1 IGBP land cover classes used for the present study	24
Table 2-2: Biomes aggregation scheme of the Olson et al. (2001) biomes used for the present study.	30
Table 3-1: The three MODIS cloud fraction definitions derived by different combinations of MODIS cloud mask (MOD35) detection confidence states.....	65
Table 3-2: Regression results of the Landsat 7 and MODIS cloud fraction scatterplots shown in Figure 3-4	69
Table 3-3: Median, 25 th and 75 th quantiles of the residuals for the linear (top) and logistic (bottom) models. The residuals were calculated globally and divided in MODIS definition #1 cloud fraction intervals	72
Table 3-4: Median, 5 th , 25 th , 75 th and 95 th quantiles of the residuals for the linear (top) and logistic (bottom) models. The residuals were calculated globally and divided in quarterly acquisition date intervals.	76
Table 4-1: Land cover aggregation scheme of the MODIS MCD12Q1 IGBP land cover classes used in this study. From Melchiorre and Boschetti (2018).....	89
Table 4-2: Average annual burned area omission error for each of the aggregated Biomes and land cover considered.....	98

Chapter 1: Introduction

1.1. Fire in the earth system

Fire is one of the most important disturbances for terrestrial ecosystems altering vegetation, soil, water and atmospheric composition (Andela et al. 2017; Bowman et al. 2009). Fire disturbances are included in global vegetation dynamics models, surface energy budget accounting and global emission inventories (Langmann et al. 2009; Mouillot et al. 2014; van der Werf et al. 2017). Although global, fire activity varies greatly depending on the ecosystem in terms of fire frequency and fire regime (Archibald et al. 2013; Skinner and Chang 1996) which, in the course of the centuries, have been altered by human activities (Bowman et al. 2011). On average $\sim 3.5 - 4.5 \cdot 10^6$ km² of burned area are detected each year globally by satellite burned area products, although the global estimations vary depending on the product considered (Andela et al. 2017; Humber et al. 2018). The majority of burning events are concentrated in tropical savannas and other ecosystems characterized by alternating dry and rainy seasons (Archibald et al. 2013; Moritz et al. 2012) where fine fuel and rapid growth of vegetation sustains frequent burning events (Archibald et al. 2018; Krawchuk and Moritz 2011).

Fire activity is limited by a combination of fuel, climate, and ignition sources (Moritz et al. 2012). Fuel conditions such as fuel moisture influence fire activity in subtropical/tropical biomes with mid-high net primary productivity whereas it has a lower influence in deserts, xeric shrublands, or grasslands and savannas where fuel load largely limits fire activity and the antecedent wet growing seasons shapes the fire activity in the following fire season (Krawchuk and Moritz 2011). In temperate and boreal biomes, fire activity has been linked to early snowmelt and precipitation amount in the winter months due to their influence on fuel load and conditions (Westerling et al. 2006). Climate influences fire activity directly, driving intra- to inter-annual variability in fuel moisture, and fine-fuel production through the growth of annual grasses (Littell et al. 2009), and indirectly influencing vegetation dynamics, and thus fuel type and fuel loading on large scale (Pausas and Fernández-Muñoz 2012). Human activities such as land conversion and agriculture practices alter fire occurrence and timing at small scales (Hall et al. 2016) whereas, at large scales, they alter the fire regimes (Bowman et al. 2011).

Depending on fuel, climate, and ignition sources, fire disturbances effects on vegetation are highly heterogeneous and have been quantified using several metrics and parameters such as fire intensity and burn severity (Keeley 2009). Measurable post-fire effects can be spatially ranging from

single burned perimeter scales to continental / global scales, which make remotely sensed data widely applied resources for fire science and management (Lentile et al. 2006). In particular, spaceborne Earth Observation (EO) data have greatly enhanced the possibility of including fire disturbances for a wide variety of global applications including studies on the global carbon cycle (Bowman et al. 2009) where fire activity information is used to model global vegetation distribution and succession (Goetz et al. 2012; Thonicke et al. 2001). Global burned area estimates, together with combustion completeness, biomass load and emission factors (Seiler and Crutzen 1980), are primary inputs in biogeochemical models used for the estimation of global emissions of greenhouse gases and aerosols from biomass burning (van der Werf et al. 2017). Finally, several studies have used post-fire albedo observations to investigate fire-induced vegetation change influences on radiative forcing and local effects on climate (Jin et al. 2012; Jin and Roy 2005; Lyons et al. 2008; Randerson et al. 2006).

1.2. Earth observation burned area maps

Satellite remote sensing of burned area is mainly based on passive optical wavelength remotely sensed measure of reflected and emitted surface electromagnetic radiation captured by specific wavelength band-pass calibrated sensors. The spectral radiance sensed at wavelengths spanning from $\sim 0.3 \mu\text{m}$ to $\sim 3.5 \mu\text{m}$ is due to reflected radiations from the surface and it can be converted to the spectral reflectance, i.e., the ratio of reflected to incident radiation (Roy et al. 2010). Spectral reflectance measures ability to discriminate between burned and unburned pixels was examined for a variety of satellite sensors such as Landsat (Chuvieco and Congalton 1988; Key and Benson 2002), AVHRR (Pereira et al. 1999), MODIS (Roy et al. 2005), and more recently Sentinel 2 (Huang et al. 2016). Spectral reflectance in the Near and Shortwave Infrared wavelength bands is more sensitive to the land cover change caused by fire events and spectral indices based on these bands, such as the Normalized Burn Ratio (NBR) (Key and Benson 2002), were designed to improve the detection of burned areas.

A wide variety of algorithms for the systematic production of global burned area maps have been tested and developed based on the temporally persistent spectral reflectance changes before and after the fire event using different satellite dataset. Global burned area maps have been produced in the past 20 years from several coarse spatial resolution (250 m - 1 km) Earth Observation (EO) systems, including MERIS, Terra and Aqua MODIS, and SPOT-VGT (Alonso-Canas and Chuvieco 2015; Giglio et al. 2018; Giglio et al. 2009; Roy et al. 2005; Simon et al. 2004; Tansey et al. 2004; Tansey et al. 2008). Thanks to the significant effects of fire disturbances on terrestrial ecosystems and atmospheric processes, burned area maps are included in the list of the required Essential Climate Variables (ECV), defined by the World Meteorological Organization (WMO) as “*a physical, chemical*

or biological variable or a group of linked variables that critically contributes to the characterization of Earth's climate” (GCOS 2011) and are a primary variable for modeling fire disturbances in support of the work of the United Nations Framework Convention on Climate Change (UNFCCC) and the Intergovernmental Panel on Climate Change (IPCC) (GCOS 2011). Global coarse spatial resolution burned area maps improved the global long-term estimates of fire activity trends (Andela et al. 2017), however, total area burned, location, and timing of burning greatly vary between the different EO products available, especially at regional scales (Humber et al. 2018).

Among others, the spatial resolution of EO data is a source of uncertainties in the currently available global burned area products because coarse spatial resolution leads to a trade-off between omission and commission errors on partially burned pixels (Boschetti et al. 2004). Pixels classified as burned are considered completely burned, which is often not the case. Due to the heterogeneous nature of fire, sub-pixel unburned patches may exist and result in overestimations of the burned area. Conversely, partially burned pixels can be undetected, resulting in underestimation of the burned area, especially in areas where small fires are common (Boschetti et al. 2019; Randerson et al. 2012; van der Werf et al. 2017). Different validation exercises evaluated the absolute uncertainties of EO burned area products comparing the burned area maps with collocated reference burned area maps generated independently from two or more moderate spatial resolution images (10 m – 30 m) (Boschetti et al. 2019; Padilla et al. 2014). The independent reference maps are a result of semi-automatic classification algorithms refined by visual interpretation and should have minimal error (Boschetti et al. 2006; Padilla et al. 2014; Padilla et al. 2015). A high negative bias (~50%) was found in the global burned area estimate of the MODIS burned area product (MCD64A1) caused by the systematic under-detection of small fires (Boschetti et al. 2019; Giglio et al. 2018; Roy et al. 2019). Consequently, there is a strong need for moderate resolution (10 - 30 m) global burned area maps for fire and ecosystem management, improved emission estimations, and carbon accounting (Hyer and Reid 2009; Mouillot et al. 2014; Randerson et al. 2012).

Automated mapping algorithms using Landsat and Sentinel data are currently being prototyped at continental scales (Boschetti et al. 2015; Hawbaker et al. 2017; Roteta et al. 2019; Roy 2015; Roy et al. 2019). Higher spatial resolution data have the benefits of increasing the detections of burned area (Roteta et al. 2019; Roy et al. 2019), reducing the occurrence of partially burned pixels and increasing the spectral separation of burned / unburned areas (Huang et al. 2016), but one of the challenges in generating a moderate resolution global burned area product is the limited revisit frequency of the sensors. Due to their narrow field of view (180 km – 300 km), moderate resolution sensors are characterized by low revisit frequency (10 - 16 days) which can translate into burned area potential

omission errors (Boschetti et al. 2015; Hawbaker et al. 2017) and can exacerbate the issues related to designing a robust and effective detection algorithm for mapping burned areas globally because the effects of fire on the vegetation spectral signal are not permanent (Melchiorre and Boschetti 2018). The persistence of the spectral signal of a burned area varies greatly across different ecosystems depending on climate and fuel burned. In the days or weeks after a fire, depending on vegetation type, phenology and climate, charcoal and ashes are removed and vegetation recovers (Chu et al. 2016; Fraser et al. 2000; Pereira et al. 1997; Solans Vila and Barbosa 2010; Trigg and Flasse 2000). For example, annual grasses have the physiological capacity to regrow lost biomass quickly enough to sustain short fire return intervals (i.e., every 1–3 years) (Archibald et al. 2013), and the rapid regrowth will shorten the period in which the spectral reflectance sensed by optical instruments can be associated with burned areas.

Previously published *in situ* spectral radiometer measurements indicate that the period of detectability of burned areas ranges from about two weeks in African savannas (Frederiksen et al. 1990; Langaas and Kane 1991; Trigg and Flasse 2000) to more than two years in boreal forests (Fuller and Rouse 1979). Similarly, studies using post-fire satellite observations indicate that the spectral changes induced by fire disappear rapidly in grasslands and savannas (Bowman et al. 2003; Pereira 2003; Trigg and Flasse 2000) but persist longer in forested ecosystems (Chen et al. 2011; Chu and Guo 2014; Röder et al. 2008).

Additionally to the low revisit frequency, clouds, smoke, and other optically thick aerosols further limit the number of valid acquisitions within a defined period (Roy et al. 2008; Smith and Wooster 2005). Coarse resolution global burned area products rely on the availability of daily or near-daily observations (Chuvieco et al. 2019) and it is generally assumed that a sufficient number of cloud-free observations are available globally to map burned areas, with few exceptions in known locations of persistent cloud cover such as Indonesia and tropical areas in Brazil (Giglio et al. 2009), or at high latitudes locations such as Siberia (Chu and Guo 2014). Burned area mapping algorithms developed for moderate resolution sensors also generally detect temporal spectral signal changes due to fire and require at least one cloud-free acquisition before and after the fire event. However, the combined effect of the impermanent nature of burned areas, lower revisit frequency and cloud cover was not analyzed on global burned area mapping at moderate spatial resolution.

1.3. Objective and goals

The systematic generation of moderate spatial resolution burned area products would improve the understanding of fire's effects for a variety of applications. For example, for improved estimates of carbon emissions in the atmosphere due to biomass burning (van der Werf et al. 2017), for an improved assessment of fire feedbacks on global vegetation dynamics (Bowman et al. 2015), for improved environmental post-fire assessment and remediation decision support (Mouillot et al. 2014; Trigg and Roy 2007), and to improve land cover / land use change research (Archibald et al. 2009; Hantson et al. 2015). However, the impermanent nature of the spectral signal associated with burned areas, combined with the lower revisit frequency of moderate resolution satellites (10 m – 30 m) and the effect of clouds, poses a major challenge for designing a global burned area mapping algorithms from moderate resolution satellites (Boschetti et al. 2015; Roy et al. 2019).

My dissertation focuses on estimating the combined effect of the impermanent nature of land cover change typical of burning events and the cloud cover, which reduces the number of valid acquisitions sensed to detect burns, on global burned area mapping using Landsat data. The dissertation pursues three main objectives: the first objective is the analysis of the temporal persistence time of the signal associated with burned areas, stratified by ecosystem and land cover type, making use of the global, multiyear MODIS data record. The second objective is to evaluate the suitability of the MODIS-derived cloud mask as a proxy of Landsat 7 cloud observations. The third objective is to estimate the potential omission errors in a hypothetical global Landsat burned area product compared to the MODIS burned area product, generated using daily observations, due to the combined effect of reduced revisit frequency and cloud contamination, considering the limited persistence time of burned areas. The acquisition calendar and viewing geometry of Landsat 7 were used because Landsat 7 and MODIS-Terra are positioned on the same orbit, the time lag of their acquisitions is constant, and sufficiently small (15-30 minutes) (Chander et al. 2010) to make it reasonable to use the MODIS-Terra cloud mask as an approximation for the cloud cover affecting Landsat-7 data. There is no equivalent global dataset acquired within a constant time window for other moderate resolution sensors. Effects of cloud cover on data availability can also be estimated using cloud probability derived from MODIS observations (Whitcraft et al. 2015); however, it implicitly requires that, at each location, the average cloud probability is estimated using observation sensed at or around the overpass time of moderate resolution instruments. Satellite cloud cover detections vary depending on the overpass time (e.g., morning versus afternoon overpass) and they increase with the off-nadir sensor view angle (King et al. 2013; Maddux et al. 2010), hence the statistical analysis of clouds have reduced accuracy if the overpass time and the observation geometry (i.e., sun and sensor zenith and azimuth angles) are not similar.

The findings of this research have implications for the future development of a global burned area product generated using moderate resolution EO data such as Landsat. The burned area persistence times provides an estimation of the period after the burning date in which burned areas can be mapped reliably and the potential omission error of a Landsat burned area product highlight the locations in which the revisit frequency of Landsat is not sufficient for accurate global burned area mapping.

The second chapter of my work presents the first global, systematic temporal analysis of the spectral signal persistence time associated with burned areas to determine the post-fire period in which burned pixels are detectable. The burned area persistence time was estimated as the maximum number of days in which the values of the NBR associated with burned and unburned areas are separable and, therefore, burned areas can be mapped reliably. The analysis was stratified spatially by ecoregions and land cover, and temporally to consider different post-fire effects occurring in different ecosystems and at different times of the year.

The third chapter of the dissertation presents the systematic comparison of Landsat- and MODIS-derived cloud cover data, to assess whether the long-term record of daily MODIS cloud detections can be used as a proxy for cloud cover observed by Landsat. One year of global Landsat 7 Enhanced Thematic Mapper Plus (ETM+) image cloud cover fractions over land were compared with contemporaneous, collocated MODIS cloud fractions, generated by combining the MODIS-Terra global daily cloud mask product (MOD35) with the Landsat 7 ETM+ image footprints and acquisition calendar. The results indicate that, globally, the two datasets have high correlation, and low bias and low residuals when using linear and logistic models to predict the Landsat cloud fractions from MODIS cloud fractions.

The fourth chapter of my dissertation builds on the results of the previous two chapters, estimating the omission errors of a hypothetical global Landsat 7 ETM+ burned area product, due to the combined effect of cloud cover, limited persistence time of burned areas, and revisit frequency of the sensor. The analysis was performed using the MODIS global burned area product to provide timing and location of the burned areas, and the MODIS daily cloud mask product to compute the number of cloud-free observations that are available post-fire on Landsat 7 ETM+ overpass days. This chapter provides a preliminary assessment of the potential limitations of a global Landsat burned area product.

Finally, the last chapter summarizes and discusses the findings of the work, highlights some limitations in the analysis, and discusses future research and applications.

1.4. References

- Alonso-Canas, I., & Chuvieco, E. (2015). Global burned area mapping from ENVISAT-MERIS and MODIS active fire data. *Remote Sensing of Environment*, *163*, 140-152
- Andela, N., Morton, D.C., Giglio, L., Chen, Y., van der Werf, G.R., Kasibhatla, P.S., DeFries, R.S., Collatz, G.J., Hantson, S., Kloster, S., Bachelet, D., Forrest, M., Lasslop, G., Li, F., Mangeon, S., Melton, J.R., Yue, C., & Randerson, J.T. (2017). A human-driven decline in global burned area. *Science*, *356*, 1356-1362
- Archibald, S., Lehmann, C.E., Belcher, C.M., Bond, W.J., Bradstock, R.A., Daniau, A., Dexter, K., Forrester, E., Greve, M., & He, T. (2018). Biological and geophysical feedbacks with fire in the Earth system. *Environmental Research Letters*, *13*, 033003
- Archibald, S., Lehmann, C.E.R., Gómez-Dans, J.L., & Bradstock, R.A. (2013). Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences*, *110*, 6442-6447
- Archibald, S., Roy, D.P., Van Wilgen, B.W., & Scholes, R.J. (2009). What limits fire? An examination of drivers of burnt area in Southern Africa. *Global Change Biology*, *15*, 613-630
- Boschetti, L., Brivio, P.A., Eva, H.D., Gallego, J., Baraldi, A., & Gregoire, J. (2006). A sampling method for the retrospective validation of global burned area products. *IEEE transactions on geoscience and remote sensing*, *44*, 1765-1773
- Boschetti, L., Flasse, S.P., & Brivio, P.A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment*, *91*, 280-292
- Boschetti, L., Roy, D.P., Giglio, L., Huang, H., Zubkova, M., & Humber, M.L. (2019). Global validation of the collection 6 MODIS burned area product. *Remote Sensing of Environment*, *235*, 111490
- Boschetti, L., Roy, D.P., Justice, C.O., & Humber, M.L. (2015). MODIS–Landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, *161*, 27-42
- Bowman, D.M., Balch, J.K., Artaxo, P., Bond, W.J., Carlson, J.M., Cochrane, M.A., D'Antonio, C.M., DeFries, R.S., Doyle, J.C., & Harrison, S.P. (2009). Fire in the Earth system. *Science*, *324*, 481-484
- Bowman, D.M.J.S., Balch, J., Artaxo, P., Bond, W.J., Cochrane, M.A., D'Antonio, C.M., DeFries, R., Johnston, F.H., Keeley, J.E., Krawchuk, M.A., Kull, C.A., Mack, M., Moritz, M.A., Pyne, S., Roos, C.I., Scott, A.C., Sodhi, N.S., Swetnam, T.W., & Whittaker, R. (2011). The human dimension of fire regimes on Earth. *Journal of Biogeography*, *38*, 2223-2236

- Bowman, D.M.J.S., Perry, G.L.W., & Marston, J.B. (2015). Feedbacks and landscape-level vegetation dynamics. *Trends in Ecology & Evolution*, *30*, 255-260
- Bowman, D.M.J.S., Zhang, Y., Walsh, A., & Williams, R.J. (2003). Experimental comparison of four remote sensing techniques to map tropical savanna fire-scars using Landsat-TM imagery. *International Journal of Wildland Fire*, *12*, 341-348
- Chander, G., Xiong, X., Choi, T., & Angal, A. (2010). Monitoring on-orbit calibration stability of the Terra MODIS and Landsat 7 ETM+ sensors using pseudo-invariant test sites. *Remote Sensing of Environment*, *114*, 925-939
- Chen, X., Vogelmann, J.E., Rollins, M., Ohlen, D., Key, C.H., Yang, L., Huang, C., & Shi, H. (2011). Detecting post-fire burn severity and vegetation recovery using multitemporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest. *International Journal of Remote Sensing*, *32*, 7905-7927
- Chu, T., & Guo, X. (2014). Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. *Remote Sensing*, *6*, 470-520
- Chu, T., Guo, X., & Takeda, K. (2016). Remote sensing approach to detect post-fire vegetation regrowth in Siberian boreal larch forest. *Ecological Indicators*, *62*, 32-46
- Chuvieco, E., & Congalton, R.G. (1988). Mapping and inventory of forest fires from digital processing of TM data. *Geocarto International*, *3*, 41-53
- Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanase, M., Koutsias, N., García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., & Giglio, L. (2019). Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sensing of Environment*, *225*, 45-64
- Fraser, R.H., Li, Z., & Landry, R. (2000). SPOT VEGETATION for characterizing boreal forest fires. *International Journal of Remote Sensing*, *21*, 3525-3532
- Frederiksen, P., Langaas, S., & Mbaye, M. (1990). NOAA-AVHRR and GIS-based monitoring of fire activity in Senegal—a provisional methodology and potential applications. *Fire in the Tropical Biota* (pp. 400-417): Springer
- Fuller, S.P., & Rouse, W.R. (1979). Spectral reflectance changes accompanying a post-fire recovery sequence in a subarctic spruce lichen woodland. *Remote Sensing of Environment*, *8*, 11-23
- GCOS (2011). Systematic observation requirements for satellite-based products for climate in support of the UNFCCC. In: WMO GCOS Tech. Doc. GCOS-138 WMO-TD 1523, 23 pp
- Giglio, L., Boschetti, L., Roy, D.P., Humber, M.L., & Justice, C.O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, *217*, 72-85

- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., & Justice, C.O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, *113*, 408-420
- Goetz, S.J., Bond-Lamberty, B., Law, B.E., Hicke, J.A., Huang, C., Houghton, R.A., McNulty, S., O'Halloran, T., Harmon, M., Meddens, A.J.H., Pfeifer, E.M., Mildrexler, D., & Kasischke, E.S. (2012). Observations and assessment of forest carbon dynamics following disturbance in North America. *Journal of Geophysical Research: Biogeosciences*, *117*, n/a-n/a
- Hall, J.V., Loboda, T.V., Giglio, L., & McCarty, G.W. (2016). A MODIS-based burned area assessment for Russian croplands: Mapping requirements and challenges. *Remote Sensing of Environment*, *184*, 506-521
- Hantson, S., Pueyo, S., & Chuvieco, E. (2015). Global fire size distribution is driven by human impact and climate. *Global Ecology and Biogeography*, *24*, 77-86
- Hawbaker, T.J., Vanderhoof, M.K., Beal, Y.-J., Takacs, J.D., Schmidt, G.L., Falgout, J.T., Williams, B., Fairaux, N.M., Caldwell, M.K., Picotte, J.J., Howard, S.M., Stitt, S., & Dwyer, J.L. (2017). Mapping burned areas using dense time-series of Landsat data. *Remote Sensing of Environment*, *198*, 504-522
- Huang, H., Roy, D., Boschetti, L., Zhang, H., Yan, L., Kumar, S., Gomez-Dans, J., & Li, J. (2016). Separability Analysis of Sentinel-2A Multi-Spectral Instrument (MSI) Data for Burned Area Discrimination. *Remote Sensing*, *8*, 873
- Humber, M.L., Boschetti, L., Giglio, L., & Justice, C.O. (2018). Spatial and temporal intercomparison of four global burned area products. *International Journal of Digital Earth*, 1-25
- Hyer, E.J., & Reid, J.S. (2009). Baseline uncertainties in biomass burning emission models resulting from spatial error in satellite active fire location data. *Geophysical Research Letters*, *36*
- Jin, Y., Randerson, J.T., Goulden, M.L., & Goetz, S.J. (2012). Post-fire changes in net shortwave radiation along a latitudinal gradient in boreal North America. *Geophysical Research Letters*, *39*, n/a-n/a
- Jin, Y., & Roy, D. (2005). Fire-induced albedo change and its radiative forcing at the surface in northern Australia. *Geophysical Research Letters*, *32*
- Keeley, J.E. (2009). Fire intensity, fire severity and burn severity: a brief review and suggested usage. *International Journal of Wildland Fire*, *18*, 116-126
- Key, C., & Benson, N. (2002). Measuring and remote sensing of burn severity. In, *US Geological Survey wildland fire workshop* (pp. 02-11): US Geological Survey Washington, DC, USA
- King, M.D., Platnick, S., Menzel, W.P., Ackerman, S.A., & Hubanks, P.A. (2013). Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites. *IEEE transactions on geoscience and remote sensing*, *51*, 3826-3852

- Krawchuk, M.A., & Moritz, M.A. (2011). Constraints on global fire activity vary across a resource gradient. *Ecology*, *92*, 121-132
- Langaas, S., & Kane, R. (1991). Temporal spectral signatures of fire scars in Savanna Woodland. In, *Geoscience and Remote Sensing Symposium, 1991. IGARSS '91. Remote Sensing: Global Monitoring for Earth Management., International* (pp. 1157-1160)
- Langmann, B., Duncan, B., Textor, C., Trentmann, J., & van der Werf, G.R. (2009). Vegetation fire emissions and their impact on air pollution and climate. *Atmospheric Environment*, *43*, 107-116
- Lentile, L.B., Holden, Z.A., Smith, A.M., Falkowski, M.J., Hudak, A.T., Morgan, P., Lewis, S.A., Gessler, P.E., & Benson, N.C. (2006). Remote sensing techniques to assess active fire characteristics and post-fire effects. *International Journal of Wildland Fire*, *15*, 319-345
- Littell, J.S., McKenzie, D., Peterson, D.L., & Westerling, A.L. (2009). Climate and wildfire area burned in western U.S. ecoprovinces, 1916–2003. *Ecological applications*, *19*, 1003-1021
- Lyons, E.A., Jin, Y., & Randerson, J.T. (2008). Changes in surface albedo after fire in boreal forest ecosystems of interior Alaska assessed using MODIS satellite observations. *Journal of Geophysical Research: Biogeosciences*, *113*, n/a-n/a
- Maddux, B., Ackerman, S., & Platnick, S. (2010). Viewing geometry dependencies in MODIS cloud products. *Journal of Atmospheric and Oceanic Technology*, *27*, 1519-1528
- Melchiorre, A., & Boschetti, L. (2018). Global Analysis of Burned Area Persistence Time with MODIS Data. *Remote Sensing*, *10*, 750
- Moritz, M.A., Parisien, M.-A., Batllori, E., Krawchuk, M.A., Van Dorn, J., Ganz, D.J., & Hayhoe, K. (2012). Climate change and disruptions to global fire activity. *Ecosphere*, *3*, art49
- Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., & Chuvieco, E. (2014). Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 64-79
- Padilla, M., Stehman, S.V., & Chuvieco, E. (2014). Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling. *Remote Sensing of Environment*, *144*, 187-196
- Padilla, M., Stehman, S.V., Ramo, R., Corti, D., Hantson, S., Oliva, P., Alonso-Canas, I., Bradley, A.V., Tansey, K., Mota, B., Pereira, J.M., & Chuvieco, E. (2015). Comparing the accuracies of remote sensing global burned area products using stratified random sampling and estimation. *Remote Sensing of Environment*, *160*, 114-121

- Pausas, J.G., & Fernández-Muñoz, S. (2012). Fire regime changes in the Western Mediterranean Basin: from fuel-limited to drought-driven fire regime. *Climatic Change*, *110*, 215-226
- Pereira, J., Chuvieco, E., Beaudoin, A., & Desbois, N. (1997). Remote sensing of burned areas: a review. *A review of remote sensing methods for the study of large wildland fires*, 127-184
- Pereira, J.C., Sá, A.L., Sousa, A.O., Silva, J.N., Santos, T., & Carreiras, J.B. (1999). Spectral characterisation and discrimination of burnt areas. In E. Chuvieco (Ed.), *Remote Sensing of Large Wildfires* (pp. 123-138): Springer Berlin Heidelberg
- Pereira, J.M. (2003). Remote sensing of burned areas in tropical savannas. *International Journal of Wildland Fire*, *12*, 259-270
- Randerson, J.T., Chen, Y., van der Werf, G.R., Rogers, B.M., & Morton, D.C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, *117*, G04012
- Randerson, J.T., Liu, H., Flanner, M.G., Chambers, S.D., Jin, Y., Hess, P.G., Pfister, G., Mack, M.C., Treseder, K.K., Welp, L.R., Chapin, F.S., Harden, J.W., Goulden, M.L., Lyons, E., Neff, J.C., Schuur, E.A.G., & Zender, C.S. (2006). The Impact of Boreal Forest Fire on Climate Warming. *Science*, *314*, 1130-1132
- Röder, A., Hill, J., Duguay, B., Alloza, J.A., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment*, *112*, 259-273
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., & Chuvieco, E. (2019). Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, *222*, 1-17
- Roy, D., Huang, H., Kumar, S., Zhang, H., Li, J., Gomez-Dans, J., Lewis, P., Boschetti, L. (2015). Towards prototyping a global Landsat-8 Sentinel-2 Burned Area Product. *EARSEL SIG Fire 2015, Cyprus, 2-5 November 2015*
- Roy, D., Jin, Y., Lewis, P., & Justice, C. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, *97*, 137-162
- Roy, D.P., Boschetti, L., & Giglio, L. (2010). Remote Sensing of Global Savanna Fire Occurrence, Extent, and Properties. *Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales*, 239
- Roy, D.P., Boschetti, L., Justice, C.O., & Ju, J. (2008). The collection 5 MODIS burned area product—Global evaluation by comparison with the MODIS active fire product. *Remote Sensing of Environment*, *112*, 3690-3707

- Roy, D.P., Huang, H., Boschetti, L., Giglio, L., Yan, L., Zhang, H.H., & Li, Z. (2019). Landsat-8 and Sentinel-2 burned area mapping - A combined sensor multi-temporal change detection approach. *Remote Sensing of Environment*, 231, 111254
- Seiler, W., & Crutzen, P.J. (1980). Estimates of gross and net fluxes of carbon between the biosphere and the atmosphere from biomass burning. *Climatic Change*, 2, 207-247
- Simon, M., Plummer, S., Fierens, F., Hoelzemann, J., & Arino, O. (2004). Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 109
- Skinner, C.N., & Chang, C. (1996). Fire regimes, past and present. In *Sierra Nevada Ecosystem Project: Final report to Congress. Vol. II. Assessments and Scientific Basis for Management Options. Wildland Resources Center Report No. 37. Centers for Water and Wildland Resources, University of California, Davis. 1041-1069* (pp. 1041-1069)
- Smith, A.M., & Wooster, M.J. (2005). Remote classification of head and backfire types from MODIS fire radiative power and smoke plume observations. *International Journal of Wildland Fire*, 14, 249-254
- Solans Vila, J.P., & Barbosa, P. (2010). Post-fire vegetation regrowth detection in the Deiva Marina region (Liguria-Italy) using Landsat TM and ETM+ data. *Ecological Modelling*, 221, 75-84
- Tansey, K., Grégoire, J.-M., Stroppiana, D., Sousa, A., Silva, J., Pereira, J.M.C., Boschetti, L., Maggi, M., Brivio, P.A., Fraser, R., Flasse, S., Ershov, D., Binaghi, E., Graetz, D., & Peduzzi, P. (2004). Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data. *Journal of Geophysical Research: Atmospheres*, 109, D14S03
- Tansey, K., Gregoire, J.M., Defourny, P., Leigh, R., Pekel, J.F.O., van Bogaert, E., & Bartholome, E. (2008). A new, global, multi-annual (2000-2007) burnt area product at 1 km resolution. *Geophysical Research Letters*, 35, 6
- Thonicke, K., Venevsky, S., Sitch, S., & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. *Global Ecology and Biogeography*, 10, 661-677
- Trigg, S., & Flasse, S. (2000). Characterizing the spectral-temporal response of burned savannah using in situ spectroradiometry and infrared thermometry. *International Journal of Remote Sensing*, 21, 3161-3168
- Trigg, S.N., & 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

- van der Werf, G.R., Randerson, J.T., Giglio, L., van Leeuwen, T.T., Chen, Y., Rogers, B.M., Mu, M., van Marle, M.J., Morton, D.C., & Collatz, G.J. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, 9, 697
- Westerling, A.L., Hidalgo, H.G., Cayan, D.R., & Swetnam, T.W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science*, 313, 940-943
- Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I., & Justice, C.O. (2015). Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sensing of Environment*, 156, 438-447

Chapter 2: Global Analysis of the Burned Area Persistence Time With MODIS Data

Published in *Remote Sensing* as:

Melchiorre, A., & Boschetti, L. (2018). Global Analysis of Burned Area Persistence Time with MODIS Data. *Remote Sensing*, 10, 750

2.1. Abstract

Biomass burning causes a non-permanent land cover change (burned area), through the removal of vegetation, the deposition of charcoal and ashes, and the exposure of soil; the temporal persistence of these changes is highly variable, ranging from a few weeks in savannas to years in forests. Algorithms for the generation of moderate resolution (10 - 30 m) continental and global burned area maps have been prototyped in the effort of meeting the needs of diverse users of fire information. Nevertheless, moderate resolution sensors have reduced temporal resolution (e.g. 16 days for Landsat), which could potentially lead to omission errors, especially in ecosystems where the spectral signal associated with burning disappears quickly and cloud cover limits the number of valid observations. This study presents a global analysis of the burned area persistence time defined as the duration of the spectral separability of the burned / unburned areas mapped by the MODIS MCD64 Global Burned Area Product. The separability was computed by analyzing time series of normalized burn ratio (NBR) from nadir BRDF-adjusted MODIS reflectances (MCD43 product). Results showed that, globally, the median burned area persistence time was estimated in 29 days and 86.6% of the global area, as detected by MODIS, can be detected accurately only for up to 48 days. Thus, results indicate that burned area persistence time can be a limiting factor for global burned area mapping from moderate resolution satellite sensors, which have a low temporal resolution (e.g. Landsat 16 days, Sentinel-2A/B 5 days).

2.2. Introduction

Fire is a natural component of any ecosystem, and it has effects on vegetation, soil, water and atmospheric composition (Bowman et al. 2009; Certini 2005; DeBano et al. 1998). Fire contributes to the global carbon cycle through the emission of greenhouse gases and aerosols from biomass burning (van der Werf et al. 2010; van der Werf et al. 2006), and as an agent of ecological change influencing

the global vegetation dynamics (Bond et al. 2005; Goetz et al. 2012; Thonicke et al. 2001) and the surface energy budget (Jin and Roy 2005; Randerson et al. 2006). Earth Observation (EO) data allow the analysis of biodiversity, dynamics of biomass, productivity, and disturbances spatially from regional to global scales (Pfeifer et al. 2012). The application of satellite data to study vegetation fires greatly enhanced the possibility to introduce the effect of fire disturbances in global models of climate and atmospheric composition and dynamics estimation (Langmann et al. 2009). Several studies have used post-fire albedo observations to investigate fire-induced vegetation change impacts on radiative forcing and local effects on climate (Jin et al. 2012; Jin and Roy 2005; Lyons et al. 2008). Due to their significant effects on terrestrial ecosystems and atmospheric processes, fire disturbances are included in the list of the required Essential Climate Variables (ECV) in support of the work of the United Nation Framework Convention on Climate Change (UNFCCC) and the Intergovernmental Panel on Climate Change (IPCC) (GCOS 2011).

Burned area maps are defined as the primary variable of the Fire ECV (GCOS 2011), and are currently one of the main inputs for the estimation of atmospheric emissions due to biomass burning (GCOS 2011; Giglio et al. 2013; van der Werf et al. 2010). Systematic global burned area maps at coarse spatial resolution (350 m - 1 km) have been produced in the past 20 years from several EO systems, including ATSR, MERIS, MODIS, and SPOT-VGT (Alonso-Canas and Chuvieco 2015; Giglio et al. 2009; Roy et al. 2005; Simon et al. 2004; Tansey et al. 2004; Tansey et al. 2008). Burned area coarse resolution maps improved the global long-term estimates of fire activity trends (Andela et al. 2017) and they were used to build fire-related databases of aggregated burned area and emission estimates to help understand the interconnections between fire activity and the global carbon cycle (van der Werf et al. 2017). Such databases include the Global Fire Emissions Database (GFEDv4) (Giglio et al. 2013), the Global Fire Assimilation System (GFAS) (Kaiser et al. 2012), the Emission for Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP) (Lamarque et al. 2010) and the United Nation's Food Agriculture Organization (UN-FAO) FAOSTAT Emission database (Rossi et al. 2016; Tubiello et al. 2013).

Among others, the coarse resolution of the burned area products is a source of uncertainties in these emission databases (van der Werf et al. 2010) because it leads to a trade-off between omission and commission errors on partially burned pixels (Boschetti et al. 2004). Pixels classified as burned are considered completely burned, which is often not the case. Due to the heterogeneous nature of fire, sub-pixel unburned patches may exist and result in overestimations of the burned area. Conversely, partially burned pixels can be undetected, resulting in underestimations of the burned area, especially in areas where small fires are common (Randerson et al. 2012; van der Werf et al. 2017). Consequently, there

is a strong need for moderate resolution (10 - 30 m) global burned area maps for fire and ecosystem management, improved emission estimations, and carbon accounting (Hyer and Reid 2009; Mouillot et al. 2014; Randerson et al. 2012).

Automated mapping algorithms using Landsat and Sentinel data are currently being prototyped (Boschetti et al. 2015; Hawbaker et al. 2017; Roy 2015) but, among others, one of the challenges of designing a global burned area detection algorithm is the limited temporal resolution of moderate resolution sensors due to their narrow field of view. The limited number of acquisitions obtained by moderate resolution satellite translates into potential omission errors (Boschetti et al. 2015) because the effects of fire on the spectral signal of vegetation are not permanent. In the days or weeks after a fire, depending on vegetation type, phenology and climate, charcoal and ashes are removed and vegetation recovers (Chu et al. 2016; Fraser et al. 2000; Pereira et al. 1997; Solans Vila and Barbosa 2010; Trigg and Flasse 2000), decreasing the differences of the pre- and post-fire satellite observations over burned areas. Consequently, it is possible to detect burned areas within a limited period that varies across different ecosystems, and within the same ecosystem, it depends on the timing of the fire. Previously published *in situ* spectral measurements indicate that the period of detectability of burned areas ranges from about two weeks in African savannas (Frederiksen et al. 1990; Langaas and Kane 1991; Trigg and Flasse 2000) to more than two years in boreal forests (Fuller and Rouse 1979). Similarly, studies using post-fire satellite observations indicate that the spectral changes induced by fire disappear rapidly in grasslands and savannas (Bowman et al. 2003; Pereira 2003; Trigg and Flasse 2000) but are generally persistent in forested ecosystems (Chen et al. 2011; Chu and Guo 2014; Röder et al. 2008).

The impermanent nature of the spectral signal associated with burned areas poses a major challenge for designing global burned area mapping algorithms from moderate resolution satellites (Boschetti et al. 2015). For example, Landsat satellites have a 16-days revisit time and the newly available Sentinel-2A and Sentinel-2B have 10-days revisit time (5 days when both satellites are used in combination) which can be a limiting factor for global burned area mapping from moderate resolution satellite sensors. Burned area mapping algorithms require at least one cloud-free acquisition before and after the fire event; and, using moderate resolution data, omission errors might result in areas where the spectral signal associated with burning disappears quickly due to the lower revisit frequency. This work presents the first global, systematic temporal analysis of the spectral signal associated with burned areas and the relative persistence time to determine the length of the period in which the burn class is spectrally distinct from the unburned and, therefore, detectable.

The burned area persistence time was estimated as the maximum number of days in which the values of the spectral index Normalized Burn Ratio associated with burned and unburned areas are

separable and, therefore, burned areas can be mapped reliably. The primary input of the study was provided by 14 years of the MODIS burned area and surface reflectance products, from January 2003 to December 2016, in order to use every full year where both MODIS-Terra and MODIS-Aqua data are available. The analysis was stratified spatially by ecoregions and land cover, and temporally considering the period of the year in which the fire occurred. Nadir Bidirectional Reflectance Distribution Function (BRDF)-adjusted reflectance product (MCD43A4) time series of burned pixels, as detected by the MODIS global burned area product (MCD64A1), within the same ecoregion and land cover were used to characterize the pre- post-fire temporal variations with the passing of time due to charcoal and ashes dissipation, vegetation regrowth and snow cover. Reflectance time series of unburned pixels within the same ecoregion and land cover were used to characterize the variations in reflectance due to vegetation phenology and other disturbances. Summary metrics were defined for reporting the results at the global scale, aggregating by global biome and realm.

The paper is organized as follows. Section 2 describes the satellite datasets used for the analysis, and the ancillary datasets used as stratification variables. Section 3 describes the methods, establishing a rigorous probabilistic framework for the definition of the persistence time of burned areas at different scales, and presenting the formulae needed for a global, multiyear analysis. Section 4 presents the results and section 5 discusses the differences observed across land cover types, biomes and realms. The paper concludes with recommendations for future research and application.

2.3. Data

2.3.1. MODIS global burned area product

The most recent Collection 6 MODIS Global Burned Area Product (MCD64A1) provides the estimated date of burn for the 500 m MODIS pixels that are classified as burned within a calendar month (Giglio et al. 2009; Roy et al. 2005). The global, monthly MCD64A1 data record from 2003 to 2016 was used, to include every full year in which both MODIS instruments – on the Terra and Aqua platforms - were operating. The algorithm (Giglio et al. 2009) is designed to be extremely tolerant of cloud and aerosol contamination, which affected the Collection 5 MODIS 500 m burned area product (Roy et al. 2008). The algorithm applies dynamic thresholds to composite MODIS Terra and Aqua imagery generated from a burn-sensitive spectral band index derived from MODIS 1240 nm and 2130 nm Terra and Aqua bands, and a measure of temporal variability. Cumulative MODIS 1 km active fire detections are used to guide the selection of burned and unburned training samples and to guide the specification of prior burned and unburned probabilities. The MCD64A1 product is distributed in the standard MODIS Level 3 $10^\circ \times 10^\circ$ Land tile format in the sinusoidal projection (Wolfe et al. 1998).

2.3.2. MODIS global nadir BRDF-adjusted reflectance product

The MODIS Nadir BRDF-Adjusted Reflectance (NBAR) product provides estimates of the nadir surface reflectance at local solar noon performing a Bidirectional Reflectance Distribution Function (BRDF) inversion of MODIS surface reflectance for each day using a 16-days moving window centered on the nominal product date (Schaaf et al. 2002; Wang et al. 2018). The Collection 6, Level 3 daily MODIS Terra and Aqua combined product (MCD43A4) was used in this work. The MCD43A4 product is defined in the MODIS Level 3 Land tile format in sinusoidal projection at 500 m resolution. Per-pixel quality assessment information is provided by the MCD43A2 product, also defined in the MODIS Level 3 Land tile format in sinusoidal projection at 500 m resolution and, for each reflective band of the MCD43A4 product, reports the quality of the BRDF inversion for each pixel of the relative MCD43A4 product, and flags the presence of water and snow pixels.

2.3.3. MODIS land cover product

The MODIS Land Cover Type product (MCD12Q1) provides five land classification schemes, which describe land cover properties derived from one year of observations from Terra- and Aqua-MODIS (Friedl et al. 2010). The Collection 5.1, Level 3 yearly land cover product is defined in the MODIS Level 3 Land tile format in sinusoidal projection at 500 m resolution. The International Geosphere and Biosphere Programme (IGBP) scheme, which identifies 16 land cover classes including 11 natural vegetation classes, 3 developed and 2 non-vegetated land classes and has a reported 75% overall land cover classification accuracy (Friedl et al. 2010) was used in this work.

2.3.4. Terrestrial ecoregions of the worlds

The Terrestrial Ecoregions of the World (TEOW) map is a biogeographic division of the Earth's terrestrial biodiversity in 867 ecoregions, which belong to 14 biomes and 8 realms (Olson et al. 2001). Ecoregions are defined as biogeographic units containing a homogeneous population of natural communities (flora and fauna) sharing a large majority of species, dynamics, and environmental conditions.

2.4. Methods

2.4.1. Theoretical basis: definition of the burned area persistence time

The burned areas persistence time was defined as the maximum time after a fire, during which burned pixels can be unambiguously separated from unburned ones (i.e., the length of the period when burned areas can be reliably mapped).

The spectral signature of vegetation is altered by burning events, and the magnitude of these changes is greater in some portions of the electromagnetic spectrum than in others. It has been demonstrated that the spectral bands more suitable for burned area mapping are the near infrared (NIR) and shortwave infrared (SWIR) for both coarse (Pereira et al. 1999; Roy et al. 2002; Stroppiana et al. 2002) and moderate resolution satellites (Huang et al. 2016; Key and Benson 2006; Koutsias and Karteris 1998; Pereira and Setzer 1993). The spectral changes induced by fire are non-permanent: charcoal and ashes are removed by atmospheric agents exposing the bare soil, and vegetation regrows over time (Roy et al. 2010). Additionally, the spectral signature of unburned pixels also changes over time: vegetation phenology, senescence and other disturbances, such as forest management thinning and clear cuts, mortality due to insect outbreaks and land use conversion, cause spectral changes that can be confused with those due to fire (Goodwin et al. 2008; Wang et al. 2012). For these reasons, burned area spectral separability is sensitive to the time elapsed since the burning event (Bowman et al. 2003; Loboda et al. 2013; Roy et al. 2005): the detectability of burned areas is a function – among other factors – of the burned area persistence time.

A wide variety of burned area mapping algorithms have been tested and developed for moderate resolution sensors including techniques that exploit single image analysis (Bastarrika et al. 2011; Chuvieco and Congalton 1988; Chuvieco et al. 2002; Koutsias and Karteris 2000; Mitri and Gitas 2004) and multi-temporal analysis (Boschetti et al. 2015; Chuvieco et al. 2002; Hudak and Brockett 2004; Koutsias and Karteris 1998; Miller and Yool 2002; Silva et al. 2005; Smith et al. 2007; Stroppiana et al. 2009). Although there is no agreement on the optimal burned area mapping algorithm, spectral indices are widely used in many studies. Among them, the Normalized Burn Ratio (NBR (Key and Benson 2002)) is widely used for burned area detection with moderate resolution satellite data (Bastarrika et al. 2011; Fraser et al. 2003; Henry 2008; Koutsias and Karteris 2000; Miller and Yool 2002; Silva et al. 2005; Smith et al. 2007; Stroppiana et al. 2009). The index range is between -1 and 1 and it is defined as:

$$NBR = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (1)$$

where ρ_{NIR} is the reflectance at $0.8 \mu m$ and ρ_{SWIR} is the reflectance at $2.1 \mu m$. The difference between pre-fire and post-fire NBR (dNBR) is an indication of the spectral changes due to fire. dNBR can assume values between -2 and 2. Because NBR drops after a fire due to the removal of vegetation, soil exposure and charcoal and ash deposition (Key and Benson 2006), positive dNBR values are associated with burning, thus making dNBR thresholding a simple burned area detection strategy (Escuin et al. 2008; Key and Benson 2002; Stroppiana et al. 2009).

In this study, daily time series of dNBR were generated for burned and unburned pixels from the Nadir BRDF-Adjusted Reflectance Albedo product (MCD43A4) to minimize the variability in reflectance due to different acquisition view geometry. Missing data, water and snow-covered pixels were discarded from further analysis.

For burned pixels, at a generic location i , detected as burned on Julian day of the year $jday$ in the MODIS global burned area product (MCD64A1), the post-fire dNBR yearly time series was defined as:

$$dNBR_B(\Delta t)_{i,jday} = [NBR_{i,jday-8} - NBR_{i,jday+\Delta t}] \quad (2)$$

where $NBR_{i,jday-8}$ is the pre-fire NBR value, obtained from the MCD43A4 product with the nominal date 8 days before the day $jday$ of burning, and $NBR_{i,jday+\Delta t}$ is the value observed Δt days after the fire, with $\Delta t \geq 8$. The 16-days minimum time difference between pre-fire and post-fire observation was needed to account for the 16-days inversion moving window used in the MCD43A4 product and to ensure that the NBR pre-fire and post-fire values were respectively calculated only from the inversion of pre-fire and post-fire reflectances.

Generally, vegetation spectral signature varies during the year depending on phenology (e.g. senescence, leaf-off) and stress conditions; as a result, the NBR of unburned vegetation is variable during the year. Equation 2 generates yearly time series of dNBR measuring only the changes from the pre-fire conditions regardless of the initial NBR value.

Similarly, a generic unburned pixel j , observed synchronously with the burned pixel i of Equation 2, will describe spectral temporal variations determined by vegetation phenology and other non-fire related factors:

$$dNBR_{UB}(\Delta t)_{j,jday} = [NBR_{j,jday-8} - NBR_{j,jday+\Delta t}] \quad (3)$$

where $NBR_{j,jday-8}$ is the NBR value of the unburned pixel j on the same pre-fire date of Equation 2, and $NBR_{j,jday+\Delta t}$ is the value observed Δt days after the starting day $jday$.

Considering a set of burned pixels ba , all burning on the same day $jday$, the individual trends of Equation 2 were aggregated in a probability distribution function:

$$P_B(dNBR)_{\Delta t,jday} = P(dNBR_B(\Delta t)_{i,jday} = dNBR, \quad i \in ba) \quad (4)$$

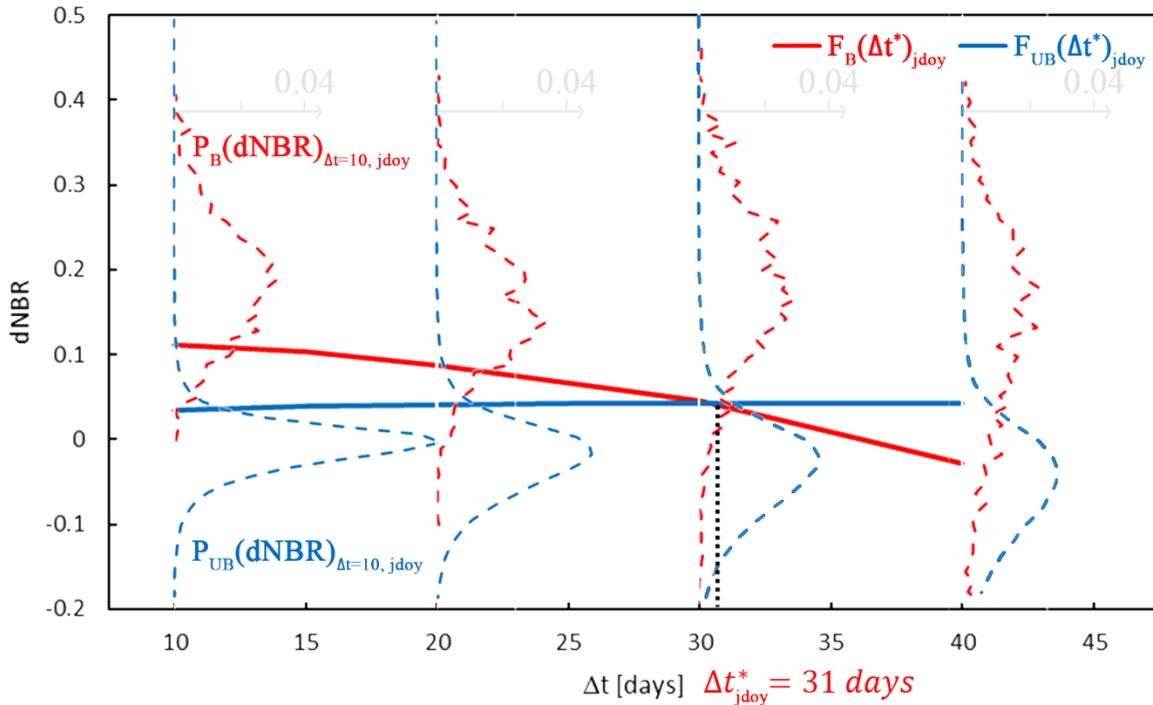


Figure 2-1: Illustration of the method for the estimation of the burned area persistence time. The plot is generated using observations of a set of burned pixels (*ba*) and unburned pixels (*ub*) in the Savanna land cover of the “Victoria Plains tropical savanna” ecoregion in northern Australia, starting on the day of detection of the burned pixels ($jday=329$, November 25th). The red dotted lines show the distribution of dNBR of the burned pixels ($P_B(dNBR)_{\Delta t, jday}$), observed at $\Delta t = 10, 20, 30$ and 40 days after burning; the blue dotted lines show the distribution of dNBR of the unburned pixels ($P_{UB}(dNBR)_{\Delta t, jday}$) observed at the same time. The red and blue solid lines represent respectively the 10th percentile of the dNBR distribution of burned pixels (i.e., $F_B(\Delta t)_{jday}$) and the 90th percentile of the dNBR distribution of the unburned pixels (i.e., $F_{UB}(\Delta t)_{jday}$). The persistence time is calculated as the time Δt^*_{jday} when the two distributions overlap by more than 20%, i.e., when $F_B(\Delta t)_{jday}$ is equal to $F_{UB}(\Delta t)_{jday}$ ($\Delta t^*_{jday} = 31$ days in this example).

This is illustrated in Figure 2-1. Considering a set of burned pixels, detected on the same day $jday$ in one spatial stratum of the analysis (‘Victoria Plains tropical savanna’ ecoregion in Northern Australia, savanna land cover), the histogram of the distribution $P_B(dNBR)_{\Delta t, jday}$ is represented by the dashed red lines. Notably, the range of dNBR values assumed by the burned area increases as time from the fire (i.e., Δt), indicating heterogeneous post-fire trajectories, and the median value decreases, indicating a tendency to return to pre-fire conditions. The post-fire spectral signature varies depending on atmospheric events and vegetation regrowth following the fire, for example wind and rain can scatter the charcoal and ashes or a new layer of grasses can grow rapidly after a fire, these effects can have great variability also within the same burned area, resulting in a wider range of post-fire dNBR values.

Similarly, considering a set of unburned pixels *ua*, the individual trajectories of Equation 3 were also aggregated into a probability distribution:

$$P_{UB}(dNBR)_{\Delta t, jday} = P(dNBR_{UB}(\Delta t)_{j, jday} = dNBR, j \in ub) \quad (5)$$

In Figure 2-1, the histogram of the distribution of $P_{UB}(dNBR)_{\Delta t, jday}$ for a set of unburned pixels in the same analysis stratum, and observed from the same day $jday$ as the burned pixels, is represented by the dashed blue line. The range of dNBR values assumed by the unburned pixels also increases with Δt while the median remains stable around 0, indicating that the set of the unburned pixels shows increasingly heterogeneous trends as time passes: while some pixels show trends that cannot be confused with burning (i.e., negative dNBR changes), others do (i.e., positive dNBR changes), likely due to senescence and phenology.

The separability of burned / unburned pixels was measured in terms of overlapping portion of the $P_B(dNBR)_{\Delta t, jday}$ and $P_{UB}(dNBR)_{\Delta t, jday}$ distributions. As a result of the temporal variation, the overlap of the two distributions is minimal immediately after the fire, and increases as time passes, reflecting the fact that burned areas can be mapped reliably only for a limited time after the fire. The persistence time of the burned area was, therefore, estimated as the maximum number of days in which the two distributions have minimal or no overlap, and more specifically the number of days in which at least 90% of the burned area pixels have dNBR higher than 90% of the unburned pixels. Using this empirical threshold, the burned area persistence time was estimated as the number of days when the two distributions overlap by 20%.

By definition, 90% of the burned area pixels have a dNBR value higher than the 10th percentile of the $P_B(dNBR)_{\Delta t, jday}$ distribution, which expressed as a function of Δt is defined as:

$$F_B(\Delta t)_{jday} = dNBR: \left\{ \int_{-2}^{dNBR} P_B(dNBR)_{\Delta t, jday} = 0.10 \right\} \quad (6)$$

Similarly, 90% of the unburned pixels have a dNBR value lower than the 90th percentile of the $P_{UB}(dNBR)_{\Delta t, jday}$:

$$F_{UB}(\Delta t)_{jday} = dNBR: \left\{ \int_{-2}^{dNBR} P_{UB}(dNBR)_{\Delta t, jday} = 0.90 \right\} \quad (7)$$

Finally, the persistence time of the burned areas was estimated as the number of days Δt_{jday}^* in which $F_B(\Delta t)_{jday}$ is greater or equal to $F_{UB}(\Delta t)_{jday}$:

$$\Delta t_{jday}^* = \max(\Delta t: F_B(\Delta t)_{jday} \geq F_{UB}(\Delta t)_{jday}) \quad (8)$$

In Figure 2-1, $F_B(\Delta t)_{j_{do y}}$ and $F_{UB}(\Delta t)_{j_{do y}}$ are represented by the solid red and blue lines respectively, and the persistence time $\Delta t_{j_{do y}}^*$ is defined by the intersection of the two lines.

2.4.2. Global implementation

The probabilistic method described in Section 3.1 was implemented globally using the 14 years of MODIS datasets described in Section 2 as shown in Figure 2-2. Because of the analysis scale, it was first necessary to stratify the analysis spatially and temporally using strata sufficiently fine to capture the variability of burned areas, limiting it to the times of the year where fire occurs (Section 3.2.1), and to define a sampling strategy to extract a representative set of burned and unburned pixels within each stratum (Section 3.2.2). The formulae of Section 3.1 were then used to build a set of summary metrics at the level of each ecoregion (Section 3.2.3) and globally at the biome scale (Section 3.2.4).

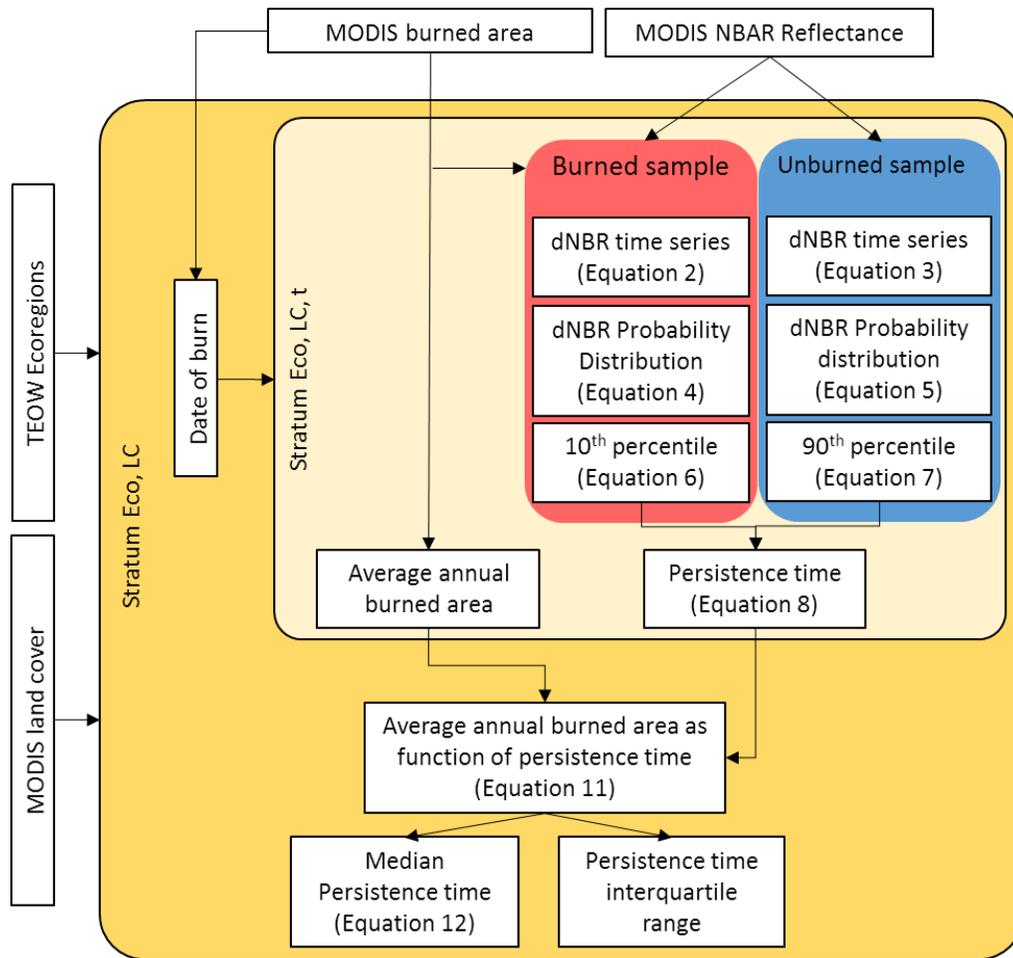


Figure 2-2: Illustration of the procedure for the estimation of the burned area persistence time. The input data for the analysis are the MODIS burned area and NBAR product. The MODIS land cover product and the TEOW ecoregions are used to define each spatial stratum defined by ecoregion and land cover. The estimated burned date extracted from the MODIS burned area product are used to allocate the pixels in the different 16-days periods (Section 3.2.1). The median persistence time for each ecoregion and land cover is computed as the median value of the average annual burned area as a function of the persistence time (Section 3.2.3).

2.4.2.1. Spatial and temporal stratification of the analysis

Spatial Stratification

Burned areas have different spectral responses depending on the type and condition of fuels, and more broadly depending on the ecosystem and time of burning (Jin and Roy 2005; Krawchuk and Moritz 2011; Pereira 2003), the analysis was performed using spatial and temporal units sufficiently fine to capture such variability. The analysis was stratified spatially adopting a two-level stratification. The Terrestrial Ecoregions of the World (TEOW) (Olson et al. 2001) were used to define the first stratification level. Ecoregions are delineated according to endemic genera and families (higher taxa), distinct assemblages of species, and the influence of geological history, such as past glaciations, on the distribution of plants and animals (Olson et al. 2001). Consequently, they are more likely to accurately reflect the distribution of species and communities than alternative stratifications derived solely from biophysical variables, such as rainfall and temperature, vegetation structure, or spectral signatures from remote sensing data (Olson et al. 2001).

The TEOW map is a broad simplification of the variability of habitats, and all contain a variety of vegetation types (Olson et al. 2001): for example, boreal ecoregions include forests, shrublands, and grasslands. Consequently, the MODIS MCD12Q1 land cover product was used to define a second level of stratification within each TEOW ecoregion. The land cover classes of the IGBP classification scheme were aggregated into three major classes of interest (Forest, Shrubland, Grassland & Savanna), or masked out and removed from the subsequent analysis (Urban areas, croplands, and miscellaneous non-burnable surfaces) (Table 2-1).

Table 2-1: Land cover aggregation scheme of the MODIS MCD12Q1 IGBP land cover classes used for the present study

IGBP Land cover	Aggregated land cover
Evergreen Needleleaf forest	Forest
Evergreen Broadleaf forest	
Deciduous Needleleaf forest	
Deciduous Broadleaf forest	
Mixed forest	
Closed shrublands	Shrubland
Open Shrublands	
Woody savannas	Grassland & Savanna
Savannas	
Grasslands	
Permanent wetlands	Not considered
Croplands	
Urban and built-up	
Cropland/Natural vegetation mosaic	
Snow and ice	
Barren and sparsely vegetated	

The spatial strata {Eco, LC}, defined by the generic land cover class LC within the generic ecoregion Eco, where at least 250 km² of burned areas were cumulatively detected by the MCD64A1 product over the entire 14 year study period (2003-2016) were considered, thus excluding land cover class / ecoregion combination with negligible fire activity.

Temporal stratification and temporal extent of the analysis

Fire activity is mainly driven by fuel abundance and structure and by climatic conditions that result in changes in fuel moisture and lightning activity (Moritz et al. 2012). Consequently, different regions of the globe have distinct fire seasons. At the global scale, the fire season peak months in the Southern hemisphere are August–September, primarily driven by the extensive burning in Southern Africa and Australia while in the Northern hemisphere the peak months are December–January due primarily to burning across Northern and Central Africa (Boschetti and Roy 2008). At the regional scale, fire seasons vary in terms of length, peak activity and numbers of peaks (Giglio et al. 2006). These differences are linked mainly to the different climatic and the pre-burning fuel conditions, which vary during the fire season (Archibald et al. 2013; Lambin et al. 2003). Generally, early season fires burn living, photosynthetically active vegetation, and dead vegetation that has a higher moisture content. During the dry season, vegetation senescence and climatic conditions lead to dryer living and dead vegetation (Schwartz and Reed 1999; White et al. 1997). As a result, fire-induced spectral changes for a given ecosystem and land cover vary within the fire season (Bucini and Lambin 2002; Dwyer et al. 2000; Lambin et al. 2003).

In order to account for these variations, a temporal stratification grid was created by dividing the January-December calendar year into 23 periods of 16 days each (*t*) (the last period being 13 days). This 16-days temporal grid – which coincides with the return interval of the Landsat satellite – was sufficiently fine to characterize the spectral temporal variations of burned areas at different times of the year, while significantly reducing the computational load of a daily temporal grid.

To further reduce the computational load, only burned areas occurring during the fire season of each ecoregion were considered. The fire season of each ecoregion was defined by adapting the method proposed by Archibald et al. (2013). All the 16-days periods of the year were ranked based on the average annual burned area detected by the MCD64A1 product, and the fire season was defined as the union of the periods in which 90% of the annual burned area is detected. This definition requires no assumption on the continuity of the fire season and is applicable to ecoregions having a variety of temporal patterns of fire activity (Figure 2-3).

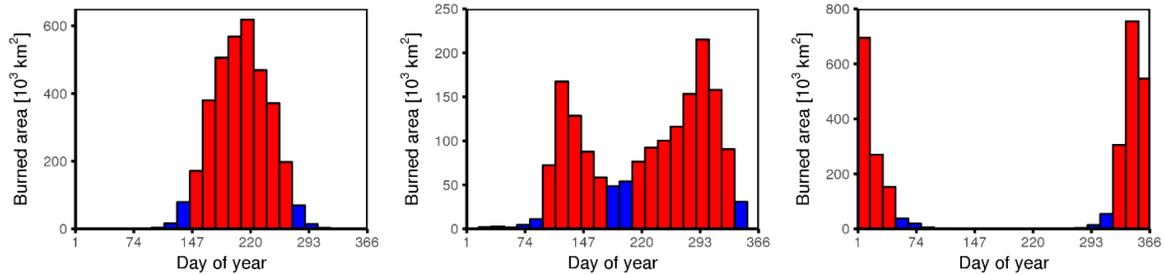


Figure 2-3: Illustration of the procedure for the definition of the fire season, showing the burned area histogram, binned in 16-days intervals t , of three representative examples of fire season. Left, unimodal fire season fully contained in a single calendar year (“Angolan woodland” ecoregion in Southern Africa); center, bimodal fire season (“Victoria Plains tropical savanna” ecoregion in northern Australia); right, unimodal fire season spanning across calendar years (“Northern Congolian tropical forest-savanna mosaic” ecoregion in northern Africa). The fire season is identified by ranking the 16-days bins in decreasing order by burned area and including in the fire season the bins necessary to reach 90% of the burned area (shown in red). The 16-days periods outside the fire season (shown in blue) were not considered in the analysis.

2.4.2.2. Sampling design

The MODIS MCD64A1 global burned area product detects on average $\sim 4.0 \cdot 10^6 \text{ km}^2$ per year (Giglio et al. 2013) of burned areas, corresponding to $\sim 16 \cdot 10^6$ pixels per year. To make the analyses logistically possible, a sampling strategy was employed to sample both burned and unburned pixels in the spatial and temporal strata $\{\text{Eco}, \text{LC}, t\}$ defined above as a combination of ecoregions, land cover classes, and time of the year quantized in 16-days periods.

The total sample size in each ecoregion was 100’000 burned and 100’000 unburned pixels (or the entire population if fewer than 100’000 pixels exist in either class). This total sample size was then allocated to the spatial sub-strata (i.e., land cover classes with fire activity in the ecoregion) and temporal strata (i.e., number of 16-days periods of the fire season) proportionally to the area burned, and the sample was extracted randomly with no replacement.

Under stratified random sampling, the probability of inclusion is constant within each stratum and is:

$$\pi_h = \frac{n_h}{N_h} \quad (9)$$

Where n_h and N_h are respectively the sample size and the population size in stratum h .

It follows that within each stratum (i.e., each combination of ecoregion Eco , land cover LC , and 16-days period t) the inclusion probability was:

$$\pi_{Eco,LC,t} = \frac{ba_{Eco,LC,t}}{BA_{Eco,LC,t}} \quad (10)$$

where ba is the extracted sample, and BA is the total population of the burned area in the stratum. It should be noted that the sample $ba_{Eco,LC,t}$ was drawn from the total population of burned areas detected in ecoregion Eco and land cover LC , burned in the 14 years of the study period in the same 16 days of the calendar year defined by the interval t .

2.4.2.3. Ecoregion level estimation of the persistence time

The persistence time is estimated for each stratum, by applying equations 1 through 8 using the sample of burned pixels $ba_{Eco,LC,t}$ and unburned pixels $ub_{Eco,LC,t}$. It should be noted that the 16-days grid was used solely only for the stratification of the sample, whereas the variation over time of the dNBR values (Equations 2 and 3) was evaluated from the daily MCD43A4 time series. As a consequence, for each land cover and ecoregion up to 23 persistence times (one for each 16-days period t) were obtained, each of them estimated from Equation 8 with daily resolution. This is exemplified in Figure 2-4 (*top*), showing the burned area persistence time for the TEOW “Victoria plains tropical savanna” in Northern Australia. The 16-days periods included in the fire season are shaded in gray, and for each of them the estimated persistence times $\Delta t^*_{Eco,LC,t}$ in “Shrubland” and “Grassland and Savannah” land cover classes are shown in gray and orange respectively. The persistence time was not computed for the ‘Forest’ land cover, because no time period t had the minimum number of burned area detections defined in Section 3.2.1.

In order to obtain meaningful summary metrics of the persistence time $\Delta t^*_{Eco,LC,t}$, it was necessary to first take into consideration the actual area burned in each time period t (reported in Figure 2-4, bottom) for the same ecoregion and land cover, and define an appropriate set of weights. This was done by estimating – in each ecoregion and land cover class - which proportion of the total burned area had a given persistence time $\Delta t^*_{Eco,LC,t}$.

The average annual burned area with persistence time Δt^* was estimated as follows:

$$\widehat{BA}_{Eco,LC}(\Delta t^*) = \frac{1}{n} \sum_{t=1}^{23} \frac{ba_{Eco,LC,t}}{\pi_{Eco,LC,t}} \delta_{\Delta t^*_{Eco,LC,t}}(\Delta t^*) \quad (11)$$

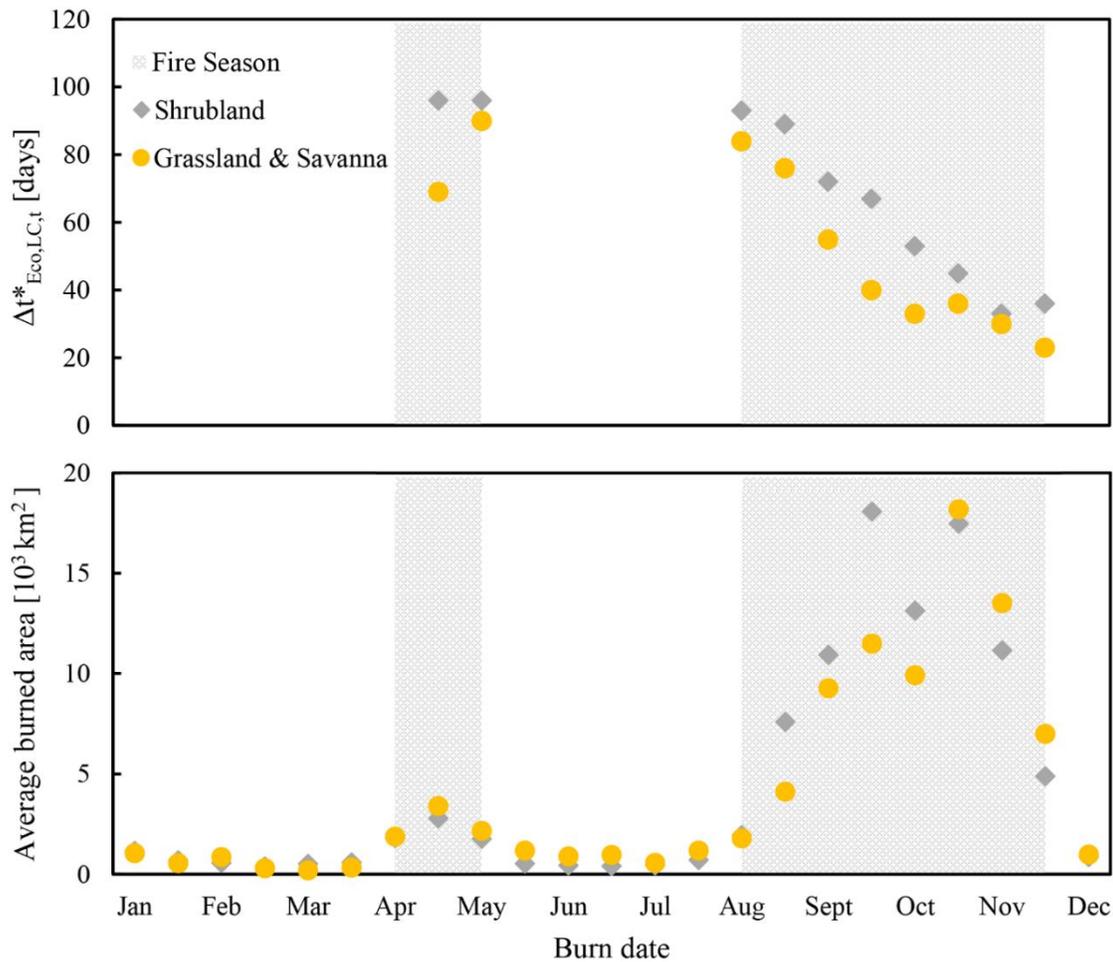


Figure 2-4: Burned area persistence time as a function of the day of burning $\Delta t^*_{Eco,LC,t}$ (*top*) and average annual burned area for each 16-days period of the year (*bottom*), in the “Victoria Plains tropical savanna” ecoregion in Northern Australia, stratified by aggregated land cover class. The persistence time was defined only for the 16-days periods of the fire season (hatched gray area). The Forest land cover class was not considered in this ecoregion because of negligible fire activity (i.e., less than 250 km² over the entire study period).

Where n is the number of years used in the analysis, $\delta_{\Delta t^*_{Eco,LC,t}}(\Delta t^*)$ is the Dirac measure assuming value 1 if the persistence time for stratum $\{Eco, LC, t\}$ is equal to Δt^* and 0 elsewhere, $ba_{Eco,LC,t}$ is the area of the sample extracted from each stratum, and $\pi_{Eco,LC,t}$ is the inclusion probability defined as in Equation 10. Figure 2-5 (*top*) presents the histogram of $\widehat{BA}_{Eco,LC}(\Delta t^*)$ obtained from the persistence times and average area burned of Figure 2-4.

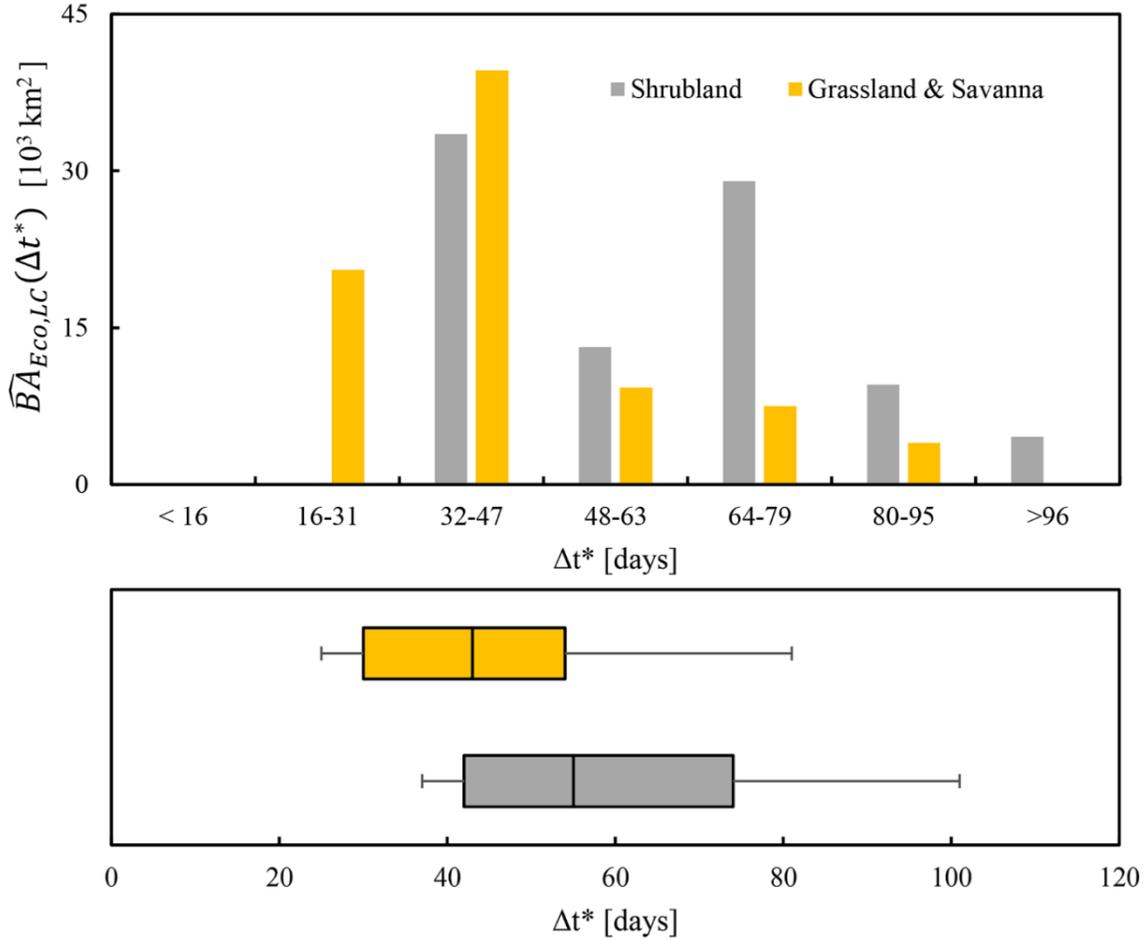


Figure 2-5: Summary persistence time metrics at the ecoregion level, in the “Victoria Plains tropical savanna” ecoregion in Northern Australia, stratified by aggregated land cover class. *Top*: Histogram of $\widehat{BA}_{Eco,LC}(\Delta t^*)$, i.e., the average annual burned area as a function of the burned area persistence time. The histogram is computed using 16 day bins. *Bottom* boxplot of $\widehat{BA}_{Eco,LC}(\Delta t^*)$, showing the median (i.e., $\widehat{\Delta t^*}_{Eco,LC}$), interquartile range, minimum and maximum of the distribution.

Finally, the estimated persistence time for each spatial stratum {Eco, LC} was the time corresponding to the 50th percentile of $\widehat{BA}_{Eco,LC}(\Delta t^*)$:

$$\widehat{\Delta t^*}_{Eco,LC} = \Delta t^* : \left\{ \int_0^{\Delta t^*} \widehat{BA}_{Eco,LC}(\Delta t^*) = 0.5 \right\} \quad (12)$$

Additionally, the interquartile range was computed as the difference between the 25th and 75th percentile of $\widehat{BA}_{Eco,LC}(\Delta t^*)$, and used to measure the variability of the burned area persistence time within each spatial stratum. Figure 2-5 (*bottom*) shows the estimated median persistence time (Equation 12) and the interquartile range for the ecoregion Victoria plains tropical savanna for the two land cover Shrubland and Grassland & savanna. In this ecoregion, the estimated persistence time $\widehat{\Delta t^*}_{Eco,LC}$ was

longer for Shrubland land cover (55 days) and 50% of the burned area had an estimated persistence within 42 and 74 days. In comparison, Grassland & savanna burned area had a shorter persistence (43 days) and variability (30 – 54 days of interquartile range).

2.4.2.4. Spatial aggregation by land cover and Biome/Realm

Because of the large number of spatiotemporal strata, it was necessary to summarize the persistence time results meaningfully in order to present and discuss them. The results were first summarized at the ecoregion level by use of the estimated median persistence time ($\widehat{\Delta t}_{ECO,LC}^*$, Equation 12), and interquartile range, and then aggregating them spatially by merging the ecoregions into larger spatial units based on the Olson’s biomes and realms (Australasia, Antarctic, Afrotropic, Indo-Malay, Neoarctic, Neotropic, Oceania, Palearctic) (Figure 2-6).

Realms provide a subdivision of the main landmasses, and biomes are a convenient stratification unit because of their homogeneity of climate and vegetation and because “they provide a framework for comparisons among units and the identification of representative habitats and species assemblages” (Olson et al. 2001). The TEOW biomes have been used to stratify studies for vegetation mapping (Lefsky 2010) and animal species distributions (Schipper et al. 2008), climatic models (Guenther 2006; Loarie et al. 2009), anthropogenic urban growth (Schneider et al. 2009), deforestation (DeFries et al. 2010) and, recently, for remote sensing global burned area products validation (Boschetti et al. 2016; Padilla et al. 2014). Following the approach from Boschetti et al. (2016), the 14 Olson biomes were aggregated into 5 more general biomes: Tropical, Temperate, Boreal, Mediterranean and Desert/Xeric biomes (Table 2-2).

Table 2-2: Biomes aggregation scheme of the Olson et al. (2001) biomes used for the present study.

Olson biome	Aggregated biome
Tropical and subtropical moist broadleaf forest	Tropical
Tropical and subtropical dry broadleaf forest	
Tropical and subtropical coniferous forest	
Tropical and subtropical grasslands, savanna	
Flooded grasslands, savanna (Latitude < 23°)	
Mangroves	
Temperate broadleaf and mixed forest	Temperate
Temperate coniferous forest	
Temperate grasslands, savanna	
Flooded grasslands, savanna (Latitude > 23°)	
Montane grasslands, savanna	
Boreal forest / Taiga	Boreal
Tundra	
Mediterranean forests, woodlands, and shrublands	Mediterranean
Deserts and xeric shrublands	Desert/Xeric

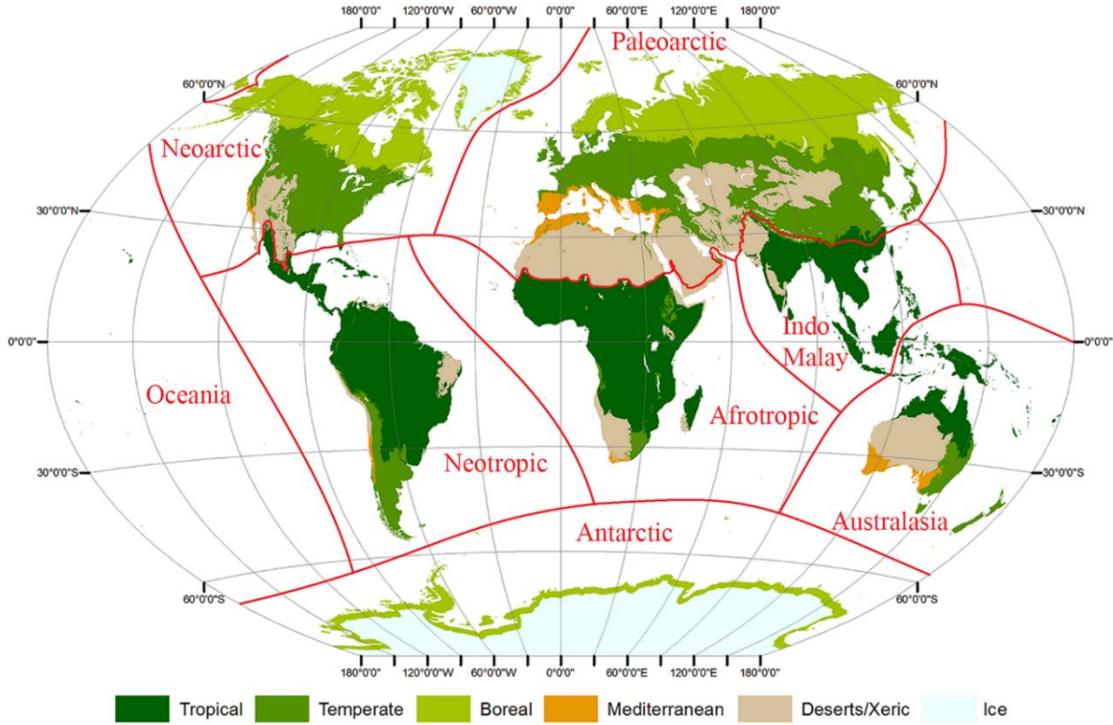


Figure 2-6: Realms and aggregated biomes used as spatial units to present the global summary metrics persistence time. Oceania and Antarctic realms (in gray) were not considered, because of negligible fire activity.

The realms of Oceania and Antarctica were excluded from the analysis because of the lack of significant fire activity. Of the 30 possible combinations of the remaining 6 realms and 5 biomes, 25 were valid (the Boreal biome is present only in the Palearctic and Neartic realms and the Mediterranean biome is not present in the Indo-Malay realm) and were adopted as aggregated spatial units for the summary of the analysis results (Figure 2-6).

The distribution of annual burned area as a function of the spectral signal persistence time Δt^* for each spatial unit defined by Biomes (B), Realms (R), and land cover (LC) was simply calculated by summation of all the ecoregions belonging to the spatial unit, and globally by summation over the realms:

$$\widehat{BA}_{B,R,LC}(\Delta t^*) = \sum_{Eco \in R,B} \widehat{BA}_{Eco,LC}(\Delta t^*) \quad (13)$$

The median persistence time for each Biome, Realm, and land cover $\widehat{\Delta t}_{B,R,LC}^*$ is the time corresponding to the 50th percentile of $\widehat{BA}_{B,R,LC}(\Delta t^*)$, and is computed as in Equation 12.

$$\widehat{\Delta t}_{B,R,LC}^* = \Delta t^* : \left\{ \int_0^{\Delta t^*} \widehat{BA}_{B,R,LC}(\Delta t^*) = 0.5 \right\} \quad (14)$$

2.5. Results

2.5.1. Fire activity distribution

Globally, from 2003 to 2016, the MODIS MCD64A1 product detected on average $4.0 \cdot 10^6 \text{ km}^2$ of burned area per year: $0.19 \cdot 10^6 \text{ km}^2$ in Forest, $0.31 \cdot 10^6 \text{ km}^2$ in Shrubland, $3.1 \cdot 10^6 \text{ km}^2$ in Grassland & Savanna, and the remaining $0.40 \cdot 10^6 \text{ km}^2$ occurring in croplands and other land covers not considered in this study. The average annual area burned, reported by land cover and biome and the geographic distribution of area burned by land cover and ecoregion are displayed in Figure 2-7. The full result tables are presented as supplementary online material (Table S1). Out of the 867 ecoregions defined by the TEOW map, 511 ecoregions had more than 250 km^2 of burned area detected in at least one land cover class over the study period, and were considered in the analysis (Section 3.2.1): 300 ecoregions were considered for the Forest land cover class, 135 for Shrubland and 392 for Grassland & Savanna (Figure 2-7, *right column*). The majority of the burned area in Forest land cover was detected in the Tropical (65.9%), Boreal (19.1%) and Temperate (13.8%) biomes; the majority of the burned area in Shrubland was detected in the Desert/Xeric (50.0%) and Tropical biomes (41.6%). The majority of the burned area in Grasslands & Savannas was detected in the Tropical (93.4%) and Temperate (3.9%) biomes (Figure 2-7, *left column*).

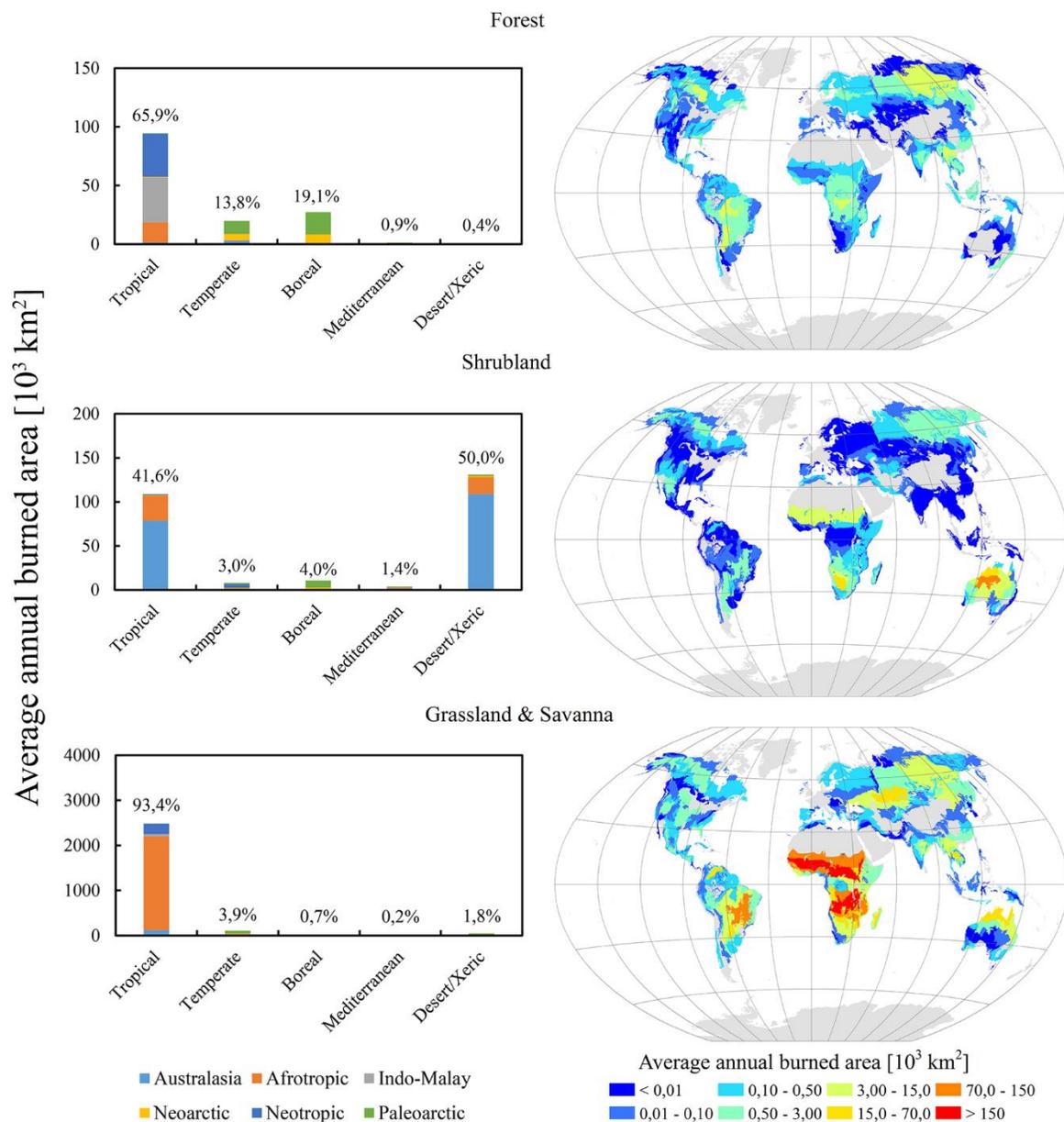


Figure 2-7: Global burned area distribution, as detected by the MODIS global burned area product (MCD64A1) in the 2003-2016 period, stratified by aggregated land cover class. *Left*: histogram of the partition of the average annual burned area by biome and realm. *Right*: average annual burned area detected in each ecoregion. The ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.

The duration of the fire season was calculated in each ecoregion as described in section 3.2.1. (Table S2). Figure 2-8 shows the global distribution of the fire season duration at the ecoregion level. The median duration of the fire season of the ecoregions considered was 112 days (7 16-days periods), the shortest 32 days (2 periods) and the longest 304 days (19 periods).

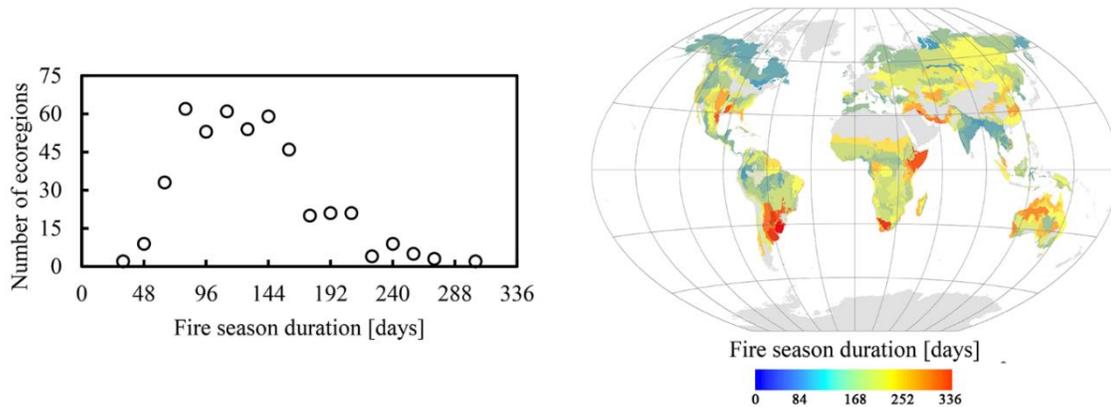


Figure 2-8: Histogram (*Left*) and map (*Right*) of the fire season duration, computed for the ecoregions considered in the analysis. The ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.

2.5.2. Ecoregion level burned area persistence time

The burned area persistence time $\Delta t_{Eco,LC,t}^*$ for each ecoregion (*Eco*), land cover class (*LC*) and 16-days period (*t*) of the fire season (Table S3), as well as the summary metrics $\widehat{BA}_{Eco,LC}(\Delta t^*)$ and $\widehat{\Delta t}_{Eco,LC}^*$ (Table S4) were estimated following the methodology described in Section 3.2.3. The median persistence time $\widehat{\Delta t}_{Eco,LC}^*$ in each ecoregion and land cover, and the corresponding interquartile ranges, are displayed Figure 2-9.

In Forest, 101 ecoregions (34%), 195 ecoregions (66%), and 230 ecoregions (77%) had a median persistence time shorter than 16, 32, and 48 days respectively (Figure 2-10, *orange bars*). Figure 2-9, *top row*, shows that the ecoregions with the shortest median persistence times are generally in the tropics and those with the longest persistence times are in North America, Western Mediterranean, and South Eastern Australia. The ecoregions with the highest interquartile range, which represents the persistence time variability during the fire season, are mostly located in Eurasia.

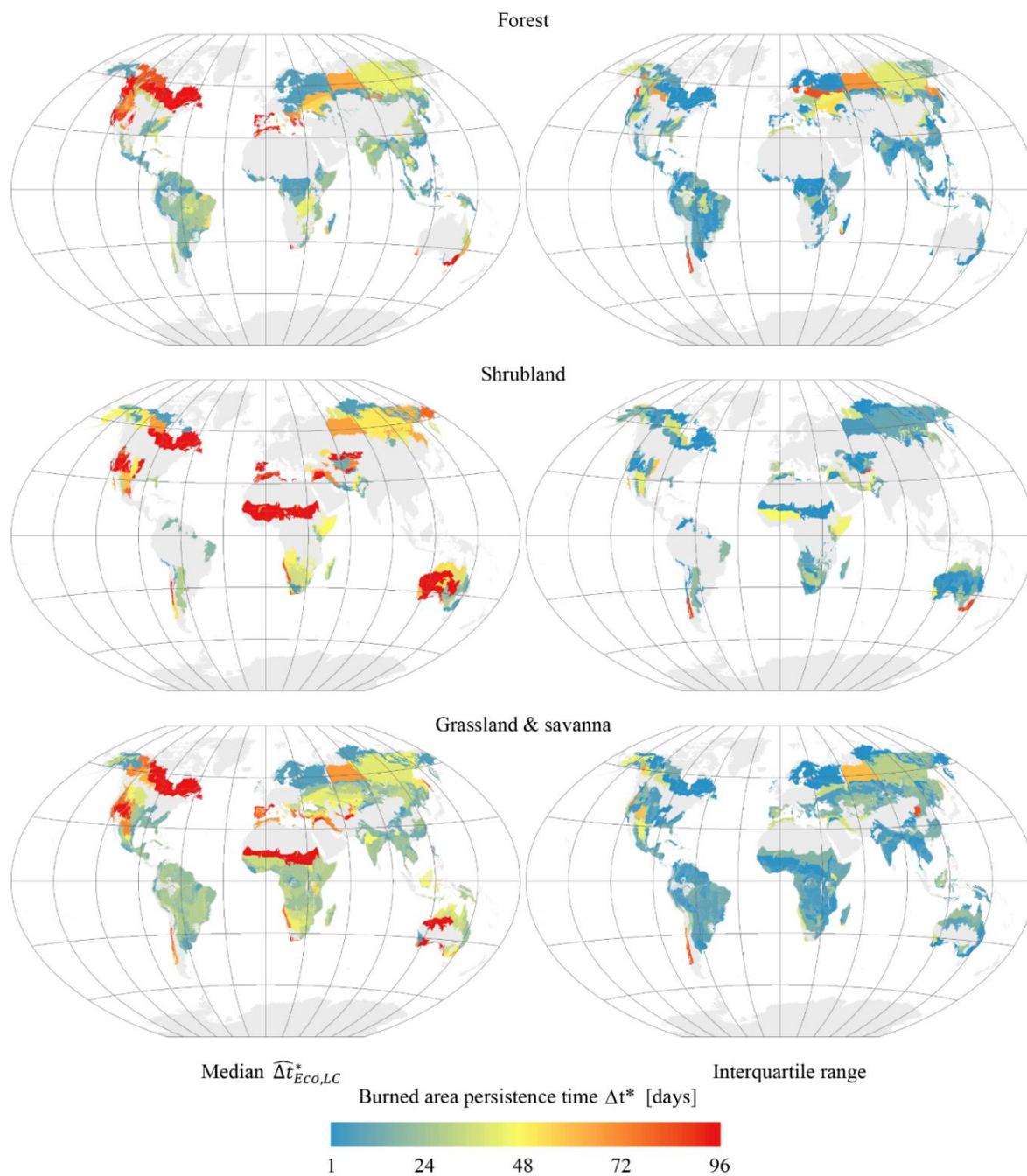


Figure 2-9: Burned area persistence time calculated for each ecoregion, stratified by aggregated land cover class: Forest (*top row*), Shrubland (*middle row*), and Grassland & Savanna (*bottom row*). The median persistence time $\widehat{\Delta t}_{Eco,LC}^*$ (*left column*) and interquartile range (*right column*) are shown using a rainbow color scale; the ecoregions not considered in the analysis because of negligible fire activity are displayed in grey.

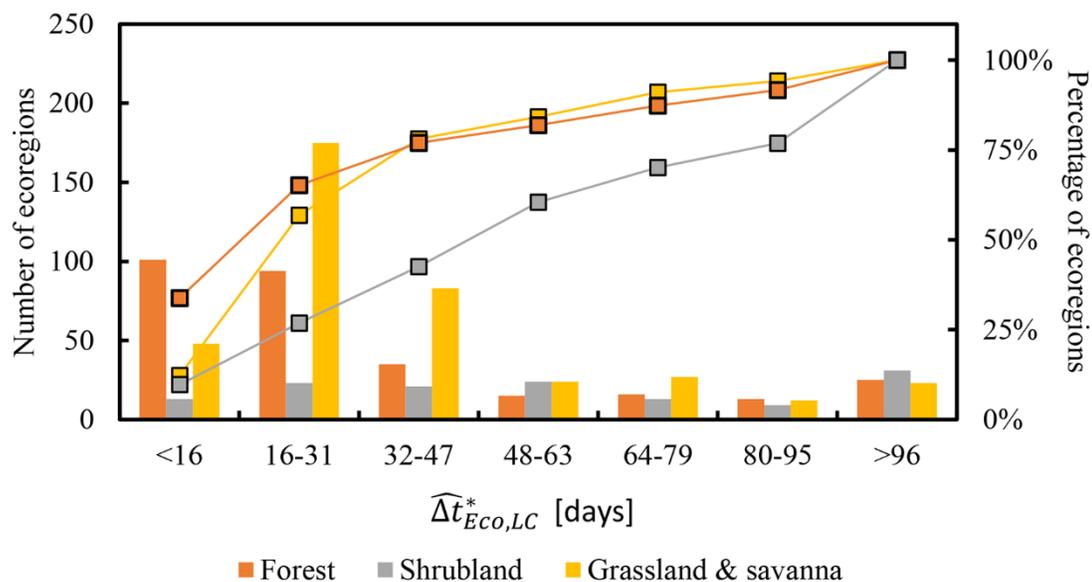


Figure 2-10: Histogram bars and cumulative lines of the estimated persistence time $\widehat{\Delta t}_{Eco,LC}^*$ calculated for all the ecoregions and stratified by aggregated land cover class. The histograms were computed using 16-days bins, up to 96 days after the date of burn.

In Shrubland, only 13 (10%), 36 (27%) and 57 (42%) ecoregions had a median persistence time under 16, 32, and 48 days respectively, whereas 31 ecoregions (23%) have a persistence time greater or equal than 96 days (Figure 2-10, *gray bars*). Figure 2-9, *middle row*, shows that the ecoregions with the shortest median persistence time are located in boreal regions, Southern Africa and South America. The longest persistence times correspond to sub-Saharan and Mediterranean ecoregions, North America, Western and Central Australia.

Finally, in Grassland and Savanna 48 (12%), 223 (57%) and 306 (78%) ecoregions had a median persistence time under 16, 32, and 48 days respectively (Figure 2-10, *yellow bars*). Figure 2-9, *bottom row*, shows that the ecoregions with the shortest median persistence time are located in Boreal regions and South-East Asia; those with the longest persistence time are located in North Africa, North America and some regions in central Australia. For this land cover, the great majority of the burned area was detected in the tropical ecoregions of Africa, South America and Australia (Figure 2-7, *bottom row*). All these ecoregions, with few exceptions, have a median persistence time between 16 and 48 days, with small variability (interquartile range) (Figure 2-9, *bottom row*).

2.5.3. Biome/Realm aggregation of the burned area persistence time

The ecoregion-level results were aggregated into summary metrics for each land cover and Realm/Biome spatial unit, applying the methods described in Section 3.2.4. (Table S5). These results are presented in the following section, which is organized by land cover class.

2.5.3.1. Forest land cover

Forests are among the largest carbon pools of the globe (Dixon et al. 1994) and forest fires are one of the main carbon sources for the atmosphere, contributing to 24.8% of the total global carbon emission from fires (van der Werf et al. 2017) even though forests contribute to just 4.75% ($0.19 \cdot 10^6 \text{ km}^2$) of the MODIS global annual area burned. Of this area, the majority was detected in Tropical (65.9%), Temperate (13.8%) and Boreal (19.1%) biomes, with a negligible contribution of the other biomes (Figure 2-7, *top left*). Although from an ecological perspective fire effects persist for decades in most forests, the separability of the spectral signature of burned and unburned areas declines quickly. The global distribution of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ (including all the 5 aggregated biomes) (Figure 2-11, *bottom right*) indicated that 66.4% of the annual burned area ($0.09 \cdot 10^6 \text{ km}^2$ per year) had a persistence time of less than 32 days, and less than 48 days for 79.4% of the annual burned area ($0.10 \cdot 10^6 \text{ km}^2$ per year); only 7.2% had a persistence time greater than 96 days. The biome level distribution of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ is also displayed in Figure 2-11 and complemented by the corresponding box plots reporting the median time $\widehat{\Delta t}_{B,R,LC}^*$ and interquartile range (Figure 2-12). In Tropical biomes the distribution of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ was unimodal, with short persistence times (80.5% of the area has a persistence under 32 days). In temperate biomes the distribution was largely bimodal; the first mode is largely due to burned areas detected in the Paleoarctic realm (where 28.0% of the area has a persistence of less than 16 days), the second mode, at $\Delta t^* \geq 96$ days, was mainly due to burned area detected in Australasia and Neoarctic realm. In boreal biomes, the distribution is trimodal, due to different behavior in the Neoarctic and Paleoarctic realms. The Paleoarctic realm was characterized by shorter persistence time ($\widehat{\Delta t}_{B,R,LC}^*$ is 44 days) and bimodal distribution with a first peak at $\Delta t^* < 16$ and a second peak at $48 \leq \Delta t^* < 64$. The Neoarctic had longer persistence time ($\widehat{\Delta t}_{B,R,LC}^*$ is >96 days) and unimodal distribution. As a result, the global distribution of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ (Figure 2-11, *bottom right*) was a combination of short persistences, mainly from Tropical and Paleoarctic Temperate and Boreal burned areas, and long persistences, mainly from Australasia Temperate and Neoarctic Temperate and Boreal burned areas.

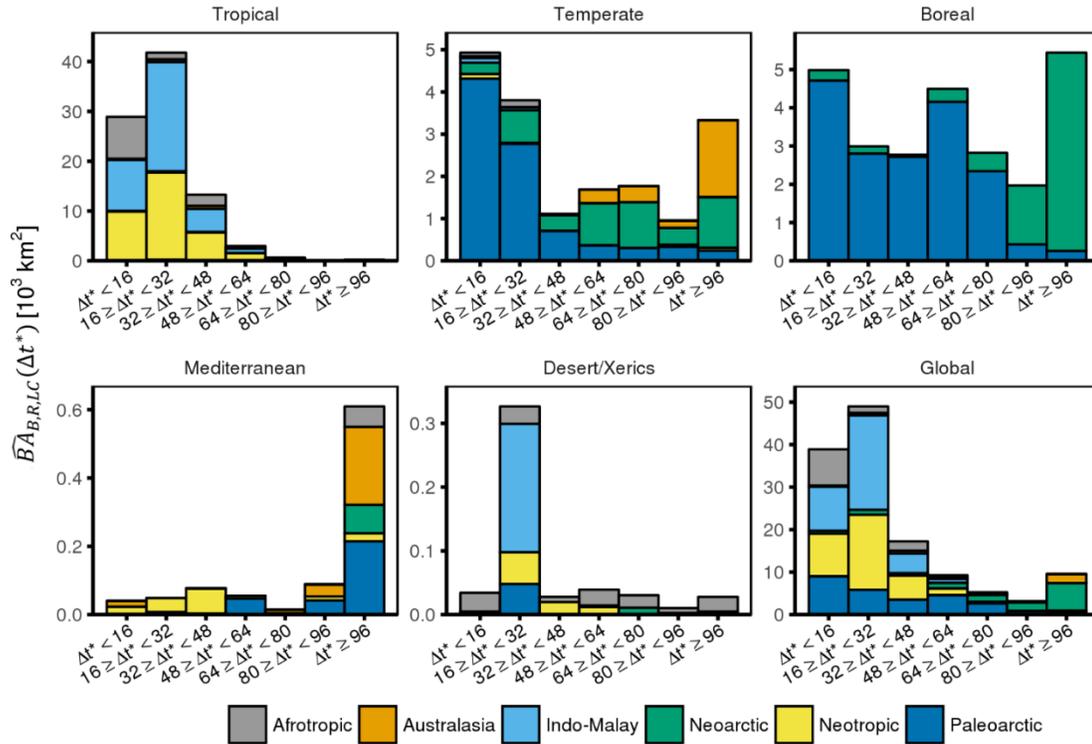


Figure 2-11: Histograms of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Forest aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.

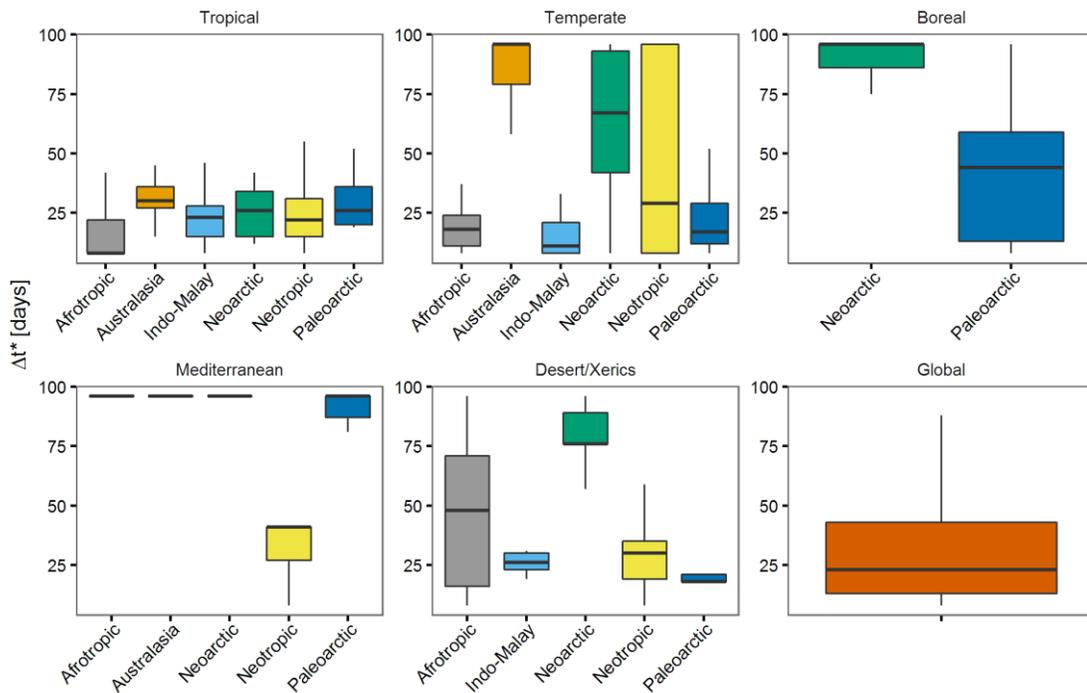


Figure 2-12: Boxplots of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Forest aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\widehat{\Delta t}_{B,R,LC}^*$), interquartile range, minimum and maximum of the distribution.

2.5.3.2. Shrubland land cover

The vast majority of the burned area in the Shrubland aggregated land cover class occurred in the Australasia (73.1%) and Afrotropic (18.7%) realms. In particular, it was a very significant class in the Australasian realm, where it contributed to 46.7% of the average annual burned area. The area burned was almost evenly split between the Tropical (41.6%) and Desert/Xeric (50.0%) biomes, with a negligible contribution of the other biomes (Figure 2-7, *center left*). Different persistence times were found in different biomes: shorter in the Tropical biomes, where the mode of the $\widehat{BA}_{B,R,LC}(\Delta t^*)$ histogram was between 16 and 32 days; and longer in the Desert/Xeric biomes, where the mode was greater than 96 days. The global histogram is bimodal, and reflects the distinct distribution of these two dominant biomes (Figure 2-13).

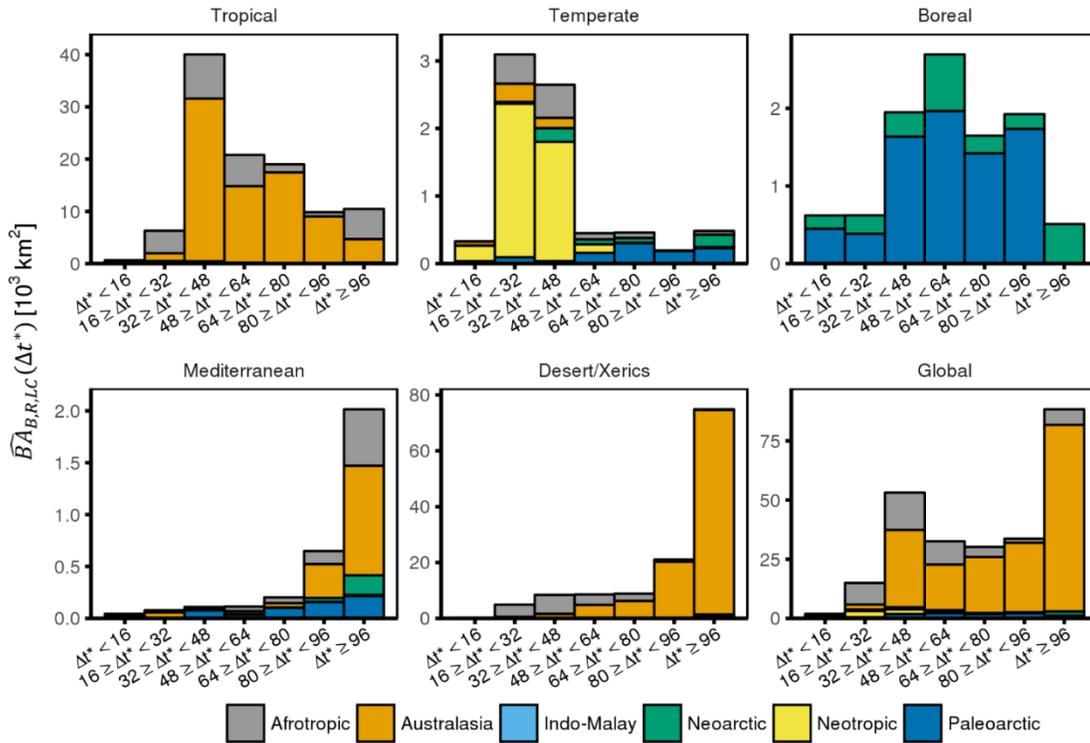


Figure 2-13: Histograms of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Shrubland aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.

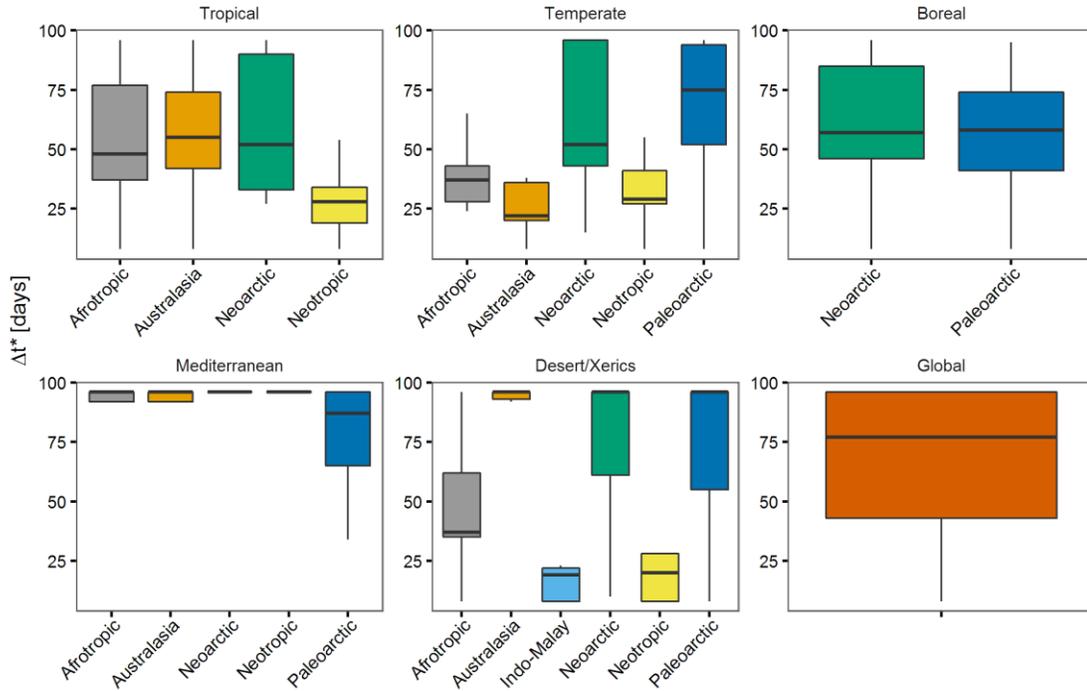


Figure 2-14: Boxplots of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Shrubland aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\widehat{\Delta t^*}_{B,R,LC}$), interquartile range, minimum and maximum of the distribution.

2.5.3.3. Grassland & savanna land cover

The greater majority (93.4 %) of the burned area for Grassland & savanna was detected in the Tropical biome. Only 3.9% was detected in the Temperate biome, with a negligible contribution of the other biomes (Figure 2-7, *bottom row*). In particular, the majority of the global burned area detections were concentrated in the Tropical biome of the Afrotropic realm (76%). Globally, the mode of the $\widehat{BA}_{B,R,LC}(\Delta t^*)$ histogram (Figure 2-15, bottom right) was between 16 and 32 days (60.5% of the annual burned area). Overall, the persistence time was short: 91% of the area burned had a persistence time within 16 and 48 days, resulting in a median estimated time $\widehat{\Delta t^*}$ of 29 days (Figure 2-15, Figure 2-16). At biome level, the histograms of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ for Tropical and Temperate biomes were both unimodal. The histogram of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ had the maximum within 16 and 32 days for Tropical biomes and within 32 and 48 days for Temperate biomes. A total of 92.7% of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ for Tropical biomes and 80.5% for Temperate biomes was associated with Δt^* within 16 and 48 days. The global histogram was unimodal, and reflects the distribution of the dominant Tropical biomes and a small contribution from the Temperate biomes (Figure 2-15).

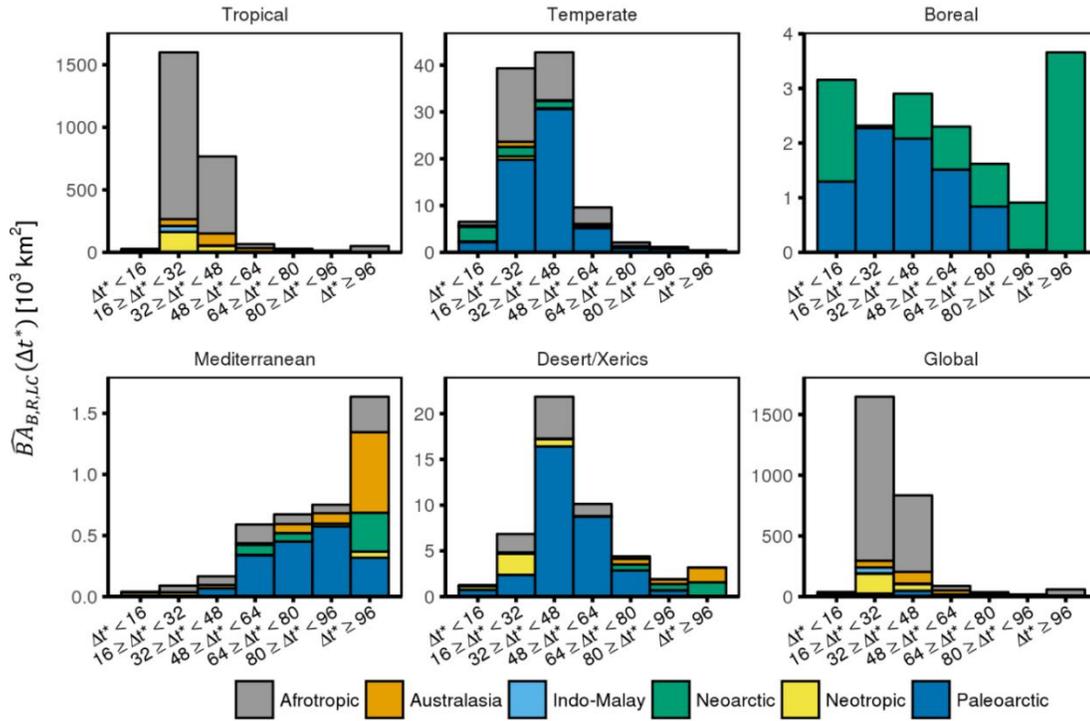


Figure 2-15: Histograms of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Grassland & savanna aggregated land cover class, presented for each aggregated biome and globally. The colors represent, in each bar of the histograms, the stratification by realm.

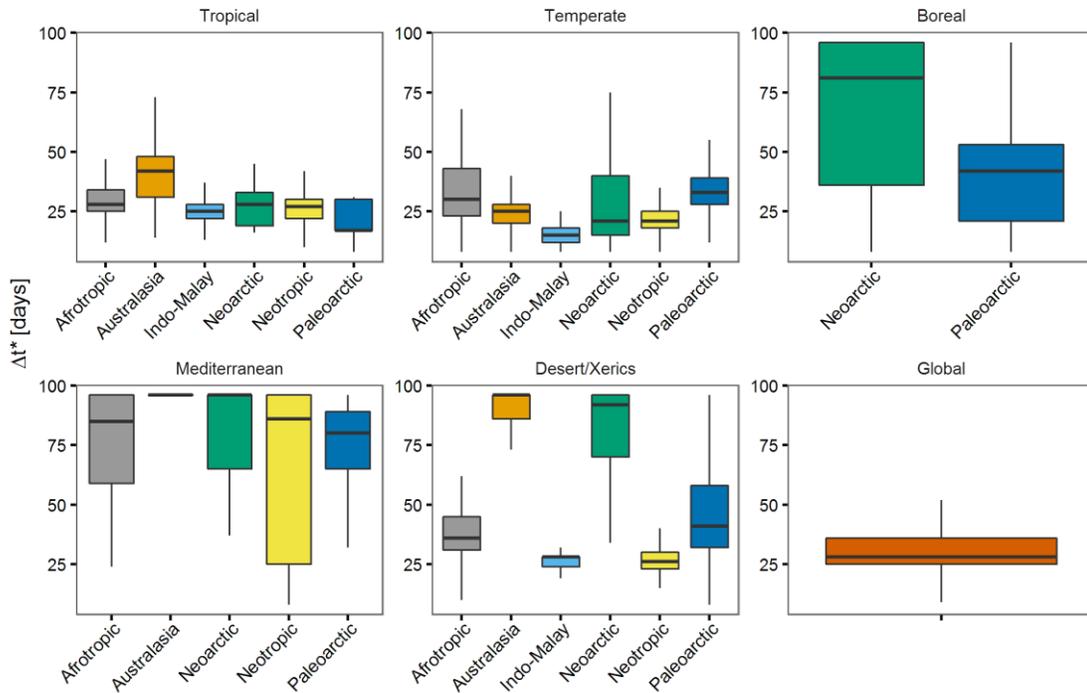


Figure 2-16: Boxplots of $\widehat{BA}_{B,R,LC}(\Delta t^*)$ in the Grassland & savanna aggregated land cover, organized by aggregated biome and globally. The boxplots show the median (i.e., $\widehat{\Delta t}_{B,R,LC}^*$), interquartile range, minimum and maximum of the distribution.

2.6. Discussion

2.6.1. Ecoregion level burned area persistence time

The results at ecoregion level confirmed that the separability of burned and unburned surfaces quickly decreases after burning, even in ecosystems where the fire effects on the vegetation are persistent for several years. In addition to the dissipation of the charcoal and ashes, vegetation regrowth in burned areas, phenology of unburned vegetation, and several other factors might have contributed to these results.

In this study, the post-fire spectral temporal variations were analyzed by using NBR time series of burned and unburned pixels, as identified by the MCD64A1 global burned area product. Detection errors in the MCD64A1 product influenced the results, by reducing the separability of the burned (Equation 4) and unburned (Equation 5) probability distributions. More specifically, omission errors resulted in the inclusion of burned pixels in the unburned population ub , and biased positively the function F_{UB} (Equation 7). Conversely, commission errors caused the inclusion of unburned pixels in the burned population ba , biasing negatively the function F_B (Equation 6). As a result, both omission and commission errors in the MCD64A1 product resulted in a shorter estimated persistence time (Equation 8).

Furthermore, the maximum persistence time Δt^*_{jday} (Equation 8) is limited by the length of the post-fire time series of nadir BRDF-adjusted reflectances (MCD43A4). Because at least 7 cloud-free MODIS observations within 16-days period are needed to perform a valid BRDF inversion (Wang et al. 2018), in regions with clearly defined dry and wet seasons, or with persistent snow cover in the winter months, no MCD64A4 data were available for part of the year. If the fire season was close to the beginning of the wet or snow season, the estimated persistence time was effectively limited to the duration of the cloud-free period.

Cropland land cover was not considered in our analysis since the MODIS burned area product is not suited to capture the size and heterogeneity of cropland burned area (Giglio et al. 2009; Lasko et al. 2017) because of its geometrical resolution, and because other agricultural practices (e.g. tilling) often follow immediately the fires (Hall et al. 2016) and have similar spectral signatures (Vanderhoof et al. 2017). Omission errors in the cropland class of the MCD12Q1 land cover product, therefore, led to the inclusion of MCD64A1 detections over agricultural areas in the burned population ba . Some of these were genuine detections, but had a very short persistence time; some were commission errors and – as detailed above – negatively biased the F_B function; in either case, the result was an underestimation

of the persistence time. Additional errors in the MCD12 land cover map also influenced the estimated persistence time, by reducing the effectiveness of the stratification by aggregated land cover type.

Finally, caution should be used in considering the estimated burned area persistence time, as defined in the present work: it should not be confused with the ecological concept of vegetation recovery to pre-fire level (Engel and Abella 2011; Meney et al. 1994; West and Hassan 1985). The duration of the persistence time can be limited by the recovery of NBR values to pre-fire levels after rapid re-vegetation or other factors, such as snowfall, that occlude the charcoal and bare soil and mask the spectral signature associated with burned areas. Additionally, errors in the MODIS burned area and land cover product could have influenced the stratification of the analysis and the division in the two population of burned and unburned pixels reducing the estimated persistence time.

2.6.2. *Biome/Realm level burned area persistence time*

2.6.2.1. Forest land cover

The majority of the burned area in Tropical biomes was detected in the Afrotropic, Indo-Malay and Neotropical realms (Figure 2-7, *top left*). In these three realms, the median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ was respectively 8, 24 and 26 days (Figure 2-11 and 2-12, *top left*). Compared to the other biomes, tropical forests had shorter persistence times, arguably due to several factors. Many tropical forests have dense canopies that keep surface fuels moist and humidity high, limiting fire spread and duration, and allowing vegetation to recover quickly (Cochrane 2003). New photo-synthetically active vegetation can alter the post-fire spectral signature sensed by EO satellites and deteriorates their ability to detect burned areas. For example, after a tropical forest fire in the Congo Basin, a new surface layer of vegetation established (*Marantaceae*) and covered the burned area in less than 3 months after the burning event (Verhegghen et al. 2016).

Additionally, the estimated persistence time was influenced by the nature of tropical forest fires. Anthropogenic activities account for the majority of burning events in the tropical Amazon rainforest (Cochrane et al. 1999; Kumar et al. 2014), African rainforests (Bucini and Lambin 2002; Eva and Lambin 2000; Tovar et al. 2014) and Indo-Australian rainforests (Hope et al. 2004). In central Africa, human-ignited rainforest fires are strongly associated with land cover changes and they are influenced by fire activity along the edges of the forest (Bucini and Lambin 2002). In Amazonia, fire activity is connected to other disturbances having spectral similarities to burned areas such as logging and deforestation which only complex algorithms can discriminate (Morton et al. 2011). In both these cases, post-fire NBR may recover rapidly in areas planted immediately after clearing, where vegetation regrows within weeks (Cochrane 2003). In Southern China, various forest fire regimes were found

(Chen et al. 2017); regions characterized by long fire season and frequent fires, mainly due to human activity, show lower estimates of persistence time compared to regions characterized by shorter fire season and infrequent fires.

The majority of the burned area in Temperate biomes was detected in Australasia, Neartic and Neotropic realms (Figure 2-7, *top left*). In these three realms, the median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ was respectively >96, 67 and 29 days (Figure 2-11 and 2-12, *top center*). In Australasia, these results were consistent with the occurrence of high intensity, low frequency fires in dense, shade-tolerant understory vegetation in many temperate forests in Australia (Archibald et al. 2013; Close et al. 2009). For the Neartic realm temperate forests, the median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ was the result of variable $\widehat{\Delta t}_{Eco,LC}^*$ at ecoregion level. The longest persistence time estimates occurred in mid-elevation, Northern Rockies ecoregions typically affected by large crown fires (Clark et al. 2016; Westerling et al. 2006). The shortest persistence times were in the southeast US, which is consistent with frequent surface fires and rapid post-fire grass regeneration observed in that region (Glitzenstein et al. 1995; Slocum et al. 2003). $\widehat{\Delta t}_{B,R,LC}^*$ estimated for the Paleartic realm temperate forests was largely influenced by the presence of anthropogenic fires linked to land conversion for roads and croplands in the western China and Russia-China borders regions (Li et al. 2012), in which the MODIS MCD64A1 product detects the majority of burned areas for this combination of realm and realm (Figure 2-7, *top right*).

The Boreal biome is present only in the Neartic and Paleartic realms, accounting for 29.9% and 70.1% percent of the annual burned area respectively (Figure 2-7, *top left*). The estimated median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ was 44 days in the Paleartic realm, and > 96 days in the Neartic (Figure 2-11 and 2-12, *top right*). This large difference was likely due to the different burning conditions in these two realms. Fires in Paleartic boreal forests are dominated by low intensity surface fires, and typically result in smaller sized fires than Neartic boreal forests fires, where instead high intensity crown fires are predominant (de Groot et al. 2013; Wierzchowski et al. 2002). Furthermore, the fire temporal distribution in the Paleartic realm was consistent with anthropogenic agricultural fires that typically occur in spring and fall to clear extensive areas of surface fuels (de Groot et al. 2013; Hall et al. 2016). The size and timing of Paleartic fire reduced the spectral separability of burned/unburned pixels in boreal biomes because of limited data quality (e.g. mixed pixels resulting from land-snow patterns and vegetation phenology) and availability (e.g. spatial resolution, cloud-free data availability) (Chu and Guo 2014). For the Boreal biomes, fire season is often followed closely by the snow season. As discussed in Section 4.2, the estimated persistence time was influenced by missing data due to snow

cover, especially when the majority of burned areas were detected at the end of the fire season, such as in the Paleoarctic Boreal realm.

2.6.2.2. Shrubland land cover

The median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ for the Australasia realm was 55 days for the Tropical biome and >96 days for the Desert/Xeric biome. Tropical fires in Shrubland land cover are typically characterized by a shorter persistence of char on the ground due to removal by wind and atmospheric agents; the observability is also limited by cloud cover immediately following the fire season (Pereira 2003). For example, Bowman et al. (2003) found the temporal persistence of the burned areas in tropical regions of the Australian northern territories to be less than 100 days. Conversely, fire scars in arid zones are typically more persistent as – once the charcoals and ash are dissipated – there is little vegetation regrowth until the wet season, leading to persistent soil exposure. The NBR value of soil is significantly different from unburned vegetation (Burrows et al. 2006), leading to a longer estimated persistence time.

The median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ in the Afrotropic realm was 48 days and 37 days for the Tropical and the Desert/Xeric biomes respectively. In both biomes, the persistence time had a large variability with 37-77 days and 35-62 days interquartile ranges (Figure 2-14, grey boxplot of *Tropical* and *Desert/Xeric* plots). Tropical shrubland fires are concentrated in the Sahelian zone in North equatorial Africa, including areas in East Africa, and the Kalahari region in the South equatorial Africa (Houghton and Hackler 2006). In the central Kalahari, the majority of fire events can interest only the herbaceous vegetation while woody fuels are unburned; and only higher intensity fires are able to burn even the woody-cover dominated areas (Mishra et al. 2016). The NBR time series over burned areas showed variable changes in magnitude depending on the amount of unscorched vegetation within a MODIS pixel, and, as a result, the persistence time duration was variable. Finally, the persistence time large variability is potentially due to the different anthropogenic influence of fire activity over managed and protected land (Grégoire and Simonetti 2010).

2.6.2.3. Grassland & savanna land cover

Grasslands and savannas are fire-prone ecosystems and fire is an essential factor contributing to the maintenance of the ecosystem: where resources availability and climatic control would allow tree seedling establishment, fire decreases tree cover density and maintains grassland and savanna ecosystems (Bond et al. 2005; Bond and Keeley 2005; Staver et al. 2011). The short estimate and the small variability of the burned area persistence time for this land cover (Figure 2-16, *bottom right*) could be explained by the common features that grassland and savanna share across the continents. Fine

fuels and low fuel loadings result in a shorter persistence of the char residue signal; cloud cover and the frequent presence of extensive smoke aerosol layers result in a limited availability of satellite optical data, and the dominance of surface fires exacerbate the problem of detecting burns in woody savannas (Pereira 2003).

The majority of the burned area in Tropical biomes was detected in the Afrotropic, Australasia and Neotropical realms (Figure 2-7, *bottom left*). In these three realms, the median persistence time $\widehat{\Delta t}_{B,R,LC}^*$ was respectively 28, 42 and 27 days (Figure 2-15 and 2-16, *top left*). In the tropics and subtropics, the total amount of burned area and the interannual variability are controlled by the complex interaction between climate, fuel, and human activity (Giglio et al. 2013; van der Werf et al. 2006; van der Werf et al. 2008). The spectral characteristics of vegetation and burned areas change depending on the acquisition date (Roy and Landmann 2005) and they are highly dynamic: rapid re-vegetation after a fire and the presence of unburned fuel loads can significantly impact the spectral signal (Pereira 2003; Trigg and Flasse 2000).

The persistence time for Grasslands & savanna was in accordance with other studies in tropical savannas. Scholes and Walker (2004) reported a recovery to pre-fire albedo values 42 days after a fire in a southern Africa savanna. Likewise, ground-based reflectance was observed to be within 20% of pre-fire values within 14 days of burning events in western African savannas (Frederiksen et al. 1990; Langaas and Kane 1991). Eva and Lambin (1998) analyzed the spectral reflectance temporal dynamics of burned and unburned woodland savannas in Central Africa using ATSR data and concluded that the discrimination between burned and unburned areas was affected by an error of 20% after 35 days from the fire event. Majority of the burned area persistence time for Tropical Australasia realm was longer than the Afrotropic realm (Figure 2-16, *top left*). This difference was likely caused by the different dominant fire regimes of the two realms. In both realms, savanna's fire frequency is annual/biannual; Australia has larger, more intense fires whereas Africa is characterized by smaller, less intense fires (Archibald et al. 2013). Australia tropical savanna fire intensity rises as the dry season progresses due to more extreme fire weather and fire-prone fuel conditions in the later season (Cheney et al. 1993; Gill et al. 1996; Williams et al. 1999). A variable degree of change in the spectral signal caused by the action of fire was reported in Tropical grasslands and savanna (Roy and Landmann 2005), which is consistent with the variability of Grassland & savanna persistence time results over Afrotropic and Australasia Tropical biomes (Figure 2-16, *top left*).

The majority of the burned area in Temperate biomes was detected in the Afrotropic, Neartic, and Palearctic realms (Figure 2-7, *bottom left*). In these three realms, the median persistence time

$\widehat{\Delta t}_{B,R,LC}^*$ was respectively 30, 21 and 33 days (Figure 2-15 and 2-16, *top center*). Like other biomes, Grassland & Savanna fire activity in temperate biomes is mainly driven by climate and human activity. In North American temperate grassland ecoregions, antecedent precipitation amount and the drought index significantly influence fire activity as high-precipitation years can increase fuel load and subsequent fire intensity in drought years that follow (Littell et al. 2009).

Agricultural fires detected in the MODIS land cover product were discarded in this work. Nevertheless, misclassified crop, managed prairie and pasture fires were likely included in the analysis and could reduce the estimated burned area persistence time, especially in ecoregions of the Paleoarctic realm where the use of fire is widespread as agricultural management tool for removal of excess residue and the control of diseases and pests (McCarty et al. 2012). Typically, temperate grasslands are intensively managed, including fire and grazing practices that alter nutrient cycling and the distribution of organic matter (Bond and Keeley 2005; Wessman et al. 1997). The relatively short burned area persistence time for temperate grassland was also observed in Canadian grassland fires. In these grasslands, Lu and He (2014) estimated the post-fire vegetation recovery using multi-temporal Landsat imagery (one, two and three months after the fire) and concluded that grassland has a strong post-fire recovery capacity, especially if there is an adequate water availability. Prescribed fires are commonly used in tallgrass prairie ecosystem and they are characterized by a relatively small size high degree of combustion heterogeneity, for this reason, the MODIS coarse resolution could be not sufficient for an adequate mapping of this type of fires (Mohler and Goodin 2012).

2.7. Conclusions

Coarse resolution sensors typically have daily temporal resolution, and it is generally assumed that a sufficient number of observations will be available to map burned areas globally. However, moderate resolution satellites have reduced temporal resolution (e.g. 16 days for Landsat, Sentinel-2A/B 5 days), which could potentially lead to large burned area omission errors in ecosystems where the spectral signal associated with burning disappears quickly, especially if more than one post-fire observation is required to detect the burn. This paper attempted to calculate the burned area persistence time to estimate the period of time after the burn date in which burned areas are detectable using remotely sensed data change detection.

The presented methodology estimated the burned area persistence time defined as the duration of the spectral separability of burned / unburned areas, using time series of Normalized Burn Ratio (NBR) spectral index computed from MCD43A4 BRDF-adjusted MODIS reflectances. Burned and

unburned areas were defined based on the MCD64A1 MODIS global burned area product. The full global MODIS record for the years 2003-2016 was used.

Among the novel aspects of the proposed methods, a probabilistic method for the analysis of spectral separability over time and space was developed, which can be extended to the analysis of other disturbances and, in general, of other non-permanent land use or land cover changes. The methods were stratified spatially (ecoregion and land cover) and temporally (time of the year) and provided for a rigorous generalization at coarser units.

The results showed that the persistence time was highly variable in time and space. Not only it was different across land cover classes, but within the same class, it depended on biome and realm. The shortest burned area persistence times were found over Tropical Forests confirming that the separability of burned and unburned surfaces quickly decreases after burning, even in ecosystems where the fire effects on the vegetation are persistent for several years. The persistence times in Tropical forests were mostly limited by the rapid surface vegetation regrowth and the influence on the spectral signal of unscorched canopies. The longest persistence times were found in Desert/Xeric Shrublands, characterized by little vegetation regrowth until the wet season, leading to persistent soil exposure spectrally different from the unburned. Burned area persistence times over Grassland & savanna, mostly detected in the African continent, were short and with small variability (16 – 48 days) due to the fine fuels and fuel loadings typical of savanna fires.

Ultimately, the results indicated that, globally, burned areas can be detected for a limited time by optical satellite sensors: globally the median burned area persistence time was estimated in 29 days and 86.6% of the area detected as burned in the MODIS MCD64A1 product can be mapped accurately only for a maximum of 48 days. Thus, results indicate that burned area persistence time can be a limiting factor for global burned area mapping from moderate resolution satellite sensors, which have a low temporal resolution. Future research is suggested to investigate the effect of cloud cover which further limits the burned area mapping capability of moderate resolution sensor for the majority of burned areas that can be detected reliably only for a relatively short time.

2.8. References

- Alonso-Canas, I., & Chuvieco, E. (2015). Global burned area mapping from ENVISAT-MERIS and MODIS active fire data. *Remote Sensing of Environment*, 163, 140-152
- Andela, N., Morton, D.C., Giglio, L., Chen, Y., van der Werf, G.R., Kasibhatla, P.S., DeFries, R.S., Collatz, G.J., Hantson, S., Kloster, S., Bachelet, D., Forrest, M., Lasslop, G., Li, F., Mangeon,

- S., Melton, J.R., Yue, C., & Randerson, J.T. (2017). A human-driven decline in global burned area. *Science*, *356*, 1356-1362
- Archibald, S., Lehmann, C.E.R., Gómez-Dans, J.L., & Bradstock, R.A. (2013). Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences*, *110*, 6442-6447
- Bastarrika, A., Chuvieco, E., & 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*, 1003-1012
- Bond, W., Woodward, F., & Midgley, G. (2005). The global distribution of ecosystems in a world without fire. *New Phytologist*, *165*, 525-538
- Bond, W.J., & Keeley, J.E. (2005). Fire as a global 'herbivore': the ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution*, *20*, 387-394
- Boschetti, L., Flasse, S.P., & Brivio, P.A. (2004). Analysis of the conflict between omission and commission in low spatial resolution dichotomic thematic products: The Pareto Boundary. *Remote Sensing of Environment*, *91*, 280-292
- Boschetti, L., & Roy, D.P. (2008). Defining a fire year for reporting and analysis of global interannual fire variability. *Journal of Geophysical Research: Biogeosciences*, *113*, G03020
- Boschetti, L., Roy, D.P., Justice, C.O., & Humber, M.L. (2015). MODIS–Landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, *161*, 27-42
- Boschetti, L., Stehman, S.V., & Roy, D.P. (2016). A stratified random sampling design in space and time for regional to global scale burned area product validation. *Remote Sensing of Environment*, *186*, 465-478
- Bowman, D.M., Balch, J.K., Artaxo, P., Bond, W.J., Carlson, J.M., Cochrane, M.A., D'Antonio, C.M., DeFries, R.S., Doyle, J.C., & Harrison, S.P. (2009). Fire in the Earth system. *Science*, *324*, 481-484
- Bowman, D.M.J.S., Zhang, Y., Walsh, A., & Williams, R.J. (2003). Experimental comparison of four remote sensing techniques to map tropical savanna fire-scars using Landsat-TM imagery. *International Journal of Wildland Fire*, *12*, 341-348
- Bucini, G., & Lambin, E.F. (2002). Fire impacts on vegetation in Central Africa: a remote-sensing-based statistical analysis. *Applied Geography*, *22*, 27-48
- Burrows, N.D., Burbidge, A.A., Fuller, P.J., & Behn, G. (2006). Evidence of altered fire regimes in the Western Desert region of Australia. *Conservation Science Western Australia*, *5*, 14
- Certini, G. (2005). Effects of fire on properties of forest soils: a review. *Oecologia*, *143*, 1-10

- Chen, D., Pereira, J.M.C., Masiero, A., & Pirotti, F. (2017). Mapping fire regimes in China using MODIS active fire and burned area data. *Applied Geography*, *85*, 14-26
- Chen, X., Vogelmann, J.E., Rollins, M., Ohlen, D., Key, C.H., Yang, L., Huang, C., & Shi, H. (2011). Detecting post-fire burn severity and vegetation recovery using multitemporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest. *International Journal of Remote Sensing*, *32*, 7905-7927
- Cheney, N., Gould, J., & Catchpole, W. (1993). The Influence of Fuel, Weather and Fire Shape Variables on Fire-Spread in Grasslands. *International Journal of Wildland Fire*, *3*, 31-44
- Chu, T., & Guo, X. (2014). Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. *Remote Sensing*, *6*, 470-520
- Chu, T., Guo, X., & Takeda, K. (2016). Remote sensing approach to detect post-fire vegetation regrowth in Siberian boreal larch forest. *Ecological Indicators*, *62*, 32-46
- Chuvieco, E., & Congalton, R.G. (1988). Mapping and inventory of forest fires from digital processing of TM data. *Geocarto International*, *3*, 41-53
- Chuvieco, E., Martin, M.P., & Palacios, A. (2002). Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, *23*, 5103-5110
- Clark, J.S., Iverson, L., Woodall, C.W., Allen, C.D., Bell, D.M., Bragg, D.C., D'Amato, A.W., Davis, F.W., Hersh, M.H., Ibanez, I., Jackson, S.T., Matthews, S., Pederson, N., Peters, M., Schwartz, M.W., Waring, K.M., & Zimmermann, N.E. (2016). The impacts of increasing drought on forest dynamics, structure, and biodiversity in the United States. *Global Change Biology*, *22*, 2329-2352
- Close, D.C., Davidson, N.J., Johnson, D.W., Abrams, M.D., Hart, S.C., Lunt, I.D., Archibald, R.D., Horton, B., & Adams, M.A. (2009). Premature Decline of Eucalyptus and Altered Ecosystem Processes in the Absence of Fire in Some Australian Forests. *The Botanical Review*, *75*, 191-202
- Cochrane, M.A. (2003). Fire science for rainforests. *Nature*, *421*, 913-919
- Cochrane, M.A., Alencar, A., Schulze, M.D., Souza, C.M., Nepstad, D.C., Lefebvre, P., & Davidson, E.A. (1999). Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests. *Science*, *284*, 1832-1835
- de Groot, W.J., Cantin, A.S., Flannigan, M.D., Soja, A.J., Gowman, L.M., & Newbery, A. (2013). A comparison of Canadian and Russian boreal forest fire regimes. *Forest Ecology and Management*, *294*, 23-34

- DeBano, L.F., Neary, D.G., & Ffolliott, P.F. (1998). *Fire effects on ecosystems*. John Wiley & Sons Incorporated
- DeFries, R.S., Rudel, T., Uriarte, M., & Hansen, M. (2010). Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geosci*, 3, 178-181
- Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexler, M.C., & Wisniewski, J. (1994). Carbon Pools and Flux of Global Forest Ecosystems. *Science*, 263, 185-190
- Dwyer, E., Pereira, J., Grégoire, J.M., & DaCamara, C.C. (2000). Characterization of the spatio-temporal patterns of global fire activity using satellite imagery for the period April 1992 to March 1993. *Journal of Biogeography*, 27, 57-69
- Engel, E.C., & Abella, S.R. (2011). Vegetation recovery in a desert landscape after wildfires: influences of community type, time since fire and contingency effects. *Journal of Applied Ecology*, 48, 1401-1410
- Escuin, S., Navarro, R., & Fernández, P. (2008). Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *International Journal of Remote Sensing*, 29, 1053-1073
- Eva, H., & Lambin, E. (1998). Burnt area mapping in Central Africa using ATSR data. *International Journal of Remote Sensing*, 19, 3473-3497
- Eva, H., & Lambin, E.F. (2000). Fires and land-cover change in the tropics: a remote sensing analysis at the landscape scale. *Journal of Biogeography*, 27, 765-776
- Fraser, R., Fernandes, R., & Latifovic, R. (2003). Multi-temporal Mapping of Burned Forest over Canada Using Satellite-based Change Metrics. *Geocarto International*, 18, 37-47
- Fraser, R.H., Li, Z., & Landry, R. (2000). SPOT VEGETATION for characterizing boreal forest fires. *International Journal of Remote Sensing*, 21, 3525-3532
- Frederiksen, P., Langaas, S., & Mbaye, M. (1990). NOAA-AVHRR and GIS-based monitoring of fire activity in Senegal—a provisional methodology and potential applications. *Fire in the Tropical Biota* (pp. 400-417): Springer
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182
- Fuller, S.P., & Rouse, W.R. (1979). Spectral reflectance changes accompanying a post-fire recovery sequence in a subarctic spruce lichen woodland. *Remote Sensing of Environment*, 8, 11-23
- GCOS (2011). Systematic observation requirements for satellite-based products for climate in support of the UNFCCC. In: WMO GCOS Tech. Doc. GCOS-138 WMO-TD 1523, 23 pp

- Giglio, L., Csiszar, I., & Justice, C.O. (2006). Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *Journal of Geophysical Research: Biogeosciences*, *111*, n/a-n/a
- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., & Justice, C.O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, *113*, 408-420
- Giglio, L., Randerson, J.T., & van der Werf, G.R. (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences*, *118*, 317-328
- Gill, A.M., Moore, P.H.R., & Williams, R.J. (1996). Fire weather in the wet-dry tropics of the World Heritage Kakadu National Park, Australia. *Australian Journal of Ecology*, *21*, 302-308
- Glitzenstein, J.S., Platt, W.J., & Streng, D.R. (1995). Effects of Fire Regime and Habitat on Tree Dynamics in North Florida Longleaf Pine Savannas. *Ecological monographs*, *65*, 441-476
- Goetz, S.J., Bond-Lamberty, B., Law, B.E., Hicke, J.A., Huang, C., Houghton, R.A., McNulty, S., O'Halloran, T., Harmon, M., Meddens, A.J.H., Pfeifer, E.M., Mildrexler, D., & Kasischke, E.S. (2012). Observations and assessment of forest carbon dynamics following disturbance in North America. *Journal of Geophysical Research: Biogeosciences*, *117*, n/a-n/a
- Goodwin, N.R., Coops, N.C., Wulder, M.A., Gillanders, S., Schroeder, T.A., & Nelson, T. (2008). Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sensing of Environment*, *112*, 3680-3689
- Grégoire, J.-M., & Simonetti, D. (2010). Interannual changes of fire activity in the protected areas of the SUN network and other parks and reserves of the West and Central Africa region derived from MODIS observations. *Remote Sensing*, *2*, 446-463
- Guenther, C.C. (2006). Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature). In
- Hall, J.V., Loboda, T.V., Giglio, L., & McCarty, G.W. (2016). A MODIS-based burned area assessment for Russian croplands: Mapping requirements and challenges. *Remote Sensing of Environment*, *184*, 506-521
- Hawbaker, T.J., Vanderhoof, M.K., Beal, Y.-J., Takacs, J.D., Schmidt, G.L., Falgout, J.T., Williams, B., Fairaux, N.M., Caldwell, M.K., Picotte, J.J., Howard, S.M., Stitt, S., & Dwyer, J.L. (2017). Mapping burned areas using dense time-series of Landsat data. *Remote Sensing of Environment*, *198*, 504-522
- Henry, M.C. (2008). Comparison of single-and multi-date Landsat data for mapping wildfire scars in Ocala National Forest, Florida. *Photogrammetric engineering and remote sensing*, *74*, 881-891

- Hope, G., Kershaw, A.P., Kaars, S.v.d., Xiangjun, S., Liew, P.-M., Heusser, L.E., Takahara, H., McGlone, M., Miyoshi, N., & Moss, P.T. (2004). History of vegetation and habitat change in the Austral-Asian region. *Quaternary International*, 118–119, 103-126
- Houghton, R.A., & Hackler, J.L. (2006). Emissions of carbon from land use change in sub-Saharan Africa. *Journal of Geophysical Research: Biogeosciences*, 111, n/a-n/a
- Huang, H., Roy, D., Boschetti, L., Zhang, H., Yan, L., Kumar, S., Gomez-Dans, J., & Li, J. (2016). Separability Analysis of Sentinel-2A Multi-Spectral Instrument (MSI) Data for Burned Area Discrimination. *Remote Sensing*, 8, 873
- Hudak, A., & Brockett, B. (2004). Mapping fire scars in a southern African savannah using Landsat imagery. *International Journal of Remote Sensing*, 25, 3231-3243
- Hyer, E.J., & Reid, J.S. (2009). Baseline uncertainties in biomass burning emission models resulting from spatial error in satellite active fire location data. *Geophysical Research Letters*, 36
- Jin, Y., Randerson, J.T., Goulden, M.L., & Goetz, S.J. (2012). Post-fire changes in net shortwave radiation along a latitudinal gradient in boreal North America. *Geophysical Research Letters*, 39, n/a-n/a
- Jin, Y., & Roy, D. (2005). Fire-induced albedo change and its radiative forcing at the surface in northern Australia. *Geophysical Research Letters*, 32
- Kaiser, J., Heil, A., Andreae, M., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M., & Suttie, M. (2012). Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9, 527-554
- Key, C., & Benson, N. (2002). Measuring and remote sensing of burn severity. In, *US Geological Survey wildland fire workshop* (pp. 02-11): US Geological Survey Washington, DC, USA
- Key, C.H., & Benson, N.C. (2006). Landscape assessment. *FIREMON: Fire effects monitoring and inventory system*, 1-55
- Koutsias, N., & Karteris, M. (1998). Logistic regression modelling of multitemporal Thematic Mapper data for burned area mapping. *International Journal of Remote Sensing*, 19, 3499-3514
- Koutsias, N., & Karteris, M. (2000). Burned area mapping using logistic regression modeling of a single post-fire Landsat-5 Thematic Mapper image. *International Journal of Remote Sensing*, 21, 673-687
- Krawchuk, M.A., & Moritz, M.A. (2011). Constraints on global fire activity vary across a resource gradient. *Ecology*, 92, 121-132
- Kumar, S.S., Roy, D.P., Cochrane, M.A., Souza, C.M., Barber, C.P., & Boschetti, L. (2014). A quantitative study of the proximity of satellite detected active fires to roads and rivers in the Brazilian tropical moist forest biome. *International Journal of Wildland Fire*, 23, 532-543

- Lamarque, J.F., Bond, T.C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M.G., Shindell, D., Smith, S.J., Stehfest, E., Van Aardenne, J., Cooper, O.R., Kainuma, M., Mahowald, N., McConnell, J.R., Naik, V., Riahi, K., & van Vuuren, D.P. (2010). Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application. *Atmos. Chem. Phys.*, *10*, 7017-7039
- Lambin, E.F., Goyvaerts, K., & Petit, C. (2003). Remotely-sensed indicators of burning efficiency of savannah and forest fires. *International Journal of Remote Sensing*, *24*, 3105-3118
- Langaas, S., & Kane, R. (1991). Temporal spectral signatures of fire scars in Savanna Woodland. In, *Geoscience and Remote Sensing Symposium, 1991. IGARSS '91. Remote Sensing: Global Monitoring for Earth Management.*, International (pp. 1157-1160)
- Langmann, B., Duncan, B., Textor, C., Trentmann, J., & van der Werf, G.R. (2009). Vegetation fire emissions and their impact on air pollution and climate. *Atmospheric Environment*, *43*, 107-116
- Lasko, K., Vadrevu, K.P., Tran, V.T., Ellicott, E., Nguyen, T.T., Bui, H.Q., & Justice, C. (2017). Satellites may underestimate rice residue and associated burning emissions in Vietnam. *Environmental Research Letters*, *12*, 085006
- Lefsky, M.A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*, n/a-n/a
- Li, M.S., Mao, L.J., Lu, C., & Wang, Y. (2012). Temporal Change in Fragmentation in China's Primary Forest Ecoregions. In, *Advanced Materials Research* (pp. 1286-1296): Trans Tech Publ
- Littell, J.S., McKenzie, D., Peterson, D.L., & Westerling, A.L. (2009). Climate and wildfire area burned in western U.S. ecoregions, 1916–2003. *Ecological applications*, *19*, 1003-1021
- Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B., & Ackerly, D.D. (2009). The velocity of climate change. *Nature*, *462*, 1052-1055
- Loboda, T.V., French, N.H.F., Hight-Harf, C., Jenkins, L., & Miller, M.E. (2013). Mapping fire extent and burn severity in Alaskan tussock tundra: An analysis of the spectral response of tundra vegetation to wildland fire. *Remote Sensing of Environment*, *134*, 194-209
- Lu, B., & He, Y. (2014). Analyzing a North American prairie wildfire using remote sensing imagery. In, *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 832-835)
- Lyons, E.A., Jin, Y., & Randerson, J.T. (2008). Changes in surface albedo after fire in boreal forest ecosystems of interior Alaska assessed using MODIS satellite observations. *Journal of Geophysical Research: Biogeosciences*, *113*, n/a-n/a

- McCarty, J.L., Ellicott, E.A., Romanenkov, V., Rukhovitch, D., & Koroleva, P. (2012). Multi-year black carbon emissions from cropland burning in the Russian Federation. *Atmospheric Environment*, *63*, 223-238
- Meney, K.A., Nielsens, G.M., & Dixon, K.W. (1994). Seed bank patterns in Restionaceae and Epacridaceae after wildfire in kwongan in southwestern Australia. *Journal of vegetation science*, *5*, 5-12
- Miller, J.D., & Yool, S.R. (2002). Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. *Remote Sensing of Environment*, *82*, 481-496
- Mishra, N.B., Mainali, K.P., & Crews, K.A. (2016). Modelling spatiotemporal variability in fires in semiarid savannas: a satellite-based assessment around Africa's largest protected area. *International Journal of Wildland Fire*, *25*, 730-741
- Mitri, G., & Gitas, I. (2004). A semi-automated object-oriented model for burned area mapping in the Mediterranean region using Landsat-TM imagery. *International Journal of Wildland Fire*, *13*, 367-376
- Mohler, R.L., & Goodin, D.G. (2012). Identifying a suitable combination of classification technique and bandwidth(s) for burned area mapping in tallgrass prairie with MODIS imagery. *International Journal of Applied Earth Observation and Geoinformation*, *14*, 103-111
- Moritz, M.A., Parisien, M.-A., Batllori, E., Krawchuk, M.A., Van Dorn, J., Ganz, D.J., & Hayhoe, K. (2012). Climate change and disruptions to global fire activity. *Ecosphere*, *3*, art49
- Morton, D.C., DeFries, R.S., Nagol, J., Souza Jr, C.M., Kasischke, E.S., Hurtt, G.C., & Dubayah, R. (2011). Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data. *Remote Sensing of Environment*, *115*, 1706-1720
- Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., & Chuvieco, E. (2014). Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 64-79
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., & Morrison, J.C. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*, 933-938
- Padilla, M., Stehman, S.V., & Chuvieco, E. (2014). Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling. *Remote Sensing of Environment*, *144*, 187-196

- Pereira, J., Chuvieco, E., Beaudoin, A., & Desbois, N. (1997). Remote sensing of burned areas: a review. *A review of remote sensing methods for the study of large wildland fires*, 127-184
- Pereira, J.C., Sá, A.L., Sousa, A.O., Silva, J.N., Santos, T., & Carreiras, J.B. (1999). Spectral characterisation and discrimination of burnt areas. In E. Chuvieco (Ed.), *Remote Sensing of Large Wildfires* (pp. 123-138): Springer Berlin Heidelberg
- Pereira, J.M. (2003). Remote sensing of burned areas in tropical savannas. *International Journal of Wildland Fire*, 12, 259-270
- Pereira, M., & Setzer, A. (1993). Spectral characteristics of fire scars in Landsat-5 TM images of Amazonia. *Remote Sensing*, 14, 2061-2078
- Pfeifer, M., Disney, M., Quaife, T., & Marchant, R. (2012). Terrestrial ecosystems from space: a review of earth observation products for macroecology applications. *Global Ecology and Biogeography*, 21, 603-624
- Randerson, J.T., Chen, Y., van der Werf, G.R., Rogers, B.M., & Morton, D.C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, 117, G04012
- Randerson, J.T., Liu, H., Flanner, M.G., Chambers, S.D., Jin, Y., Hess, P.G., Pfister, G., Mack, M.C., Treseder, K.K., Welp, L.R., Chapin, F.S., Harden, J.W., Goulden, M.L., Lyons, E., Neff, J.C., Schuur, E.A.G., & Zender, C.S. (2006). The Impact of Boreal Forest Fire on Climate Warming. *Science*, 314, 1130-1132
- Röder, A., Hill, J., Duguay, B., Alloza, J.A., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment*, 112, 259-273
- Rossi, S., Tubiello, F.N., Prosperi, P., Salvatore, M., Jacobs, H., Biancalani, R., House, J.I., & Boschetti, L. (2016). FAOSTAT estimates of greenhouse gas emissions from biomass and peat fires. *Climatic Change*, 135, 699-711
- Roy, D., Huang, H., Kumar, S., Zhang, H., Li, J., Gomez-Dans, J., Lewis, P., Boschetti, L. (2015). Towards prototyping a global Landsat-8 Sentinel-2 Burned Area Product. *EARSEL SIG Fire 2015, Cyprus, 2-5 November 2015*
- Roy, D., Jin, Y., Lewis, P., & Justice, C. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, 97, 137-162
- Roy, D., & Landmann, T. (2005). Characterizing the surface heterogeneity of fire effects using multi-temporal reflective wavelength data. *International Journal of Remote Sensing*, 26, 4197-4218

- Roy, D., Lewis, P., & Justice, C. (2002). Burned area mapping using multi-temporal moderate spatial resolution data—A bi-directional reflectance model-based expectation approach. *Remote Sensing of Environment*, 83, 263-286
- Roy, D.P., Boschetti, L., & Giglio, L. (2010). Remote Sensing of Global Savanna Fire Occurrence, Extent, and Properties. *Ecosystem Function in Savannas: Measurement and Modeling at Landscape to Global Scales*, 239
- Roy, D.P., Boschetti, L., Justice, C.O., & Ju, J. (2008). The collection 5 MODIS burned area product—Global evaluation by comparison with the MODIS active fire product. *Remote Sensing of Environment*, 112, 3690-3707
- Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., & Muller, J.-P. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, 83, 135-148
- Schipper, J., Chanson, J.S., Chiozza, F., Cox, N.A., Hoffmann, M., Katariya, V., Lamoreux, J., Rodrigues, A.S.L., Stuart, S.N., Temple, H.J., Baillie, J., Boitani, L., Lacher, T.E., Mittermeier, R.A., Smith, A.T., Absolon, D., Aguiar, J.M., Amori, G., Bakkour, N., Baldi, R., Berridge, R.J., Bielby, J., Black, P.A., Blanc, J.J., Brooks, T.M., Burton, J.A., Butynski, T.M., Catullo, G., Chapman, R., Cokeliss, Z., Collen, B., Conroy, J., Cooke, J.G., da Fonseca, G.A.B., Derocher, A.E., Dublin, H.T., Duckworth, J.W., Emmons, L., Emslie, R.H., Festa-Bianchet, M., Foster, M., Foster, S., Garshelis, D.L., Gates, C., Gimenez-Dixon, M., Gonzalez, S., Gonzalez-Maya, J.F., Good, T.C., Hammerson, G., Hammond, P.S., Happold, D., Happold, M., Hare, J., Harris, R.B., Hawkins, C.E., Haywood, M., Heaney, L.R., Hedges, S., Helgen, K.M., Hilton-Taylor, C., Hussain, S.A., Ishii, N., Jefferson, T.A., Jenkins, R.K.B., Johnston, C.H., Keith, M., Kingdon, J., Knox, D.H., Kovacs, K.M., Langhammer, P., Leus, K., Lewison, R., Lichtenstein, G., Lowry, L.F., Macavoy, Z., Mace, G.M., Mallon, D.P., Masi, M., McKnight, M.W., Medellín, R.A., Medici, P., Mills, G., Moehlman, P.D., Molur, S., Mora, A., Nowell, K., Oates, J.F., Olech, W., Oliver, W.R.L., Oprea, M., Patterson, B.D., Perrin, W.F., Polidoro, B.A., Pollock, C., Powel, A., Protas, Y., Racey, P., Ragle, J., Ramani, P., Rathbun, G., Reeves, R.R., Reilly, S.B., Reynolds, J.E., Rondinini, C., Rosell-Ambal, R.G., Rulli, M., Rylands, A.B., Savini, S., Schank, C.J., Sechrest, W., Self-Sullivan, C., Shoemaker, A., Sillero-Zubiri, C., De Silva, N., Smith, D.E., Srinivasulu, C., Stephenson, P.J., van Strien, N., Talukdar, B.K., Taylor, B.L., Timmins, R., Tirira, D.G., Tognelli, M.F., Tsytsulina, K., Veiga, L.M., Vié, J.-C., Williamson, E.A., Wyatt, S.A., Xie, Y., & Young, B.E. (2008). The Status of the World's Land and Marine Mammals: Diversity, Threat, and Knowledge. *Science*, 322, 225-230

- Schneider, A., Friedl, M.A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters*, 4, 044003
- Scholes, R.J., & Walker, B.H. (2004). *An African savanna: synthesis of the Nylsvley study*. Cambridge University Press
- Schwartz, M.D., & Reed, B.C. (1999). Surface phenology and satellite sensor-derived onset of greenness: an initial comparison. *International Journal of Remote Sensing*, 20, 3451-3457
- Silva, J.M., Sá, A.C., & Pereira, J.M. (2005). Comparison of burned area estimates derived from SPOT-VEGETATION and Landsat ETM+ data in Africa: Influence of spatial pattern and vegetation type. *Remote Sensing of Environment*, 96, 188-201
- Simon, M., Plummer, S., Fierens, F., Hoelzemann, J., & Arino, O. (2004). Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 109
- Slocum, M.G., Platt, W.J., & Cooley, H.C. (2003). Effects of Differences in Prescribed Fire Regimes on Patchiness and Intensity of Fires in Subtropical Savannas of Everglades National Park, Florida. *Restoration Ecology*, 11, 91-102
- Smith, A.M.S., Drake, N.A., Wooster, M.J., Hudak, A.T., Holden, Z.A., & Gibbons, C.J. (2007). Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: comparison of methods and application to MODIS. *International Journal of Remote Sensing*, 28, 2753-2775
- Solans Vila, J.P., & Barbosa, P. (2010). Post-fire vegetation regrowth detection in the Deiva Marina region (Liguria-Italy) using Landsat TM and ETM+ data. *Ecological Modelling*, 221, 75-84
- Staver, A.C., Archibald, S., & Levin, S.A. (2011). The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. *Science*, 334, 230-232
- Stroppiana, D., Boschetti, M., Zaffaroni, P., & Brivio, P.A. (2009). Analysis and Interpretation of Spectral Indices for Soft Multicriteria Burned-Area Mapping in Mediterranean Regions. *Geoscience and Remote Sensing Letters, IEEE*, 6, 499-503
- Stroppiana, D., Pinnock, S., Pereira, J.M.C., & Grégoire, J.-M. (2002). Radiometric analysis of SPOT-VEGETATION images for burnt area detection in Northern Australia. *Remote Sensing of Environment*, 82, 21-37
- Tansey, K., Grégoire, J.-M., Stroppiana, D., Sousa, A., Silva, J., Pereira, J.M.C., Boschetti, L., Maggi, M., Brivio, P.A., Fraser, R., Flasse, S., Ershov, D., Binaghi, E., Graetz, D., & Peduzzi, P. (2004). Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data. *Journal of Geophysical Research: Atmospheres*, 109, D14S03

- Tansey, K., Gregoire, J.M., Defourny, P., Leigh, R., Pekel, J.F.O., van Bogaert, E., & Bartholome, E. (2008). A new, global, multi-annual (2000-2007) burnt area product at 1 km resolution. *Geophysical Research Letters*, *35*, 6
- Thonicke, K., Venevsky, S., Sitch, S., & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. *Global Ecology and Biogeography*, *10*, 661-677
- Tovar, C., Breman, E., Brncic, T., Harris, D.J., Bailey, R., & Willis, K.J. (2014). Influence of 1100 years of burning on the central African rainforest. *Ecography*, *37*, 1139-1148
- Trigg, S., & Flasse, S. (2000). Characterizing the spectral-temporal response of burned savannah using in situ spectroradiometry and infrared thermometry. *International Journal of Remote Sensing*, *21*, 3161-3168
- Tubiello, F.N., Salvatore, M., Rossi, S., Ferrara, A., Fitton, N., & Smith, P. (2013). The FAOSTAT database of greenhouse gas emissions from agriculture. *Environmental Research Letters*, *8*, 015009
- van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G., Mu, M., Kasibhatla, P.S., Morton, D.C., DeFries, R., Jin, Y.v., & van Leeuwen, T.T. (2010). Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric Chemistry and Physics*, *10*, 11707-11735
- van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Kasibhatla, P.S., & Arellano Jr, A.F. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmospheric Chemistry and Physics*, *6*, 3423-3441
- van der Werf, G.R., Randerson, J.T., Giglio, L., Gobron, N., & Dolman, A.J. (2008). Climate controls on the variability of fires in the tropics and subtropics. *Global Biogeochemical Cycles*, *22*, n/a-n/a
- van der Werf, G.R., Randerson, J.T., Giglio, L., van Leeuwen, T.T., Chen, Y., Rogers, B.M., Mu, M., van Marle, M.J., Morton, D.C., & Collatz, G.J. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, *9*, 697
- Vanderhoof, M., Brunner, N., Beal, Y.-J., & Hawbaker, T. (2017). Evaluation of the U.S. Geological Survey Landsat Burned Area Essential Climate Variable across the Conterminous U.S. Using Commercial High-Resolution Imagery. *Remote Sensing*, *9*, 743
- Verhegghen, A., Eva, H., Ceccherini, G., Achard, F., Gond, V., Gourlet-Fleury, S., & Cerutti, P. (2016). The Potential of Sentinel Satellites for Burnt Area Mapping and Monitoring in the Congo Basin Forests. *Remote Sensing*, *8*, 986

- Wang, Z., Schaaf, C.B., Chopping, M.J., Strahler, A.H., Wang, J., Román, M.O., Rocha, A.V., Woodcock, C.E., & Shuai, Y. (2012). Evaluation of Moderate-resolution Imaging Spectroradiometer (MODIS) snow albedo product (MCD43A) over tundra. *Remote Sensing of Environment*, *117*, 264-280
- Wang, Z., Schaaf, C.B., Sun, Q., Shuai, Y., & Román, M.O. (2018). Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. *Remote Sensing of Environment*, *207*, 50-64
- Wessman, C.A., Bateson, C.A., & Benning, T.L. (1997). Detecting fire and grazing patterns in tallgrass prairie using spectral mixture analysis. *Ecological applications*, *7*, 493-511
- West, N.E., & Hassan, M.A. (1985). Recovery of Sagebrush-Grass Vegetation following Wildfire. *Journal of Range Management*, *38*, 131-134
- Westerling, A.L., Hidalgo, H.G., Cayan, D.R., & Swetnam, T.W. (2006). Warming and earlier spring increase western US forest wildfire activity. *Science*, *313*, 940-943
- White, M.A., Thornton, P.E., & Running, S.W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, *11*, 217-234
- Wierzchowski, J., Heathcott, M., & Flannigan, M.D. (2002). Lightning and lightning fire, central cordillera, Canada. *International Journal of Wildland Fire*, *11*, 41-51
- Williams, R.J., Cook, G.D., Gill, A.M., & Moore, P.H.R. (1999). Fire regime, fire intensity and tree survival in a tropical savanna in northern Australia. *Australian Journal of Ecology*, *24*, 50-59
- Wolfe, R.E., Roy, D.P., & Vermote, E. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *Geoscience and Remote Sensing, IEEE Transactions on*, *36*, 1324-1338

Chapter 3: Global Evaluation of the Suitability of MODIS-Terra Detected Cloud Cover as a Proxy for Landsat 7 Cloud Conditions

Published in *Remote Sensing* as:

Melchiorre, A., Boschetti, L. & Roy, D.P. (2020). Global evaluation of the suitability of MODIS-Terra detected cloud cover as a proxy for Landsat 7 cloud conditions. *Remote Sensing*, 12(2), 202

3.1. Abstract

Clouds limit the quality and availability of optical wavelength surface observations from Earth Observation (EO) satellites. This limitation is particularly relevant for the generation of systematic thematic products from EO medium spatial resolution polar orbiting sensors, such as Landsat, which have reduced temporal resolution compared to coarser resolution polar orbiting sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS). MODIS on the Terra satellite is in the same orbit as Landsat 7 with an approximately 30 minute overpass difference. In this study, one year of global Landsat 7 Enhanced Thematic Mapper Plus (ETM+) image cloud fractions over land are compared with collocated MODIS cloud fractions, generated by combining the MODIS-Terra global daily cloud mask product (MOD35) with the Landsat 7 ETM+ image footprints and acquisition calendar. The results show high correlation between the MODIS and Landsat 7 ETM+ cloud fractions ($R^2 = 0.83$), negligible bias (median difference: < 0.01) and low dispersion around the median (interquartile range: $[-0.02, 0.06]$). These results indicate that, globally, the cloud cover detected by MODIS Terra data can be used as a proxy for Landsat 7 ETM+ cloud cover.

3.2. Introduction

Earth Observation (EO) data sensed in the optical portion of the electromagnetic spectrum are widely used for studying the Earth's biosphere, its dynamics and disturbances from regional to global scale. Cloud obscuration affects the quality and availability of EO optical data over land which presents a challenge for global land monitoring. Thanks to the systematic acquisition strategy and free data access policy of currently available moderate resolution optical data such as Landsat (Roy et al. 2014) and Sentinel-2 (Drusch et al. 2012), global systematic land monitoring products at moderate resolution are being developed (Wulder et al. 2019). Landsat-8, Sentinel-2A, and Sentinel-2B provide 16-day, 10-

day and 10-day global median average revisit intervals, respectively, and 2.9 days when combined together (Li and Roy 2017). These revisit cycles are not particularly high, particularly for individual sensors, and consequently cloud obscuration may remain an issue for monitoring applications using these data, for example, for monitoring features that evolve with a fast temporal dynamics, such as certain burned areas (Boschetti et al. 2015; Melchiorre and Boschetti 2018; Roy et al. 2019), ephemeral water bodies (Pekel et al. 2016), or agricultural crops (Roy and Yan 2018).

The near daily global coverage 1 km cloud product derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Ackerman et al. 2010) has been used in feasibility studies to assess the potential impact of cloud obscuration on other land monitoring satellites, using the MODIS time series to generate a spatially explicit climatology of cloud cover probabilities (Whitcraft et al. 2015a). Moderate resolution data such as Landsat cannot be directly used for generating such a climatology because of the limited number of available acquisitions, due to the low revisit frequency and to the fragmentary state of the archive outside the United States (Wulder et al. 2016). For example, researchers have used the MODIS global cloud mask product as a proxy to estimate the effect of cloud cover on the HypsIRI mission design (Mercury et al. 2012), and to define the observation requirements for operational cropland monitoring from medium resolution systems (Whitcraft et al. 2015b).

However, the use of MODIS cloud data as a proxy assumes that the MODIS 1 km cloud mask is representative of cloud conditions experienced by other satellites. A recent regional study on the probability of Landsat cloud-free over the Eastern United States found a strong correlation between the cloud cover observed by Landsat 7 ETM+ and MODIS, but the intercomparison was limited to three Landsat path/row locations (Goward et al. 2019).

Several factors can contribute to differences among satellite cloud products. Wind can displace clouds and other atmospheric constituents in the time between different satellite overpasses (Feidas and Cartalis 2005; Li et al. 2019). Clouds evolve over time and in many parts of the world cloud cover increases during the day over land (King et al. 2013). The MODIS Terra and Aqua satellites have 10.30 am and 1.30 pm equatorial overpass times respectively and although globally they observe similar cloud amounts, in many regions they reveal different cloud amounts at the time of overpass (King et al. 2013; Roy et al. 2006). Observations from the afternoon-train (A-train) polar orbiting CloudSat and CALIPSO sensors, that are designed for cloud monitoring, show good agreement of total horizontal cloud fraction retrievals at 2° resolution with the A-train MODIS-Aqua 1 km cloud mask product (Hagihara et al. 2010). This is likely because of the small 3 minute (CloudSat) and 1 minute (CALIPSO) overpass time differences relative to MODIS-Aqua. Other differences are due to the size and spatial distribution of clouds relative to the sensor spatial resolutions. For example, the 1 km MODIS cloud

product may not detect sub-pixel or spatially fragmented clouds that are detectable at 30 m Landsat resolution. Differences in sensor spectral resolution may also introduce cloud detection differences. For example Zhao and Di Girolamo (2006) compared 15 m cloud detections from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data with the MODIS cloud product and found that MODIS systematically overestimated the cloud fraction relative to ASTER, with a mean difference of 0.18 and large variations across scenes associated with the spatial cloud patterns. Because ASTER is onboard the same Terra platform as MODIS, these discrepancies were attributed to the differences in spatial and spectral resolution of the two instruments. Nominally, clouds have high reflectance at visible and SWIR wavelengths, and generally are cold, and so sensors with visible to SWIR bands and thermal bands provide more reliable cloud detection than sensors that have only a subset of these bands (Ackerman et al. 1998). Finally, cloud detection algorithm differences may result in different cloud masks at coarse (Holz et al. 2008) and at medium resolution (Foga et al. 2017).

In this study, we evaluate the assumption that the MODIS-Terra 1 km cloud product can be used globally to predict Landsat 7 ETM+ image cloud fractions over land. This is a reasonable expectation as both sensors are in the same orbit and have an overpass time difference of approximately 30 minutes (Chander et al. 2010). The fraction of cloud cover in every global Landsat image acquired in 2002 is compared to cloud cover fractions in the contemporaneous MODIS Terra cloud cover product (MOD35). The relationship between the MODIS and Landsat 7 ETM+ cloud fractions is characterized using both linear and logistic models, and the corresponding regression coefficients and residual errors are examined.

3.3. Data

3.3.1. Landsat cloud data

The Landsat 7 is in a polar sun-synchronous orbit with approximately a 705 km altitude, 98.2° inclination, and 10:00 a.m. ± 15 min descending equatorial overpassing local time (Goward et al. 2001). The Landsat 7 carries the Enhanced Thematic Mapper Plus (ETM+) instrument which senses multispectral data over a 15° field of view. The orbit altitude and sensor field of view result in an approximately 185 km swath and 16-day nadiral overpass revisit frequency. Landsat data are distributed as $185 \text{ km} \times 170 \text{ km}$ images defined in the Worldwide Reference System (WRS) that divides the globe into 233 paths (orbital ground-tracks) and 248 rows (latitude parallels) (Arvidson et al. 2001).

The most recent Landsat Collection 1 cloud mask information was used; it is generated using the Fmask algorithm (Zhu et al. 2015; Zhu and Woodcock 2012) that was found to have better performance than other algorithms (Foga et al. 2017). Fmask is a two-step object-based algorithm. First,

a potential cloud layer is generated by combining spectral tests based on cloud physical properties, and a cloud probability mask based on temperature, brightness and spectral variability. The final cloud mask is subsequently generated through segmentation of the potential cloud layer. The percentage of clouds over only the land pixels in each image defined by the CLOUD_COVER_LAND metadata was used and scaled in the range 0 to 1. This is henceforth referred to as the “Landsat 7 cloud fraction”.

3.3.2. *MODIS cloud mask and geolocation product*

The MODIS Terra satellite is in the same descending orbit as Landsat 7 with a nominal period of 99 minutes, and a 16-days nadiral overpass revisit frequency, and trails Landsat 7 by approximately 30 minutes (Chander et al. 2010). MODIS has a 110° field of view and senses a 2330 km swath with overlapping swaths at latitudes $>30^\circ$ (Wolfe et al. 1998). The Collection 6 MODIS-Terra 1 km cloud mask product (MOD35) (Ackerman et al. 2010) was used. The MOD35 cloud detection algorithm involves five steps: (1) a series of spectral and spatial variability tests to detect the presence or absence of clouds are applied independently; (2) confidence scores are computed for each test applied; (3) the confidence scores of each test are combined into a preliminary overall confidence of the presence or absence of clouds for each pixel; (4) additional spatial and spectral tests (‘restoral tests’) are applied to reduce commission errors due to particular scene conditions (e.g. sun-glint, presence of bright surfaces); (5) a final cloud detection confidence level is generated, combining the results of the all the cloud detection tests.

The MOD35 product classifies each 1 km MODIS observation into four possible states of decreasing confidence of cloud detection: “Cloudy” (confidence clear $\leq 66\%$), “Probably Cloudy” ($66\% < \text{confidence clear} \leq 95\%$), “Probably Clear” ($95\% < \text{confidence clear} \leq 99\%$), and “Clear” (confidence clear $> 99\%$). In addition, ancillary data layers that describe the viewing and illumination geometry, and the land/water status of each 1 km observation are included in the product (Ackerman et al. 2010). The MOD35 product is a Level 2 product and so is defined in the MODIS orbit swath geometry. The MODIS Geolocation product (MOD03) that defines the geographic location of the MODIS Level 2 swath products (Masuoka et al. 1998; Wolfe et al. 2002) was also used in this study.

3.3.3. *Spatial and temporal extent of the analysis*

The analysis was performed using all available Landsat and MODIS data acquired from 1 January to 31 December 2002. This study period was selected because year 2002 was the last complete year of Landsat 7 ETM+ acquisitions before the Scan Line Corrector (SLC) failure that resulted in a systematic failure to sense 22% of each image (Markham et al. 2004) and reduced the reliability of SLC-off Landsat image cloud metadata information (Kovalskyy and Roy 2013). In 2002 the MODIS-Terra data were globally available as MODIS-Terra was launched in 1999. The greater majority of the

global Landsat image archive is acquired over land masses between 70° S and 70° N (Wulder et al. 2016) and so the study was restricted to this latitudinal range.

3.4. Methods

3.4.1. Computation of MODIS cloud fractions for each Landsat image

The MOD35 1 km cloud and land/water masks for each orbit swath for each global day of 2002 were reprojected into the 1 km MODIS sinusoidal equal area projection (Wolfe et al. 1998) using the corresponding MOD03 swath geolocation files. The data were reprojected by nearest neighbor resampling to preserve the input 1 km data values. The Landsat 7 ETM+ ground swaths coincide with the central nadir-looking portion of the MODIS-Terra swath acquired on the same day from the same orbit. The footprint of each Landsat image was reprojected into the gridded 1 km MODIS sinusoidal equal area projection. The 1 km land pixels falling under each Landsat image for each orbit were derived from the reprojected 1 km MOD35 MODIS land/water mask, and the 1 km MODIS cloud mask values were summed to derive the MODIS land cloud fraction for each Landsat image.

The MOD35 land/water mask has four possible values, namely “Land”, “Desert”, “Coastal area”, and “Water”. To ensure consistency with the land pixel definition used to compute the Landsat 7 cloud fractions, only the MOD35 1 km pixels labelled as “Land”, “Desert” or “Coastal area” in the MOD35 land/water mask were used to derive the MODIS cloud fractions.

The MODIS cloud fractions were defined for each Landsat image in three ways using different combinations of the MODIS cloud mask cloud detection confidence states (Table 3-1). For each definition the MODIS cloud fraction was derived as the number of 1 km pixels labeled as cloud, divided by the total number of 1 km land pixels (labelled as “Land”, “Desert” or “Coastal area” in the MOD35 land/water mask) encompassing the Landsat image. The three definitions include a conservative cloud definition (#1), an intermediated definition (#2) and a much less conservative definition (#3).

Table 3-1: The three MODIS cloud fraction definitions derived by different combinations of MODIS cloud mask (MOD35) detection confidence states

Definition	MOD35 confidence state combination
#1	Cloudy
#2	Cloudy, Probably Cloudy
#3	Cloudy, Probably Cloudy, Probably Clear

3.4.2. MODIS-Landsat cloud fraction comparison methodology

First, histograms of the Landsat cloud fractions and the MODIS cloud fractions were computed to verify that the whole range [0,1] of cloud fractions was represented in the dataset. Subsequently, the

cloud fraction difference, defined as the Landsat cloud fraction minus the MODIS cloud fraction, was computed for all the global 2002 study data. The distribution of the cloud differences was summarized through non-parametric summary statistics (median, interquartile range, and 5th-95th quantile range).

The relationship between the Landsat and MODIS cloud fractions was investigated using linear and logistic models, where the Landsat cloud fraction is the response variable and the MODIS cloud fraction is the predictor variable. A logistic regression was used because, in general, coarse resolution detection products, compared to higher resolution products, are affected by saturation issues. Thus a logistic regression would be appropriate if the cloud fraction is over-estimated at high cloud proportions and under-estimated at low cloud proportions. The regression residuals were analyzed to investigate the presence of heteroscedasticity in the relationship between Landsat and MODIS cloud fractions.

Standard metrics and coefficients were reported. For the linear models, these include the linear regression β_0 (intercept) and β_1 (slope) coefficients, the standard error of the regression coefficients, the coefficient of determination of the linear regression (R^2), and the significance of the regression (p-value). In addition, the standard error of the residuals, i.e., the observed minus the predicted Landsat 7 cloud fraction, were derived. For the logistic models, these include the logistic regression coefficients β_0 (intercept term of the log-odds function) and β_1 (slope term of the log-odds function), the standard error of the regression coefficients, and the significance of the regression (p-value). Since there is no direct analog of R^2 in a logistic regression, the residual deviance was reported as a goodness-of-fit statistic for the model. In addition, the null deviance was reported, i.e., the deviance of a model when only a single coefficient (the intercept term of the log-odds function) is estimated. The difference between the null deviance and the residual deviance provides an indication of the overall performance of the model, with large differences indicating good performance and small differences indicating poor performance.

The presence of spatial patterns in the residuals were investigated by analyzing the distribution of the residuals at each Landsat WRS path/row location. Summary parameters of the distribution (median, interquartile range and 5th - 95th quantile range) were calculated as a function of the land fraction, defined as the number of 1 km land pixels over the total footprint of the image, to verify whether the behavior of linear and logistic models was significantly different in coastal and inland areas. Finally, the distribution of the residuals was calculated as a function of the time of acquisition, using three-month seasonal intervals.

3.5. Results

3.5.1. Landsat 7 ETM+ image availability and global cloud histograms

A total of 109,814 Landsat 7 ETM+ images acquired in 2002 between 70° S and 70° N were considered in this study (Figure 3-1). Nominally, the Landsat 7 ETM+ overpasses each WRS path/row 22 or 23 times per year but in many regions fewer images are available in the Landsat archive for a variety of reasons (Ju and Roy 2008; Wulder et al. 2016). In 2002, most acquisitions occur over the Conterminous United States (CONUS), much of South America, Central and Southern Europe, Australia and Eastern Asia, whereas the coverage is reduced in parts of Africa, Central America and Central, and Northern Asia.

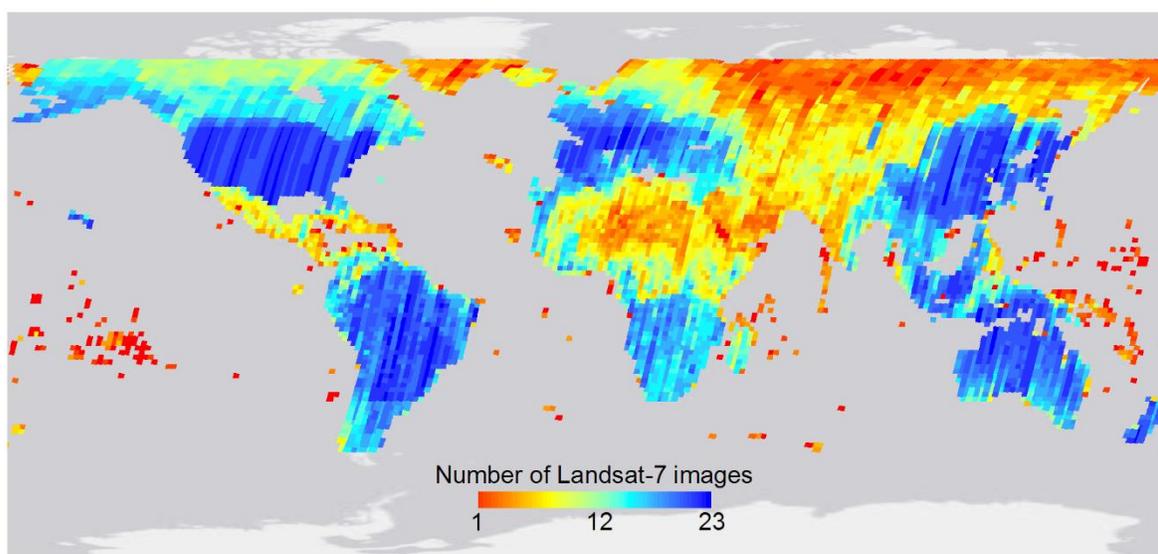


Figure 3-1: Spatial distribution and number of Landsat 7 ETM+ images acquired from 1 January to 31 December 2002 between 70° S and 70° N. The number of images for each Landsat WRS path/row is depicted with a rainbow color scale.

Figure 3-2 shows histograms of the Landsat 7 and also the MODIS cloud fractions (defined for the three MODIS cloud definitions) for the 109,814 Landsat 7 ETM+ images acquired in 2002 and illustrated in Figure 3-1. The histograms exhibit a full range of cloud fraction values, i.e., from 0 to 1, and have bimodal distributions. Except for the least conservative MODIS cloud definition (#3) the first mode corresponds to images with 0% cloud fractions and the second mode correspond to 100% cloud fractions. The MODIS cloud fraction histograms have progressively fewer cloud-free images for the less conservative cloud definitions.

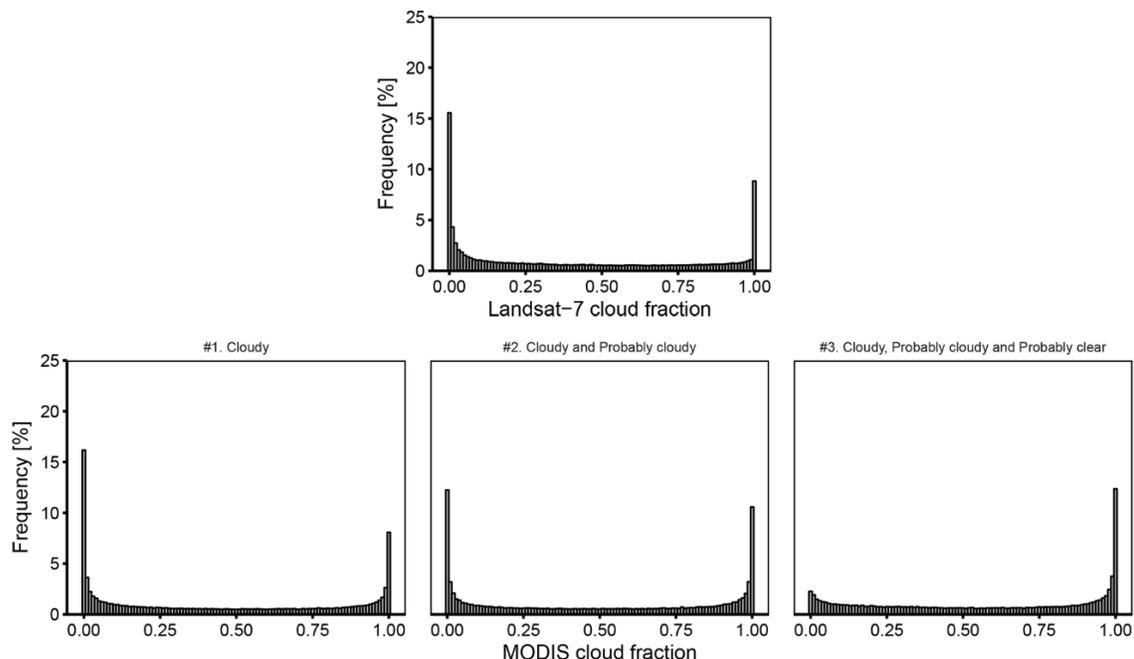


Figure 3-2: Histograms of Landsat 7 (top) and MODIS (bottom row) cloud fractions, observed in all the 2002 study data (Figure 3-1). The histograms are quantized using 100 bins of equal width. The three MODIS cloud fraction histograms are for the three MODIS cloud definitions (Table 3-1).

3.5.2. MODIS and Landsat cloud fraction differences

Figure 3-3 shows summary statistics of the MODIS cloud fraction minus the Landsat 7 cloud fraction derived for all the global 2002 study data. The smallest differences are observed for the most conservative MODIS cloud fraction definition (#1, median difference = 0.00, interquartile range = [-0.02, 0.06]), and increases for the less conservative MODIS cloud fraction definitions (#2: median difference = 0.04, interquartile range = [0.00, 0.14]; #3: median difference = 0.12, interquartile range = [0.03, 0.25]).

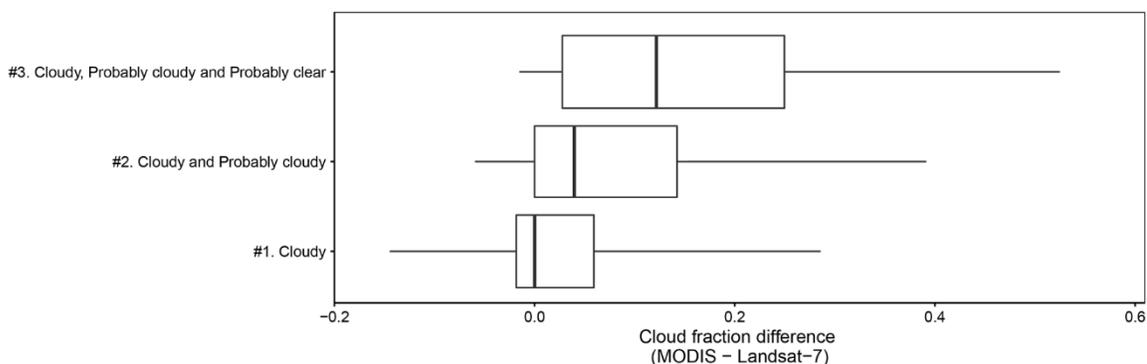


Figure 3-3: Box and whisker plots summarizing the distribution of the difference between the MODIS and Landsat 7 cloud fractions for all the global 2002 data (Figure 3-1) considering the three MODIS cloud fraction definitions (Table 3-1). The thick vertical lines show the median of the distribution, the boxes show the interquartile range, and the whiskers show the 5th to 95th quantile range.

3.5.3. Regression analysis

Figure 3-4 shows scatterplots comparing the Landsat 7 cloud fractions and the MODIS cloud fractions for the three MODIS cloud fraction definitions and considering all the study data. The point density distribution, calculated using a 100 x 100 quantization of the axes, is plotted using a rainbow color scale. Linear and logistic regressions are shown plotted as dotted and continuous lines respectively. The correlation and regression coefficients are summarized in Table 3-2. Consistent with the boxplot results presented in Figure 3-3, the scatterplots indicate that the MODIS cloud fraction is closest to the Landsat cloud fractions when the most conservative MODIS cloud fraction definition is used (#1). Generally, MODIS overestimates the cloud fraction compared to Landsat and there is a progressive overestimation from definition #1 to #3, as shown by the majority of the point distribution (green to purple colors) positioned below the identity line (dashed line) for #2 and #3. This is also reflected by the linear regression lines: the regression lines of #2 and #3 are noticeably further apart from the identity line than the regression line for #1.

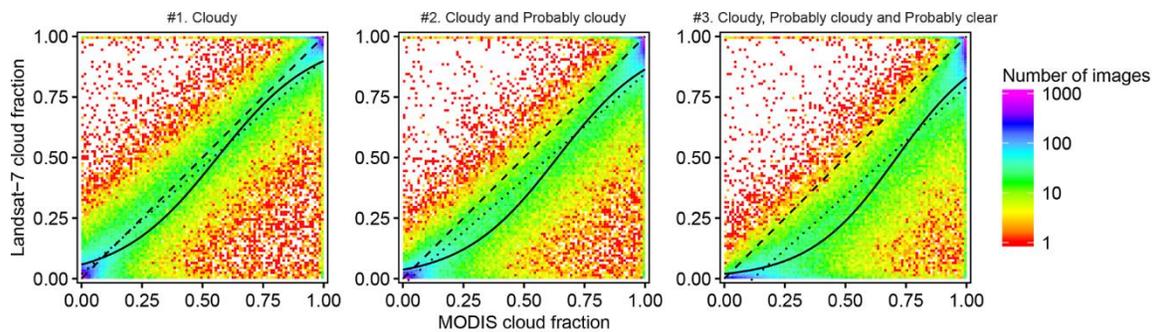


Figure 3-4: Scatterplot of the Landsat 7 cloud fraction plotted against the MODIS cloud fraction, considering all the 2002 study data, plotted for each MODIS cloud fraction definition (Table 3-1). The logistic regression (solid) linear regression (dotted) and the identity line (dashed) are shown superimposed on point density distributions that are generated using a 100 × 100 quantization of the axes, and are displayed with a rainbow logarithmic color scale.

Table 3-2: Regression results of the Landsat 7 and MODIS cloud fraction scatterplots shown in Figure 3-4

	Definition	Coefficients	Standard Error of the Coefficients	Standard Error of the Residuals	R ²	p-value
Linear Model	#1	$\beta_0: 0.02, \beta_1: 0.89$	$\beta_0: 0.0007, \beta_1: 0.0012$	0.149	0.83	<2 ⁻¹⁶
	#2	$\beta_0: -0.02, \beta_1: 0.87$	$\beta_0: 0.0008, \beta_1: 0.0012$	0.157	0.82	<2 ⁻¹⁶
	#3	$\beta_0: -0.10, \beta_1: 0.90$	$\beta_0: 0.0011, \beta_1: 0.0016$	0.186	0.74	<2 ⁻¹⁶
	Definition	Coefficients	Standard Error of the Coefficients	Residual deviance	Null deviance	p-value
Logistic Model	#1	$\beta_0: -2.79, \beta_1: 4.97$	$\beta_0: 0.016, \beta_1: 0.026$	18194	76678	<2 ⁻¹⁶
	#2	$\beta_0: -3.26, \beta_1: 5.11$	$\beta_0: 0.019, \beta_1: 0.027$	18199	76678	<2 ⁻¹⁶
	#3	$\beta_0: -3.91, \beta_1: 5.50$	$\beta_0: 0.023, \beta_1: 0.031$	22146	76678	<2 ⁻¹⁶

The logistic regression results are consistent with the results of the linear regression, with a similar residual deviance for definitions #1 (18194) and #2 (18199) and considerably higher residual deviance for #3 (22146). For all three MODIS cloud definitions the residual deviance is considerably smaller than the null deviance, indicating that the MODIS cloud fraction is an effective predictor for the Landsat 7 cloud fraction. For all three MODIS cloud definitions the standard errors of the coefficients are smaller than the estimated values of the coefficients, and the regressions are significant ($p < 0.05$).

For all three MODIS cloud definitions, the linear regressions between the MODIS and the Landsat 7 cloud fractions are significant ($p < 0.05$). The highest coefficient of determination (R^2) is observed for MODIS cloud definition #1 ($R^2 = 0.83$), with a progressively lower coefficient of determination for definition #2 ($R^2=0.82$) and #3 ($R^2 = 0.74$). While the β_1 (slope) coefficient is similar in the three definitions, with values of 0.89 (#1), 0.87 (#2) and 0.90 (#3), the β_0 (intercept) coefficient decreases from 0.02 (#1), to -0.02 (#2) and -0.10 (#3), reflecting that the MODIS cloud fractions become systematically higher than the Landsat cloud fractions when the less conservative cloud definitions #2 and #3 are considered. The standard error of the residuals is also lower for definition #1 (0.149) compared to #2 (0.157) and #3 (0.186).

To further investigate the behavior of the linear and logistic models, the regression residuals, i.e., the observed minus the predicted Landsat 7 cloud fractions, are summarized in Figure 3-5 for the global year of study data. The scatter plots in Figure 3-5 were obtained by plotting the regression residuals as a function of the MODIS cloud fraction. The scatter plots separately consider the linear (*top row*) and logistic models (*bottom row*), for MODIS cloud definitions #1 (*left column*), #2 (*center column*) and #3 (*right column*). In all plots the majority of the data points (green to purple colors) corresponds to low residuals (i.e., y axis values close to 0), and only isolated data points (red colors) correspond to large residuals (up to 1.0 and -1.0). Consistent with the results of Figure 3-4 and Table 3-2, the most conservative MODIS cloud fraction definition (#1) has residuals that are clustered around 0 on the y axis, whereas definitions #2 and #3 have more dispersed residuals.

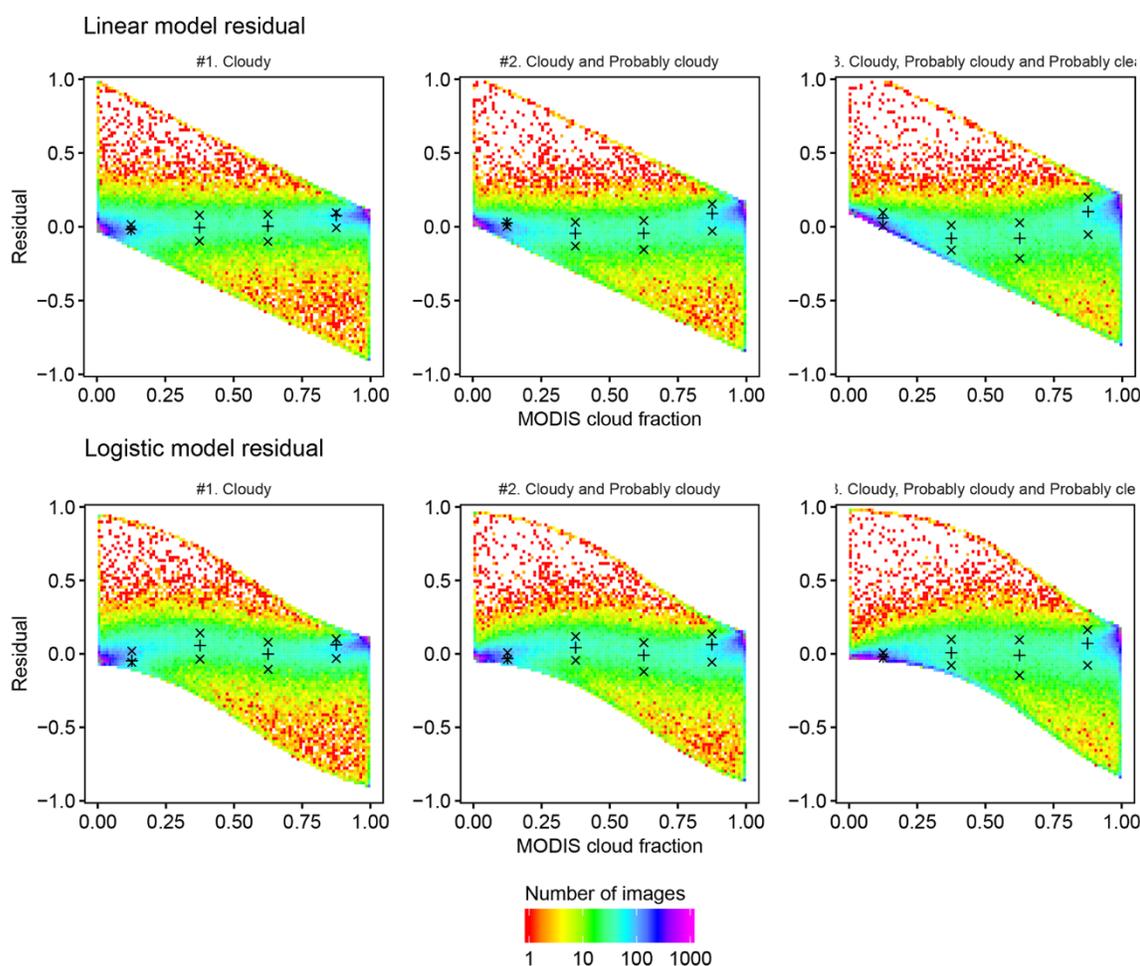


Figure 3-5: Scatterplot of the regression model residuals, defined as the observed minus the predicted Landsat 7 cloud fraction, plotted against the MODIS cloud fraction for the linear (top row) and the logistic regression model (bottom row) considering the 3 MODIS cloud fraction definitions (Table 3-1). The median (+) and interquartile range (x) of the residuals are shown for four MODIS cloud fraction bins: [0, 0.25], [0.25, 0.50], [0.5, 0.75], and [0.75, 1].

The median (denoted by + symbols) and interquartile range (x symbols) of the residuals over four subdivisions of the MODIS cloud fraction range from [0, 0.25], [0.25, 0.50], [0.5, 0.75], and [0.75, 1] are shown in Figure 3-5. These four ranges are used to broadly verify whether the distribution of the residuals changes as a function of the MODIS cloud fraction, i.e., to identify any potential heteroscedasticity of the data, which would violate some of the underlying assumptions of the regression analyses. The residual median and interquartile range values do not significantly change across the MODIS cloud fraction range, indicating that no marked heteroscedasticity is present for both the linear and logistic models and for any of the three MODIS cloud mask definitions.

The most conservative MODIS cloud fraction definition (#1) has the highest linear regression R^2 , and lowest logistic regression residual deviance (Table 3-2). Table 3-3 summarizes for this

definition the median and interquartile range of the residuals for both regressions. The linear model has lower residuals than the logistic model for very low cloud fraction (<0.25), but slightly higher residuals at higher cloud fraction; overall, however, the two models have similar performance, with identical value of the median error (-0.003 linear and logistic) and comparable interquartile ranges ($[-0.032, 0.085]$ linear model, $[-0.058, 0.091]$ logistic model). Thus, the most conservative MODIS cloud fraction definition (#1) is an effective predictor of the Landsat 7 cloud fraction, both when fitting a linear and a logistic model.

Table 3-3: Median, 25th and 75th quantiles of the residuals for the linear (top) and logistic (bottom) models. The residuals were calculated globally and divided in MODIS definition #1 cloud fraction intervals

Linear model		MODIS cloud definition #1 Cloud fraction interval				
Residual value	[0, 0.25]	[0.25, 0.50]	[0.50, 0.75]	[0.75, 1]	[0, 1]	
25 th quantile	-0.022	-0.097	-0.100	-0.007	-0.032	
Median	-0.018	-0.005	0.004	0.077	-0.003	
75 th quantile	0.015	0.079	0.084	0.096	0.085	
Logistic model		MODIS cloud definition #1 Cloud fraction interval				
Residual value	[0, 0.25]	[0.25, 0.50]	[0.50, 0.75]	[0.75, 1]	[0, 1]	
25 th quantile	-0.058	-0.036	-0.105	-0.031	-0.058	
Median	-0.047	0.057	-0.001	0.062	-0.003	
75 th quantile	0.019	0.142	0.079	0.101	0.091	

Figure 3-6 shows the geographic distribution of the minimum (*top row*), median (*middle row*) and maximum (*bottom row*) residuals for the linear (*left column*) and logistic (*right column*) regression using the most conservative MODIS cloud definition (#1). Negative residuals (red, yellow colors) indicate that the linear and logistic models overestimate the observed Landsat cloud fraction, whereas positive residuals (cyan and blue colors) indicate the linear and logistic models underestimate the Landsat cloud fraction. Large minimum and maximum residuals (*top and bottom row*) are predominantly observed at WRS path/rows locations over coastlines, or at high latitude, or over snow prone mountain ranges. This is discussed further in Section 3.6.

The minimum and maximum residual values are driven by outliers whereas the median values correspond to the 50th percentile of the residual distribution and are robust to outliers; as indicated by the prevalence of green colors in the Figure 3-6 middle row figures, the median of the residuals is close to 0 at most locations. This is further illustrated by Figure 3-7, which summarizes the distribution of the median residuals shown in the middle row of Figure 3-6. The global median residual is slightly negative (-0.002 in both models), with a small interquartile range ($[-0.019; 0.043]$ for the linear model, $[-0.048; 0.047]$ for the logistic model) and small 5th-95th quantile range ($[-0.065; 0.095]$ for the linear model, $[-0.060, 0.101]$ for the logistic model). This indicates that, while large median residuals

occasionally occur, the most conservative MODIS cloud fraction definition (#1) is usually an effective predictor of the median Landsat 7 cloud fraction not only globally, but also at each WRS path/row location.

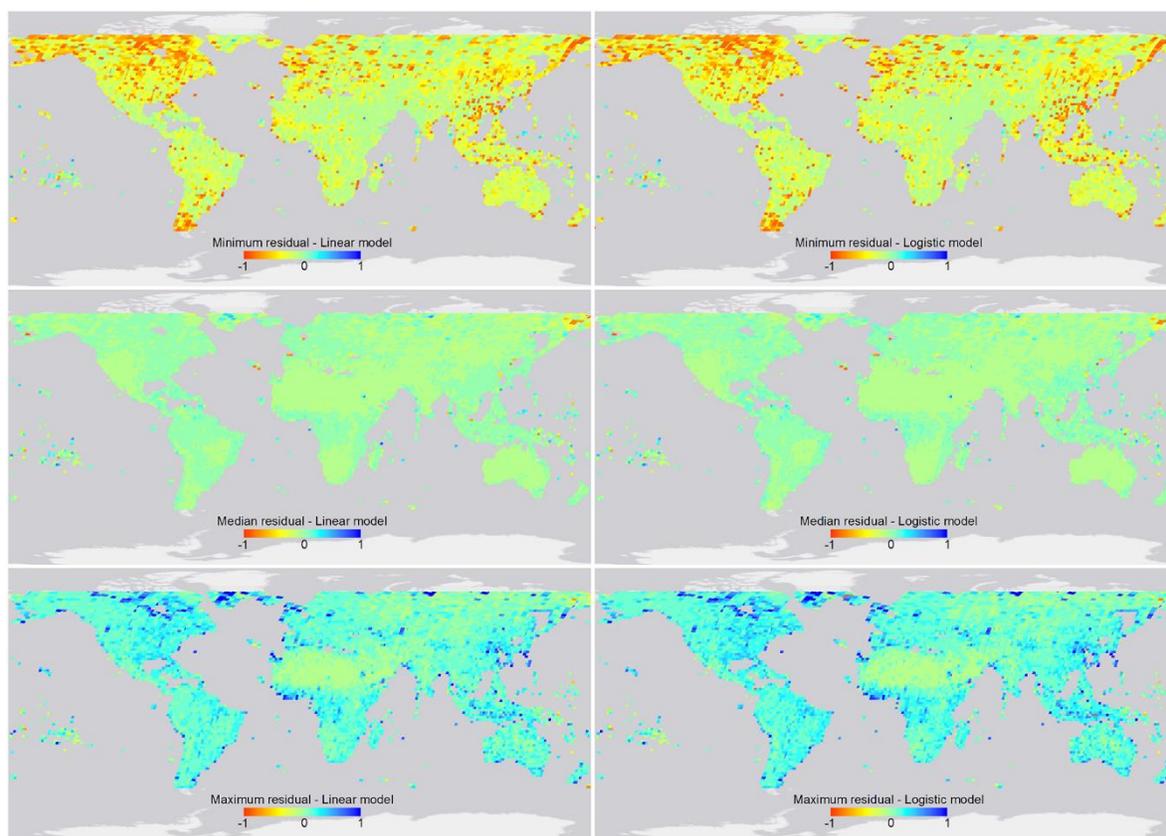


Figure 3-6: Spatial distribution of the minimum (*top row*), median (*middle row*) and maximum (*bottom row*) residuals observed at each WRS path/row location for the linear (*left column*) and logistic (*right column*) regression using the most conservative MODIS cloud definition (#1). The residuals are depicted using the same rainbow color scale for easier comparison.

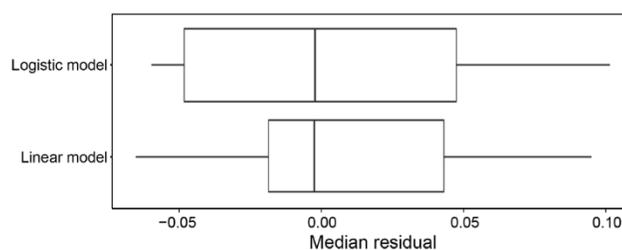


Figure 3-7: Box and whisker plots summarizing the distribution of the median residual of the linear and logistic models at each WRS path/row location (i.e., Figure 3-6, middle row). The thick vertical lines show the median of the distribution, the boxes show the interquartile range, and the whiskers show the 5th- 95th quantile range.

Figure 3-8 shows scatterplots of the linear and logistic regression residuals as a function of the land fraction in each Landsat image. The median, interquartile range and 5th- 95th quantile range of the distribution summarize the distribution of the residuals. The residual median and interquartile ranges

are stable, albeit with a small positive median residual (0.042 with both models) at small land fractions (0 to 0.05), and a small negative median residual (-0.009 linear, -0.014 logistic) for large land fractions (0.95 to 1.00). Consistent with Figure 3-6, the 5th–95th quantile range indicates that at small land fractions (0 to 0.05) there are large outliers in the residuals (range [-0.556, 0.394] linear, [-0.573, 0.412] logistic). The 5th–95th quantile range, however, drops quickly as the land fraction increases, and it remains substantially stable for land fractions >0.25.

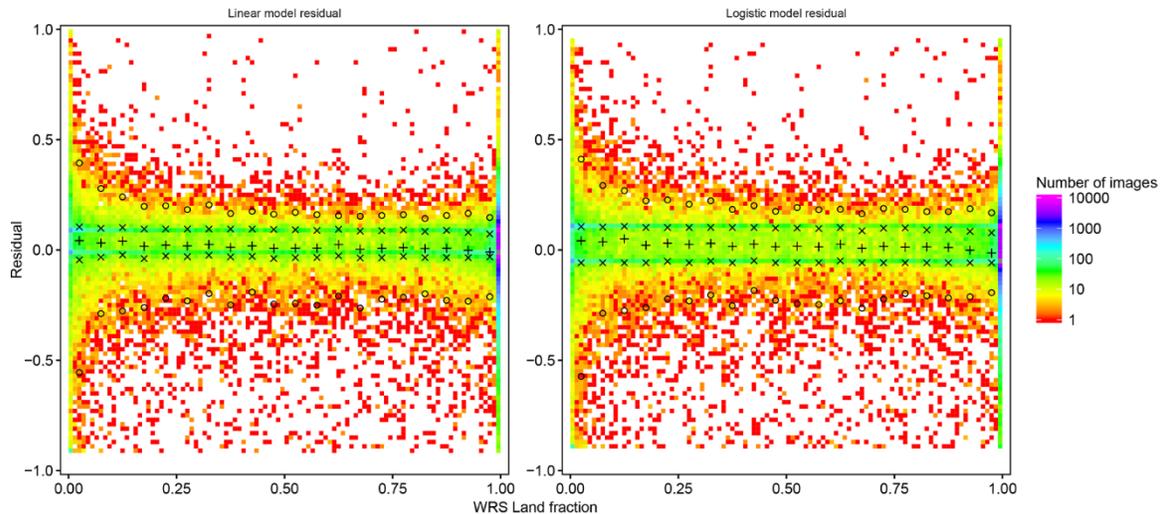


Figure 3-8: Scatterplot of the linear (left) and logistic (right) regression residuals plotted against the land fraction in each Landsat path/row, considering all the 2002 study data, for the most conservative MODIS cloud fraction definition (#1). The median (+), interquartile range (x) and 5th- 95th quantile range (°) of the residuals are shown for 0.05 land fraction bins. The point density distributions are generated using a 100 × 100 quantization of the axes, and are displayed with a rainbow logarithmic color scale.

Finally, in order to verify the presence of seasonal pattern, Figure 3-9 shows the residuals of the linear (*top row*) and logistic (*bottom row*) models as function of the acquisition date for the Landsat images in the northern (*left column*) and southern (*right column*) hemispheres. Median, interquartile range, and 5th- 95th quantile ranges of the residuals are plotted for quarterly intervals, defined as {[January - March], [April - June], [July - September], [October - December]}, and are summarized in Table 3-4.

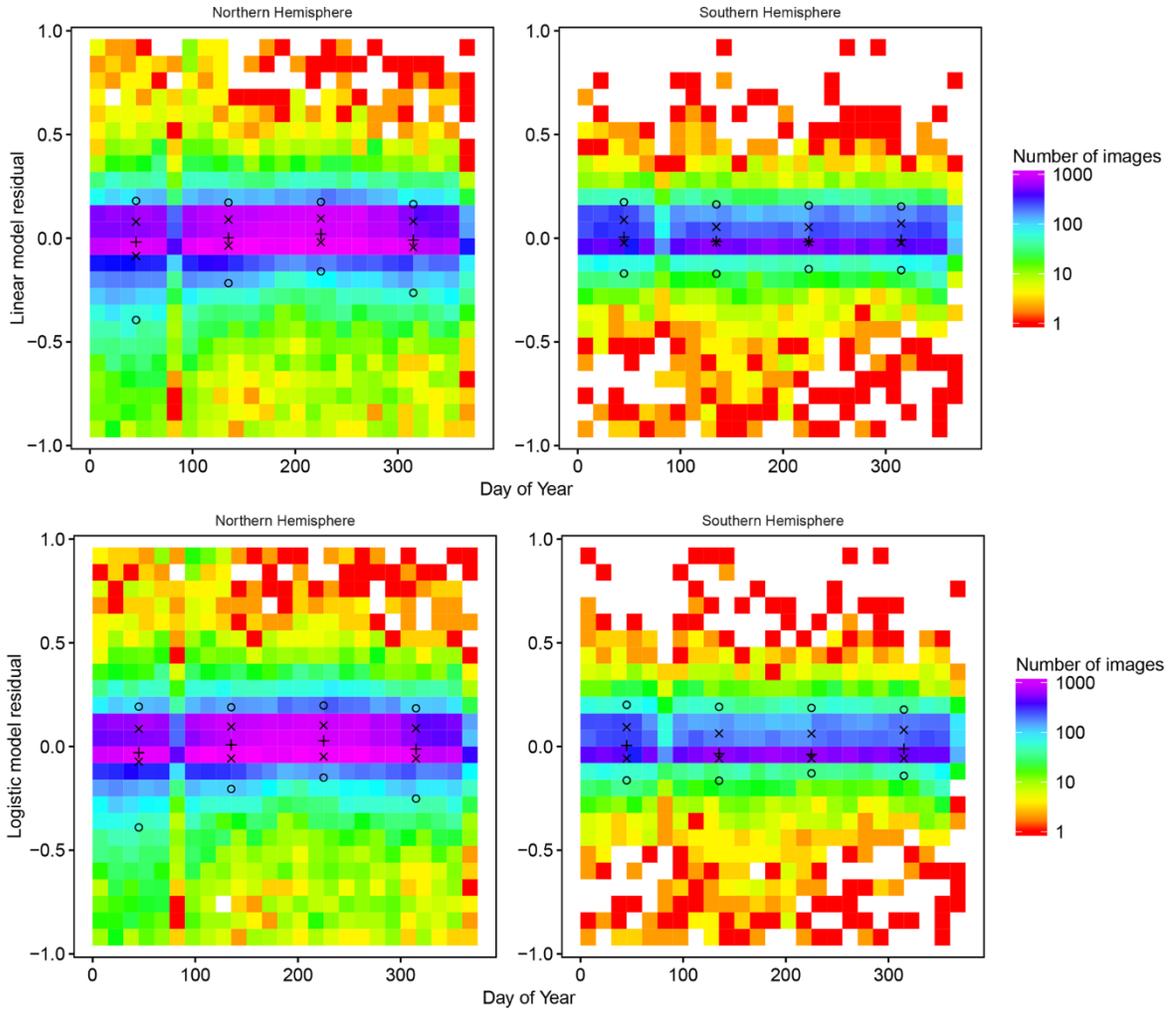


Figure 3-9: Scatterplot of the linear (top row) and logistic (bottom row) regression residuals divided between the northern (left column) and southern hemisphere (right column) plotted against the Landsat acquisition day of year, considering all the 2002 study data, plotted for the most conservative MODIS cloud fraction definition (#1) (Table 3-1). The point density distributions are generated using a 25 x 25 quantization of the axes, and are displayed with a rainbow logarithmic color scale. It should be noted that the artifact visible in all plots for days 75-90 is due to a gap from March 19th to 28th in the MODIS Terra data acquisition. The median (+), interquartile range (x) and the 5th-95th quantile range (°) of the residuals are shown for quarterly acquisition intervals, defined as {[Jan - Mar], [Apr - Jun], [Jul - Sep], [Oct - Dec]}.

While no seasonal pattern is detected in the southern Hemisphere, in the Northern Hemisphere the residuals exhibit a slight seasonal pattern, with higher residual in the winter months. The median of the residuals changes from 0.020 (linear) and 0.027 (logistic) in July – September to -0.019 (linear) and -0.030 (logistic) in January – March. Furthermore, while the 25th, 75th and 95th quantile are substantially stable, the 5th percentile shows the largest seasonal trend, with a variation from -0.161 (linear) and -0.150 (logistic) in July-September to -0.394 (linear) and -0.390 (logistic). These results indicate that the large minimum residuals at high latitude shown in Figure 3-6, *top row*, are predominantly observed on images acquired in the Northern Hemisphere in January-March, i.e., in the winter months.

Table 3-4: Median, 5th, 25th, 75th and 95th quantiles of the residuals for the linear (top) and logistic (bottom) models. The residuals were calculated globally and divided in quarterly acquisition date intervals.

		Acquisition date interval					
		Jan–Mar	Apr–Jun	Jul–Sep	Oct–Dec	Year	
Linear model	Northern Hemisphere	5 th quantile	-0.394	-0.217	-0.161	-0.264	-0.258
		25 th quantile	-0.086	-0.036	-0.020	-0.044	-0.038
		Median	-0.019	0.002	0.020	-0.009	0.001
		75 th quantile	0.079	0.089	0.095	0.081	0.089
		95 th quantile	0.180	0.172	0.175	0.164	0.173
	Southern Hemisphere	5 th quantile	-0.170	-0.172	-0.147	-0.154	-0.159
		25 th quantile	-0.022	-0.019	-0.019	-0.022	-0.020
		Median	0.005	-0.015	-0.016	-0.008	-0.009
		75 th quantile	0.088	0.054	0.053	0.071	0.069
		95 th quantile	0.174	0.163	0.158	0.152	0.161
Logistic model	Northern Hemisphere	5 th quantile	-0.390	-0.205	-0.150	-0.251	-0.244
		25 th quantile	-0.073	-0.058	-0.049	-0.058	-0.058
		Median	-0.030	0.008	0.027	-0.012	0.003
		75 th quantile	0.085	0.095	0.101	0.087	0.094
		95 th quantile	0.192	0.189	0.197	0.184	0.192
	Southern Hemisphere	5 th quantile	-0.164	-0.165	-0.129	-0.141	-0.150
		25 th quantile	-0.058	-0.058	-0.058	-0.058	-0.058
		Median	0.004	-0.034	-0.038	-0.012	-0.021
		75 th quantile	0.093	0.063	0.062	0.079	0.075
		95 th quantile	0.201	0.191	0.186	0.178	0.188

3.6. Discussion

Linear and logistic models were used to predict the Landsat 7 ETM+ cloud fraction from MODIS MOD35 observations, and a non-parametric analysis of the residuals was performed. Overall, the analysis indicates a high degree of correspondence between the Landsat and MODIS cloud fractions, with the best results observed when only the pixels confidently labeled as cloudy in the MOD35 product are used. This case resulted in the highest coefficient of determination and lowest residuals, when fitting the linear and the logistic models. The difference between the MODIS and Landsat cloud fractions increases when less conservative MOD35 cloud definitions are used, i.e., when MODIS pixels flagged as potentially cloudy and potentially clear are counted in the cloud fraction. Arguably, this can be explained considering that the MOD35 cloud masking algorithm makes use of the multiple thermal bands and the water vapor/CO₂ absorption bands which are available on the MODIS instrument. These bands are missing from the Landsat ETM+ instrument, resulting in low sensitivity to thin clouds (Zhu and Woodcock 2012).

The analysis of the residuals indicates that the performance of the linear and logistic models were substantially similar, with the linear model marginally better at low cloud fraction conditions, and the logistic model marginally better for higher cloud fraction conditions. In the case of the linear model, the intercept of the regression line is small (0.02), the slope is close to unity (0.87), and the median of the residuals is also small (-0.002); this small positive error indicates that the use of MODIS to predict the Landsat cloud fraction results in a small overestimation of the cloud cover. With either model, the MODIS-Terra cloud mask underestimates the Landsat cloud fraction at low cloud levels, and overestimates it at high cloud levels. Arguably this is due to the MODIS coarse resolution: small, sparse clouds are not detected, and similarly small gaps in cloud cover are missed. These findings are consistent with previous studies reporting that the MODIS cloud masking algorithm results in some degree of overestimation on fragmented and optically thin clouds (King et al. 2013; Zhao and Di Girolamo 2006). Further, this is a well known scale issue whereby coarse resolution detection products tend to over- and under-estimate the area detected, compared to higher resolution products, when the areal proportions being detected are high and low respectively. For example this has been observed in comparison of MODIS and Landsat derived burned areas (Roy and Boschetti 2009), snow cover (Hall and Riggs 2007), and comparison of AVHRR and Landsat derived forest maps (Mayaux and Lambin 1997).

The geographic analysis of the residuals revealed that the largest differences between the MODIS and Landsat cloud fraction occur along the coastlines, and at high latitude. In the case of coastal areas, further analysis indicated that the errors are mostly observed at WRS path/row locations with land fractions ≤ 0.05 . In these cases, cloud movement in the 30 minutes between the MODIS-Terra and Landsat 7 overpasses may result in large discrepancies of the estimated cloud percentage over land. Furthermore, false positives in the MOD35 cloud mask over coastlines, rivers and lakes regions are known issues of the MOD35 product (Ackerman et al. 2010; Ault et al. 2006). Additionally, the temporal analysis of the residuals showed the presence of a seasonal effect in the northern hemisphere: the overestimation of the Landsat cloud fraction by the MODIS cloud mask is greater in winter than in summer. Visual inspection of the MODIS and Landsat images where large discrepancies were observed confirmed that, at high latitudes and over mountain ranges, ice and snow were flagged as clouds by the MODIS MOD35 product, also a known issue of the MOD35 product (Ackerman et al. 2010; Li et al. 2004; Stillinger et al. 2019).

The results of this study indicate that the MODIS-Terra 1 km cloud product can be used to predict Landsat 7 ETM+ image cloud fractions over land. Thus, the consistent, long term (2000 onwards) daily cloud observations from the MODIS-Terra instrument can be used to assess the impact

of cloud on moderate resolution land monitoring products from the Landsat satellites. Recommendations for the next (after Landsat 9) mission design are currently being developed (Wulder et al. 2019). Due to the high cost of thermal wavelength sensors a possible design option may be to have separate satellites carrying thermal infrared (TIR) and visible to shortwave infrared wavelength (VSWIR) detectors. The degree of temporal overpass separation between the TIR and the VSWIR observations will be considered in application specific terms, but also to minimize overpass cloudiness differences. The results of this study indicate that, in the morning sun-synchronous orbit, an approximately 30 minute overpass time difference does not, globally on average, result in significant differences in observed cloud fraction over land. We finally note that some cautions must be used in considering these results, as high residuals of the linear and logistic regression models were observed on individual images, especially in the presence of snow or on coastal areas.

3.7. Conclusions

In this paper a systematic global comparison between the Landsat 7 ETM+ image cloud fraction over land, and equivalent cloud fractions derived from contemporaneous MODIS-Terra (MOD35) cloud observations was undertaken. Three MODIS cloud fraction definitions were considered, including a conservative cloud definition, an intermediated definition and a much less conservative definition. Due to the similar acquisition orbits and overpass times, the MODIS MOD35 global cloud mask product is a good predictor of the Landsat 7 ETM+ image cloud fraction over land. The most conservative MODIS cloud fraction definition was the best predictor of the Landsat 7 ETM+ cloud fraction, resulting in high coefficient of determination ($R^2 = 0.83$), negligible bias (median difference: <0.01) and low dispersion around the median (inter-quartile range: $[-0.02, 0.06]$) of the estimated linear model. These results suggest that it is possible to use the daily probability of cloud cover, as observed by MODIS-Terra, as a proxy of the cloud cover observed by Landsat 7.

3.8. References

- Ackerman, S., Strabala, K., Menzel, P., Frey, R., Moeller, C., & Gumley, L. (2010). Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35. In, *MODIS Cloud Mask Team, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin*: Citeseer
- Ackerman, S.A., Strabala, K.I., Menzel, W.P., Frey, R.A., Moeller, C.C., & Gumley, L.E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research: Atmospheres*, 103, 32141-32157

- Arvidson, T., Gasch, J., & Goward, S.N. (2001). Landsat 7's long-term acquisition plan — an innovative approach to building a global imagery archive. *Remote Sensing of Environment*, 78, 13-26
- Ault, T.W., Czajkowski, K.P., Benko, T., Coss, J., Struble, J., Spongberg, A., Templin, M., & Gross, C. (2006). Validation of the MODIS snow product and cloud mask using student and NWS cooperative station observations in the Lower Great Lakes Region. *Remote Sensing of Environment*, 105, 341-353
- Boschetti, L., Roy, D.P., Justice, C.O., & Humber, M.L. (2015). MODIS–Landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, 161, 27-42
- Chander, G., Xiong, X., Choi, T., & Angal, A. (2010). Monitoring on-orbit calibration stability of the Terra MODIS and Landsat 7 ETM+ sensors using pseudo-invariant test sites. *Remote Sensing of Environment*, 114, 925-939
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., & Martimort, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36
- Feidas, H., & Cartalis, C. (2005). Application of an automated cloud-tracking algorithm on satellite imagery for tracking and monitoring small mesoscale convective cloud systems. *International Journal of Remote Sensing*, 26, 1677-1698
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., & Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment*, 194, 379-390
- Goward, S.N., Loboda, T.V., Williams, D.L., & Huang, C. (2019). Landsat Orbital Repeat Frequency and Cloud Contamination: A Case Study for Eastern United States. *Photogrammetric Engineering & Remote Sensing*, 85, 109-118
- Goward, S.N., Masek, J.G., Williams, D.L., Irons, J.R., & Thompson, R. (2001). The Landsat 7 mission: Terrestrial research and applications for the 21st century. *Remote Sensing of Environment*, 78, 3-12
- Hagihara, Y., Okamoto, H., & Yoshida, R. (2010). Development of a combined CloudSat-CALIPSO cloud mask to show global cloud distribution. *Journal of Geophysical Research: Atmospheres*, 115, n/a-n/a
- Hall, D.K., & Riggs, G.A. (2007). Accuracy assessment of the MODIS snow products. *Hydrological Processes*, 21, 1534-1547

- Holz, R., Ackerman, S., Nagle, F., Frey, R., Dutcher, S., Kuehn, R., Vaughan, M., & Baum, B. (2008). Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP. *Journal of Geophysical Research: Atmospheres*, *113*
- Ju, J., & Roy, D.P. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment*, *112*, 1196-1211
- King, M.D., Platnick, S., Menzel, W.P., Ackerman, S.A., & Hubanks, P.A. (2013). Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites. *IEEE transactions on geoscience and remote sensing*, *51*, 3826-3852
- Kovalskyy, V., & Roy, D.P. (2013). The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*, *130*, 280-293
- Li, J., & Roy, D. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, *9*, 902
- Li, Z., Cribb, M., Chang, F., & Trishchenko, A. (2004). Validation of MODIS-retrieved cloud fractions using whole sky imager measurements at the three ARM sites. In, *Proc. 14th ARM Science Team Meeting* (pp. 1-6)
- Li, Z., Roy, D.P., Zhang, H.K., Vermote, E.F., & Huang, H. (2019). Evaluation of Landsat-8 and Sentinel-2A Aerosol Optical Depth Retrievals across Chinese Cities and Implications for Medium Spatial Resolution Urban Aerosol Monitoring. *Remote Sensing*, *11*, 122
- 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
- Masuoka, E., Fleig, A., Wolfe, R.E., & Patt, F. (1998). Key characteristics of MODIS data products. *IEEE transactions on geoscience and remote sensing*, *36*, 1313-1323
- Mayaux, P., & Lambin, E.F. (1997). Tropical forest area measured from global land-cover classifications: Inverse calibration models based on spatial textures. *Remote Sensing of Environment*, *59*, 29-43
- Melchiorre, A., & Boschetti, L. (2018). Global Analysis of Burned Area Persistence Time with MODIS Data. *Remote Sensing*, *10*, 750
- Mercury, M., Green, R., Hook, S., Oaida, B., Wu, W., Gunderson, A., & Chodas, M. (2012). Global cloud cover for assessment of optical satellite observation opportunities: A HypsIRI case study. *Remote Sensing of Environment*, *126*, 62-71
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A.S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, *540*, 418

- Roy, D., Lewis, P., Schaaf, C., Devadiga, S., & Boschetti, L. (2006). The global impact of clouds on the production of MODIS bidirectional reflectance model-based composites for terrestrial monitoring. *Geoscience and Remote Sensing Letters, IEEE*, 3, 452-456
- Roy, D.P., & Boschetti, L. (2009). Southern Africa validation of the MODIS, L3JRC, and GlobCarbon burned-area products. *Geoscience and Remote Sensing, IEEE Transactions on*, 47, 1032-1044
- Roy, D.P., Huang, H., Boschetti, L., Giglio, L., Yan, L., Zhang, H.H., & Li, Z. (2019). Landsat-8 and Sentinel-2 burned area mapping - A combined sensor multi-temporal change detection approach. *Remote Sensing of Environment*, 231, 111254
- Roy, D.P., Wulder, M.A., Loveland, T.R., C.E. W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalsky, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, 145, 154-172
- Roy, D.P., & Yan, L. (2018). Robust Landsat-based crop time series modelling. *Remote Sensing of Environment*
- Stillinger, T., Roberts, D.A., Collar, N.M., & Dozier, J. (2019). Cloud Masking for Landsat 8 and MODIS Terra Over Snow-Covered Terrain: Error Analysis and Spectral Similarity Between Snow and Cloud. *Water Resources Research*, 55, 6169-6184
- Whitcraft, A.K., Becker-Reshef, I., Killough, B.D., & Justice, C.O. (2015a). Meeting Earth Observation Requirements for Global Agricultural Monitoring: An Evaluation of the Revisit Capabilities of Current and Planned Moderate Resolution Optical Earth Observing Missions. *Remote Sensing*, 7, 1482-1503
- Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I., & Justice, C.O. (2015b). Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sensing of Environment*, 156, 438-447
- Wolfe, R.E., Nishihama, M., Fleig, A.J., Kuyper, J.A., Roy, D.P., Storey, J.C., & Patt, F.S. (2002). Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, 83, 31-49
- Wolfe, R.E., Roy, D.P., & Vermote, E. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *Geoscience and Remote Sensing, IEEE Transactions on*, 36, 1324-1338

- Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., McCorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., & Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, 225, 127-147
- Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., Fosnight, E.A., Shaw, J., Masek, J.G., & Roy, D.P. (2016). The global Landsat archive: Status, consolidation, and direction. *Remote Sensing of Environment*, 185, 271-283
- Zhao, G., & Di Girolamo, L. (2006). Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. *Geophysical Research Letters*, 33
- Zhu, Z., Wang, S., & 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., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83-94

Chapter 4: Potential Landsat Global Burned Area Omission Error due to Reduced Revisit Frequency and Cloudy Observations

4.1. Abstract

Global burned area products have been generated in the past 20 years from daily coarse resolution satellite data. The transition to moderate resolution (10-30 m) global burned area products would lead to very significant improvements for a variety of fire information users, from global applications to regional and local studies. Higher spatial resolution clearly has benefits, reducing the occurrence of partially burned pixels and increasing the spectral separation of burned / unburned areas compared to coarse resolution sensors. However, moderate resolution satellites have reduced temporal resolution (e.g. 16 days for Landsat), which could potentially lead to large omission errors in ecosystems where the spectral signal associated with burns fades rapidly and there is a high occurrence of clouds.

The objective of this study was to estimate the potential burned area omission error of a hypothetical Landsat 7 global burned area product due to the combined effect of the limited persistence time of the burned area spectral signal, and cloud cover. The simulation was informed by MODIS observations, and by the Landsat acquisition calendar and ground swath footprints. The MODIS global burned area product (MCD64A1) was used as a fire mask, defining the location and timing of burning, whereas the MODIS cloud product (MOD35) was used to determine the number of post-fire cloud-free observations available on Landsat 7 overpass days within the burned area spectral signal persistence time, estimated in Chapter 2. MODIS burned pixels with no cloud-free post-fire observations on Landsat overpass days within the persistence time were considered as potential omission errors, regardless of the detection algorithm adopted.

The result is a spatially explicit map of the burned area omission error. The terrestrial ecoregions of the world (Olson et al. 2001) were used as spatial analysis unit, and within each ecoregion the analysis was further stratified by land cover class (Forest, Shrubland, Grassland & Savanna). Globally, the omission error of the hypothetical Landsat global burned area product was 19%, and presented large variability across the different land covers, with a maximum error over Forest (33%) and minimum over Shrubland (5%).

4.2. Introduction

Global burned area maps have been produced in the past 20 years from several coarse spatial resolution (250 m - 1 km) Earth Observation (EO) systems, including ATSR, MERIS, Terra and Aqua MODIS, and SPOT-VGT (Alonso-Canas and Chuvieco 2015; Giglio et al. 2018; Giglio et al. 2009; Roy et al. 2005; Simon et al. 2004; Tansey et al. 2004; Tansey et al. 2008). Burned area products have been used for a wide variety of applications including global carbon cycle studies (Bowman et al. 2009), estimation of the emission of greenhouse gases and aerosols from biomass burning (van der Werf et al. 2017), modelling ecological changes for global vegetation dynamics (Goetz et al. 2012; Thonicke et al. 2001) and for surface energy budgets (Jin and Roy 2005; Randerson et al. 2006), and studies on fire disturbances effects in global climate models on atmosphere composition and dynamics (Langmann et al. 2009).

Although there is no agreement on an optimal burned area mapping algorithm, coarse resolution global burned area products rely on the availability of daily or near-daily observations (Chuvieco et al. 2019). It is generally assumed that a sufficient number of cloud-free observations are available globally to map burned areas, with few exceptions in known locations of persistent cloud cover such as Indonesia and tropical areas in Brazil (Giglio et al. 2009), or at high latitudes locations such as Siberia (Chu and Guo 2014). There is, however, a strong need for moderate resolution (10 - 30 m) global burned area maps for fire management and environmental restoration, in order to improve emissions estimations, and in support of carbon accounting (Bowman et al. 2015; Hyer and Reid 2009; Mouillot et al. 2014; Randerson et al. 2012; van der Werf et al. 2017). A wide variety of burned area mapping algorithms have been tested and developed for moderate resolution sensors including techniques that exploit single image analysis (Bastarrika et al. 2011; Chuvieco and Congalton 1988; Chuvieco et al. 2002; Koutsias and Karteris 2000; Mitri and Gitas 2004) and multi-temporal analysis (Boschetti et al. 2015; Chuvieco et al. 2002; Hudak and Brockett 2004; Koutsias and Karteris 1998; Miller and Yool 2002; Silva et al. 2005; Smith et al. 2007; Stroppiana et al. 2009). Moderate resolution sensors are characterized by lower revisit frequency (10 - 16 days) due to their narrow field of view (180 – 300 km), which can translate into burned area potential omission errors (Boschetti et al. 2015; Hawbaker et al. 2017) and exacerbates the issues related to designing a robust and effective detection algorithm for mapping burned areas globally.

Recently prototypes of continental scale burned area algorithms were developed using Sentinel 2 and Landsat 8 data (Roteta et al. 2019, Roy et al. 2019). Detection issues linked to observations sparse time series were mitigated using datasets with higher revisit frequency, such as the MODIS daily active fire (Roteta et al. 2019), or using temporal consistency checks between pre- and post-fire acquisitions

(Roy et al. 2019). The extent of the burned area detected with Sentinel 2 and Landsat 8 data is larger than the area detected by MODIS. Roteta et al. (2019) and Roy et al. (2019) estimated a burned area extent increase greater than 70% of the total MODIS (MCD64A1) product burned area due to the detection of smaller and more fragmented fires, but also noted the presence of omission errors in areas of persistent cloud cover.

Among others, two main factors limiting the accuracy of burned area mapping with moderate resolution sensors are the impermanent nature of fire effects on the spectral signature of land (Frederiksen et al. 1990; Langaas and Kane 1991; Melchiorre and Boschetti 2018; Trigg and Flasse 2000) and the availability of cloud-free acquisitions in a limited period (Kovalsky and Roy 2013). Fire effects are generally persistent in forested ecosystems (Chen et al. 2011; Chu and Guo 2014; Fraser et al. 2000; Loboda et al. 2011; Röder et al. 2008) but can disappear rapidly in grasslands and savannas (Bowman et al. 2003; Laris 2005; Pereira 2003; Trigg and Flasse 2000), due to vegetation phenology and vegetation post-fire recovery, and as a result of the removal of charcoal and ashes by atmospheric agents. Additionally to the lower revisit frequency, clouds, smoke, and other optically thick aerosols further limit the number of valid acquisitions within a defined period (Roy et al. 2008; Smith and Wooster 2005).

As detailed in Chapter 2, the burned area spectral signal persistence time, as observed by MODIS, was estimated globally across different ecosystems and periods of the year: 58% of the MODIS global annual burned area can be accurately detected for up to 32 days after the day of burning and 29% for a period between 32 and 48 days after burning (Melchiorre and Boschetti 2018). When considering the 16-days Landsat revisit time, this means that 87% of the annual global burned area is imaged by Landsat only up to three times after burning, and before the burned area spectral signal fades. Hence, the limited availability of data due to low temporal revisit frequency and cloud obscuration, in combination with the limited burned area spectral signal persistence time, could potentially result in large burned area omission errors over areas of short persistence time and high cloud occurrence. This work builds on Chapter 2 and calculates the potential burned area omission error for a hypothetical global burned area product informed only by observations acquired on Landsat 7 overpass days (up to 16-days apart), within the burned area spectral signal persistence time, compared to the MODIS global burned area product, generated with daily observations.

The chapter is structured as follows. Section 4.3 describes the datasets used to stratify and perform the analysis. Section 4.4 describes the definition of the omission error for a global Landsat burned area mapping product. Section 4.5 describes the results at two different scales to highlight

regional and global differences. The chapter concludes with discussion and recommendations for future research and applications.

4.3. Data

4.3.1. MODIS global burned area product

The Collection 6 Terra and Aqua MODIS Global Burned Area Product (MCD64A1) provides the estimated day of burning for the 500 m MODIS pixels that are classified as burned within a calendar month (Giglio et al. 2018). The algorithm (Giglio et al. 2018) is designed to be extremely tolerant of cloud and aerosol contamination, which affected the Collection 5 and 5.1 MCD45A1 MODIS 500 m burned area product (Roy et al. 2008). The algorithm applies dynamic thresholds to composite MODIS Terra and Aqua imagery generated from a burn-sensitive spectral band index derived from MODIS shortwave infrared bands, and a measure of temporal variability. Cumulative MODIS 1 km active fire detections are used to guide the selection of burned and unburned training samples and to guide the specification of prior burned and unburned probabilities. The MCD64A1 product is distributed in the standard MODIS Level 3 $10^\circ \times 10^\circ$ land tile format in the sinusoidal projection (Wolfe et al. 1998).

4.3.2. MODIS land cover product

The MODIS Land Cover Type product (MCD12Q1) provides five land classification schemes, which describe land cover properties derived from one year of observations from Terra and Aqua MODIS (Friedl et al. 2010). The Collection 5.1, Level 3 yearly MCD12Q1 is distributed in the standard MODIS Level 3, $10^\circ \times 10^\circ$ Land tile format in the sinusoidal projection at 500 m resolution. In this work, the International Geosphere and Biosphere Programme (IGBP) scheme was used, which identifies 16 land cover classes including 11 natural vegetation classes, 3 developed and 2 non-vegetated land classes and has a reported 75% overall land cover classification accuracy (Friedl et al. 2010).

4.3.3. MODIS cloud mask

The Collection 6, Level 2, MODIS-Terra cloud mask product (MOD35) (Ackerman et al. 2010) provides for each 1km pixel four possible values in decreasing confidence of cloud detection: “Cloudy”, “Probably Cloudy”, “Probably Clear” and “Clear”. In this study a conservative definition of the cloud mask was used where only pixels labeled as “Cloudy” in the MOD35 product were considered as cloud covered since it provides the best correlation between MODIS and Landsat 7 cloud detections (see Chapter 3 for more details). The MOD35 product also provides information about the set of multispectral test results informed by the 19 MODIS spectral bands, the decision tree used to generate

the cloud mask, and limited ancillary information such as day/night, land/ocean, sun glint, and snow pixel flags.

Terra and Landsat 7 are on the same sun-synchronous 16-days revisit period orbit and their overpasses lag only 15-30 minutes globally. Because of this orbit configuration, the Landsat 7 ETM+ ground swaths coincide with the central nadir-looking portion of the MODIS-Terra swath acquired on the same day from the same orbit (Figure 4-1), therefore, the acquisition geometry (sensor's zenith and azimuth angles) is similar and the overpass time lag is within 30 minutes globally.

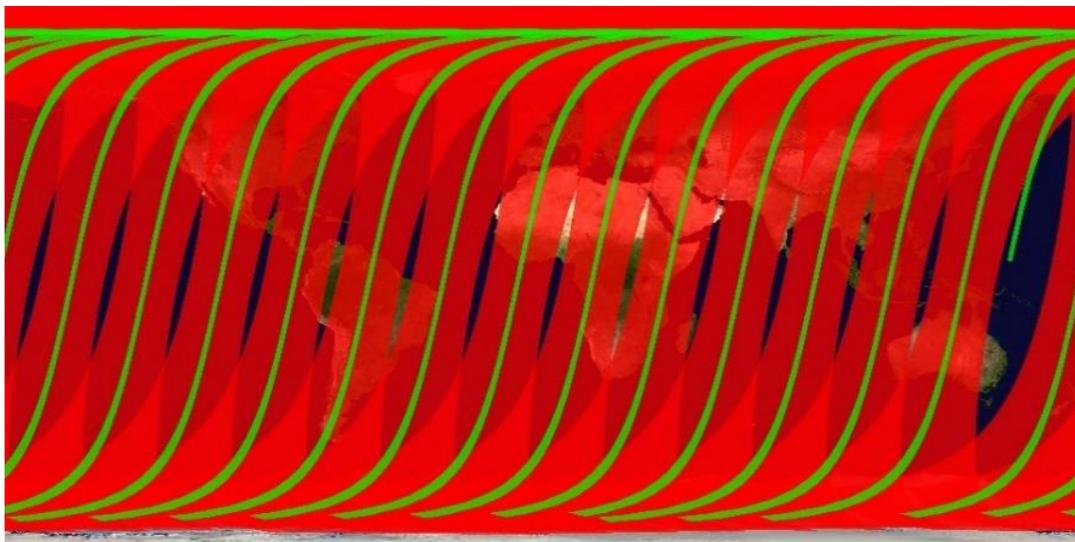


Figure 4-1: MODIS-Terra ground swaths for 03/21/2016 (red) and the daily overlapping Landsat 7 ETM+ ground swaths within 30 minutes (green). The ground swaths were generated using the Committee on Earth Observation System (CEOS) COVE Tool (Kessler et al. 2013). The figure illustrates that the nadir ($< 8^\circ$) portion of the MODIS-Terra swaths always coincides with the Landsat 7 ETM+ swaths and there is a 30 minutes lag between Terra and Landsat 7 overpasses (Mercury et al. 2012).

The MOD35 product is a Level 2 product and so is defined in the MODIS orbit swath geometry. For each orbit swath of the MOD35 1 km cloud mask, the central portion corresponding to the Landsat 7 ETM+ swath was extracted (Figure 4-1) and reprojected into the 1 km MODIS sinusoidal equal-area projection (Wolfe et al. 1998) using the corresponding MOD03 swath geolocation files (Masuoka et al. 1998; Wolfe et al. 2002). Nearest neighbor resampling was used, to preserve the input 1 km data values. It should be noted that, while in general the reprojection of MODIS swaths to a fixed-Earth projection requires compositing because daily swaths overlap at latitudes greater than 30° N/S, the central portion, corresponding to the daily Landsat-7 ETM+ swaths, only overlap at latitudes over 80° N/S (Figure 4-1). Because the MODIS MCD64A1 burned area product is only defined between the latitudes of 70° N/S, no compositing was necessary.

The output was a daily gridded and tiled cloud mask product, defining the presence of absence of clouds as detected by the MOD35 product in correspondence with each Landsat 7 ETM+ overpass. This dataset is henceforth referred to as the MOD35-L7 cloud mask.

4.3.4. Terrestrial ecoregions of the world map

The Terrestrial Ecoregions of the World (TEOW) map is a biogeographic division of the Earth's terrestrial biodiversity in 867 ecoregions, which belong to 14 biomes and 8 realms (Olson et al. 2001). Ecoregions are defined as biogeographic units containing a homogeneous population of natural communities (flora and fauna) sharing a large majority of species, dynamics, and environmental conditions. Realms provide a subdivision of the main landmasses, and biomes are a convenient stratification unit because of their homogeneity of climate and vegetation (Olson et al. 2001).

In this work, the spatial division of the burned area persistence time, described in Chapter 2, was adopted: the TEOW were used as spatial units of the analysis whereas biomes and realms were used to present the results at a larger spatial scale. Also following the burned area persistence time analysis, the 14 Olson's biomes were aggregated in 5 larger biomes: Tropical, Temperate, Boreal, Mediterranean and Desert/Xeric biomes (Figure 4-2).

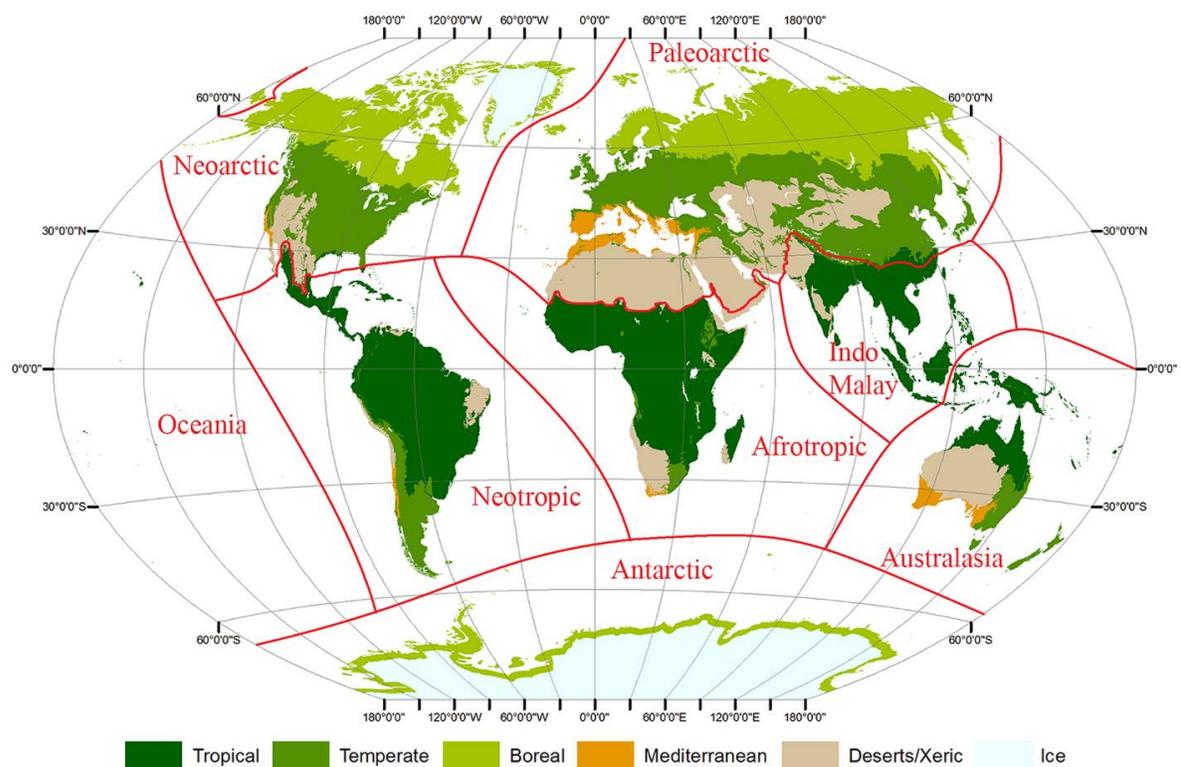


Figure 4-2: Map delineating the borders of the realm (red lines) and the aggregated biomes (color-coded). Oceania and Antarctic realms were not considered for this study. From Melchiorre and Boschetti (2018).

4.3.5. Burned area persistence time

The persistence of the spectral signal associated with burned areas was estimated by Melchiorre and Boschetti (2018) as the duration of the spectral separability of the burned/unburned areas mapped by the MODIS MCD64A1 global burned area product. The separability was computed by analyzing time series of Normalized Burn Ratio (NBR) from nadir BRDF-adjusted MODIS reflectances (MCD43A4 product). The persistence time was calculated independently for each TEOW ecoregion and a second level spatial stratification by land cover was adopted by aggregating the IGBP classes of the MODIS land cover product (MCD12Q1) in Forests, Shrublands, and Grasslands & Savanna (Table 4-1).

Table 4-1: Land cover aggregation scheme of the MODIS MCD12Q1 IGBP land cover classes used in this study. From Melchiorre and Boschetti (2018).

IGBP Land cover	Aggregated land cover
Evergreen Needleleaf forest	Forest
Evergreen Broadleaf forest	
Deciduous Needleleaf forest	
Deciduous Broadleaf forest	
Mixed forest	
Closed shrublands	Shrubland
Open Shrublands	
Woody savannas	Grassland & Savanna
Savannas	
Grasslands	
Permanent wetlands	Not considered
Croplands	
Urban and built-up	
Cropland/Natural vegetation mosaic	
Snow and ice	
Barren and sparsely vegetated	

The persistence time $\Delta t_{\text{Eco,LC},j}^*$ was estimated for each ecoregion (Eco), land cover (LC), and day of burning j (Table S3, Melchiorre and Boschetti 2018).

4.4. Methods

The potential global Landsat burned area omission error was estimated as the average annual burned area, detected by the MCD64A1 product, without a post-fire cloud-free observation on Landsat overpass days in the period starting from the estimated day of burning, and lasting the duration of the associated burned area spectral signal persistence time. At each location, the burned area omission error was estimated using the MCD64A1 burned area product and the MOD35-L7 cloud mask from January 2002 to December 2016. Because of the scale of the analysis, a spatial stratification was adopted, using

strata sufficiently fine to capture the variability of burned areas and cloud dynamics (Section 4.4.1). The potential burned area omission error was estimated for each stratum, defined by the ecoregions and land cover (Section 4.4.2). Finally, the burned area omission error was aggregated at the broader scale of biomes and realms to summarize the results (Section 4.4.3).

4.4.1. Spatial stratification

The effect of cloud cover on satellite data availability is highly variable regionally and depends on the time of the year (Ju and Roy 2008), therefore, following the stratification adopted in Chapter 2, the analysis was stratified spatially adopting a two levels stratification. At the first level, burned pixels across different ecosystems of the world were stratified by using Olson's TEOW. Only the ecoregions with a valid estimate of the burned area persistence time were considered (Figure 4-3).

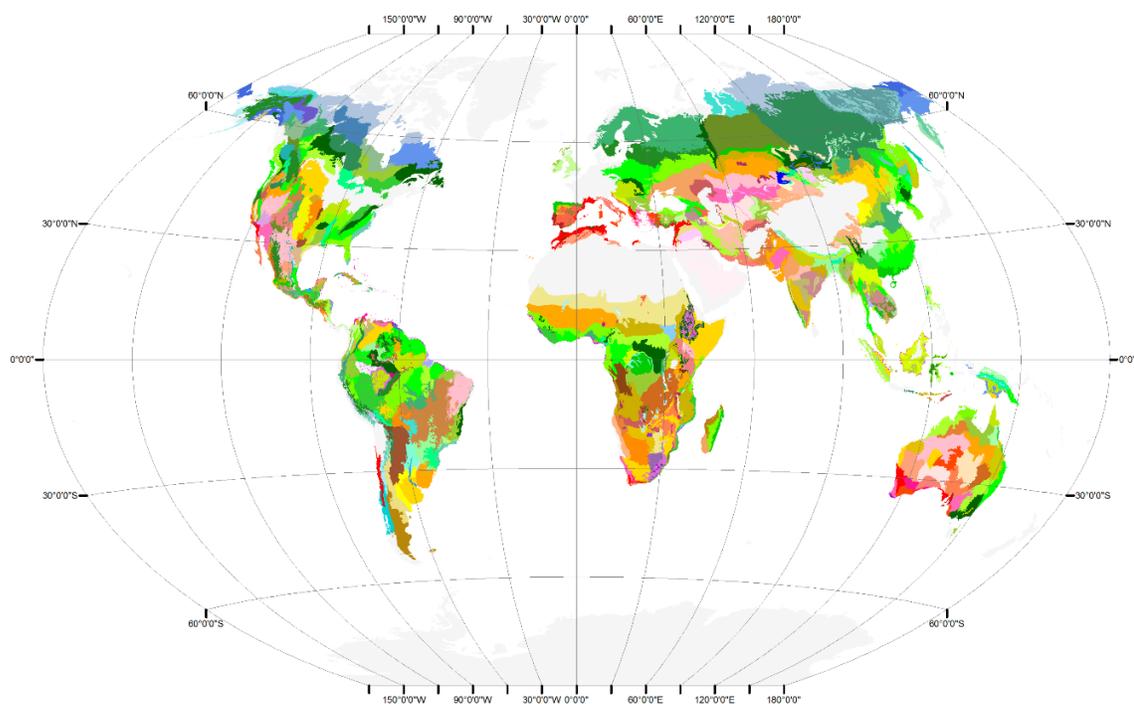


Figure 4-3: TEOW map used to stratify the analysis. Ecoregions considered in the analysis are random color-coded, ecoregions with negligible fire activity were discarded and are shown in light grey.

Within each ecoregion, a second level of stratification was introduced, based on land cover as defined by the MODIS land cover product (MCD12Q1). Following Melchiorre and Boschetti (2018), the land cover classes of the IGBP classification scheme were aggregated into three major classes of interest (Forest, Shrubland, Grassland & Savanna), or masked out and removed from the subsequent analysis (Urban areas, croplands, and miscellaneous non-burnable surfaces) (Table 4-1).

4.4.2. Burned area omission error estimation

Each MODIS MCD64A1 burned pixel ($X_{B,j}$) in each combination of Ecoregion (Eco) and land cover (LC) considered ($X_{B,j} \in [\text{Eco}, \text{LC}]$), and with an estimated day of burning j within the ecoregion fire season was analyzed. The burned area persistence time $\Delta t^*_{\text{Eco,LC},j}$ is defined for each burned pixel based on its location (i.e., ecoregion and land cover) and burning date j . The MOD35-L7 cloud mask time series for each generic burned pixel $X_{B,j} \in [\text{Eco}, \text{LC}]$ detected as burned on the generic Day of Year (DOY) j was extracted starting from the day of burning j and for $\Delta t^*_{\text{Eco,LC},j}$ days; that is, the estimated persistence time for the given stratum $[\text{Eco}, \text{LC}, j]$. Only the observations acquired on Landsat overpass days were used (Figure 4-4).

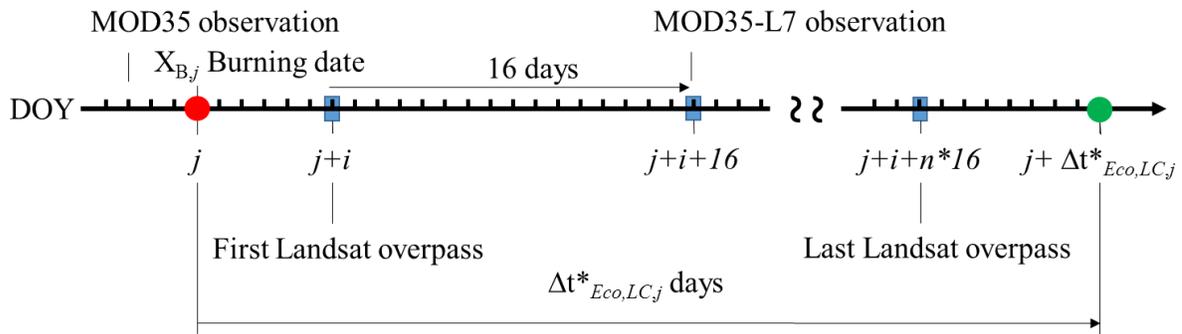


Figure 4-4: MOD35-L7 cloud mask time series extraction over each burned pixel $X_{B,j}$. The time series started on the day of burning j and lasted $\Delta t^*_{\text{Eco,LC},j}$ days; i.e., the burned area persistence time. MOD35-L7 cloud data observed on Landsat overpass days are extracted for the whole duration of the period considered.

From the observations extracted from the MOD35-L7 cloud mask time series, the number of post-fire cloud-free observations was calculated for all the burned pixels of the stratum $[\text{Eco}, \text{LC}]$. The procedure is illustrated in Figure 4-5. The number of times each burned pixel $X_{B,j}$ was cloud-free ($CF_{X_{B,j}}$) after the burning date (j) and within the burned area persistence time ($\Delta t^*_{\text{Eco,LC},j}$) was defined as (Equation 1):

$$CF_{X_{B,j}} = \sum_{DOY=j}^{j+\Delta t^*_{ba,j}} MOD35 - L7(DOY)_{X_{B,j}} \quad (1)$$

where $MOD35 - L7(DOY)_{X_{B,j}}$ assumes value 1 if the pixel $X_{B,j}$ was cloud-free in the MOD35-L7 cloud mask time series on day DOY and 0 otherwise. If there are no post-fire cloud-free observations within $\Delta t^*_{\text{Eco,LC},j}$ days (i.e., $CF_{X_{B,j}} = 0$), the burned pixel $X_{B,j}$ is considered as omission error in a hypothetical Landsat global burned area product.

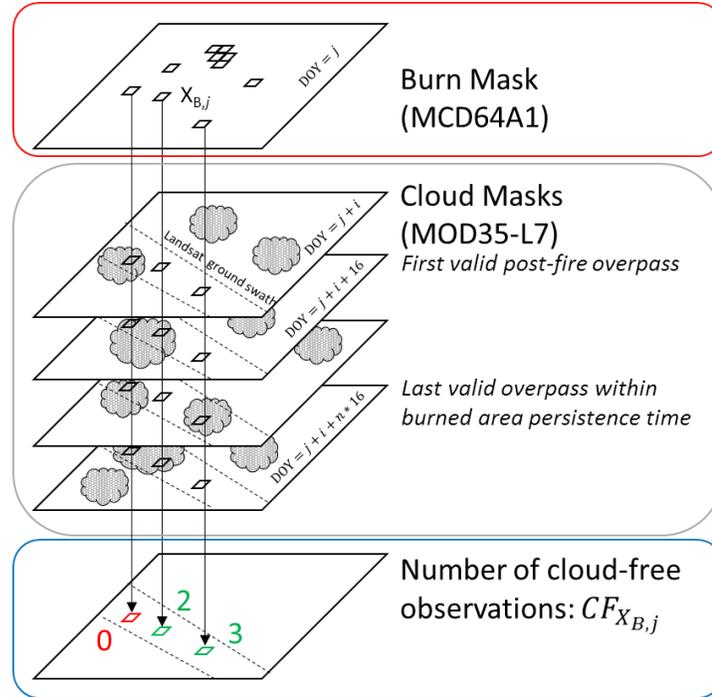


Figure 4-5: Schematic representation of the cloud-free observations count. At each location $X_{B,j}$ classified as burned in the MCD64A1 product on Day Of Year (DOY) j , cloud mask observations were extracted following the Landsat 7 ETM+ geometry and acquisition calendar within the time series defined by the day of burning and the persistence time. The burned pixel is classified as detected from Landsat if at least one cloud-free acquisition is retrieved (green pixels) or omitted if all the acquisitions extracted are flagged as cloudy (red pixel).

$CF_{X_{B,j}}$ was calculated for each pixel detected as burned by the MCD64A1 product between 2002 and 2016, and it was subsequently used to calculate the average annual area burned, where at least a given number of post-fire observations are available. The average area burned, calculated as a function of the Number of Cloud-Free (NCF) observations for each stratum [Eco, LC] is defined as (Equation 2).

$$A_{Eco,LC}(NCF) = \frac{1}{Y} \sum_{X_{B,j} \in Eco,LC} A_{X_{B,j}} \delta_{CF_{X_{B,j}}}(NCF) \quad (2)$$

Where Y is the number of years considered in the analysis, $A_{X_{B,j}}$ is the area of the generic pixel $X_{B,j}$ of a given stratum [Eco, LC], $\delta_{CF_{X_{B,j}}}(NCF)$ is the Dirac measure assuming value 1 if $CF_{X_{B,j}}$ (i.e., the number of cloud-free observations associated with the generic pixel $X_{B,j}$) is equal to NCF and 0 elsewhere.

Finally, the burned area omission error for each stratum [Eco, LC] was calculated as the average annual burned area having no cloud-free observations (Equation 3).

$$OE_{Eco,LC} = A_{Eco,LC}(NCF = 0) \quad (3)$$

Figure 4-6 shows an example of the average annual burned area calculated as a function of the NCF observations and the relative omission error for a combination of ecoregion and land cover.

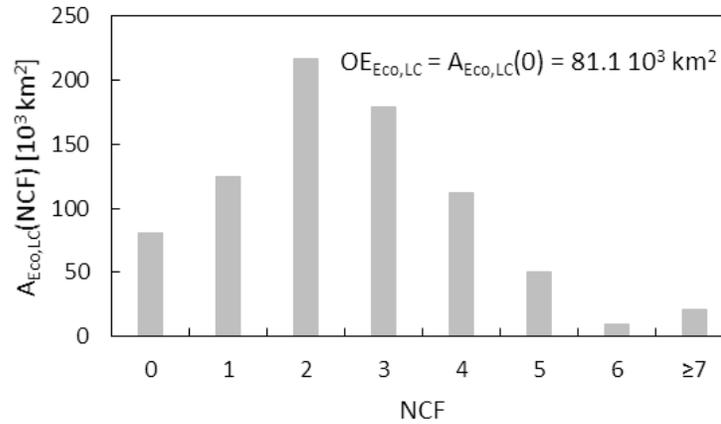


Figure 4-6: Example of Landsat omission error estimate from a set of burned pixels extracted from the TEOW Kimberly Tropical Savanna, Northwest Australia, for the Grassland & Savanna land cover. The graph shows the average annual burned area calculated as a function of the number of cloud-free (NCF) observations. The estimated omission error is the average annual burned area with no cloud-free observations (NCF = 0).

4.4.3. Biome/Realm level aggregation of the results

The average annual burned area potentially omitted by Landsat 7 ETM+ calculated in each ecoregion was aggregated at a larger spatial scale using the aggregated Biomes and Realms as defined in Section 4.3.4. (Figure 4-2), while maintaining the stratification by land cover.

The average annual burned area as a function of the number of cloud-free observations for each spatial unit defined by Biomes (B), Realms (R), and land cover (LC) was simply calculated by summation of all the ecoregions belonging to the spatial unit, and globally by summation over the realms (Equation 4):

$$A_{B,R,LC}(NCF) = \sum_{Eco \in B,R} A_{Eco,LC}(NCF) \quad (4)$$

Finally, the burned area omission error for each Biome, Realm and land cover was calculated as the average annual burned area having no cloud-free observations (Equation 5).

$$OE_{B,R,LC} = A_{B,R,LC}(NCF = 0) \quad (5)$$

4.5. Results

4.5.1. Burned area omission error at the ecoregion scale

Globally, for the 2002-2016 period and limited to the ecoregions and land cover considered, the MCD64A1 average annual burned area analyzed was $\sim 3,470 \cdot 10^3 \text{ km}^2$, of which $\sim 170 \cdot 10^3 \text{ km}^2$ (5.0

%), $\sim 290 \cdot 10^3 \text{ km}^2$ (8.6 %), and $\sim 3,000 \cdot 10^3 \text{ km}^2$ (86.4 %) over Forest, Shrubland, and Grassland & Savanna land cover, respectively. Figure 4-7 shows the estimated potential omission error of a Landsat global burned area product compared to the MCD64A1 product for each ecoregion and aggregated land cover class in terms of average annual burned area omitted (Equation 3) (left column) and in terms of average annual omission error percentage of the total MCD64A1 average annual burned area detected (right column). The full result tables are presented as supplementary material (Table S1).

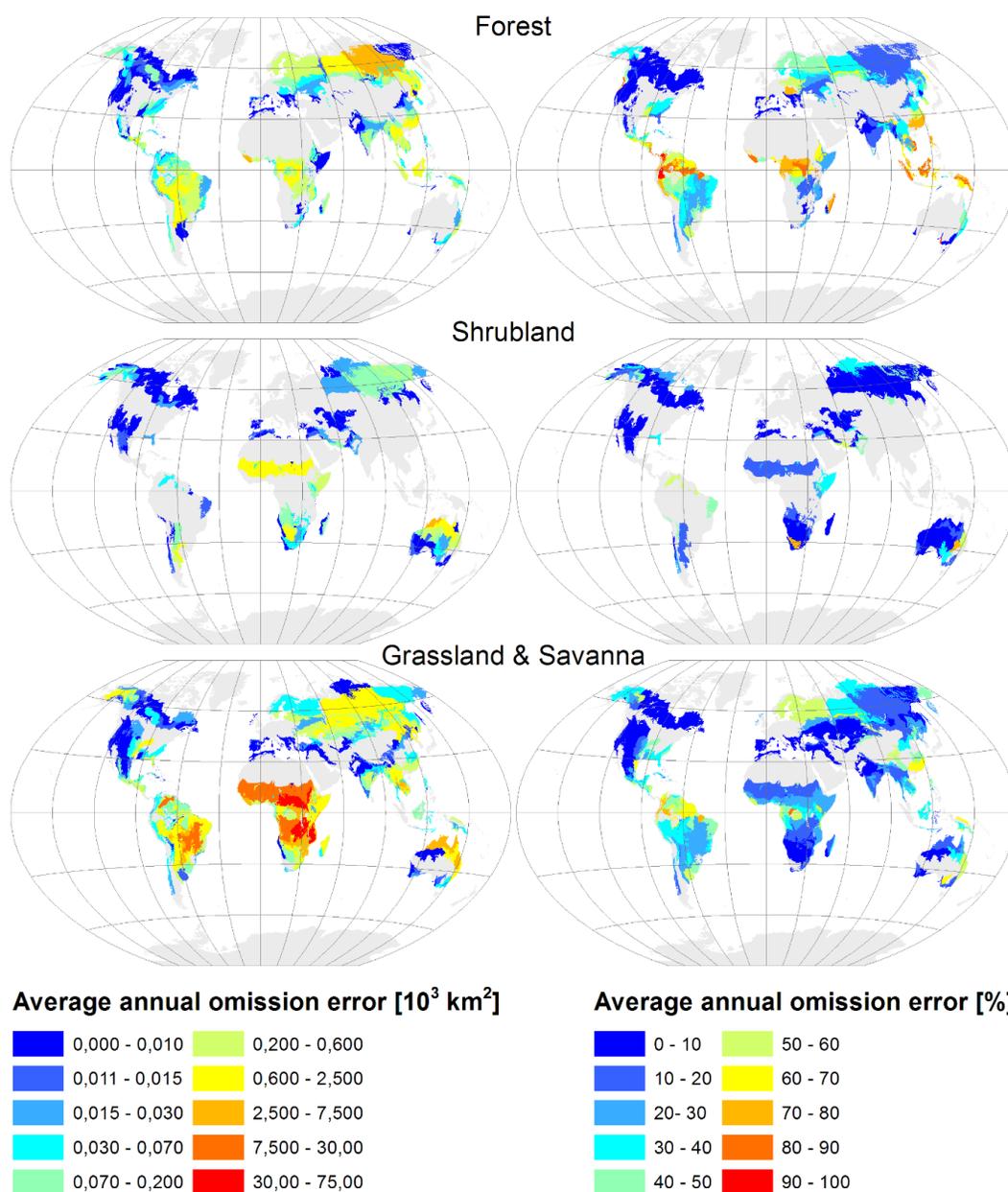


Figure 4-7: Potential average annual burned area omission error of Landsat 7 with respect to the MCD64A1 global burned area product (left column) and relative percentage (right column) for each aggregated land cover: Forest (top row), Shrubland (middle row) and Grassland & Savanna (bottom row). Ecoregions excluded from the analysis are depicted in gray for geographic reference.

The Forest land cover has the lowest amount of MCD64A1 burned area detected. The equatorial ecoregions have large omission errors with peak values of 90%, 80%, and 83% in the Peruvian, Northern Congo, and Indonesian tropical forest ecoregions respectively (Figure 4-7, *top row*). Large omission errors (18%) were estimated also in the East Siberian taiga ecoregion. The largest omission errors in the North American continent were 30% and 17% respectively in the boreal forest ecoregions of Alaska and in the South-East United States respectively.

The Shrubland land cover had the lowest omission error globally, with values ranging from 3% in the Kalahari savanna to 15% in the Sub-Saharan ecoregions.

Because of the prevalence of burned area in the Grassland & savanna land cover class in the Tropical ecoregions, this land cover had the largest omission errors, when expressed in terms of area rather than percentage (Figure 4-7, *bottom row, left column*). Some ecoregions in Zambia, Congo, and Sudan have an average annual burned area omission error over $30 \cdot 10^3 \text{ km}^2$, with the maximum value ($71 \cdot 10^3 \text{ km}^2$) observed in Sudan. Expressed in percentage, these errors range 15 % to 25 % (Figure 4-7, *bottom row, right column*). The Cerrado ecoregion in Brazil has the largest estimated burned area omission error in South America with $\sim 25 \cdot 10^3 \text{ km}^2$ (25 %) and the tropical savanna ecoregions of Northern Australia have similar results with $\sim 5 \cdot 10^3 \text{ km}^2$ average annual omission error (10 % to 15 % when expressed as percentage). Equatorial ecoregions in South America and Africa have the largest omission errors (in the order of 80% and 70% respectively). In North America, the largest errors (30%) are observed in the boreal forest ecoregions of Alaska and in the South-East United States.

4.5.2. *Burned area omission error at the Realm/Biome scale*

The results were aggregated spatially using the Biomes and Realms of the TEOW (section 4.4.3) maintaining the stratification in the three aggregated land cover classes considered (Table 4-1).

Figure 4-8 shows the histogram of the average annual burned area as a function of the number of cloud-free observations for the Forest land cover in all the biomes and realms considered (Equation 4). For the Forest land cover, the majority of the MCD64A1 burned area was detected in the Tropical, Temperate, and Boreal biomes. The histogram was unimodal for the Tropical biome with a mode of 0 cloud-free observations within the persistence time and 43.3% of the total average annual MCD64A1 burned area omission error. The histogram for the Temperate biomes was also unimodal, with a peak at 1 cloud-free observation, mainly due to burned area detected in the Paleoarctic realm and omission error equal to 24.3% of the total average MCD64A1 burned area detected. Only 7.8% of the average annual burned area had more than 7 cloud-free observations, mainly located in the Australasia and Neoarctic realm. Finally, the histogram for the Boreal biomes showed different values for the two realms of Neoarctic and Paleoarctic; the Neoarctic realm is characterized by a larger number of cloud-

free acquisitions compared to the Paleoarctic. The difference is partially explained by longer persistence times of the Neoarctic realm and, therefore, a higher probability of sensing cloud-free observations compared to the Paleoarctic realm.

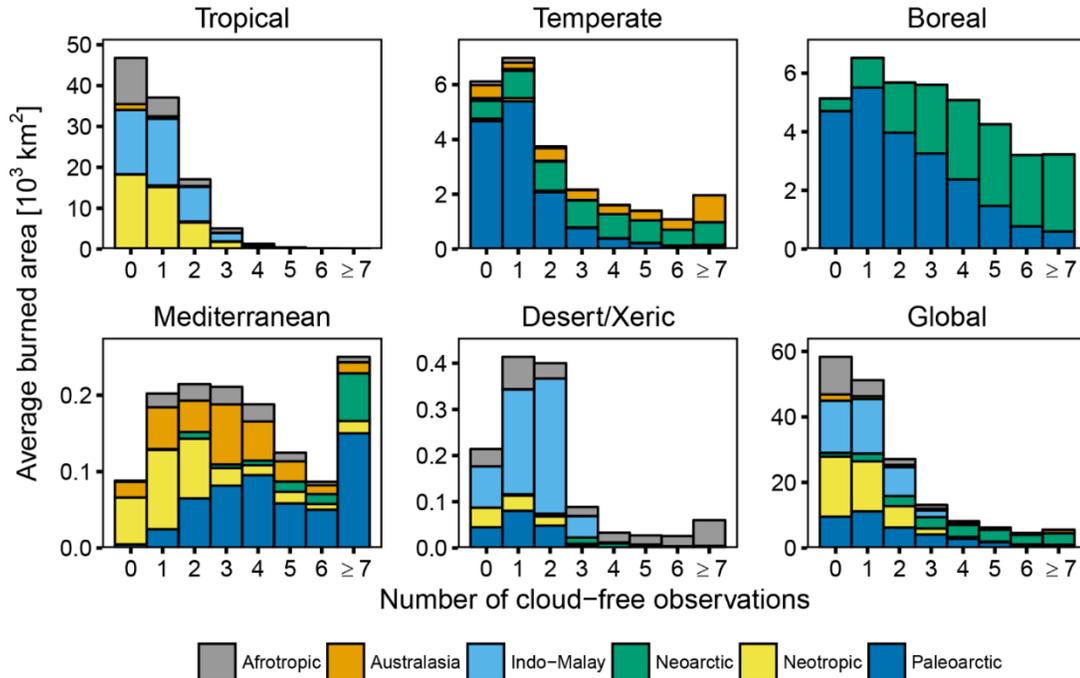


Figure 4-8: Average annual burned area detected in Forest aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm.

Figure 4-9 shows the histogram of the average annual burned area as a function of the number of cloud-free observations for the Shrubland land cover in all the biomes and realms considered (Equation 4). For the Shrubland land cover, the majority of the burned area was detected in the Tropical and Desert/Xeric biomes of the Australasia and Afrotropic realms. The histogram was unimodal for both biomes, the mode for the Tropical biome was 2 cloud-free observations and it was 4 for the Desert/Xeric biome. The largest omission error was found in Tropical biomes with $\sim 10 \cdot 10^3 \text{ km}^2$ average annual omission error ($\sim 9\%$ of the total MCD64A1 average annual burned area).

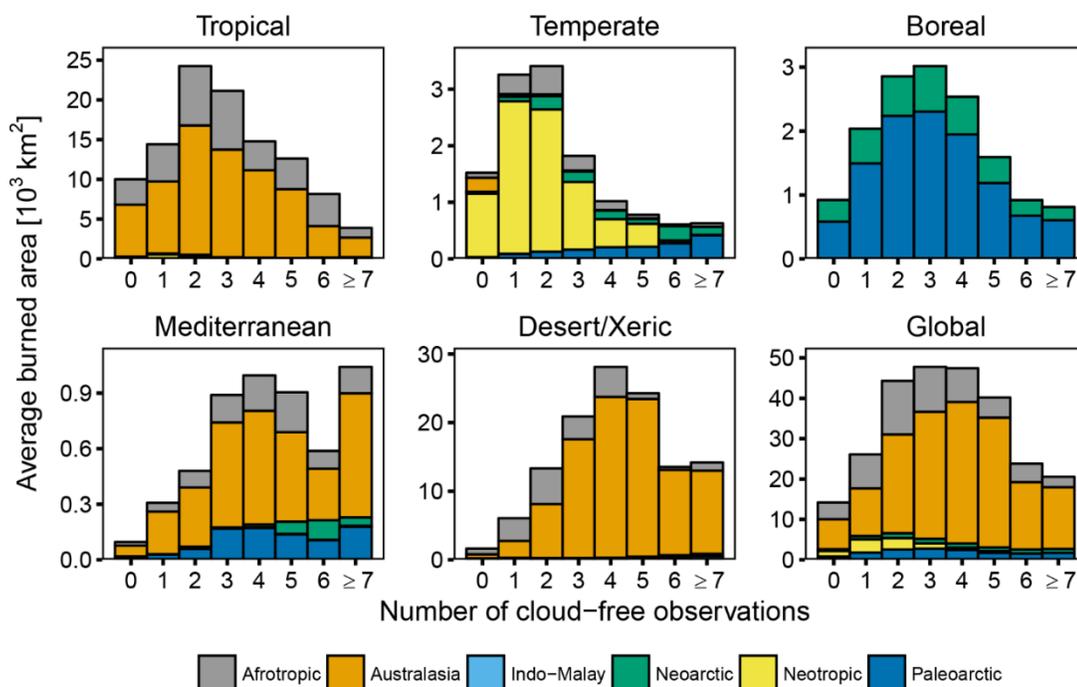


Figure 4-9: Average annual burned area detected in Shrubland aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm.

Finally, Figure 4-10 shows the histogram of the average annual burned area as a function of the number of cloud-free observations for the Grassland & Savanna land cover in all the biomes and realms considered (Equation 4). For the Grassland & savanna land cover, the majority of the burned area was detected in the Tropical and Temperate biomes. The histogram in the Tropical biomes was dominated by the detections in the Afrotropic realm. The mode was 2 cloud-free observations and only 6.0 % of the burned area analyzed had 3 or more cloud-free observations. The omission error for the Afrotropic Tropical biome was $\sim 422 \cdot 10^3 \text{ km}^2$ average annual omission error (~ 19 % of the total MCD64A1 average annual burned area). Similarly, for the Temperate biomes, the histogram of the burned area was unimodal with a mode of 1 cloud-free observation, corresponding to 27.7% of the average annual burned area; 25.6% had 2 cloud-free observations and only 14.5% had more than 3 cloud-free observations. The omission error for Temperate biomes was $\sim 20 \cdot 10^3 \text{ km}^2$ average annual omission error (~ 15 % of the total MCD64A1 average annual burned area).

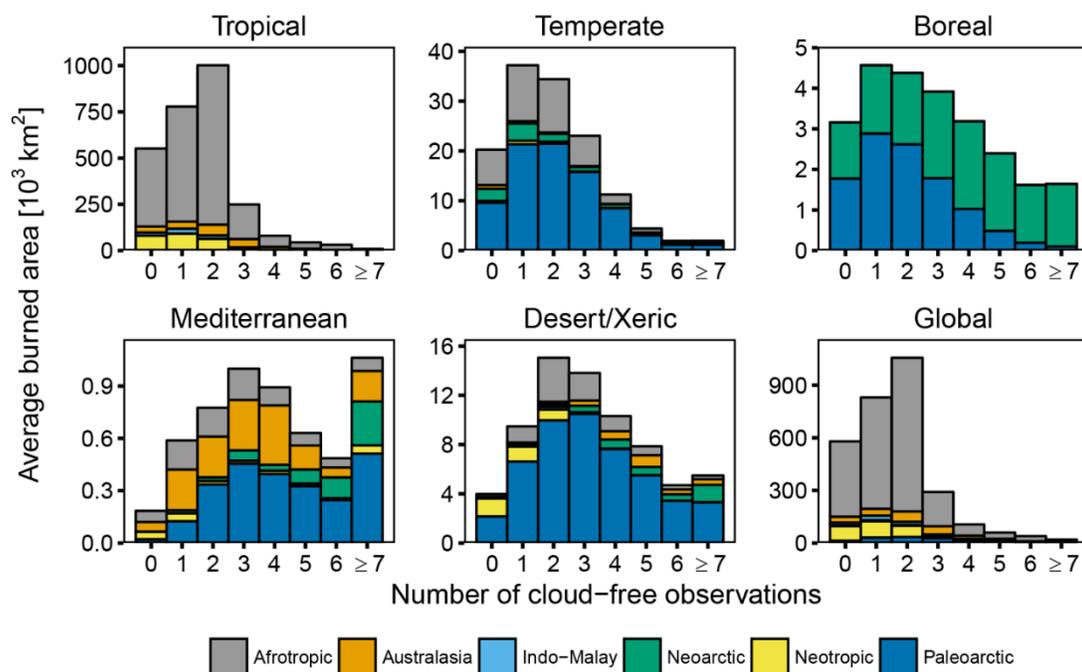


Figure 4-10: Average annual burned area detected in Grassland & Savanna aggregated land cover class as a function of the number of cloud-free observations, presented for each aggregated biome and globally. The histograms were computed with bins of 1 observation, and overflow bin of 7 cloud-free observations. The colors represent, in each bar of the histograms, the aggregation by realm.

Globally, the burned area omission error was 19% of the total average annual burned area detected by the MODIS MCD64A1 product. The omission error was largely dependent from the biome and land cover analyzed and, globally, it was 33.4 %, 5.4 % and 19.4 % in Forest, Shrubland, and Grassland & savanna, respectively (Table 4-2). In all biomes, the highest estimated omission errors were observed in Forest, ranging from 6.4% in the Mediterranean biomes to 43.3% in the Tropical biomes. Conversely, in all biomes the lowest errors were observed in the Shrubland land cover, ranging 1.3% in the Desert/Xeric biomes to 11.6% in Temperate biomes.

Table 4-2: Average annual burned area omission error for each of the aggregated Biomes and land cover considered.

10^3 km^2 (%)	Tropical	Temperate	Boreal	Mediterranean	Desert/Xeric	Global
Forest	46.8 (43.3)	6.1 (24.4)	5.1 (13.2)	0,09 (6.4)	0.21 (16.9)	58.3 (33.4)
Shrubland	10.0 (9.1)	1.5 (11.6)	0.92 (6.2)	0.09 (1.8)	1.6 (1.3)	14.2 (5.3)
Grassland & Savanna	551.9 (20.1)	20.2 (15.0)	3.1 (12.7)	0.18 (3.2)	4.0 (5.6)	579.5 (19.4)
Total	608.8 (20.5)	27.9 (16.1)	9.2 (11.8)	0.37 (3.0)	5.8 (3.0)	652.0 (19.0)

Overall, the Desert/Xeric and Mediterranean biomes had the lowest omission errors (3%). Tropical biomes were the most affected, with 20.5 % omission error. Tropical biomes, having the largest amount of MCD64A1 average annual burned area detected, had also the highest omission error when expressed in terms of area ($608.8 \cdot 10^3 \text{ km}^2$ per year) (Table 4-2).

4.6. Discussion

This Chapter estimated the potential omission error of a global burned area product informed by observations on Landsat 7 overpass days, compared to the MCD64A1 MODIS global burned area product. The Landsat geometry of acquisition and 16-days revisit calendar was used to establish the location and timing of the overpasses, the MODIS MCD64A1 monthly burned area product and daily MODIS MOD35-L7 cloud mask from January 2002 to December 2016 were used to identify burned and cloudy pixels, respectively.

Globally, 81% of the total MCD64A1 burned area analyzed had at least one cloud-free Landsat acquisition within the persistence time of burned areas, hence the global omission error was estimated as 19% of the total MCD64A1 average annual burned area detected. The maximum error was estimated for burned areas detected over Forest (33%) and the minimum over Shrublands (5.3%).

The burned area omission error was above 50% for ecoregions in equatorial Amazon forest, Western Central Africa, and Indonesia where the cloud occurrence and persistence, combined with a rapid post-fire vegetation regrowth, affect also the burned area detection capability of the MODIS product (Giglio et al. 2009), informed by daily observations. The high cloudiness of these regions also influences the availability of cloud-free observations limiting the potential ability of Landsat land monitoring applications requiring more than one cloud-free acquisition per year (Kovalskyy and Roy 2013).

4.6.1. Realm/Biome scale

For the forest land cover, the majority of the MCD64A1 burned area in Tropical biomes was detected in the Afrotropic, Indo-Malay and Neotropic realms. Anthropogenic activities account for the majority of burning events in the tropical Amazon rainforest (Cochrane et al. 1999; Kumar et al. 2014), African rainforests (Bucini and Lambin 2002; Eva and Lambin 2000; Tovar et al. 2014) and Indo-Australian rainforests (Hope et al. 2004). Due to the anthropogenic nature, these fires have larger omission error rates, due to a combined effect of post-fire treatments (e.g. land clearing, tilling) and cloud cover (Cochrane 2003), even for the MCD64A1 product informed by daily observations.

The majority of the forest burned area in Temperate biomes was detected in Australasia, Neartic and Palearctic realms.

In the Neoarctic realm, the highest omission errors were found in Alaska and the southeast US, which is a good example for the combination of short persistence of burned area spectral signal and cloud cover effect. In the southeast US, the omission error is likely linked to the short persistence time estimate found in the southeast US, which is consistent with frequent surface fires and rapid post-fire grass regeneration observed in that region (Glitzenstein et al. 1995; Slocum et al. 2003). In Alaska, early snowfall immediately after the fire season can cover the burned land which remains undetectable until snowmelt.

The large differences found in the omission error for the Neoarctic and Paleoarctic boreal forest is likely due to the different burning conditions in these two realms, which also influenced the length of the persistence time of the spectral signal (Melchiorre and Boschetti 2018). Fires in Paleoarctic boreal forests are dominated by low intensity surface fires and typically result in smaller sized fires (Li et al. 2012) than Neoarctic boreal forests fires, where high intensity crown fires are predominant instead (de Groot et al. 2013; Wierzchowski et al. 2002).

For the Shrubland land cover, the majority of the MCD64A1 burned area was detected in the Tropical and Desert/Xeric biomes of the Australasia and Afrotropic realm. In Australasia, Shrubland Tropical fires are typically characterized by a shorter persistence of char on the ground due to rapid regrowth of vegetation and due to removal by wind and atmospheric agents (Pereira 2003) which, in combination with cloud cover effect, limited the number of cloud-free observations and the resulting omission error is larger compared to Desert/Xeric fires.

The great majority of global MCD64A1 burned area in Grassland & Savanna land cover was detected in the Afrotropic realm. Cloud cover limits the number of available post-fire observations to 1 or 2 for 67% of the total MCD64A1 burned area whereas the omission error is 19%. The persistence time for the Afrotropic Grassland and Savanna had small variability (Melchiorre et al. 2018) and the omission error differences were mostly due to the different cloud cover occurrences. The highest omission error was estimated in the Afrotropic equatorial regions (>30% of the total MCD64A1 annual average burned area) and it was reduced for Tropical and subtropical regions.

4.7. Conclusions

In this work, we estimated the omission error of a hypothetical Landsat global burned area product due to the combined effect of the impermanent spectral signal associated with burned area and cloud cover. The omission error was estimated as the average annual burned area detected by the MODIS MCD64A1 product with no post-fire cloud-free observations within the burned area persistence time, estimated in Chapter 2. MCD64A1 burned pixels were considered detected by the

hypothetical Landsat 7 if at least one post-fire cloud-free observation was available, hence the omission error estimated is due solely to the limited availability of data caused by the combined effect of the lower revisit frequency of Landsat and cloud cover. Globally, the omission error was 19% of the MCD64A1 average annual burned area (33.4% in Forest, 5.4% in Shrublands and 19.4% in Grassland & Savanna).

The revisit frequency of moderate resolution satellites can be virtually increased using data fusion of different sensors (Li and Roy 2017). Thanks to the successful launches of Landsat 8 (Roy et al. 2014), Sentinel 2A and Sentinel 2B (Drusch et al. 2012) and the planned launch of Landsat 9 in 2020, the number of operative moderate resolution satellites designed for terrestrial monitoring is increasing. The techniques necessary to use time series of harmonized data from different virtual constellations of satellites are constantly improving (Claverie et al. 2018; Franch et al. 2019), in particular for the design of global burned area detection algorithms (Roy et al. 2019).

Virtual constellations are defined by the Committee on Earth Observation Satellites (CEOS) as a “set of space and ground segment capabilities that operate in a coordinated manner to meet a combined and common set of Earth Observation requirements” and, when they are composed by sensors with similar spatial and spectral characteristics, they are able to mitigate limitations of any one particular sensor (Wulder et al. 2015). A possibility for future research would be to estimate the benefit, in terms of reduction of the expected burned area omission errors, of using data from different virtual constellations of existing and new sun-synchronous, near-polar orbiting satellites.

4.8. *References*

- Ackerman, S., Strabala, K., Menzel, P., Frey, R., Moeller, C., & Gumley, L. (2010). Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35. In, *MODIS Cloud Mask Team, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin: Citeseer*
- Alonso-Canas, I., & Chuvieco, E. (2015). Global burned area mapping from ENVISAT-MERIS and MODIS active fire data. *Remote Sensing of Environment*, 163, 140-152
- Bastarrika, A., Chuvieco, E., & 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, 1003-1012
- Boschetti, L., Roy, D.P., Justice, C.O., & Humber, M.L. (2015). MODIS–Landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, 161, 27-42

- Bowman, D.M., Balch, J.K., Artaxo, P., Bond, W.J., Carlson, J.M., Cochrane, M.A., D'Antonio, C.M., DeFries, R.S., Doyle, J.C., & Harrison, S.P. (2009). Fire in the Earth system. *Science*, 324, 481-484
- Bowman, D.M.J.S., Perry, G.L.W., & Marston, J.B. (2015). Feedbacks and landscape-level vegetation dynamics. *Trends in Ecology & Evolution*, 30, 255-260
- Bowman, D.M.J.S., Zhang, Y., Walsh, A., & Williams, R.J. (2003). Experimental comparison of four remote sensing techniques to map tropical savanna fire-scars using Landsat-TM imagery. *International Journal of Wildland Fire*, 12, 341-348
- Bucini, G., & Lambin, E.F. (2002). Fire impacts on vegetation in Central Africa: a remote-sensing-based statistical analysis. *Applied Geography*, 22, 27-48
- Chen, X., Vogelmann, J.E., Rollins, M., Ohlen, D., Key, C.H., Yang, L., Huang, C., & Shi, H. (2011). Detecting post-fire burn severity and vegetation recovery using multitemporal remote sensing spectral indices and field-collected composite burn index data in a ponderosa pine forest. *International Journal of Remote Sensing*, 32, 7905-7927
- Chu, T., & Guo, X. (2014). Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. *Remote Sensing*, 6, 470-520
- Chuvieco, E., & Congalton, R.G. (1988). Mapping and inventory of forest fires from digital processing of TM data. *Geocarto International*, 3, 41-53
- Chuvieco, E., Martin, M.P., & Palacios, A. (2002). Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, 23, 5103-5110
- Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanase, M., Koutsias, N., García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., & Giglio, L. (2019). Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sensing of Environment*, 225, 45-64
- Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145-161
- Cochrane, M.A. (2003). Fire science for rainforests. *Nature*, 421, 913-919
- Cochrane, M.A., Alencar, A., Schulze, M.D., Souza, C.M., Nepstad, D.C., Lefebvre, P., & Davidson, E.A. (1999). Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests. *Science*, 284, 1832-1835

- de Groot, W.J., Cantin, A.S., Flannigan, M.D., Soja, A.J., Gowman, L.M., & Newbery, A. (2013). A comparison of Canadian and Russian boreal forest fire regimes. *Forest Ecology and Management*, 294, 23-34
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., & Martimort, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25-36
- Eva, H., & Lambin, E.F. (2000). Fires and land-cover change in the tropics: a remote sensing analysis at the landscape scale. *Journal of Biogeography*, 27, 765-776
- Franch, B., Vermote, E., Skakun, S., Roger, J.-C., Masek, J., Ju, J., Villaescusa-Nadal, J.L., & Santamaria-Artigas, A. (2019). A Method for Landsat and Sentinel 2 (HLS) BRDF Normalization. *Remote Sensing*, 11, 632
- Fraser, R., Li, Z., & Cihlar, J. (2000). Hotspot and NDVI differencing synergy (HANDS): A new technique for burned area mapping over boreal forest. *Remote Sensing of Environment*, 74, 362-376
- Frederiksen, P., Langaas, S., & Mbaye, M. (1990). NOAA-AVHRR and GIS-based monitoring of fire activity in Senegal—a provisional methodology and potential applications. *Fire in the Tropical Biota* (pp. 400-417): Springer
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114, 168-182
- Giglio, L., Boschetti, L., Roy, D.P., Humber, M.L., & Justice, C.O. (2018). The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment*, 217, 72-85
- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., & Justice, C.O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, 113, 408-420
- Glitzenstein, J.S., Platt, W.J., & Streng, D.R. (1995). Effects of Fire Regime and Habitat on Tree Dynamics in North Florida Longleaf Pine Savannas. *Ecological monographs*, 65, 441-476
- Goetz, S.J., Bond-Lamberty, B., Law, B.E., Hicke, J.A., Huang, C., Houghton, R.A., McNulty, S., O'Halloran, T., Harmon, M., Meddens, A.J.H., Pfeifer, E.M., Mildrexler, D., & Kasischke, E.S. (2012). Observations and assessment of forest carbon dynamics following disturbance in North America. *Journal of Geophysical Research: Biogeosciences*, 117, n/a-n/a
- Hawbaker, T.J., Vanderhoof, M.K., Beal, Y.-J., Takacs, J.D., Schmidt, G.L., Falgout, J.T., Williams, B., Fairaux, N.M., Caldwell, M.K., Picotte, J.J., Howard, S.M., Stitt, S., & Dwyer, J.L. (2017). Mapping burned areas using dense time-series of Landsat data. *Remote Sensing of Environment*, 198, 504-522

- Hope, G., Kershaw, A.P., Kaars, S.v.d., Xiangjun, S., Liew, P.-M., Heusser, L.E., Takahara, H., McGlone, M., Miyoshi, N., & Moss, P.T. (2004). History of vegetation and habitat change in the Austral-Asian region. *Quaternary International*, 118–119, 103-126
- Hudak, A., & Brockett, B. (2004). Mapping fire scars in a southern African savannah using Landsat imagery. *International Journal of Remote Sensing*, 25, 3231-3243
- Hyer, E.J., & Reid, J.S. (2009). Baseline uncertainties in biomass burning emission models resulting from spatial error in satellite active fire location data. *Geophysical Research Letters*, 36
- Jin, Y., & Roy, D. (2005). Fire-induced albedo change and its radiative forcing at the surface in northern Australia. *Geophysical Research Letters*, 32
- Ju, J., & Roy, D.P. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment*, 112, 1196-1211
- Kessler, P.D., Killough, B.D., Gowda, S., Williams, B.R., Chander, G., & MIN, Q. (2013). CEOS Visualization Environment (COVE) Tool for Intercalibration of Satellite Instruments. *IEEE transactions on geoscience and remote sensing*, 51, 1081-1087
- Koutsias, N., & Karteris, M. (1998). Logistic regression modelling of multitemporal Thematic Mapper data for burned area mapping. *International Journal of Remote Sensing*, 19, 3499-3514
- Koutsias, N., & Karteris, M. (2000). Burned area mapping using logistic regression modeling of a single post-fire Landsat-5 Thematic Mapper image. *International Journal of Remote Sensing*, 21, 673-687
- Kovalskyy, V., & Roy, D.P. (2013). The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*, 130, 280-293
- Kumar, S.S., Roy, D.P., Cochrane, M.A., Souza, C.M., Barber, C.P., & Boschetti, L. (2014). A quantitative study of the proximity of satellite detected active fires to roads and rivers in the Brazilian tropical moist forest biome. *International Journal of Wildland Fire*, 23, 532-543
- Langaas, S., & Kane, R. (1991). Temporal spectral signatures of fire scars in Savanna Woodland. In, *Geoscience and Remote Sensing Symposium, 1991. IGARSS '91. Remote Sensing: Global Monitoring for Earth Management., International* (pp. 1157-1160)
- Langmann, B., Duncan, B., Textor, C., Trentmann, J., & van der Werf, G.R. (2009). Vegetation fire emissions and their impact on air pollution and climate. *Atmospheric Environment*, 43, 107-116
- Laris, P.S. (2005). Spatiotemporal problems with detecting and mapping mosaic fire regimes with coarse-resolution satellite data in savanna environments. *Remote Sensing of Environment*, 99, 412-424

- Li, J., & Roy, D. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, *9*, 902
- Li, M.S., Mao, L.J., Lu, C., & Wang, Y. (2012). Temporal Change in Fragmentation in China's Primary Forest Ecoregions. In, *Advanced Materials Research* (pp. 1286-1296): Trans Tech Publ
- Loboda, T.V., Hoy, E.E., Giglio, L., & Kasischke, E.S. (2011). Mapping burned area in Alaska using MODIS data: a data limitations-driven modification to the regional burned area algorithm. *International Journal of Wildland Fire*, *20*, 487-496
- Masuoka, E., Fleig, A., Wolfe, R.E., & Patt, F. (1998). Key characteristics of MODIS data products. *IEEE transactions on geoscience and remote sensing*, *36*, 1313-1323
- Melchiorre, A., & Boschetti, L. (2018). Global Analysis of Burned Area Persistence Time with MODIS Data. *Remote Sensing*, *10*, 750
- Melchiorre, A., Boschetti, L., & Roy, D.P. (2020). Global Evaluation of the Suitability of MODIS-Terra Detected Cloud Cover as a Proxy for Landsat 7 Cloud Conditions. *Remote Sensing*, *12*, 202
- Mercury, M., Green, R., Hook, S., Oaida, B., Wu, W., Gunderson, A., & Chodas, M. (2012). Global cloud cover for assessment of optical satellite observation opportunities: A HypsIRI case study. *Remote Sensing of Environment*, *126*, 62-71
- Miller, J.D., & Yool, S.R. (2002). Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. *Remote Sensing of Environment*, *82*, 481-496
- Mitri, G., & Gitas, I. (2004). A semi-automated object-oriented model for burned area mapping in the Mediterranean region using Landsat-TM imagery. *International Journal of Wildland Fire*, *13*, 367-376
- Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., & Chuvieco, E. (2014). Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 64-79
- Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V., Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E., & Morrison, J.C. (2001). Terrestrial Ecoregions of the World: A New Map of Life on Earth A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience*, *51*, 933-938
- Pereira, J.M. (2003). Remote sensing of burned areas in tropical savannas. *International Journal of Wildland Fire*, *12*, 259-270
- Picotte, J.J., & Robertson, K. (2011). Timing constraints on remote sensing of wildland fire burned area in the southeastern US. *Remote Sensing*, *3*, 1680-1690

- Randerson, J.T., Chen, Y., van der Werf, G.R., Rogers, B.M., & Morton, D.C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, *117*, G04012
- Randerson, J.T., Liu, H., Flanner, M.G., Chambers, S.D., Jin, Y., Hess, P.G., Pfister, G., Mack, M.C., Treseder, K.K., Welp, L.R., Chapin, F.S., Harden, J.W., Goulden, M.L., Lyons, E., Neff, J.C., Schuur, E.A.G., & Zender, C.S. (2006). The Impact of Boreal Forest Fire on Climate Warming. *Science*, *314*, 1130-1132
- Röder, A., Hill, J., Duguay, B., Alloza, J.A., & Vallejo, R. (2008). Using long time series of Landsat data to monitor fire events and post-fire dynamics and identify driving factors. A case study in the Ayora region (eastern Spain). *Remote Sensing of Environment*, *112*, 259-273
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., & Chuvieco, E. (2019). Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, *222*, 1-17
- Roy, D., Jin, Y., Lewis, P., & Justice, C. (2005). Prototyping a global algorithm for systematic fire-affected area mapping using MODIS time series data. *Remote Sensing of Environment*, *97*, 137-162
- Roy, D.P., Boschetti, L., Justice, C.O., & Ju, J. (2008). The collection 5 MODIS burned area product—Global evaluation by comparison with the MODIS active fire product. *Remote Sensing of Environment*, *112*, 3690-3707
- Roy, D.P., Huang, H., Boschetti, L., Giglio, L., Yan, L., Zhang, H.H., & Li, Z. (2019). Landsat-8 and Sentinel-2 burned area mapping - A combined sensor multi-temporal change detection approach. *Remote Sensing of Environment*, *231*, 111254
- Roy, D.P., Wulder, M.A., Loveland, T.R., C.E. W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., McCorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., & Zhu, Z. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154-172
- Silva, J.M., Sá, A.C., & Pereira, J.M. (2005). Comparison of burned area estimates derived from SPOT-VEGETATION and Landsat ETM+ data in Africa: Influence of spatial pattern and vegetation type. *Remote Sensing of Environment*, *96*, 188-201

- Simon, M., Plummer, S., Fierens, F., Hoelzemann, J., & Arino, O. (2004). Burnt area detection at global scale using ATSR-2: The GLOBSCAR products and their qualification. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *109*
- Slocum, M.G., Platt, W.J., & Cooley, H.C. (2003). Effects of Differences in Prescribed Fire Regimes on Patchiness and Intensity of Fires in Subtropical Savannas of Everglades National Park, Florida. *Restoration Ecology*, *11*, 91-102
- Smith, A.M., & Wooster, M.J. (2005). Remote classification of head and backfire types from MODIS fire radiative power and smoke plume observations. *International Journal of Wildland Fire*, *14*, 249-254
- Smith, A.M.S., Drake, N.A., Wooster, M.J., Hudak, A.T., Holden, Z.A., & Gibbons, C.J. (2007). Production of Landsat ETM+ reference imagery of burned areas within Southern African savannahs: comparison of methods and application to MODIS. *International Journal of Remote Sensing*, *28*, 2753-2775
- Stroppiana, D., Boschetti, M., Zaffaroni, P., & Brivio, P.A. (2009). Analysis and Interpretation of Spectral Indices for Soft Multicriteria Burned-Area Mapping in Mediterranean Regions. *Geoscience and Remote Sensing Letters, IEEE*, *6*, 499-503
- Tansey, K., Grégoire, J.-M., Stroppiana, D., Sousa, A., Silva, J., Pereira, J.M.C., Boschetti, L., Maggi, M., Brivio, P.A., Fraser, R., Flasse, S., Ershov, D., Binaghi, E., Graetz, D., & Peduzzi, P. (2004). Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data. *Journal of Geophysical Research: Atmospheres*, *109*, D14S03
- Tansey, K., Gregoire, J.M., Defourny, P., Leigh, R., Pekel, J.F.O., van Bogaert, E., & Bartholome, E. (2008). A new, global, multi-annual (2000-2007) burnt area product at 1 km resolution. *Geophysical Research Letters*, *35*, 6
- Thonicke, K., Venevsky, S., Sitch, S., & Cramer, W. (2001). The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model. *Global Ecology and Biogeography*, *10*, 661-677
- Tovar, C., Breman, E., Brncic, T., Harris, D.J., Bailey, R., & Willis, K.J. (2014). Influence of 1100 years of burning on the central African rainforest. *Ecography*, *37*, 1139-1148
- Trigg, S., & Flasse, S. (2000). Characterizing the spectral-temporal response of burned savannah using in situ spectroradiometry and infrared thermometry. *International Journal of Remote Sensing*, *21*, 3161-3168
- van der Werf, G.R., Randerson, J.T., Giglio, L., van Leeuwen, T.T., Chen, Y., Rogers, B.M., Mu, M., van Marle, M.J., Morton, D.C., & Collatz, G.J. (2017). Global fire emissions estimates during 1997–2016. *Earth System Science Data*, *9*, 697

- Wierzchowski, J., Heathcott, M., & Flannigan, M.D. (2002). Lightning and lightning fire, central cordillera, Canada. *International Journal of Wildland Fire*, *11*, 41-51
- Wolfe, R.E., Nishihama, M., Fleig, A.J., Kuyper, J.A., Roy, D.P., Storey, J.C., & Patt, F.S. (2002). Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of Environment*, *83*, 31-49
- Wolfe, R.E., Roy, D.P., & Vermote, E. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *Geoscience and Remote Sensing, IEEE Transactions on*, *36*, 1324-1338
- Wulder, M.A., Hilker, T., White, J.C., Coops, N.C., Masek, J.G., Pflugmacher, D., & Crevier, Y. (2015). Virtual constellations for global terrestrial monitoring. *Remote Sensing of Environment*, *170*, 62-76

Chapter 5: Conclusion and Future Work

5.1. Summary and conclusion

My dissertation focused on estimating the combined effect of the impermanent nature of burned areas, revisit frequency and cloud cover on global burned area mapping using Landsat data. In particular, I first analyzed the temporal persistence of the spectral signal associated with burned areas, which provides an estimation of the period after the burning date in which burned areas can be mapped reliably. Using the results of the persistence time analysis, I then estimated the potential omission error of a hypothetical global Landsat burned area product due to low revisit frequency and cloud cover, performing a simulation where daily MODIS data are combined with the Landsat acquisition calendar and observation geometry. The findings of this simulation highlighted the locations in which the low revisit frequency of Landsat, combined with the reduced data availability due to cloud cover, is not sufficient for accurate global burned area mapping.

Chapter 2 presented a global estimation of the burned area persistence time, defined as the duration of the spectral separability of the burned / unburned areas mapped by the MODIS global burned area product. The separability was computed by analyzing time series of Normalized Burned Ratio (NBR) derived from nadir BRDF-adjusted MODIS reflectances. Results showed that, globally, the median burned area persistence time was 29 days and 86.6% of the MODIS global area can be detected accurately only for up to 48 days. The results, therefore, indicate that the persistence time can be a limiting factor for mapping burned areas using moderate resolution satellite sensors, which have a low temporal resolution (e.g. Landsat 16 days, Sentinel 2A and 2B 10 days each). The analysis highlighted the importance of the burning date on burned area mapping: early and late fires had lower persistence time compared to fires burning in the central portion of the fire season. The effect was prevalent in ecoregions with clearly defined and alternated dry and wet seasons, or with persistent snow cover in the winter months. Persistent clouds and snow can cover the burned area spectral signature for long periods, reducing the accuracy of burned area mapping algorithms based on change detection techniques over rolling periods of time.

Chapter 3 presented a global intercomparison of MODIS Terra and Landsat 7 Enhanced Thematic Mapper (ETM+) cloud detections. The results showed high correlation between MODIS and Landsat cloud fractions ($R^2 = 0.83$), negligible bias (median difference: $< 1\%$) and low dispersion around the median (inter-quartile range: $[-2\%, 6\%]$). The results indicate that the cloud cover detected by MODIS data can be used as a proxy for Landsat 7 cloud cover, globally at the Landsat WRS scale.

The geographic analysis of the results revealed that the largest differences between the MODIS (MOD35) and Landsat cloud fraction occur along the coastlines, and at high latitude. In the case of coastal areas, further analysis indicated that the errors are mostly observed at WRS path/row locations with land fractions ≤ 0.05 . In these cases, cloud movement in the 30 minutes between the MODIS-Terra and Landsat 7 overpasses may result in large discrepancies of the estimated cloud percentage over land. At high latitudes and over mountain ranges, ice and snow were flagged as clouds by the MODIS MOD35 product introducing large discrepancies with the Landsat cloud cover fraction.

Finally, Chapter 4 presented the potential omission error of a hypothetical Landsat global burned area product estimated as the amount of burned area detected by MODIS that would not be detected by Landsat 7 because of the combined effect of the impermanent spectral signal associated with burned area and missing observations due to cloud cover. Globally, the resulting omission error was estimated in 19% of the average annual burned area detected by MODIS, with a maximum error over forest land cover (33%) and minimum over shrubland land cover (5%). The potential omission error should be considered conservative for the design of a global burned area product using Landsat 7 acquisitions after 2003 because of the Scan Line Corrector (SLC) failure that resulted in a systematic failure to sense 22% of each image (Markham et al. 2004). The results were derived using the acquisition calendar of Landsat 7 only; however, thanks to the aggregation of data from over 15 years of acquisitions, the results can be extended to the other existing Landsat sensors, which are positioned on the same Landsat 7 orbit shifted by an 8-days lag, and also to Landsat 9, which is planned to be launched in the same orbit as Landsat 7 (Wulder et al. 2019) by Spring 2021.

5.2. Limitations and future research

The magnitude of the potential omission error estimated for a Landsat global burned area product suggested that the revisit frequency of Landsat is too low to generate accurate global burned area maps. However, the potential omission error analysis, described in Chapter 4, used as input the burned area persistence time estimated in Chapter 2, which has some caveats and limitations. Burned area persistence time can be limited by the recovery of NBR values to pre-fire levels after rapid re-vegetation or other factors, such as snowfall, that occlude the charcoal and bare soil masking the spectral signature associated with burned areas. However, the analysis was performed comparing the post-fire spectral signal temporal trajectories to unburned pixel trajectories within the same ecoregion and land cover assuming they had similar characteristics of burned pixels. Pre-fire fuel conditions can be different from unburned areas: different vegetation types, fuel moisture, and load can alter temporal

trends of burned and unburned pixels (Archibald et al. 2018) potentially with the effect of reducing the estimated persistence time.

The persistence time was estimated using MODIS data at 500 m spatial resolution and the effect of partially burned pixels, understory burns, land-snow patterns, and vegetation phenology can reduce the estimated persistence time (Chu and Guo 2014). Higher spatial resolution data have the benefits of increasing the detections of burned area (Roteta et al. 2019; Roy et al. 2019), reducing the occurrence of partially burned pixels and increasing the spectral separation of burned / unburned areas (Huang et al. 2016); therefore, it is reasonable to expect longer estimates of burned area persistence computed using moderate resolution data.

The persistence time analysis did not account for further post-fire land cover and land use changes. Anthropogenic activities account for the majority of burning events in tropical rainforests (Cochrane et al. 1999; Eva and Lambin 2000; Hope et al. 2004) and human-ignited rainforest fires are strongly associated with further land cover changes (Bucini and Lambin 2002). The persistence time estimates are influenced by other post-fire activities that alter burned areas spectral signatures. For example, in Amazonia fire activity is connected to other disturbances having spectral similarities to burned areas such as logging, deforestation, and land conversion which only complex algorithms can discriminate (Morton et al. 2011).

Cropland land cover was not considered in the persistence time analysis since the MODIS burned area product is not suited to capture the size and heterogeneity of cropland burned area (Giglio et al. 2009; Lasko et al. 2017) because of its spatial resolution, and because other agricultural practices (e.g. tilling) often follow immediately (Hall et al. 2016) and have similar spectral signatures (Vanderhoof et al. 2017). However, croplands and other managed land were considered in the analysis if not correctly mapped in the MODIS land cover product and the inclusion of fires over these land cover could have had a shortening effect on the estimated persistence time in ecoregions where agricultural fires are predominant.

Finally, both omission and commission errors in the MCD64A1 product, used to separate the population of burned and unburned pixels, caused the inclusion of post-fire temporal trajectories in the unburned class and the inclusion of unburned trajectories in the burned class with a shortening effect on the estimated persistence time.

The estimation of the potential burned area omission error with moderate resolution sensors could be expanded to other sensors used independently and in virtual constellations, which are defined by the Committee on Earth Observation Satellites (CEOS) as a “set of space and ground segment

capabilities that operate in a coordinated manner to meet a combined and common set of Earth Observation requirements”, since virtual constellations having similar spatial and spectral characteristics are able to mitigate the limitations of any one particular sensor (Wulder et al. 2015). It should be noted that the approach described in Chapter 4 exploits the fact that Terra trails Landsat 7 on the same orbit with 30 minutes lag. Landsat 7 ETM+ ground swaths always coincide with the central nadir-looking part of MODIS-Terra ground swaths, and the time lag between MODIS-Terra and Landsat 7 ETM+ acquisitions is constant and sufficiently small to approximate the cloud cover affecting Landsat 7 data using MODIS-Terra cloud mask data. This approximation is not realistic for other sensors with variable and longer overpass lag with MODIS-Terra.

Future research should establish a statistical framework for the analysis of cloud-free data availability to virtual constellations of moderate resolution (10 – 30 m) sensors within the burned area persistence time as calculated in Chapter 2. The goal would be to estimate the benefit, in terms of reduction of expected omission errors, of using data from different virtual constellations of existing (e.g. Landsat 8, Sentinel 2A/B) and planned (e.g., Landsat 9) sun-synchronous, near-polar orbiting satellites. The effect of cloud cover on data availability for mapping burned areas using data from different constellations of moderate resolution satellites can be assessed using statistical metrics at a lower spatial resolution, similar to the metrics proposed for defining the revisit frequency requirements for global agricultural monitoring (Whitcraft et al. 2015a; Whitcraft et al. 2015b). However, due to cloud cover detections variability with sensors’ overpass time and geometry of acquisitions, estimates of the burned area differences with global products will present bias and uncertainties difficult to estimate accurately.

When the orbits and acquisition calendars of other satellites (e.g., Landsat 8, Sentinel 2A and 2B) are compared to the MODIS Terra overpasses, the time lag is not constant, and, in extreme cases, it is longer than 1 day. The reduced number of cloud free observations for moderate resolution sensors can be estimated using cloud probability derived from MODIS observations (Whitcraft et al. 2015b); however, this approach requires additional assumptions on cloud detections (e.g., spatial and temporal independence of cloud occurrences (Roy et al. 2006)) and might produce inaccurate results due to several factors such as the dynamics of clouds. Statistical analysis of moderate resolution sensors cloud cover using MODIS data should take into account that clouds move, evolve and, on average, cloud cover increases during the day over land (Feidas and Cartalis 2005; King et al. 2013; Li et al. 2019). Additionally, satellite cloud detections increase with the off-nadir sensor view angle (King et al. 2013; Maddux et al. 2010), hence the statistical analysis of clouds have reduced accuracy if the overpass time

and the observation geometry (i.e., sun and sensor zenith and azimuth angles) of MODIS and the simulated moderate resolution sensor are not similar.

Also, satellite cloud detections vary depending on the sensor's spatial resolution. MODIS observations, having a coarser spatial resolution, fail to detect small and sparse clouds at low cloud levels and miss small gaps in cloud cover at high cloud levels, with an overall overestimation of cloud detections compared to the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (Zhao and Di Girolamo 2006) and Landsat 7 (Melchiorre et al. 2020). The MODIS cloud detection overestimation can also be partly explained by false positive detections in the MODIS cloud mask over snow, coastlines, rivers and lakes regions, known issues of the MODIS cloud product (Ackerman et al. 2010; Ault et al. 2006). Finally, multispectral bands available and different spectral resolution can result in different cloud detections (Nakajima et al. 2011; Zhao and Di Girolamo 2006). MODIS cloud masking algorithm uses multiple thermal bands and the water vapor/CO₂ absorption bands, typically not available on moderate resolution sensors. For example, these bands are missing from the Landsat 7 instrument, resulting in low sensitivity to optically thin clouds (Zhu and Woodcock 2012). In general, cloud detection algorithm differences may result in different cloud masks at coarse (Holz et al. 2008) and at moderate spatial resolution (Foga et al. 2017)

Frequent acquisitions of geostationary satellites could be used in future works to estimate the expected occurrence and variability of cloud cover at the overpass time of specific moderate resolution sensors. Cloud cover detections from geostationary satellites have been historically used to build the International Satellite Cloud Cover Project (ISCCP) dataset (Rossow and Schiffer 1999; Schiffer and Rossow 1985) and their data are used in cloud cover forecast with good accuracy (Escrig et al. 2013); however, cloud detection differences due to the different geometry of acquisition, bands availability and spatial / radiometric resolution should be addressed. For example, different cloud fractions were detected between MODIS and the ISCCP, especially over optically thin clouds (Pincus et al. 2012) and between MODIS and the geostationary Advanced Himawari Imager (AHI) over the Tibetan plateau (Shang et al. 2018).

In the future, when the moderate spatial resolution sensors period of activity will allow robust multi-year analysis globally, it will be possible to estimate burned area persistence times and perform statistical analysis of cloud-free data availability using moderate resolution data. Thanks to the open policy of the ESA/EU Copernicus Sentinels (Berger et al. 2012) and USGS Landsat (Zhu et al. 2019) data, global moderate resolution burned area mapping algorithms are being developed using different optical multi-spectral sensors (Roteta et al. 2019; Roy et al. 2019) and can include radar data

acquisitions to improve detection accuracy in areas with persistent cloud cover (Belenguer-Plomer et al. 2019; Tanase et al. 2020).

The results presented and discussed in this work have implications on the length of the rolling periods, used in change detection algorithms to detect burned areas, and identified locations where the occurrence of clouds can have degrading effects on burned areas mapping accuracy. While burned area maps generated from moderate resolution sensors data have been traditionally used for local applications (Chuvienco et al. 2019; Mouillot et al. 2014) or for the validation of global coarse resolution products (Boschetti et al. 2016; Padilla et al. 2014; Roy and Boschetti 2009), global moderate resolution burned area products will be generated in the foreseeable future. This task is facilitated by the harmonization of the data streams generated by the NASA/USGS Landsat program and by the ESA/EU Copernicus missions (Claverie et al. 2018), leading to a virtual constellation of moderate resolution satellites having less than 3 days median average revisit time (Li and Roy 2017). The first prototype algorithms exploiting the virtual constellation of Landsat 8 and Sentinel 2 for global burned area mapping are under development (Roy et al. 2019); once the products are available to the scientific community, they will enable improved estimates of global burned areas to all the fire information user.

5.3. References

- Ackerman, S., Strabala, K., Menzel, P., Frey, R., Moeller, C., & Gumley, L. (2010). Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35. In, *MODIS Cloud Mask Team, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin*: Citeseer
- Archibald, S., Lehmann, C.E., Belcher, C.M., Bond, W.J., Bradstock, R.A., Daniau, A., Dexter, K., Forrester, E., Greve, M., & He, T. (2018). Biological and geophysical feedbacks with fire in the Earth system. *Environmental Research Letters*, *13*, 033003
- Ault, T.W., Czajkowski, K.P., Benko, T., Coss, J., Struble, J., Spongberg, A., Templin, M., & Gross, C. (2006). Validation of the MODIS snow product and cloud mask using student and NWS cooperative station observations in the Lower Great Lakes Region. *Remote Sensing of Environment*, *105*, 341-353
- Belenguer-Plomer, M.A., Tanase, M.A., Fernandez-Carrillo, A., & Chuvienco, E. (2019). Burned area detection and mapping using Sentinel-1 backscatter coefficient and thermal anomalies. *Remote Sensing of Environment*, *233*, 111345
- Berger, M., Moreno, J., Johannessen, J.A., Levelt, P.F., & Hanssen, R.F. (2012). ESA's sentinel missions in support of Earth system science. *Remote Sensing of Environment*, *120*, 84-90

- Boschetti, L., Stehman, S.V., & Roy, D.P. (2016). A stratified random sampling design in space and time for regional to global scale burned area product validation. *Remote Sensing of Environment*, 186, 465-478
- Bucini, G., & Lambin, E.F. (2002). Fire impacts on vegetation in Central Africa: a remote-sensing-based statistical analysis. *Applied Geography*, 22, 27-48
- Chu, T., & Guo, X. (2014). Remote sensing techniques in monitoring post-fire effects and patterns of forest recovery in boreal forest regions: a review. *Remote Sensing*, 6, 470-520
- Chuvieco, E., Mouillot, F., van der Werf, G.R., San Miguel, J., Tanase, M., Koutsias, N., García, M., Yebra, M., Padilla, M., Gitas, I., Heil, A., Hawbaker, T.J., & Giglio, L. (2019). Historical background and current developments for mapping burned area from satellite Earth observation. *Remote Sensing of Environment*, 225, 45-64
- Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, 219, 145-161
- Cochrane, M.A., Alencar, A., Schulze, M.D., Souza, C.M., Nepstad, D.C., Lefebvre, P., & Davidson, E.A. (1999). Positive Feedbacks in the Fire Dynamic of Closed Canopy Tropical Forests. *Science*, 284, 1832-1835
- Escrig, H., Batlles, F.J., Alonso, J., Baena, F.M., Bosch, J.L., Salbidegoitia, I.B., & Burgaleta, J.I. (2013). Cloud detection, classification and motion estimation using geostationary satellite imagery for cloud cover forecast. *Energy*, 55, 853-859
- Eva, H., & Lambin, E.F. (2000). Fires and land-cover change in the tropics: a remote sensing analysis at the landscape scale. *Journal of Biogeography*, 27, 765-776
- Feidas, H., & Cartalis, C. (2005). Application of an automated cloud-tracking algorithm on satellite imagery for tracking and monitoring small mesoscale convective cloud systems. *International Journal of Remote Sensing*, 26, 1677-1698
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dille, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., & Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment*, 194, 379-390
- Giglio, L., Loboda, T., Roy, D.P., Quayle, B., & Justice, C.O. (2009). An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sensing of Environment*, 113, 408-420
- Hall, J.V., Loboda, T.V., Giglio, L., & McCarty, G.W. (2016). A MODIS-based burned area assessment for Russian croplands: Mapping requirements and challenges. *Remote Sensing of Environment*, 184, 506-521

- Holz, R., Ackerman, S., Nagle, F., Frey, R., Dutcher, S., Kuehn, R., Vaughan, M., & Baum, B. (2008). Global Moderate Resolution Imaging Spectroradiometer (MODIS) cloud detection and height evaluation using CALIOP. *Journal of Geophysical Research: Atmospheres*, *113*
- Hope, G., Kershaw, A.P., Kaars, S.v.d., Xiangjun, S., Liew, P.-M., Heusser, L.E., Takahara, H., McGlone, M., Miyoshi, N., & Moss, P.T. (2004). History of vegetation and habitat change in the Austral-Asian region. *Quaternary International*, *118–119*, 103-126
- Huang, H., Roy, D., Boschetti, L., Zhang, H., Yan, L., Kumar, S., Gomez-Dans, J., & Li, J. (2016). Separability Analysis of Sentinel-2A Multi-Spectral Instrument (MSI) Data for Burned Area Discrimination. *Remote Sensing*, *8*, 873
- King, M.D., Platnick, S., Menzel, W.P., Ackerman, S.A., & Hubanks, P.A. (2013). Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites. *IEEE transactions on geoscience and remote sensing*, *51*, 3826-3852
- Lasko, K., Vadrevu, K.P., Tran, V.T., Ellicott, E., Nguyen, T.T., Bui, H.Q., & Justice, C. (2017). Satellites may underestimate rice residue and associated burning emissions in Vietnam. *Environmental Research Letters*, *12*, 085006
- Li, J., & Roy, D. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, *9*, 902
- Li, Z., Roy, D.P., Zhang, H.K., Vermote, E.F., & Huang, H. (2019). Evaluation of Landsat-8 and Sentinel-2A Aerosol Optical Depth Retrievals across Chinese Cities and Implications for Medium Spatial Resolution Urban Aerosol Monitoring. *Remote Sensing*, *11*, 122
- Maddux, B., Ackerman, S., & Platnick, S. (2010). Viewing geometry dependencies in MODIS cloud products. *Journal of Atmospheric and Oceanic Technology*, *27*, 1519-1528
- 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
- Melchiorre, A., Boschetti, L., & Roy, D.P. (2020). Global Evaluation of the Suitability of MODIS-Terra Detected Cloud Cover as a Proxy for Landsat 7 Cloud Conditions. *Remote Sensing*, *12*, 202
- Morton, D.C., DeFries, R.S., Nagol, J., Souza Jr, C.M., Kasischke, E.S., Hurtt, G.C., & Dubayah, R. (2011). Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data. *Remote Sensing of Environment*, *115*, 1706-1720
- Mouillot, F., Schultz, M.G., Yue, C., Cadule, P., Tansey, K., Ciais, P., & Chuvieco, E. (2014). Ten years of global burned area products from spaceborne remote sensing—A review: Analysis of user needs and recommendations for future developments. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 64-79

- Nakajima, T.Y., Tsuchiya, T., Ishida, H., Matsui, T.N., & Shimoda, H. (2011). Cloud detection performance of spaceborne visible-to-infrared multispectral imagers. *Applied Optics*, *50*, 2601-2616
- Padilla, M., Stehman, S.V., & Chuvieco, E. (2014). Validation of the 2008 MODIS-MCD45 global burned area product using stratified random sampling. *Remote Sensing of Environment*, *144*, 187-196
- Pincus, R., Platnick, S., Ackerman, S.A., Hemler, R.S., & Hofmann, R.J.P. (2012). Reconciling Simulated and Observed Views of Clouds: MODIS, ISCCP, and the Limits of Instrument Simulators. *Journal of Climate*, *25*, 4699-4720
- Rossow, W.B., & Schiffer, R.A. (1999). Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, *80*, 2261-2287
- Roteta, E., Bastarrika, A., Padilla, M., Storm, T., & Chuvieco, E. (2019). Development of a Sentinel-2 burned area algorithm: Generation of a small fire database for sub-Saharan Africa. *Remote Sensing of Environment*, *222*, 1-17
- Roy, D., Lewis, P., Schaaf, C., Devadiga, S., & Boschetti, L. (2006). The global impact of clouds on the production of MODIS bidirectional reflectance model-based composites for terrestrial monitoring. *Geoscience and Remote Sensing Letters, IEEE*, *3*, 452-456
- Roy, D.P., & Boschetti, L. (2009). Southern Africa validation of the MODIS, L3JRC, and GlobCarbon burned-area products. *Geoscience and Remote Sensing, IEEE Transactions on*, *47*, 1032-1044
- Roy, D.P., Huang, H., Boschetti, L., Giglio, L., Yan, L., Zhang, H.H., & Li, Z. (2019). Landsat-8 and Sentinel-2 burned area mapping - A combined sensor multi-temporal change detection approach. *Remote Sensing of Environment*, *231*, 111254
- Schiffer, R., & Rossow, W. (1985). ISCCP global radiance data set: A new resource for climate research. *Bulletin of the American Meteorological Society*, *66*, 1498-1505
- Shang, H., Letu, H., Nakajima, T.Y., Wang, Z., Ma, R., Wang, T., Lei, Y., Ji, D., Li, S., & Shi, J. (2018). Diurnal cycle and seasonal variation of cloud cover over the Tibetan Plateau as determined from Himawari-8 new-generation geostationary satellite data. *Scientific Reports*, *8*, 1105
- Tanase, M.A., Belenguer-Plomer, M.A., Roteta, E., Bastarrika, A., Wheeler, J., Fernández-Carrillo, Á., Tansey, K., Wiedemann, W., Navratil, P., Lohberger, S., Siegert, F., & Chuvieco, E. (2020). Burned Area Detection and Mapping: Intercomparison of Sentinel-1 and Sentinel-2 Based Algorithms over Tropical Africa. *Remote Sensing*, *12*, 334

- Vanderhoof, M., Brunner, N., Beal, Y.-J., & Hawbaker, T. (2017). Evaluation of the U.S. Geological Survey Landsat Burned Area Essential Climate Variable across the Conterminous U.S. Using Commercial High-Resolution Imagery. *Remote Sensing*, 9, 743
- Whitcraft, A.K., Becker-Reshef, I., & Justice, C.O. (2015a). A framework for defining spatially explicit earth observation requirements for a global agricultural monitoring initiative (GEOGLAM). *Remote Sensing*, 7, 1461-1481
- Whitcraft, A.K., Vermote, E.F., Becker-Reshef, I., & Justice, C.O. (2015b). Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sensing of Environment*, 156, 438-447
- Wulder, M.A., Hilker, T., White, J.C., Coops, N.C., Masek, J.G., Pflugmacher, D., & Crevier, Y. (2015). Virtual constellations for global terrestrial monitoring. *Remote Sensing of Environment*, 170, 62-76
- Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., & Cohen, W.B. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, 225, 127-147
- Zhao, G., & Di Girolamo, L. (2006). Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. *Geophysical Research Letters*, 33
- Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83-94
- Zhu, Z., Wulder, M.A., Roy, D.P., Woodcock, C.E., Hansen, M.C., Radeloff, V.C., Healey, S.P., Schaaf, C., Hostert, P., & Strobl, P. (2019). Benefits of the free and open Landsat data policy. *Remote Sensing of Environment*, 224, 382-385

Appendix A: Publishing Rights & Permissions

Chapter 2 and Chapter 3 are published on MDPI Remote Sensing journal with the following references:

Melchiorre, A., & Boschetti, L. (2018). Global Analysis of Burned Area Persistence Time with MODIS Data. *Remote Sensing*, 10, 750

Melchiorre, A., Boschetti, L. & Roy, D.P. (2020). Global evaluation of the suitability of MODIS-Terra detected cloud cover as a proxy for Landsat 7 cloud conditions. *Remote Sensing*, 12(2), 202

For all articles published in MDPI journals, copyright is retained by the authors. Articles are licensed under an open access Creative Commons CC BY 4.0 license, meaning that anyone may download and read the paper for free. In addition, the article may be reused and quoted provided that the original published version is cited. These conditions allow for maximum use and exposure of the work, while ensuring that the authors receive proper credit.

More information can be found at <https://www.mdpi.com/authors/rights>.