Geo-information Science and Remote Sensing

Thesis Report GIRS-2018-12

Analysis of the use of spatio-temporal early warning metrics to estimate fire vulnerability resilience in the Amazon rainforest

Elke Hendrix

April 4, 2018

WAGENINGEN
UNIVERSITY & RESEARCH

Analysis of the use of spatio-temporal early warning metrics to estimate fire vulnerability resilience in the Amazon rainforest

Elke Hendrix

Registration number 94 07 25 326 060

Supervisors:

Dr. Jan Verbesselt
Simon Besnard, M.Sc.

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

April 4, 2018
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2018-12
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Foreword and Acknowledgement

I would like to thank all the people who supported me during my thesis project. First of all my parents, brother, boyfriend and friends, thank you for the delicious home cooked meals and your patience. Thank you Bram for always letting me borrow your computer.

Special thanks go to Jan Verbesselt and Simon Besnard, my supervisors. Thank you Jan for your expertise, kindness and enthusiasm which motivated and excited me about how amazing research can be. Thank you Simon for your critical and kind help. Your door was always open for a short discussion or help with R programming. Without your help my scripts would still be running, I owe you a lot.

I would also like to thank Imke and Juultje for accompanying me in the thesis room and taking me on much needed hikes (and occasional ice skating) around the University. Finally, I want to thank all my fellow MGI master student for accompanying me during the countless breaks.

Abstract

The current Amazon tropical rainforest's *"climate-vegetation"* equilibrium is one of the ecosystems at risk of entering an alternative stable state. Heat, drought and tree mortality can cause the current equilibrium as a carbon (C) sink to shift to the *"fire dominated tropical savanna"* equilibrium, as a C source. The critical transition of the C balance can eventually have a huge impact on the global C cycle. However, it is poorly understood what processes drive the critical transition in ecosystem states. As a result, there is a need to identify early warning signals to help us have a better understanding of critical transition mechanisms.

Early warning metrics (EWMs) are commonly used to characterize resilience under stressed regimes, by measuring the properties of time series. Therefore it is important to understand whether critical transitions are preceded by any spatio-temporal ecosystem dynamics. In this study we aimed to evaluate whether a series of EWMs, derived from remote sensing time series (i.e. LST, vegetation indices and VOD), can provide relevant knowledge for better predicting both forest fires and drought vulnerability. To do so, we implemented a baseline random forest (RF) model, constrained by the aforementioned remote sensing variables, which only characterize the contemporary conditions of the ecosystems. Further, we benchmark the baseline model against a RF model constrained both by the current condition and EWMs. It allowed us to better understand the importance of EWMs for predicting fire vulnerability on a local scale and drought stress on a regional scale.

We showed that adding EWMs that are consistent with theoretical expectations based on critical slowing down increased the overall accuracy of the predictions by 4.71%, from 68.52% to 73.23%, for the fire vulnerability predictions and by 1.48%, from 85.69% to 87.17%, for the drought stress predictions. Among the different predictive variables tested, spatial autocorrelation derived from LST night datasets was the most important followed by spatial variance derived from LST night datasets.

We conclude that the addition of EWMs shows potential for the prediction of fire and related drought stress. But critical note should be made that these additions are relatively small, further research is needed to evaluate whether EWMs can strengthen existing drought stress and related fire risk frameworks. The VOD dataset used in this research was too coarse to use in the RF models, but the time series showed significant signs of critical slowing down for the drought stressed areas. Further research is needed to assess whether a less coarse VOD dataset could increase the accuracy of the predictions for both drought stress and fire vulnerability.

Keywords: Tipping points, Early warning metrics, spatial early warning metrics, forest fires, NDVI, LST, VOD, NBR, critical slowing down, resilience, random forest.

Contents

List of Figures

List of Tables

Abbreviations

AMSR2: Advanced Microwave Scanning Radiometer 2

CH: Conditional Hetroscedasticity

CSD: Critical Slowing Down

ENSO: El Niño Southern Oscillation

EWM: Early Warning Metrics

EWS: Early Warning Signal

GFC: Global Forest Change

LST: Land Surface Temperature

MODIS: Moderate Resolution Imaging Spectroradiometer

NASA: National Aeronautics and Space Administration

NBR: Normalized Burn Ratio

NDVI: Normalized Different Vegetation Index

NIR: Near Infrared

ReRa: Return Rate

SPI: Standardized Precipitation Index

SWIR: Short-wave Infrared

TRMM: Tropical Rainfall Measuring Mission

VOD: Vegetation Optical Depth

Chapter 1

Introduction

1.1 Setting the scene

The Amazon rainforest is of critical importance as a carbon (C) pool, for the maintenance of tropical biodiversity and for both local and regional climate stability (Field et al., 1998, Dirzo & Raven, 2003, Nobre at al., 1991, Brienen et al., 2015). Climate extremes (e.g. heat and drought) and anthropogenic disturbances (e.g. deforestation) cause gradual changes in the Amazonian rainforest (Nobre & Borma, 2009). Over the past years the Amazon forest has dealt with several climate extreme events, e.g. the 2015-2016 El Niño Southern Oscillation (ENSO) event (Jiménez-Munoz et al., 2016). Such episodes of heat and drought negatively impact tree productivity and can eventually result in tree mortality, causing a rise in atmospheric CO2 concentration (Jiménez-Munoz et al., 2016). Different lines of evidence point towards a scenario in which these gradual changes in the ecosystem state might result in the Amazon forest transiting from a *"climate-vegetation"* equilibrium to a *"fire dominated tropical savanna"* equilibrium (Reyer et al., 2015, Nobre & Borma, 2009, Verbesselt et al., 2016, Hirota et al., 2011, Staver & Hansen, 2015, Nepstad et al., 2007, da Costa et al., 2010, Brando et al., 2014). The latter state will result in an ecosystem composed by drought- and fire-adapted vegetation, causing the loss of the Amazon rainforest's biodiversity and its function to sequester C. As a consequence, this ecosystem state transition will shift the Amazon forest's C balance from a C sink to a C source, having an eventual impact on the global C cycle (Brando et al., 2014, Nobre & Borma, 2009).

Drought and associated high temperature stresses are the primary climate characteristics associated with tree mortality (Allen et al., 2010; Williams et al., 2013). Although tree mortality from a single drought event can be severe (Michaelian et al., 2011; Anderegg et al., 2013), mortality is often related with prolonged or repeated droughts (Bigler et al., 2006; Gustafson & Sturtevant, 2013; Mitchell et al., 2013; Jump et al., 2017, Nepstad et al., 2007, Brando et al., 2008, Fisher et al., 2007, Meir et al., 2008). Severe droughts and high temperature stress can cause forest fires, resulting in a positive feedback loop. Burned areas dry out more easily and are again more susceptible to forest fires (Nepstad et al., 2008, Davidson et al., 2012, Nepstad et al., 2001). In addition, tropical rainforest species in the Amazon are poorly adapted to forest fires, leading to lower resilience against fire and making the return to the current *"climate-vegetation"* state challenging (Nobre & Borma, 2009).

The positive feedback loop of drought stress and related forest fires could lead to the transition from a *"climate-vegetation"* equilibrium to a *"fire dominated tropical savanna"* equilibrium. This transition can be induced when a critical threshold, called a tipping point is exceeded (Scheffer et al., 2001, Scheffer, 2009, Dakos et al., 2012). Over the past years the evaluation of these tipping points of Earth systems has gained scientific interest (Lenton et al., 2008). Current examples of ecosystems at risk of collapsing are the Atlantic thermohaline circulation, the decay of the Greenland ice sheet and the dieback of the Amazon

rainforest (Rahmstorf & Ganopolski, 1999, Cox et al., 2000, Huybrechts & de Wolde, 1999). Small perturbations can have long-term consequences, once the tipping point is reached the effects are irreversible (Lenton et al., 2008). Resilience towards these perturbations can be defined as: *"The ability of a forest to absorb disturbances and re-organize under change to maintain similar functioning and structure"* (Scheffer, 2009). The lower the resilience the longer it takes for a system to recover to the original state, this is due to the phenomenon of critical slowing down (CSD) (see figure 1.1). Mathematically, CSD means that the system at equilibrium vanishes when it comes to close to the dominant eigenvalue (Beisner et al., 2003).

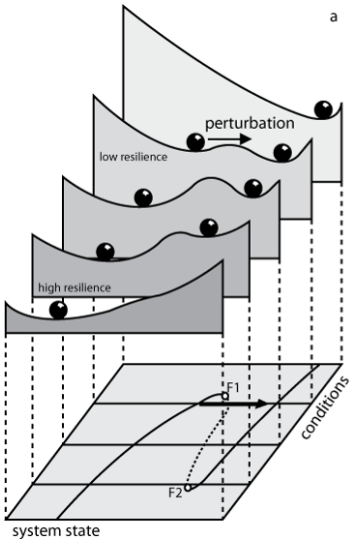


FIGURE 1.1: Shift from high resilience to low resilience due to perturbations. The steepness and the width determine the absorption capacity of a system without shifting to an alternative stable state (Beisner et al., 2003).

Potential to timely identify areas with a low resilience towards drought and related forest fires remains in the analysis of time series of ecosystem dynamics. Time series often indicate that the productivity of an ecosystem declines when the resilience towards a regime shift is lowered. The resulting temporal dynamics could precede mortality years to decades before the death of forests (Rogers et al., 2018). Typically, tree vigor gradually declines by the accumulation of drought- and/or biotic- stress until the tipping point is reached and finally triggers tree mortality, which exhibits itself in early warning signals, e.g. declining productivity rates and subtle variations (Beisner et al., 2003, Rogers et al., 2018). Previous studies have used Early Warning Metrics (EWMs) as a measures of CSD. EWMs are generic statistical measures that can detect the proximity of a system towards a critical transition to an alternative stable state in time series, without having full knowledge about the drivers of the tipping point under research (Beisner et al., 2003, Dakos et al., 2012, Scheffer et al., 2009).

Ecosystems with a high resilience will recover fast and show significantly different ecosystem properties between time steps. Ecosystems with a low resilience need more consecutive time steps to recover to the original state than the ecosystem far from transition. This results in ecosystem properties that remain relatively similar between time steps because recovery takes longer (see figure, 1.2). Direct EWMs of CSD are temporal autocorrelation and return rate (ReRa). An increase in autocorrelation shows that a system is becoming more

similar between consecutive observations suggesting that the return to the original states takes longer (see figure 1.3 a.2& b.2)(Kleinen et al., 2003). The ReRa EWM calculates the distance towards the stable state, the slower the return time the more time steps needed to recover (see figure 1.3 a.1 & b.1). Indirect measures of CSD are variance, skewness, kurtosis and conditional heteroscedasticity (see figure 1.3 a.3, a.4, b.3 & b.4). Assuming that a system close to transitions is less resilient to disruptions, it will take longer for the ecosystem to return to the original stable state. The time spent away from the stable state will increase once the system comes closer to transition. Therefore, variance, skewness, kurtosis and conditional heteroscedasticity will increase compared to the stable state (Scheffer et al., 2009, Dakos et al., 2012, Guttal., 2008, Bigg et al., 2009, Seekell et al., 2011). Studies suggest that spatial structures of ecosystems can also provide early warning signals for an impending regime shift. Due to increased recovery time to the local equilibrium the spatial coupling between neighbouring cells becomes the dominant force. This favours neighbouring units to become more like each other, i.e. they become increasingly correlated (Kefi et al., 2014, dakos et al., 2011).

FIGURE 1.2: The first time series shows the recovery time after a disruption for forest areas with a low resilience. The second time series shows the recovery time after a disruption for forest areas with a high resilience.

Remote sensing is capable of observing these changes in ecosystem dynamics, i.e. the potential decline in forest productivity or the increase in variance, on a large spatial scale and could therefore be suitable for monitoring drought stress and related fire vulnerability (Rogers et al., 2018). Areas that are at risk of forest fires are currently monitored with wireless sensor networks or can be predicted based on satellite images, wind speeds and weather forecasts, which is on a small temporal scale (Hefeeda & Bagheri, 2007, Oštir et al., 2003).

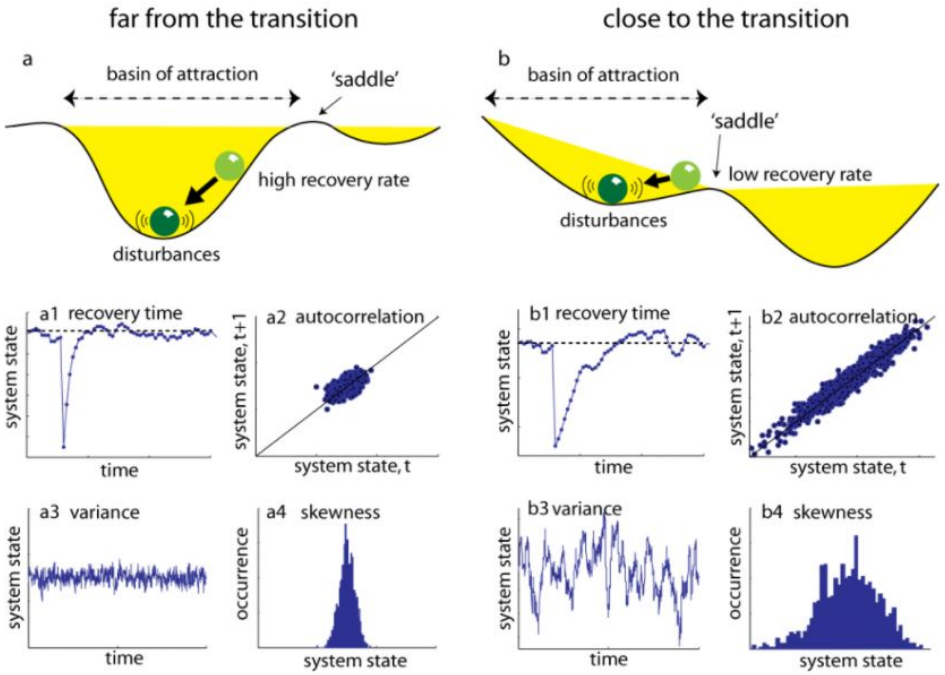


FIGURE 1.3: EWMs far from transition and EWMs close to transition (Early Warning Signals Toolbox, 2014)

1.2 Problem statement

The trees of the Amazonia contain 90-140 billion tons of C, which is comparable to 9-14 decades of current human-induced C emissions (Soare-Filho et al., 2006, Canadell et al., 2007). By reaching the tipping point the rainforest could potentially lose its role as a giant C reservoir (Nepstad et al., 2008). The prospect of reducing global warming and keeping global average temperatures from rising 2° C will be near impossible when the current *"climate vegetation state"* collapses (Gullison et al., 2007). Therefore, insight into indicators that estimate the resilience of an ecosystem towards drought stress and related forest fires would add to prevention and adaptation strategies.

What precisely degrades the resilience of a system during drought and heat waves remains a topic of debate, but the emerging view is we need to focus on systematic resilience of trees rather than focusing on isolated elements (Hartman et al., 2015, Anderegg et al., 2012). To overcome these challenges previous studies have used generic temporal EWMs in time series to detect the proximity of a tipping point (Scheffer et al., 2009, Kefi et al., 2014, Verbesselt et al., 2016). Detecting and characterizing change over time is the first step towards identifying and understanding change mechanisms (Nepstad et al., 2008). The temporal EWMs have been widely examined in literature, however theoretical studies suggest spatial structures of ecosystems can also provide information about the ecosystem degradation level (Kefi et al., 2014). Hence, one can wonder whether the use of spatio-temporal EWMs can provide a better understanding of drought and fire processes.

1.3 Research objectives and research questions

Given the lack of knowledge in understanding forest resilience and tipping points, this study aims to evaluate whether the use of early warning metrics could help identifying drought stressed and related vulnerable forest fire areas. The overall objective of this thesis is to analyze whether surface temperature and vegetation related EWMs can differentiate between fire versus non-fire conditions on a local scale and drought vs non-drought conditions on a regional scale. Such metrics would allow us to characterize forest resilience under fire and drought regimes as ecosystems approach thresholds for critical transitions. This study aims to answer the following research questions:

1. Can spatio-temporal early warning metrics improve the predictions of fire vulnerability estimates in a drought stressed forest ecosystem on a local scale?

2. Can spatio-temporal early warning metrics help differentiate between drought stressed and non-drought stressed forest scenari on a regional scale?

1.4 Hypothesis

The overarching hypothesis is that characterizing the properties of time series using spatio-temporal EWMs is important to understand drought and fire regimes.

We hypothesize that properties related to temporal autocorrelation matter for predicting drought and forest fires (Verbesselt et al., 2016, Dakos et al., 2012, Pace et al., 2013). In addition, we expect that spatial autocorrelation related properties are also important for predicting drought and forest fires (Dakos et al., 2010).

Chapter 2

Materials

2.1 Study area

In order to test our hypothesis, we selected two areas in the region of the amazon rainforest; that is (i) a region around Santarem prone to forest fires (Fig. 2.1) and (ii) a subset of the Amazon region with drought and non-drought stressed areas (Fig. 2.2). Santarem lies next to the banks of the river the Amazone and the area is covered by tropical rainforest. The rainforest is struck by forest fires due to the dry conditions of the 2015-2016 ENSO event (Salisbury, 2017). The region of Santarem of 740.000 hectares was burned between August-November 2015 and January-May 2016. Satellite surveillance was unable to detect the forest fires until June 2016 due to heavy cloud cover (Pegurier & Vieira, 2016). The subset of the Amazon region covers an area of 12.601.400 hectares.

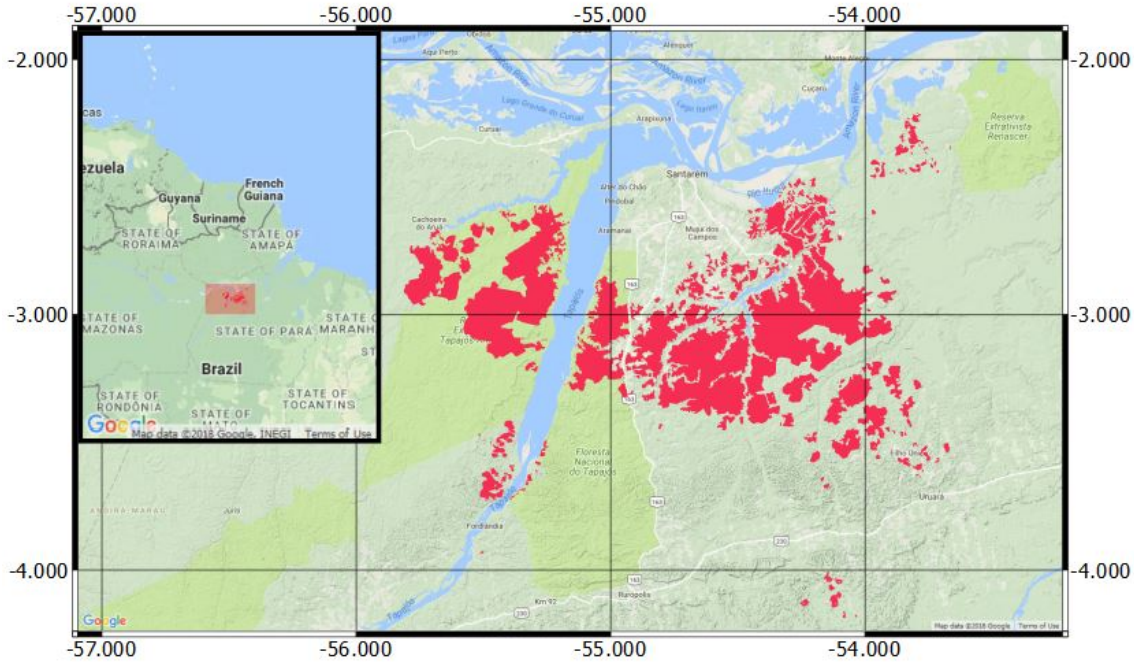


FIGURE 2.1: The locations of the fire areas are shown in the red polygons. The red box on the left shows the location of the fire vs non-fire analysis within the Amazon.

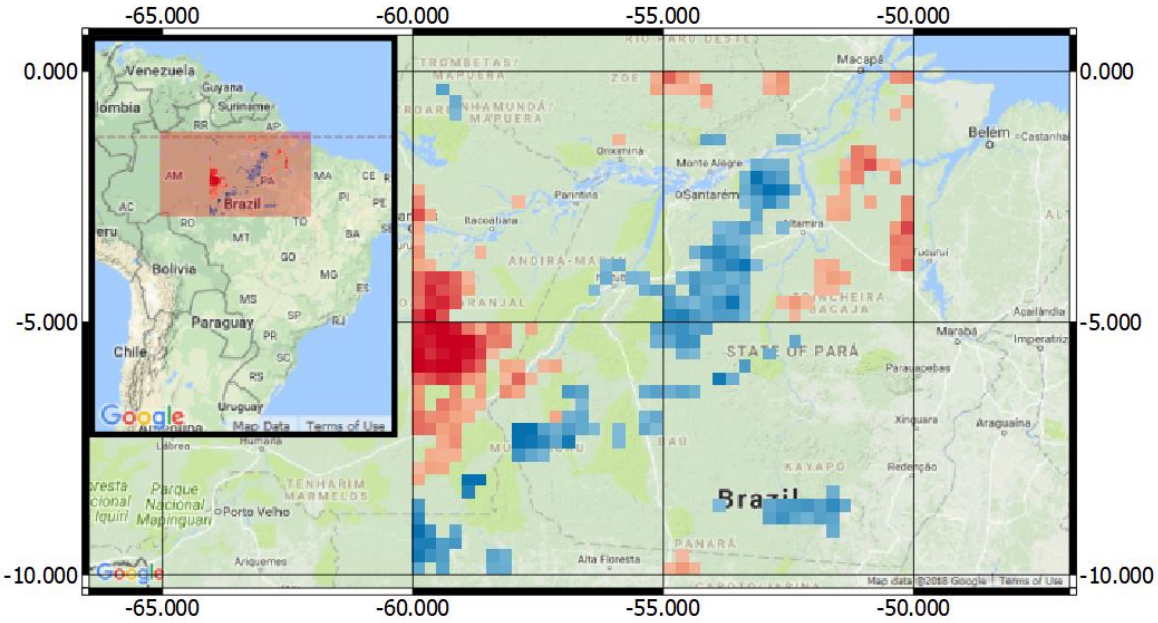


FIGURE 2.2: The locations of the drought areas are shown in the red pixels while the locations of the non-drought areas are shown by the blue pixels. The red box on the left shows the location of the drought/non-drought analysis within the Amazon.

2.2 Data

2.2.1 Fire data

A dataset with the locations of the fires that took place in the area of Santarem was used. The locations of the fire patches were collected between August-November 2015 and January-May 2016. The dataset provides the locations and the extent of the fire patches within the area (Pegurier & Vieira, 2016).

2.2.2 Climate data

Tropical rainfall measuring mission's (TRMM) monthly precipitation data (TMPA/3B43) was used to differentiate between drought vs non-drought stressed areas. The data was obtained from the Goddard Earth Sciences Data and Information Services Center. See Methods section 3.1.2 and Appendix A for more information on how we defined drought stress.

2.2.3 Remote sensing data

- Normalized Difference Vegetation Index (NDVI). In this study the MODIS NDVI 16-day product (MOD13Q1) from the Terra satellite is used at 250 - by 250- meters for the period 2002-2015. The NDVI is sensitive to chlorophyll abundance and photosynthetically active biomass of the leaves (Liu et al., 2013)

- Vegetation Optical Depth (VOD). The VOD represents water content in above ground woody and leaf biomass and is sensitive to long-term climate change (Liu et al., 2013). The AMSR2 2/3-day VOD (x-band) product is used in this research. The dataset is distributed with a spatial resolution of 0.25 degree ($\approx$ 27.64 km) and

covers a period of 2007-2015. Previous studies indicate that VOD values correlate to precipitation variations, and that the mutually independent VOD and NDVI values do not necessarily correspond in an identical manner (Zhou et al., 2014). Considering both products together provides a more robust assessment also because the VOD has the ability to penetrate clouds, resulting in less no-data values in the dataset (Liu et al.,2013).

- Land Surface Temperature (LST). The MODIS LST 8-day product (MOD11A2) from the terra satellite is used. The dataset is distributed with a spatial resolution of 1- by 1- kilometre and covers a period of 2000-2015. Both the LST day and LST night datasets are used. The LST dataset is an indirect measure of soil moisture conditions and yields a distinctive thermal signature. Moisture deficiencies in the root zone lead to vegetation stress and elevated canopy temperatures (Anderson et al., 2007). During dry conditions, rising leaf temperatures are good indicators of plant moisture stress and precede the onset of drought, this thermal response can occur even when the plants are green (Wan et al., 2004).

- Normalized Burn Ratio (NBR). In this study the MODIS surface reflectance 8-day product (MOD09A1) from both the Terra and Aqua satellite is used at 500- by 500-meters for the period 2002-2015. The NBR is sensitive to the changes in the amount of live green vegetation, moisture content and soil conditions which may occur after fire (Miller & Thode, 2007).

- Global Forest Change database (GFC). The GFC database is based on Landsat data (i.e. Landsat 4, 5, 7, 8) and was freely downloaded from the Global Forest Change (GFC) database (Hansen et al., 2013).The database contains a dataset with the forest cover in 2000 and the stable forest cover from 2001 till 2015. The dataset is distributed with a spatial resolution of 30 meters (Hansen et al., 2013).

2.3 Software

The software packages to carry out this research will be R for the statistical analysis, Earth Engine for the downloading of the data, and Quantum GIS (2.18.12) for visualization.

Chapter 3

Methods

Figure 3.1 gives an overview of the methods used to answer the research questions in section 1.3. First the datasets used to (i) calculate the EWMs and contemporary conditions of the ecosystem, (ii) create the forest mask and (iii) validate the results were pre-processed. To evaluate whether a series of EWMs derived from aforementioned remote sensing variables could provide knowledge for better predictions of both forest fires and drought vulnerability a baseline random forest (RF) model, constrained by variables characterizing the contemporary conditions of the ecosystem, was implemented. The predictive power of our baseline RF model was further compared to a second RF model, which was constrained by both the contemporary conditions of the ecosystems and the EWMs.

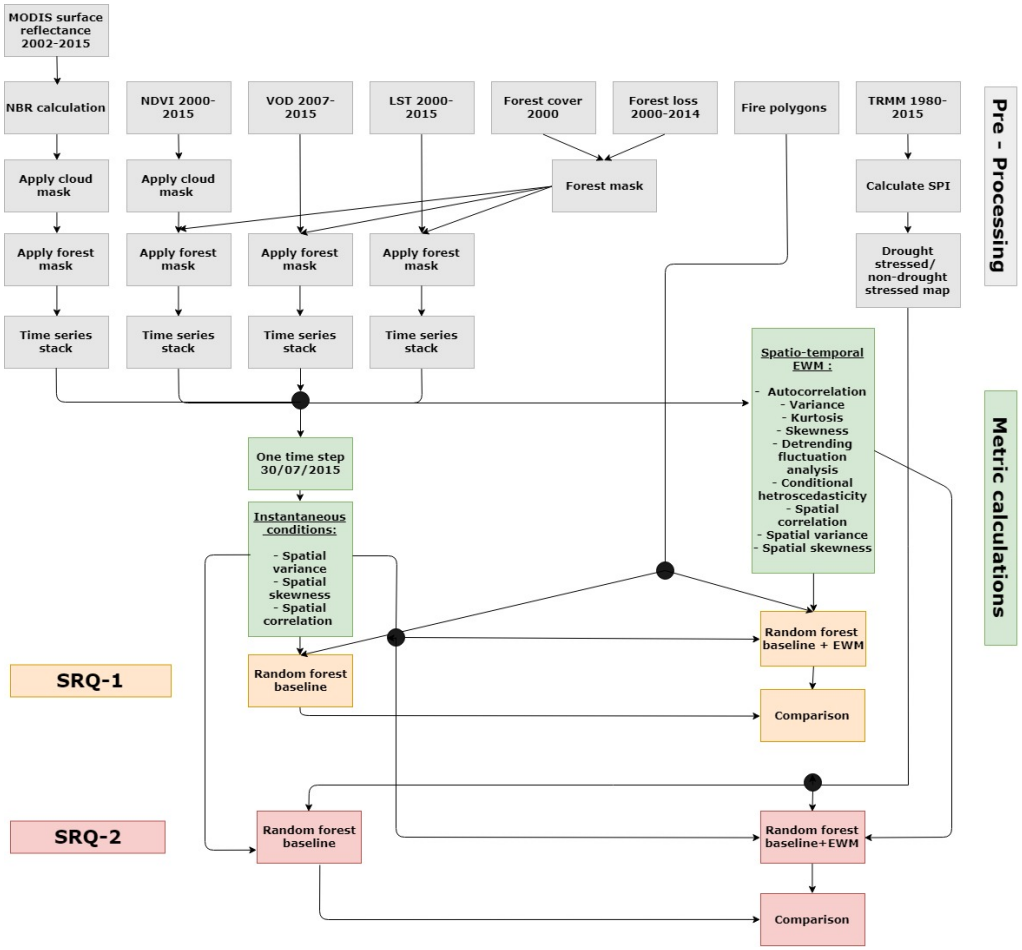


FIGURE 3.1: Schematic framework of the methodology section, which indicates pre-processing steps, metric calculations and sub-research questions (SRQ).

3.1 Pre-processing

3.1.1 Vegetation indices

The NDVI was based on the ratio between the red 0.66 micrometer wavelength and the near infrared (NIR) 0.83 micrometer wavelength. The NDVI was calculated based on formula 3.1 (Tucker et al., 1979).

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{3.1}$$

The NBR index was based on the NIR 0.79-0.90 micrometer wavelengths and the short-wave infrared (SWIR) 2.08-2.35 micrometer wavelengths and was used to estimate burn severity. For the MOD09A1 imagery the NBR was calculated based on formula 3.2, using band 2 as the NIR and band 7 as the SWIR (van Wagtendonk et al., 2004).

$$NBR = \frac{NIR - SWIR}{NIR + SWIR} \tag{3.2}$$

3.1.2 SPI calculation

Mckee's Standardized Precipitation Index (SPI) was used to differentiate between drought and non-drought stressed areas. For this research a moving window of 6 months was used to calculate the SPI index for the Amazonia. The SPI was calculated for the dry season before occurrences of the first fires (30/07/2015). We assumed the dry season to be between July and September, a simplification often used by other studies (Samanta et al., 2012). The 10% of the lowest SPI pixel values were used to identify drought stress and the 10% of the highest SPI pixel values were used to identify non-drought stress. The TRMM dataset was used to calculate the SPI and stores the precipitation in millimetres per hour-1. The cumulative monthly precipitation was estimated in mm per month considering a 30-day month for all the datasets (Aragao et al., 2007).

3.1.3 Cloud filter

A cloud mask was applied to the NDVI, LST and NBR datasets. For the NDVI and LST datasets the pixel reliability layers were used to mask out the clouds. For the NBR datasets the surface reflectance quality assurance (QA) layers were used to mask out the clouds. All the time series that were cloud contaminated for more than 50% of the total time of the time series were removed from the analysis.

3.1.4 Forest mask

To exclude areas that were not stable forest a forest mask was applied. To create the mask the GFC dataset with the (i) percentage of forest cover in 2000 and the GFC dataset with (ii) forest loss from 2000 to 2015 was used. The pixels that represented a loss of forest from 2000 till 2015 were selected and merged with the forest cover from 2000. The new mask represents the stable forest cover from 2000 till 2015 right before the forest fires occurred. Gain of forest is not considered in this research because these forest areas were not stable during the time period of this research. We filtered out the forest pixels with a

forest coverage below 60% (Baumann et al., 2012, Gregorio, 2005). See Appendix G for the local and regional forest masks.

3.2 Early warning calculations

The temporal and spatial EWMs are calculated with the aforementioned remote sensing data: NDVI, LST, VOD and NBR, see material section 2.2.3.

3.2.1 Temporal early warning metrics

Figure 3.2 summarizes the different steps followed to calculate the temporal EWMs (Dakos et al., 2012, Dakos et al., 2015). The *preprocessing* phase involved the linear interpolation of the missing data in the time series (Dakos et al., 2012). The *Detrending and Filtering* phase involved the application of a Gaussian filter to the rolling window to remove seasonal periodicities (Dakos et al., 2012). The length of the *Rolling window* was set to half the length of the time series, which is the most commonly used window size (Dakos et al., 2012). The *calculation of the metrics* phase involves the calculation of six temporal EWMs based on CSD: autocorrelation, return time, skewness, kurtosis, variance and conditional heteroscedasticity, using the *earlywarning* package in R(Dakos et al., 2012, Dakos et al., 2015). The *quantifying trends* phase involved the quantification of the trend by using Kendall's tau. Kendall's Tau rank correlation coefficient was used to check against the null hypothesis of randomness (Dakos et al., 2008). Kendall's tau ranges between 1, for an indicator that is always increasing to -1 for an indicator that is always decreasing. A Kendall's tau of 0 implies that there is no net trend in the indicator or that the trend increases as much as it decreases (Boulton et al., 2013).

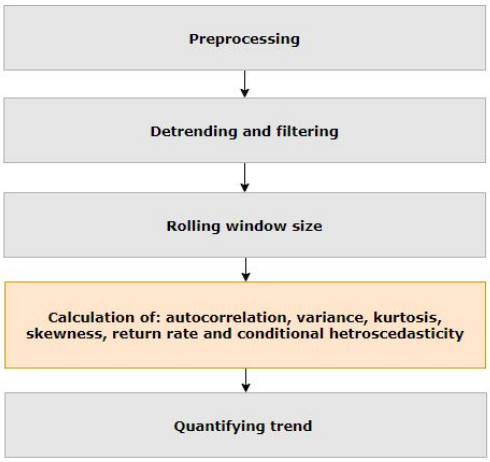


FIGURE 3.2: Schematic framework for the calculation of the temporal early warning metrics.

3.2.2 Spatial early warning metrics

Figure 3.3 summarizes the different steps followed to calculate the spatial EWMs (Kefi et al., 2014). The *preprocessing* phase involved the linear interpolation of the missing data in the time series. The *calculation of metrics* phase involved the calculation of the three simplest spatial EWMs: spatial autocorrelation, spatial skewness and spatial variance. The

amount in which neighbouring cells look and behave alike is measured with Moran's I to calculate spatial autocorrelation (Moran, 1950, Geary, 1954, Anselin, 1995). The spatial EWMs autocorrelation, variance and skewness were calculated within a kernel. The kernel size used in this research was 3x3. The last step was the *quantifying trends* phase. After the metrics were applied to every time step the results were stacked and quantified by using Kendall's Tau, as explained in section 3.2.1.

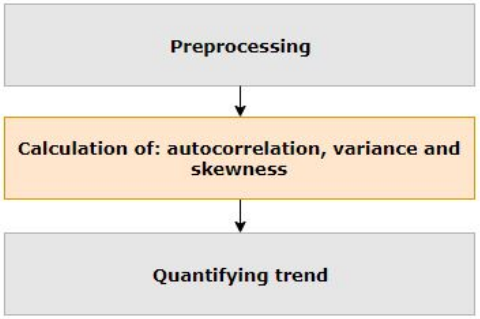


FIGURE 3.3: Schematic framework for the calculation of the spatial early warning metrics.

3.3 Model development and assessment

3.3.1 Input variables

Input features for the baseline model were calculated based on one-time step before the occurrence of the fires in the remote sensing images for both the local and the regional scale (30/07/2015). The contemporary conditions of the ecosystem calculations were spatial variance, spatial skewness and spatial correlation. For the temporal model the contemporary conditions of the ecosystem were added to the temporal EWMs: autocorrelation, variance, kurtosis, skewness, return time, conditional heteroscedasticity, and spatial EWMs: correlation, variance and skewness. See table 3.1 for all the input variables.

TABLE 3.1: Input variables for the random forest predictions for both the RF baseline model and the RF EWMs model. The input variables were composed from NDVI, LST, VOD and NBR time series.

contemporary conditions and EWMs	Metric
Baseline	Contemporary spatial variance
Baseline	Contemporary spatial skewness
Baseline	Contemporary spatial autocorrelation
Spatial EWM	Autocorrelation
Spatial EWM	Variance
Spatial EWM	Skewness
Temporal EWM	Autocorrelation
Temporal EWM	Variance
Temporal EWM	Kurtosis
Temporal EWM	Skewness
Temporal EWM	Return time
Temporal EWM	Conditional hetroscedasticity

3.3.2 Variable selection

Only metrics that are consistent with theoretical considerations based on CSD were considered useful in this research. To identify whether a metric was consistent with EWM theory the mean Kendall's tau for both the fire/drought and non-fire/non-drought areas were compared. For example the slope of temporal autocorrelation should rise more for the fire/drought areas than for the non-fire/non-drought areas. A two-sampled, two-tailed t-test was used to determine whether the difference between the fire/drought and non-fire/non-drought was significant per metric.

3.3.3 Model training

After the features were calculated a random forest classification is constructed to predict which areas were drought stressed, non-drought stressed, fire and non-fire (either 1 for fire/drought stress and 0 for no-fire/no-drought stress) (Breiman, 2001, Kuhn, 2017). The random forest classifier was based on a machine learning algorithm that constructs decision tree classifiers based on bootstrapped samples (Breiman, 2001).

Two model set-ups were defined: the RF baseline model and RF EWMs models. The RF baseline model was constructed with the contemporary conditions of the ecosystem while the RF EWMs model was constructed with the baseline + spatio-temporal EWMs.

3.3.4 Model evaluation

The performance of the random forest was evaluated by directly comparing model estimates with observed fire/drought occurrence for each fire/drought patch. We corroborated the models by using a k-fold cross-validation. The results of the RF baseline model predictions were compared to the RF EWMs model predictions to see if the addition of EWMs add to drought stress and fire vulnerability estimations.

3.3.5 Variable importance

To asses which input variable had the biggest importance in the random forest predictions the Boruta package was used (Kursa & Rudnicki, 2017). The Boruta algorithm introduced noise to one input variable at the time and assessed whether that influences the accuracy of the predictions of the random forest. The variable importance score is based on the ratio between the mean accuracy loss and the σ of accuracy loss.

Chapter 4

Results

4.1 Spatial autocorrelation for fire and drought

We hypothesized that spatial autocorrelation would be among the best performing EWMs in the predictions of fire/drought and non-fire/non-drought areas. Figure 4.1 shows the distribution for the fire/drought and non-fire/non-drought pixels for spatial autocorrelation derived from the LST night, LST day, NDVI and NBR datasets. For the local scale it is visible that the boxplots are relatively similar for the fire and non-fire areas, but the LST night boxplots shows a difference. The mean Kendall's tau for the LST night fire and non-fire areas are 0.01605 and -0.00431 respectively with a p-value lower than 0.001, suggesting that the difference between fire and non-fire patches is significant. A positive Kendall's tau for the fire areas suggests that spatial autocorrelation is increasing over time while a negative Kendall's tau for the non-fire patches suggests a decrease in spatial autocorrelation over time, which is consistent with theoretical expectations based on CSD. The other datasets have a higher Kendall's tau for the non-fire areas, which is not consistent with theoretical expectations, see Appendix C for the complete list of mean values for fire and non-fire areas for all the metrics. For the regional scale it is visible that both the NDVI and the LST night dataset show a higher median Kendall's tau values for the drought areas in the boxplots as well as in the mean Kendall's tau values, see Appendix E. Figure 4.2 shows two time series

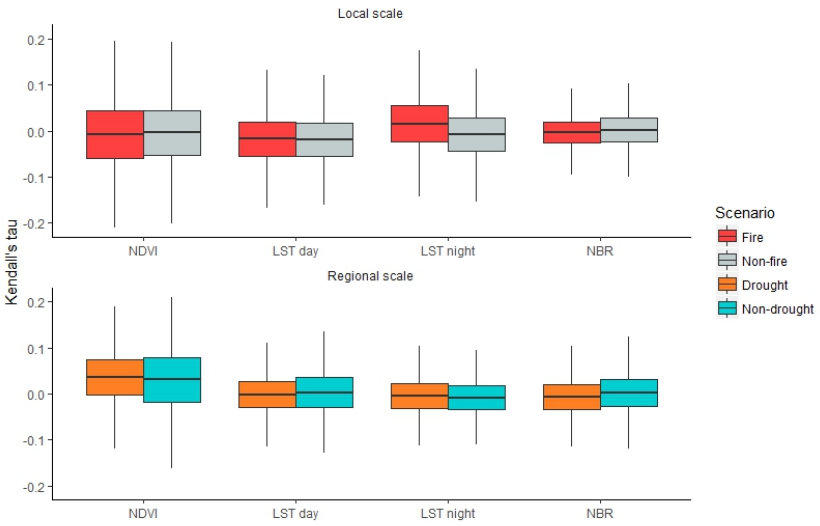


FIGURE 4.1: This figure shows the distribution of fire/drought and non-fire/non-drought pixels for spatial autocorrelation. The line in the middle of the box represents the median, the hinges of the box represent the 1st and 3rd quartiles and the whiskers represent 1.5 of the interquartile range.

of scenario fire/drought and non-fire/non-drought pixels for the LST night time series. The time series for the local scale show variation for both the fire and the non-fire time series, but the time series for the fire areas increases more, which was also visible in the boxplots.

Both of the time series for the regional scale show variation, but the drought time series appears to have higher peaks. The mean kendall's tau for drought and non-drought areas are -0.00387 and -0.00801 respectively with a p-value lower than 0.001, suggesting that the difference between drought and non-drought patches is significant. A less negative kendall's tau for the drought areas suggests that spatial correlation is decreasing less over time while the more negative kendall's tau for non-drought suggests a decrease in spatial correlation over time, which is consistent with theoretical expectations based on CSD.

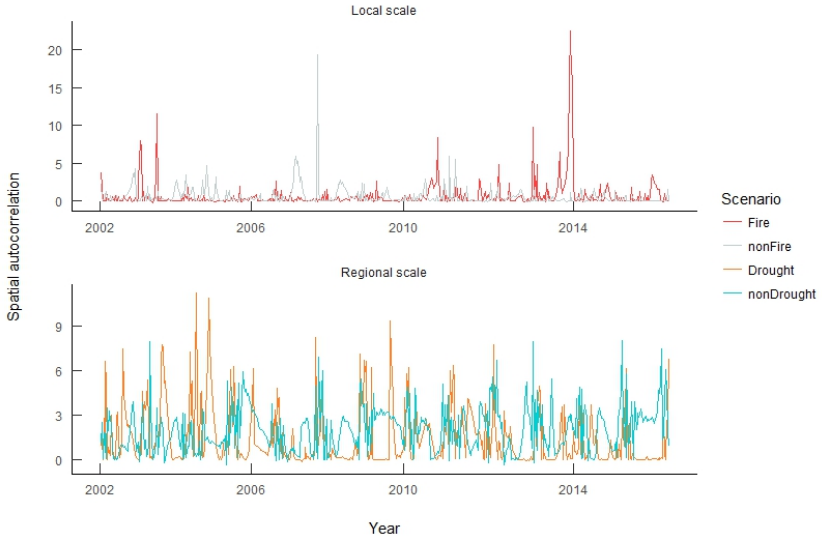


FIGURE 4.2: This figure shows two scenario time series for a fire/drought and non-fire/non-drought pixel based on spatial autocorrelation derived from the LST night dataset.

4.2 Temporal autocorrelation for fire and drought

We hypothesized that next to spatial autocorrelation, temporal autocorrelation would be among the important EWMs in the predictions of fire/drought and non-fire/non-drought areas. Figure 4.3 shows the distribution of fire and non-fire pixels for temporal autocorrelation derived from the LST night, LST day, NDVI and NBR datasets in a boxplot. It can be seen that, only the NBR dataset shows a higher Kendall's tau value for the fire areas. The boxplots for the regional scale show that the Kendall's tau is higher for the LST day and NBR dataset, which is consistent with theoretical expectations. Figure 4.4 shows two scenario time series for a fire/drought and non-fire non-drought pixel derived from LST day datasets. The time series for the local scale appear to show an increase in temporal autocorrelation for the non-fire time series and a decrease in temporal autocorrelation for the fire areas which is confirmed by the Kendall's tau. The mean Kendall's tau for fire and non-fire areas are -0.00313 and 0.08410 respectively with a p-value lower than 0.001, suggesting that the difference between the fire and non-fire patches is significant. A negative Kendall's tau for the fire areas implies that autocorrelation is decreasing while a positive Kendall's tau for the non-fire areas implies that autocorrelation is increasing, which is not consistent with

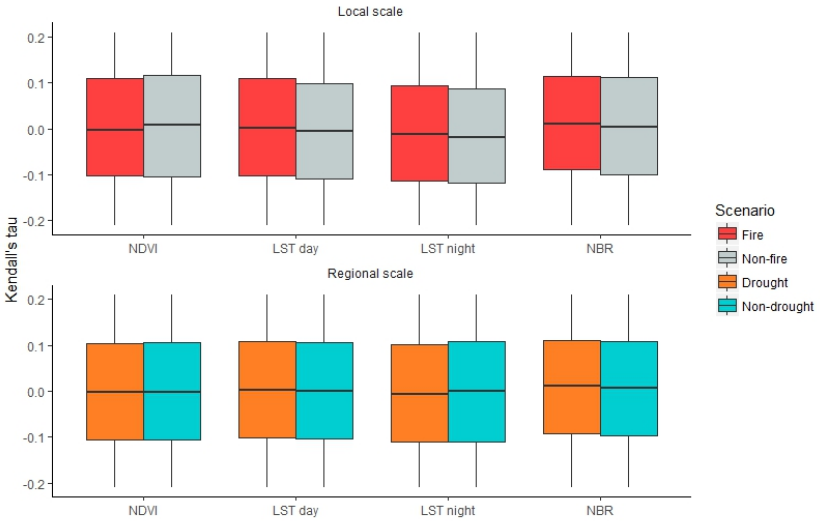


FIGURE 4.3: This figure shows the distribution of fire/drought and non-fire/non-drought pixels for temporal autocorrelation derived from the NDVI, LST day, LST night and NBR datasets.

theoretical expectations based on CSD. See Appendix C for the complete list of mean values for fire and non-fire areas for all the metrics. The time series of the regional scale shows an increase in Kendall's tau for the drought pixels and a decrease in Kendall's tau for the non-drought areas. The mean Kendall's tau for drought and non-drought areas are 0.04731 and -0.01162 respectively with a p-value lower than 0.001, suggesting that the difference between the drought and non-drought patches is significant. A positive Kendall's tau for the drought areas suggests that temporal autocorrelation is increasing while a negative Kendall's tau for the non-drought areas suggests that temporal autocorrelation is decreasing, which is consistent with theoretical expectations based on CSD.

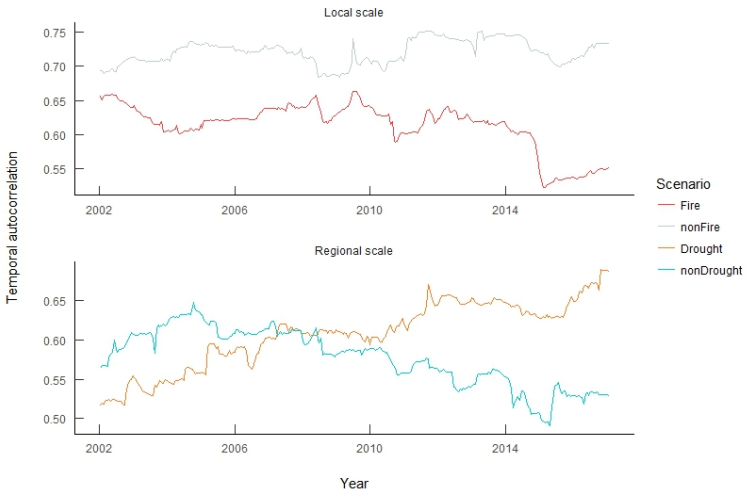


FIGURE 4.4: This figure shows two scenario time series for a fire/drought and non-fire/non-drought pixel based on temporal autocorrelation derived from the LST night dataset.

4.3 Model performance and variable importance analysis

4.3.1 Local scale: fire vulnerability

Variable selection

Table 4.1 shows which EWMs perform consistent with theoretical expectations based on CSD. The LST night derived metrics are most consistent with the expectations while the NDVI derived metrics perform worst: only 2 out of 9 metrics are consistent. The spatial variance metric performs best and is measured according to expectations in 3 different datasets followed up by spatial skewness, temporal variance, temporal skewness and temporal kurtosis. Only the metrics that perform consistent with theoretical expectations based on CSD were used as input variables in the RF EWMs model.

TABLE 4.1: Metrics that are consistent with theoretical expectations based on CSD for the local scale

Metric	NDVI	LST day	LST night	NBR	Sum
Spatial autocorrelation	-	-	X	-	1
Spatial skewness	X	X	-	-	2
Spatial variance	X	-	X	X	3
Temp variance	-	X	-	X	2
Temp skewness	-	X	X	-	2
Temp kurtosis	-	X	X	-	2
Temp CH	-	-	X	-	1
Temp autocorrelation	-	-	-	X	1
Temp return rate	-	-	-	X	1
Sum	2	4	5	4	-

Model predictions

Table 4.2 shows the results of the RF baseline model and the RF EWMs model, respectively. The addition of the spatial and temporal EWMs to the baseline model increases the overall accuracy of the model by 4.71% from 68.52% to 73.23%. The accuracy of the fire areas increased by 0.9% and the accuracy of the non-fire areas increased with 3.9%, suggesting that the addition of the EWMs mainly increases the accuracy of the predictions of the non-fire areas. The error of omission decreased by 3.9% and the error of commission decreased by 0.9%.

TABLE 4.2: Confusion matrices local scale

Prediction RF baseline model

		Fire	Non-Fire	Total
Reference data	Fire	36.7%	18.2%	-
	Non-Fire	13.3%	31.8%	-
	Total	-	-	68.52%

Prediction RF EWMs model

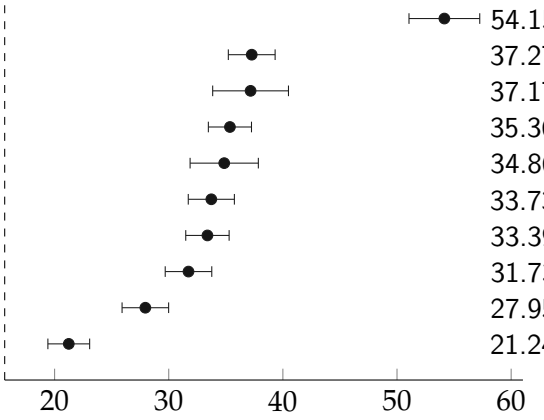
		Fire	Non-Fire	Total
Reference data	Fire	37.6%	14.3%	-
	Non-Fire	12.4%	35.7%	-
	Total	-	-	73.23%

Variable importance

The importance scores of the top 10 predictive variables are shown in table 4.3, see Appendix D for the complete list. The most important predictors are spatial autocorrelation derived from the LST night dataset, non-temporal spatial correlation derived from the LST night dataset and non-temporal spatial variance derived from the NDVI dataset. None of the variables were deemed unimportant. Overall the LST day derived metrics, the LST night derived metrics and the NDVI derived metrics had the highest importance scores per metric, the NBR derived metrics performed less. We hypothesized that temporal autocorrelation and spatial autocorrelation were the top performing EWMs in the predictions. Our findings show that spatial autocorrelation derived from the LST night dataset is the most important EWM in the predictions. The temporal autocorrelation metrics do not perform as expected in our hypothesis.

TABLE 4.3: Mean importance per metric for the RF EWMs model local scale. The EWMs are shown in bold.

Metric	Mean
LST night spatial autocorrelation	54.15
LST night contemporary spatial correlation	37.27
NDVI contemporary spatial variance	37.17
LST day contemporary spatial correlation	35.36
LST night	34.86
LST day contemporary spatial variance	33.73
LST day	33.39
LST night spatial variance	31.73
NDVI	27.95
NDVI contemporary spatial correlation	21.24



4.3.2 Regional scale: drought stress

Variable selection

Table 4.4 gives an overview of the metrics that are consistent with theoretical expectation. The LST day derived metrics performed best, with 6 out of 9 metrics that are consistent with the theory. The second best performing metrics are the NBR derived EWMs. Remarkable is the performance of the LST night dataset, only 2 out of 9 metrics are consistent with theory. The temporal kurtosis metric performs best with 3 out of 4 datasets that show consistent results with theory. Only the metrics that perform consistent with theoretical expectations based on CSD were as input variables in the RF EWMs model.

Model predictions

Table 4.5 shows the results of the RF baseline model and the RF EWMs model predictions. The addition of the spatial and temporal EWMs to the baseline model increases the overall accuracy of the model by 1.48% from 85.69% to 87.17%. The accuracy of the drought areas increased by 0.5% and the accuracy of the non-fire areas increased with 1%, suggesting that the addition of the EWMs mainly increases the accuracy of the predictions of the non-fire

TABLE 4.4: Number of EWMs that are consistent with theoretical expectations based on CSD for the regional scale.

Metric	**NDVI**	**LST day**	**LST night**	**NBR**	**Sum**
Spatial autocorrelation	X	-	X	-	2
Spatial skewness	X	-	-	-	1
Spatial variance	-	-	-	X	1
Temp variance	-	X	X	-	2
Temp skewness	X	X	-	-	2
Temp kurtosis	X	X	-	X	3
Temp CH	-	X	-	X	2
Temp autocorrelation	-	X	-	X	2
Temp return rate	-	X	-	X	2
Sum	4	6	2	5	-

areas. The error of omission decreased by 1% and the error of commission decreased by 0.5%.

TABLE 4.5: Confusion matrices regional scale

Prediction RF baseline model

		Drought	Non-Drought	Total
Reference data	Drought	42.3%	6.6%	-
	Non-Drought	7.7%	43.4%	-
	Total	-	-	85.69%

Prediction RF EWMs model

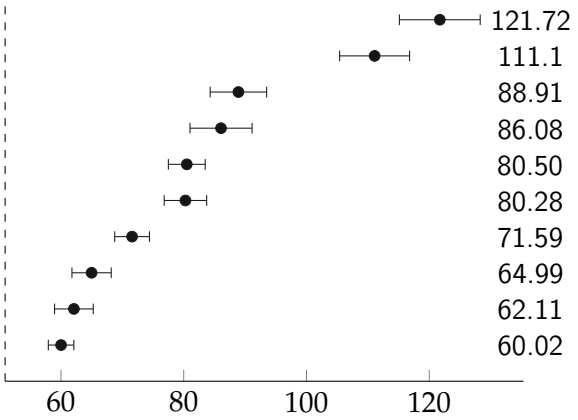
		Drought	Non-Drought	Total
Reference data	Drought	42.8%	5.6%	-
	Non-Drought	7.2%	44.4%	-
	Total	-	-	87.17%

Variable importance

The importance scores of the top 10 predictive variables are shown in table 4.6, see Appendix F for the complete table. The most important predictive variables when introducing noise to the separate variables are LST night, LST day and LST night spatial autocorrelation. None of the variables were deemed unimportant. Overall the LST day derived EWMs and LST night derived EWM had the highest importance scores per metric, while the NDVI derived EWM and the NBR derived EWM performed less. We hypothesized that temporal autocorrelation and spatial autocorrelation would be the most important EWM in the predictions. Our findings show that 2 of the 4 datasets performs consistent with theoretical expectations for the temporal autocorrelation metrics. The spatial autocorrelation is the best performing EWM in the random forest predictions. Our findings are not in line with our hypothesis, we expected temporal autocorrelation to play an important role.

TABLE 4.6: Mean importance per metric for the regional scale RF EWMs model.The EWM are in bold.

Metric	Mean
LST night	121.72
LST day	111.1
LST night spatial autocorrelation	88.91
LST day contemporary spatial variance	86.08
LST night contemporary spatial autocorrelation	80.50
LST day contemporary spatial autocorrelation	80.28
LST day contemporary spatial autocorrelation	71.59
NDVI spatial autocorrelation	64.99
LST night contemporary spatial variance	62.11
NBR contemporary spatial variance	60.02



4.4 Spatial patterns of predictions

Figure 4.5 shows the spatial distribution of the correct and incorrect predicted fire and non-fire areas for both the RF baseline model and the RF EWMs model. In the RF baseline model (figure 4.5 a.1), the false positive predictions within the fire areas were more dispersed then in the RF EWMs model (figure 4.5 a.2). After the addition of the EWMs the false positive predictions in the RF EWMs model are more clustered around the borders of the fire areas, with a few exceptions. It is visible that less non-fire areas are predicted as false negatives compared to the RF baseline model predictions. The non-fire areas that are near the fire areas are predicted as false negatives fewer times compared to the RF baseline model.

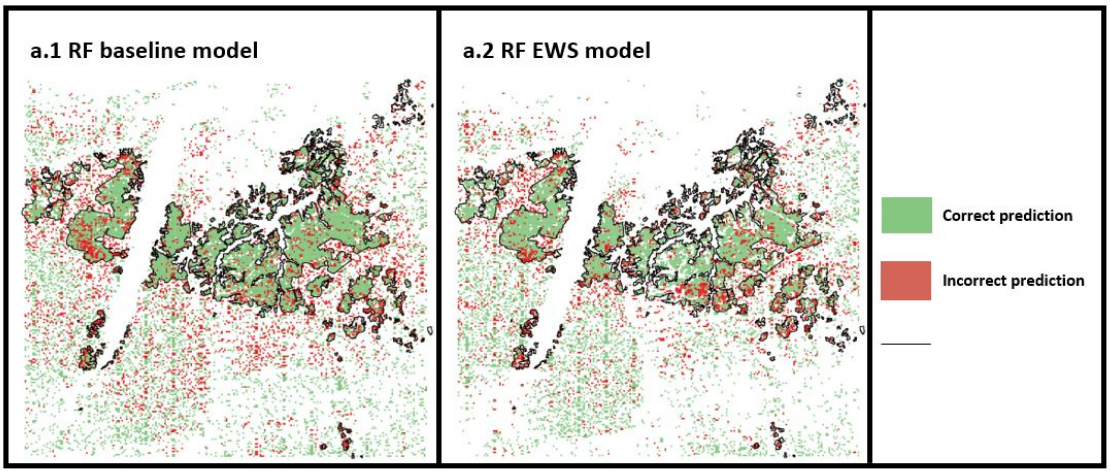


FIGURE 4.5: This figure shows the spatial results of the random forest predictions for the RF baseline model in a.1 and the results for the RF EWMs model in a.2. The correct predicted areas are green while the incorrect predicted areas are red.

Figure 4.6 shows the spatial distribution of the correct and incorrect predicted drought and non-drought areas for both the RF baseline model and the RF EWMs model. Comparing

figure 4.6 a.1 and a.2, there are no big differences visible. The spatial patterns are not changed after the addition of the EWMs.

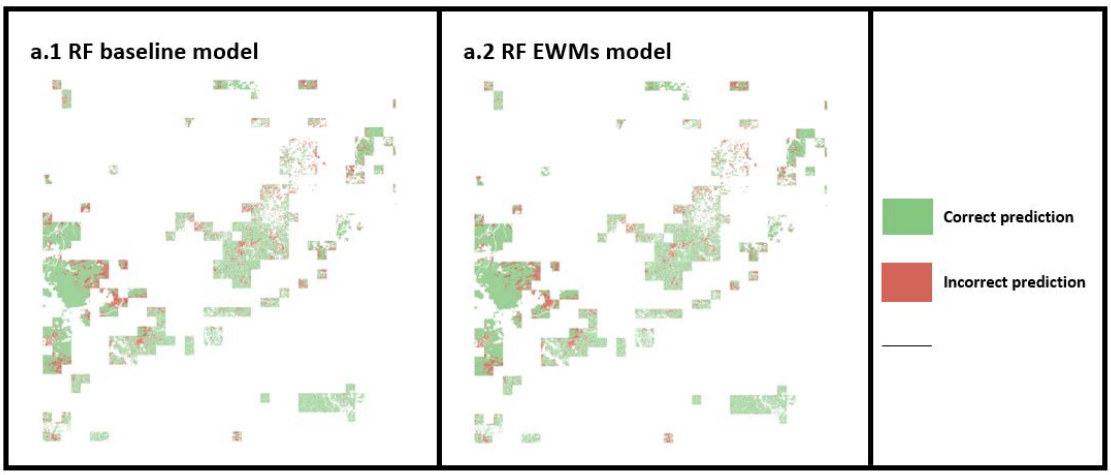


FIGURE 4.6: This figure shows the spatial results of the random forest predictions for the RF baseline model in a.1 and the results for the RF EWMs model in a.2. The correct predicted areas are green while the incorrect predicted areas are red.

4.4.1 Fire vulnerability map

Figure 4.7 shows the map with the areas that are vulnerable to forest fires based on the RF EWMs model in the year 2015, before the fires took place. Most of the vulnerable areas are the same areas as where the forest fires took place. The other vulnerable areas that are not inside the fire polygons are relatively isolated areas.

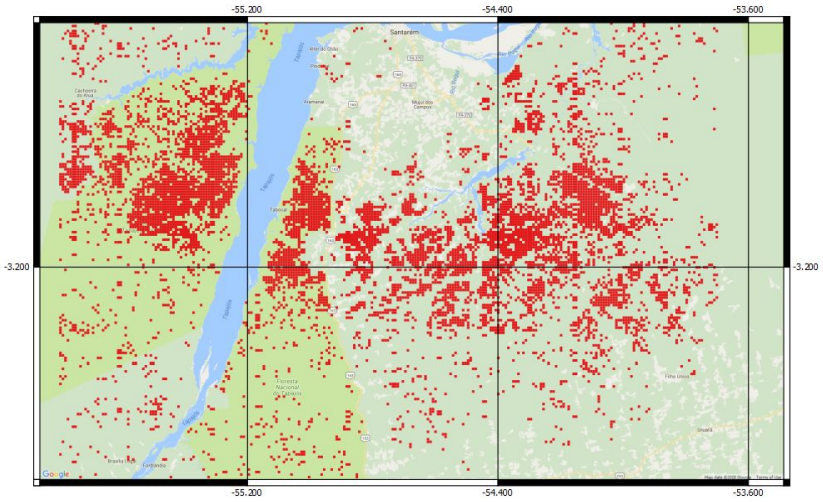


FIGURE 4.7: This figure shows which areas were predicted to be vulnerable to forest fires in 2015 before the actual fires took place.

4.5 Potential of VOD dataset

The VOD dataset is coarse and therefore too little samples were available for the random forest predictions, but the few results that we had showed potential. The time series of the VOD temporal autocorrelation and spatial autocorrelation metrics are shown in figure 4.8. The temporal autocorrelation time series shows a drop for the non-drought time series around 2011, while there is an increase for the drought time series around 2008. Mainly the spatial autocorrelation time series show big differences, with an almost flat line for the non-drought time series. Table 4.7 gives an overview of the metrics that are consistent with theoretical expectations based on CSD. The VOD derived metrics perform best with 7 out of 9 metrics that are consistent with theory. Appendix H shows the relationship of VOD and SPI values.

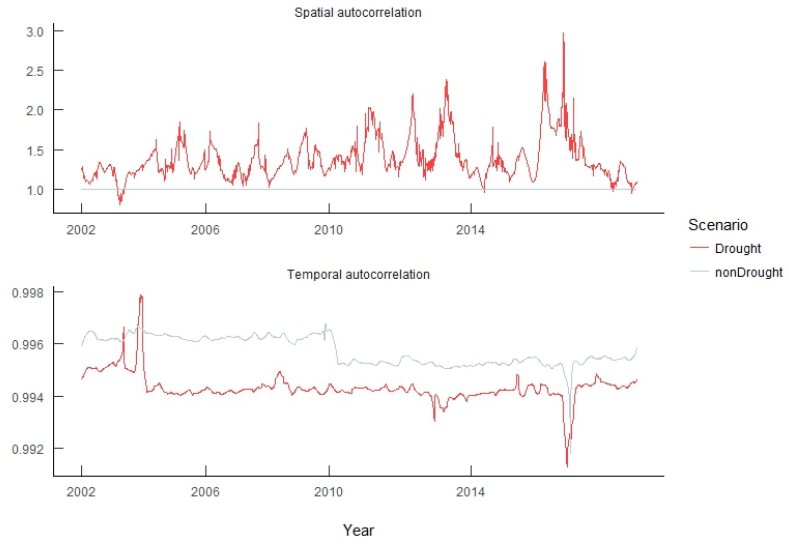


FIGURE 4.8: This figure shows two scenario time series for spatial autocorrelation and temporal autocorrelation derived from the VOD dataset.

TABLE 4.7: Number of EWMs that are consistent with theoretical expectations based on CSD for drought stress, with VOD

Metric	NDVI	LST day	LST night	NBR	VOD	Sum
Spatial autocorrelation	X	-	X	-	X	3
Spatial skewness	X	-	-	-	X	2
Spatial variance	-	-	-	X	X	2
Temp variance	-	X	X	-	X	3
Temp skewness	X	X	-	-	X	3
Temp kurtosis	X	X	-	X	-	3
Temp CH	-	X	-	X	-	2
Temp autocorrelation	-	X	-	X	X	3
Temp return rate	-	X	-	X	X	3
Sum	4	6	2	5	7	-

Chapter 5

Discussion

5.1 Performance of the EWMs

We showed that spatial autocorrelation derived from the LST night and NDVI datasets and spatial variance derived from the LST night dataset are the most important EWMs in the predictions of fire vulnerability and drought stress, see table 4.3 & 4.6. These outcomes are not in line with our hypothesis and findings of Dakos et al. (2012) & Pace et al. (2013) that temporal autocorrelation and variance are the most robust EWMs to identify an impending tipping point. The spatial EWMs have not been applied often in research and it is remarkable that both spatial autocorrelation and spatial variance perform better than the widely tested temporal EWMs (Kefi et al., 2014).

We hypothesized that temporal autocorrelation would be among the most important metrics in the predictions. Our outcomes showed that the metric had little importance in the predictions and was not consistent with theoretical expectations based on CSD. For the local scale only the NBR dataset is consistent with theory and for the regional scale the NBR and LST day datasets are consistent with theory. A potential explanation of the weak performance of the temporal autocorrelation metric is the use of interpolation on the time series. Interpolation of the datasets can cause an artificial rise in temporal autocorrelation, resulting in inaccurate resilience estimates.

Our second hypothesis that spatial autocorrelation would be one of the most important metrics in the predictions is accepted. The spatial autocorrelation derived metrics were the most important metrics for the predictions of both the local and the regional scale. A possible explanation of the good performance of spatial autocorrelation compared to temporal autocorrelation could be that temporal autocorrelation is more sensitive to interpolation and the related artificial rise of autocorrelation. Further research into the effect of interpolation on spatial autocorrelation is needed.

Analysis of temporal kurtosis and temporal skewness showed ambiguous results and had little importance in the predictions of the fire vulnerability and drought stress. Dai et al. (2012) found that skewness did not provide early warnings in a laboratory yeast population undergoing a regime shift (Pace et al., 2013), while Pace et al. (2013), found that the results of kurtosis and skewness were less useful in providing early warnings for the collapse of the zooplankton ecosystem. In this research skewness and kurtosis performed consistent with theoretical expectations for the LST day and LST night datasets for the local scale. Temporal skewnessl was the best performing metric for the regional scale: with 3 out of 4 dataset showing consistent relations with theory. Even though skewness and kurtosis showed consistent signs of CSD the variables did not have a big importance in the predictions, which is in line with findings of Pace et al. (2013) and Dai et al. (2012).

The conditional heteroscedasticity (CH) metric was not successful in this research. The differences between the fire/drought and non-fire/non-drought patches were in most cases not significant, adding little extra accuracy to the predictions. The CH test for structure in variance has proven to be useful and un-useful as an EWM in other studies (seekel et al., 2010). The usefulness of CH mostly depended on the length of the time series which was the source of failure in the zoo-plankton study (Pace et al., 2013). In this study we used time series with a length between 350 and 690 observations, suggesting that even longer time series could lead to a better result. Further research is needed to identify if longer time series are able to differentiate between fire/drought based on conditional hetroscedasticity derived variables.

Return rate (ReRa) mirrored the temporal autocorrelation statistics presented in this study. Therefore we assume that the weak performance of ReRa can also be caused by interpolation of the datasets. Further research is needed to identify is ReRa would perform better in an analysis without interpolation

In general, the LST day- and LST night derived EWMs had the highest mean importance for the predictions, suggesting that drought stress and related fire vulnerability is mainly temperature driven. The LST dataset is an indirect measure of soil moisture conditions and yields a distinctive thermal signature (Anderson et al., 2007). During dry conditions rising leaf temperatures are good indicators of plant moisture stress and precede the onset of drought, this thermal response can occur even when the plants are green (Wan et al., 2004). This suggests that LST datasets detect vegetation stress faster than NDVI and NBR dataset and are therefore more suitable for timely resilience estimates.

The weak performance of some of the EWMs could be explained by the quality of the datasets used to derive the metrics. The NDVI time series are often hindered by noise arising from varying atmospheric conditions and sun-sensor-surface viewing geometries (Hird & Mcdermid, 2009). To be usable as an EWM, system behaviour must be detectable well enough in advance of a tipping point to serve in decision making and be distinguishable from other patterns (Boettiger et al., 2013).The NDVI datasets used in this research are often noisy and sparse due to cloud cover, resulting in NA values. Figure 5.1 shows the cloud cover percentage for the NDVI regional scale dataset for the time period 2002-2015. Some areas have been covered by clouds 80% of the duration of the time series. This results in irregular time series and is a problem for the NDVI, LST and NBR datasets. Removing scenes from the time series entails the risk of reducing the ability of the algorithm to detect change. Therefore, in this research all the areas that have been covered by clouds for more than 50% of the time are removed from the analysis. But still interpolation was needed for the incomplete time series, otherwise not enough data was available for the random forest predictions. Interpolation of the missing values can cause an artificial rise in autocorrelation for both the spatial and temporal variant. Perretti and Munch (2012) identified that under common levels of noise found in data, CSD based EWMs often fail.

Another reason the NDVI dataset performed below expectations could be due to the rate at which a system is sampled. A low sampling rate could potentially cause a missed alarm (Boulton et al., 2013). The sampling rate of the system has to be small enough to capture fast dynamics in a system caused by short term perturbations (Kleinen et al., 2003). The LST datasets had a sampling rate of once every 8-days, while the NDVI datasets are sampled once every 16-days.

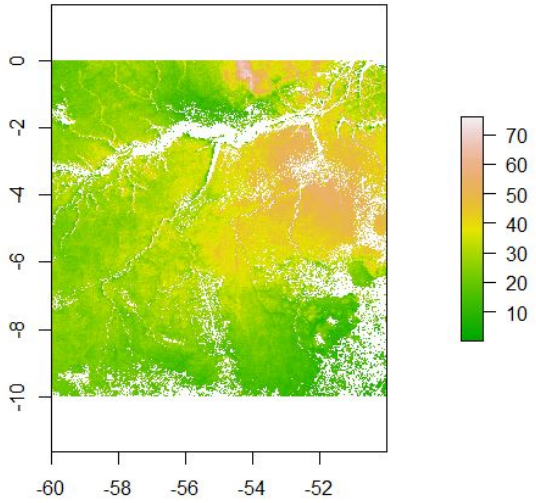


FIGURE 5.1: Cloud cover percentage for the time period 2002-2015.

5.2 Local scale predictions

We hypothesized that the addition of EWMs would increase the accuracy of the predictions and that temporal autocorrelation and spatial autocorrelation would be the most important EWM in the predictions. Our results show that the addition of consistent EWMs increases the accuracy from 68.52% to 72.23%, which is in line with our hypothesis. Remarkable is that the addition of EWMs increases the prediction accuracy for the non-fire areas more than for the fire areas, which was expected otherwise. This could be due to the ambiguity of our validation dataset. The shift from the *"climate-vegetation"* equilibrium to the *"fire dominated tropical savanna"* equilibrium is depicted by the validation dataset fire vs non-fire. This dataset implies that the proximity of a tipping point in the fire dominated areas was very close and that the proximity of a tipping point was distant in the non-fire areas. However, the true distance to a tipping point remains unknown. It could be possible that the whole local study area had a close proximity towards a tipping point, only the ignition missed. Therefore, using the fire vs non-fire dataset on a local scale entails the risk of predicting the proximity of a tipping point in an area in which the proximity was equal.

Looking at the spatial patterns of the predictions a few potential explanations of why the addition of EWMs increased the accuracy of the non-fire areas more are visible. The false negative classification of high fire vulnerability in the non-fire areas could be the isolation of the patches, see the fire vulnerability map in figure 4.5. When large areas are susceptible to forest fires the probability of a forest fire is much higher than when a susceptible area is surrounded by resilient forest patches. The low resilience to forest fires could be accurately predicted by the model, but no fires occurred because the neighbouring areas did have a high resilience. In figure 4.4 it can be seen that most of the predictions errors are made near the borders of the fire areas. It is unlikely that the difference between areas with a low resilience and areas with a high resilience is a sharp line, but more likely a continuous decline of resilience and therefore predictions errors could be expected. Another problem occurring around the edges of the fire and non-fire areas is the use of a kernel for the spatial calculations. Near the edges the kernel calculates skewness, variance and autocorrelation based on both fire and non-fire areas.

5.3 Regional scale predictions

The results of the regional scale are similar to the results of the local scale. The addition of EWMs increased the total accuracy from 85.69% to 87.17%, which is a smaller increase of accuracy than for the local scale. This small increase suggests that drought stress can already be predicted relatively accurately by using contemporary conditions of the ecosystem alone.

The spatial patterns of the predictions are also less profound than the changes on the local scale. The reason for the weaker performance of the EWMs for the regional scale could be found in the validation dataset. The drought vs non-drought validation dataset is based on the SPI index. The SPI was chosen based on its simplicity and its ability to calculate relative drought in situ (Hayes et al., 1999, McKee et al., 1993). In respect to the SPI index itself, it is important to know how well it can reproduce water stress. Currently and as far as our research stretched we could not find a systematic assessment of the power of the SPI index in the context of the tropical rain forest. However, Li et al. (2008) did a 30-year trend study of SPI in the Amazonia to assess how drought conditions develop under a variety of different scenarios of climate change. The outcome of the research was positive in the sense that the SPI was a flexible index, which adapts to local rainfall regimes. On the other hand, it should be mentioned that according to Hayer (1999) and Gutmann (1999) long time series of at least 50 years of data are needed to analyse drought periods, while in the research of Li et al. (2008) a time series of respectively 30-years was used. In this research we only have a time-series of 18 years available. It is not clear if the short time series had an impact on the final predictions of the EWMs. Another point of critique for the SPI was identified by Trenberth et al. (2014): the index is based on precipitation alone, without considering the forest's demand for water. Soil moisture conditions and topography are not taken into account in the calculations (Hayes et al., 1999).

A point of critique on the whole drought vs non-drought concept was found by Verbesselt et al. (2016). Their results showed that elevated temporal autocorrelation as a measure of CSD is sometimes found in places where one would not expect low resilience based on mean annual precipitation alone. Low resilience was found in areas with high annual precipitation levels (Verbesselt et al., 2016). By only taking annual precipitation values into account our validation dataset could potentially point out drought as an area with low resilience wrongly, while non-drought areas could also contain areas of low resilience.

Potential of VOD dataset

We found that the coarse-scaled VOD imagery showed huge potential. 7 Out of 9 EWMs performed consistent with theoretical expectations based on CSD, which is the best score for the five datasets used in this research. Even though too little sample were available for our study area the results show potential for upscaling. Verbesselt et al., (2016) showed that VOD derived temporal autocorrelation was a good indicator of the resilience of ecosystems such as the Amazon.

The VOD dataset represents water content in above ground woody and leaf biomass, suggesting that drought vs non-drought could be well predicted by the above ground water content in the vegetation (Liu et al., 2013). The potential of the VOD dataset could be in the sampling scheme. The sampling rate of a system has to be small enough to capture fast dynamics in a system caused by short term perturbations (Kleinen et al., 2003). The VOD dataset is sampled daily and therefore the chance of missing fast dynamic is a lot smaller

then for the LST dataset (8-days) or the NDVI dataset (16-days). Another explanation of the consistent performance of the VOD metrics is the lack of effect atmospheric and weather conditions have on the VOD datasets (Verbesselt et al., 2016). The LST, NBR and NDVI datasets are constantly influenced by weather conditions. Clouds are a major obstacle in the optical remote sensing of the tropical rainforest of the Amazon (Gu et al., 2012; Gu et al., 2009; Kennedy et al., 2010; Michishita et al., 2014; Poulter & Cramer., 2009). Hence the results of the VOD datasets are much more stable than the results of the other datasets.

The consistent results for the VOD datasets might be explained by the weak correlation to the SPI dataset. Appendix H shows that the datasets have an R-square of -0.1, which indicates that there is a weak negative correlation. The weak correlation between the two datasets could be the reason that the VOD derived EWMs perform well. Further research is needed to identify the magnitude of this minor correlation.

5.4 Limitations of the thesis

5.4.1 Large data volumes

Time series analysis is a powerful tool to identify changes over time. Increasing the length and resolution of time series gives lots of opportunities to identify changes over time, and offers more data to create a robust baseline. But longer and more detailed time series comes at the cost of very large data volumes that have to be (pre-)processed. For example, NDVI imagery from 2002 till 2015 has a download size of around 25 gigabyte. Overall longer and more detailed time series might give better and more accurate results, but suitable hardware is needed to accomplish that.

5.4.2 Reference data generation

Experts gave an average of 20% chance of tipping the Amazon (which means that at least half of its current area is converted from year-round forest to a tropical savanna) if global warming is between 2 and 4 degrees, and 70% if global warming exceeds a rise of 4 degrees (Kriegler et al., 2009). However, the future of the Amazon rainforest is highly uncertain (Boulton et al., 2013). In recent years, research into the field of tipping points and their predictability has yield several suggestions for generic early warning signals of an approaching tipping point (Scheffer et al., 2009, Lenton, 2011). In this research we assumed that the fire vs no-fire and the drought vs non-drought datasets (based on the 2015-2016 El Niño event) depicted the resilience and proximity of a tipping point. This research is at risk of predicting the proximity of a tipping point while the system might not be at risk of tipping.

5.4.3 Definition of resilience

We use the term resilience in the same sense Scheffer (2009) used it: "The ability of a forest to absorb disturbances and re-organize under change to maintain similar functioning and structure". This definition accounts for the fact that a system rarely recovers to the exact same pre- disturbance conditions (Scheffer, 2009). Furthermore, with this definition one can frame resilience as the rate of recovery after a disturbance and as the maximum amount of disturbances a system can absorb before reaching a tipping point (Gunderson, 2000, Scheffer, 2009). In this research the resilience of a system towards a regime shift is measured in statistical metrics based on the CSD principle. However, CSD cannot be used as evidence of regime shifts on its own. In some cases, CSD will be measured but no transition is approaching while in other cases regime shifts occur without CSD (Boettiger et al., 2013).

5.4.4 Potential failure of EWMs

Foundational research in EWS identified certain patterns that may forecast a sudden transition in a wide variety of systems (Scheffer et al., 2009). Subsequently work has identified a growing library of case studies in which the EWS are not present before a transition (Schreiber, 2003, Schreiber & Rudolf, 2008, Hastings & wysham, 2010, Bel et al., 2012, Boettiger et al., 2013).The examples are distinct for the statistical-error case, the signal is present, but too weak to detect due to insufficient available data (Dakos et al., 2008, Scheffer et al., 2009, Perretti & Munch, 2012).

Using generic EWMs one does not take into account the change of background conditions like changing CO2 concentrations and human-driven pressures such as alterations of fire regimes, deforestation, management and their interactions (Reyer et al., 2015). External drivers interacting with internal processes may also cause regime shifts and the best indicator variables are often unknown, suggesting that an absence of EWMs does not imply a high resilience (Pace et al., 2013). One can identify indicators of reduced resilience but EWMs are not capable of predicting the total resilience of a system. The processes to be predicted are too complex and resilience is dependent on heaps of identified and unidentified drivers.

Another reason the detection of EWMs might fail is the use of satellite imagery. Remote sensing registers signals from the entire plot, in which some trees may be dying and others are healthy and even benefiting from neighbouring trees dying (Rogers et al., 2018). The use of aggregated information can send the wrong signal.

5.4.5 Variable selection

In section 3.3.2 a framework to select EWMs that were consistent with theoretical considerations based on CSD was presented. The mean Kendall's tau values were used to determine whether the average trend line showed the expected difference for the fire/drought and non-fire/non-drought time series. When taking the mean of thousands of time series a lot of variation is removed from the analysis. It could be possible that half of the time series was consistent with theoretical considerations and half of it was not. By using this method of variable selection EWMs that are consistent with theory could be excluded from the analysis while they did measure CSD.

Another point of critique on the variable selection method is the interpretation of the negative trend lines. If the trend line for the fire areas was decreasing less than the trend line for the non-fire areas it was assumed that this was a sign of CSD. Stricter rules could have been applied in the variable selection method by only accepting trends that show an increasing Kendall's tau value.

Chapter 6

Conclusions and recommendations

6.1 Conclusions

The overall objective of this thesis was to analyze whether surface temperature and vegetation related EWMs can differentiate between fire versus non-fire conditions on a local scale and drought vs non-drought conditions on a regional scale. With four main research questions:

1. Can spatio-temporal early warning metrics improve the predictions of fire vulnerability estimates in a drought stressed forest ecosystem on a local scale?
2. Can spatio-temporal early warning metrics help differentiate between drought stressed and non-drought stressed forest scenari on a regional scale?

The results of the local scale have proven that fire vulnerability can be predicted with an accuracy of 68.25% in the RF baseline model and with an accuracy of 73.23% in the RF EWMs model. The addition of EWMs that are consistent with theoretical expectations increases the prediction accuracy. The LST derived spatial autocorrelation metric has proven to be the most important metric in the prediction. This suggests that low fire resilience is related to spatial autocorrelated patterns of indirect soil moisture conditions and related vegetation stress.

The results of the regional scale have proven that drought vs non-drought areas can be predicted with an accuracy of 85.69% in the RF baseline model and with an accuracy of 87.17% in the RF EWMs model. The addition of EWM that are consistent with theoretical expectations does increase the prediction accuracy. The LST derived spatial autocorrelation is again the most important EWM in the prediction for the drought vs non-drought areas. This suggests that drought resilience is also related to spatial autocorrelated patterns of indirect soil moisture conditions and related vegetation stress.

This research has shown that future tipping points of climate systems have the potential to be predicted in some cases. Mainly the spatial EWMs show potential. However there are some obstacles preventing EWMs from giving strong signals on the approach to a tipping point. These include noisy datasets, cloud contamination, external perturbations and the unknown distance towards a tipping point.

The addition of EWMs increases the accuracy of the predictions for both the local and the regional scale by 4.98% and 1.64% respectively. These results prove that EWMs help in the understanding of forest fire/drought stress resilience. But critical notes should be made that these additions are relatively small. Further research is needed to evaluate whether EWMs can strengthen existing drought stress and related fire risk frameworks.

6.2 Recommendations

1. This research has shown the potential of the VOD dataset for the identification of EWMs in drought vs non-drought areas. The upcoming higher resolution VOD dataset could be used in further research towards EWMs. In this way enough data-points are available to make a reliable random forest prediction.

2. The coarse resolution of the VOD dataset (0.25 degrees) and the LST dataset (1km) could have been a cause of missed EWS. Using datasets with a high-resolution could improve the predictions. On the other hand Rogers et al. (2018) found that coarse data did not influence their research towards EWMs. However this may still be a matter of further investigation.

3. Another data driven problem is the shortage in usable data. Due to thick cloud covers above the Amazon rainforest time series are incomplete. In further research towards the resilience of forests more complete datasets are needed. This could be accomplished by using datasets based on cloud penetrating wavelength.

4. Using datasets with a higher temporal resolution could also improve the detection of EWMs. In our results in can be seen that the NDVI (16-day product) performed less than the LST (8-day product) and VOD (daily product) dataset. Whether the lower performance is due to temporal resolution has to be assessed in further research. Different datasets like NDVI 8-day products could be used in further research (Landsat).

5. The drought vs non-drought map could be improved in further research. The main improvement could be to use another drought index. The SPI is based on precipitation alone, without considering the forest's demand for water. This could lead to the wrong assumption that areas with a high SPI value are areas with a high resilience and areas with a low SPI value are areas with a low resilience. Verbesselt et al. (2016) found that there is no direct relationship between low precipitation levels and a low resilience. Therefore a different drought index could improve the predictions e.g. a drought index that also takes topography and soil moisture into account. Further research is needed to clarify what difference it would make to use a different drought index.

Chapter 7

References

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., ... & Gonzalez, P. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. Forest ecology and management, 259(4), 660-684.

Anderegg, W. R., Berry, J. A., & Field, C. B. (2012). Linking definitions, mechanisms, and modeling of drought-induced tree death. Trends in plant science, 17(12), 693-700.

Anderegg, W. R., Plavcová, L., Anderegg, L. D., Hacke, U. G., Berry, J. A., & Field, C. B. (2013). Drought's legacy: multiyear hydraulic deterioration underlies widespread aspen forest die-off and portends increased future risk. Global change biology, 19(4), 1188-1196.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A.,& Kustas, W. P. (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation. Journal of Geophysical Research: Atmospheres, 112(D10).

Anselin, L. (1995). Local indicators of spatial association—LISA. Geographical analysis, 27(2), 93-115.

Aragao, L. E. O., Malhi, Y., Roman-Cuesta, R. M., Saatchi, S., Anderson, L. O., & Shimabukuro, Y. E. (2007). Spatial patterns and fire response of recent Amazonian droughts. Geophysical Research Letters, 34(7).

Baumann, M., Ozdogan, M., Kuemmerle, T., Wendland, K.J., Esipova, E., & Radeloff, V.C., (2012). Using the Landsat record to detect forest-cover changes during and after the collapse of the Soviet Union in the temperate zone of European Russia 124, 174–184.

Beguería, S., Vicente-Serrano, S. M., & Beguería, M. S. (2017). Package 'SPEI'.

Beisner, B. E., Haydon, D. T., & Cuddington, K. (2003). Alternative stable states in ecology. Frontiers in Ecology and the Environment, 1(7), 376-382.

Bel G, Hagberg A, & Meron E (2012) Gradual regime shifts in spatially extended ecosystems. Theor Ecol 5:591–604

Bigler, C., Bräker, O. U., Bugmann, H., Dobbertin, M., & Rigling, A. (2006). Drought as an inciting mortality factor in Scots pine stands of the Valais, Switzerland. Ecosystems, 9(3), 330-343.

Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: the charted and uncharted territories. Theoretical ecology, 6(3), 255-264.

Boulton, C. A., Good, P., & Lenton, T. M. (2013). Early warning signals of simulated Amazon rainforest dieback. Theoretical Ecology, 6(3), 373-384.

Brando, P. M., Nepstad, D. C., Davidson, E. A., Trumbore, S. E., Ray, D., & Camargo, P. (2008). Drought effects on litterfall, wood production and belowground carbon cycling in an Amazon forest: results of a throughfall reduction experiment. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1498), 1839-1848.

Brando, P. M., Balch, J. K., Nepstad, D. C., Morton, D. C., Putz, F. E., Coe, M. T., & Alencar, A. (2014). Abrupt increases in Amazonian tree mortality due to drought–fire interactions. Proceedings of the National Academy of Sciences, 111(17), 6347-6352.

Brede, B., Verbesselt, J., Dutrieux, L., & Herold, M. (2015). Performance of the Enhanced Vegetation Index to Detect Inner-annual Dry Season and Drought Impacts on Amazon Forest Canopies. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(7), 337.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brienen, R. J. W., Phillips, O. L., Feldpausch, T. R., Gloor, E., Baker, T. R., Lloyd, J., & Martinez, R. V. (2015). Long-term decline of the Amazon carbon sink. Nature, 519(7543), 344-348.

Canadell, J. G., Le Quéré, C., Raupach, M. R., Field, C. B., Buitenhuis, E. T., Ciais, P., & Marland, G. (2007). Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. Proceedings of the national academy of sciences, 104(47), 18866-18870.

da Costa, A. C. L., Galbraith, D., Almeida, S., Portela, B. T. T., da Costa, M., de Athaydes Silva Junior, J., ... & Phillips, O. L. (2010). Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. New Phytologist, 187(3), 579-591.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., & Totterdell, I. J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. Nature, 408(6809), 184-187.

Dai, L., Vorselen, D., Korolev, K. S., & Gore, J. (2012). Generic indicators for loss of resilience before a tipping point leading to population collapse. Science, 336(6085), 1175-1177.

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., & Scheffer, M. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. PloS one, 7(7), e41010.

Dakos, V., van Nes, E. H., Donangelo, R., Fort, H., & Scheffer, M. (2010). Spatial correlation as leading indicator of catastrophic shifts. Theoretical Ecology, 3(3), 163-174.

Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. Proceedings of the National Academy of Sciences, 105(38), 14308-14312.

Dakos, V., Carpenter, T. C., Lahti, L., & Dakos, M. V. (2015). Package 'earlywarnings'.

Davidson, E. A., de Araújo, A. C., Artaxo, P., Balch, J. K., Brown, I. F., Bustamante, M. M., & Munger, J. W. (2012). The Amazon basin in transition. Nature, 481(7381), 321.

Dirzo, R., & Raven, P. H. (2003). Global state of biodiversity and loss. Annual Review of Environment and Resources, 28(1), 137-167.

Early Warning Signals Toolbox. (2014). What are early warning signals. Available at: http://www.early-warning-signals.org/theory/what-are-early-warning-signals/. [Accessed 21 February 2018].

Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science, 281(5374), 237-240.

Fisher, R. A., Williams, M., COSTA, D., Lola, A., Malhi, Y., da Costa, R. F., ... & Meir, P. (2007). The response of an Eastern Amazonian rain forest to drought stress: results and modelling analyses from a throughfall exclusion experiment. Global Change Biology, 13(11), 2361-2378.

Geary, R. C. (1954). The contiguity ratio and statistical mapping. The incorporated statistician, 5(3), 115-146.

Gregorio, A. (2005). Land cover classification system: classification concepts and user manual: LCCS (No. 8). Food & Agriculture Org.

Gu, J., Fan, D., Jiang, N., & Liu, J. (2012, August). A noise detection method for NDVI time series data based on dixon test. In Agro-Geoinformatics (Agro-Geoinformatics), 2012 First International Conference on (pp. 1-5). IEEE.

Gu, J., Li, X., Huang, C., & Okin, G. S. (2009). A simplified data assimilation method for reconstructing time-series MODIS NDVI data. Advances in Space Research, 44(4), 501-509.

Gullison, R. E., Frumhoff, P. C., Canadell, J. G., Field, C. B., Nepstad, D. C., Hayhoe, K., & Nobre, C. (2007). Tropical forests and climate policy. Science- New York ,316(5827), 985.

Gunderson, L. H. (2000). Ecological resilience—in theory and application. Annual review of ecology and systematics, 31(1), 425-439.

Gustafson, E. J., & Sturtevant, B. R. (2013). Modeling forest mortality caused by drought stress: implications for climate change. Ecosystems, 16(1), 60-74.

Guttal, V., & Jayaprakash, C. (2008). Changing skewness: an early warning signal of regime shifts in ecosystems. Ecology letters, 11(5), 450-460.

Guttman, N. B. (1999). Accepting the standardized precipitation index: a calculation algorithm. JAWRA Journal of the American Water Resources Association, 35(2), 311-322.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., & Kommareddy, A. (2013). High-resolution global maps of 21st-century forest cover change. science, 342(6160), 850-853.

Hartmann, H., Adams, H. D., Anderegg, W. R., Jansen, S., & Zeppel, M. J. (2015). Research frontiers in drought-induced tree mortality: crossing scales and disciplines. New Phytologist, 205(3), 965-969.

Hastings, A., & Wysham, D. B. (2010). Regime shifts in ecological systems can occur with no warning. Ecology letters, 13(4), 464-472.

Hayes, M. J., Svoboda, M. D., Wilhite, D. A., & Vanyarkho, O. V. (1999). Monitoring the 1996 drought using the standardized precipitation index. Bulletin of the American meteorological society, 80(3), 429-438.

Hefeeda, M., & Bagheri, M. (2007). Wireless sensor networks for early detection of forest fires. In Mobile adhoc and sensor systems, 2007. MASS 2007. IEEE international conference on (pp. 1-6). IEEE.

Hird, J. N., & McDermid, G. J. (2009). Noise reduction of NDVI time series: An empirical comparison of selected techniques. Remote Sensing of Environment, 113(1), 248-258.

Hirota, M., Holmgren, M., Van Nes, E. H., & Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. Science, 334(6053), 232-235.

Huybrechts, P., & de Wolde, J. (1999). The dynamic response of the Greenland and Antarctic ice sheets to multiple-century climatic warming. Journal of Climate, 12(8), 2169-2188.

Jiménez-Muñoz, J. C., Mattar, C., Barichivich, J., Santamaría-Artigas, A., Takahashi, K., Malhi, Y., ... & Van Der Schrier, G. (2016). Record-breaking warming and extreme drought in the Amazon rainforest during the course of El Niño 2015–2016. Scientific reports, 6, 33130.

Jump, A. S., Ruiz-Benito, P., Greenwood, S., Allen, C. D., Kitzberger, T., Fensham, R., ... & Lloret, F. (2017). Structural overshoot of tree growth with climate variability and the global spectrum of drought-induced forest dieback. Global change biology, 23(9), 3742-3757.

Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., & Dakos, V. (2014). Early warning signals of ecological transitions: methods for spatial patterns. PloS one, 9(3), e92097.

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sensing of Environment, 114(12), 2897-2910.

Kleinen, T., Held, H., & Petschel-Held, G. (2003). The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. Ocean Dynamics, 53(2), 53-63.

Kriegler, E., Hall, J. W., Held, H., Dawson, R., & Schellnhuber, H. J. (2009). Imprecise probability assessment of tipping points in the climate system. Proceedings of the national Academy of Sciences, 106(13), 5041-5046.

Kuhn, M. (2017). Caret package. Journal of Statistical Software, 28(5), 1-26.

Kursa, B. M, Rudnicki, W. R. (2017). Boruta package.

Lenton, T. M. (2011). Early warning of climate tipping points. Nature Climate Change, 1(4), 201.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., & Schellnhuber, H. J. (2008). Tipping elements in the Earth's climate system. Proceedings of the national Academy of Sciences, 105(6), 1786-1793.

Li, W., Fu, R., Juarez, R. I. N., & Fernandes, K. (2008). Observed change of the standardized precipitation index, its potential cause and implications to future climate change in the Amazon region. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363(1498), 1767-1772.

Lingtong, D. U., Qingjiu, T. I. A. N., Yan, H., & Jun, L. (2012). Drought monitoring based on TRMM data and its reliability validation in Shandong province. Transactions of the Chinese Society of Agricultural Engineering, 2012(2).

Liu, Y. Y., Dijk, A. I., McCabe, M. F., Evans, J. P., & Jeu, R. A. (2013). Global vegetation biomass change (1988–2008) and attribution to environmental and human drivers. Global ecology and biogeography, 22(6), 692-705.

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In Proceedings of the 8th Conference on Applied Climatology (Vol. 17, No. 22, pp. 179-183). Boston, MA: American Meteorological Society.

Meir, P., Metcalfe, D. B., Costa, A. C. L., & Fisher, R. A. (2008). The fate of assimilated carbon during drought: impacts on respiration in Amazon rainforests. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363(1498), 1849-1855.

Michaelian M, Hogg EH, Hall RJ, Arsenault E (2011) Massive mortality of aspen following severe drought along the southern edge of the Canadian boreal forest. Global Change Biology, 17, 2084–2094.

Michishita, R., Jin, Z., Chen, J., & Xu, B. (2014). Empirical comparison of noise reduction techniques for NDVI time-series based on a new measure. ISPRS Journal of Photogrammetry and Remote Sensing, 91, 17-28.

Miller, J. D., & Thode, A. E. (2007). Quantifying burn severity in a heterogeneous landscape with a relative version of the delta Normalized Burn Ratio (dNBR). Remote Sensing of Environment, 109(1), 66-80. ISO 690

Mitchell, P. J., O'Grady, A. P., Tissue, D. T., White, D. A., Ottenschlaeger, M. L., & Pinkard, E. A. (2013). Drought response strategies define the relative contributions of

hydraulic dysfunction and carbohydrate depletion during tree mortality. New Phytologist, 197(3), 862-872.

Moran, P. A. (1950). A test for the serial independence of residuals. Biometrika, 37(1/2), 178-181.

Naresh Kumar, M., Murthy, C. S., Sesha Sai, M. V. R., & Roy, P. S. (2009). On the use of Standardized Precipitation Index (SPI) for drought intensity assessment. Meteorological applications, 16(3), 381-389.

Nepstad, D. C., Stickler, C. M., Soares-Filho, B., & Merry, F. (2008). Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363(1498), 1737-1746.

Nepstad, D. C., Tohver, I. M., Ray, D., Moutinho, P., & Cardinot, G. (2007). Mortality of large trees and lianas following experimental drought in an Amazon forest. Ecology, 88(9), 2259-2269.

Nepstad, D., Carvalho, G., Barros, A. C., Alencar, A., Capobianco, J. P., Bishop, J., & Prins, E. (2001). Road paving, fire regime feedbacks, and the future of Amazon forests. Forest ecology and management, 154(3), 395-407.

Nobre, C. A., & Borma, L. D. S. (2009). 'Tipping points' for the Amazon forest. Current Opinion in Environmental Sustainability, 1(1), 28-36.

Nobre, C. A., Sellers, P. J., & Shukla, J. (1991). Amazonian deforestation and regional climate change. Journal of Climate, 4(10), 957-988.

Oštir, K., Veljanovski, T., Podobnikar, T., & Stančič, Z. (2003). Application of satellite remote sensing in natural hazard management: the Mount Mangart landslide case study. International Journal of Remote Sensing, 24(20), 3983-4002.

Pace, M. L., Carpenter, S. R., Johnson, R. A., & Kurtzweil, J. T. (2013). Zooplankton provide early warnings of a regime shift in a whole lake manipulation. Limnology and Oceanography, 58(2), 525-532.

Pegurier, E. & Vieira, P. (2016). Incêndios na região de Santarém em 2015 superam desmatamento em toda a Amazônia. Last visited at 09-06-2017. From: http://www.oeco.org.br/reportagens/incendios-na-regiao-de-santarem-em-2015-degradaram-74-mil-km2/

Perretti, C. T., & Munch, S. B. (2012). Regime shift indicators fail under noise levels commonly observed in ecological systems. Ecological Applications, 22(6), 1772-1779.

Poulter, B., & Cramer, W. (2009). Satellite remote sensing of tropical forest canopies and their seasonal dynamics. International Journal of Remote Sensing, 30(24), 6575-6590.

Rahmstorf, S., & Ganopolski, A. (1999). Long-term global warming scenarios computed with an efficient coupled climate model. Climatic change, 43(2), 353-367.

Reyer, C. P., Brouwers, N., Rammig, A., Brook, B. W., Epila, J., Grant, R. F., & Medlyn, B.

(2015). Forest resilience and tipping points at different spatio-temporal scales: approaches and challenges. Journal of Ecology, 103(1), 5-15.

Rogers, B. M., Solvik, K., Hogg, E. H., Ju, J., Masek, J. G., Michaelian, M., ... & Goetz, S. J. (2018). Detecting early warning signals of tree mortality in boreal North America using multi-scale satellite data. Global change biology.

Salisbury, S. (2017). Record heat and drought seen in Amazon during 2015-16 El Nino. Last visited at 09-06-2017. From: https://news.mongabay.com/2017/01/record-heat-and-drought-seen-in-amazon-during-2015-16-el-nino/

Samanta, A., Knyazikhin, Y., Xu, L., Dickinson, R. E., Fu, R., Costa, M. H., & Myneni, R. B. (2012). Seasonal changes in leaf area of Amazon forests from leaf flushing and abscission. Journal of Geophysical Research: Biogeosciences, 117(G1).

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., ... & Sugihara, G. (2009). Early-warning signals for critical transitions. Nature, 461(7260), 53.

Scheffer, M. (2009). Critical transitions in nature and society. Princeton University Press.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. Nature, 413(6856), 591.

Schreiber SJ (2003) Allee effects, extinctions, and chaotic transients in simple population models. Theor Popul Biol 64:201–209

Schreiber, S., & Rudolf, V. H. (2008). Crossing habitat boundaries: coupling dynamics of ecosystems through complex life cycles. Ecology letters, 11(6), 576-587.

Seekell, D. A., Carpenter, S. R., & Pace, M. L. (2011). Conditional heteroscedasticity as a leading indicator of ecological regime shifts. The American Naturalist, 178(4), 442-451

Soares-Filho, B. S., Nepstad, D. C., Curran, L. M., Cerqueira, G. C., Garcia, R. A., Ramos, C. A., & Schlesinger, P. (2006). Modelling conservation in the Amazon basin. Nature, 440(7083), 520.

Staver, A. C., & Hansen, M. C. (2015). Analysis of stable states in global savannas: is the CART pulling the horse?–a comment. Global ecology and biogeography, 24(8), 985-987.

Trenberth, K. E., Dai, A., Van Der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. Nature Climate Change, 4(1), 17.

Tucker, C. J., Elgin Jr, J. H., McMurtrey Iii, J. E., & Fan, C. J. (1979). Monitoring corn and soybean crop development with hand-held radiometer spectral data. Remote Sensing of Environment, 8(3), 237-248.

Verbesselt, J., Umlauf, N., Hirota, M., Holmgren, M., Van Nes, E. H., Herold, M., ... & Scheffer, M. (2016). Remotely sensed resilience of tropical forests. Nature Climate Change, 6(11), 1028.

Van Wagtendonk, J. W., Root, R. R., & Key, C. H. (2004). Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. Remote Sensing of Environment, 92(3), 397-408.

Wan, Z., Wang, P., & Li, X. (2004). Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the southern Great Plains, USA. International Journal of Remote Sensing, 25(1), 61-72.

Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., ... & Dean, J. S. (2013). Temperature as a potent driver of regional forest drought stress and tree mortality. Nature Climate Change, 3(3), 292.

Zhou, L., Tian, Y., Myneni, R. B., Ciais, P., Saatchi, S., Liu, Y. Y., & Hwang, T. (2014). Widespread decline of Congo rainforest greenness in the past decade. Nature, 509(7498), 86-90

Appendix A

SPI background

The SPI was introduced by McKee et al. (1993) as a tool to identify and monitor drought for multiple time scales. The time scales used reflects different changes. Soil moisture conditions respond to precipitation anomalies on a short time scale while groundwater, streamflow and reservoir storage are reflected in long-term precipitation anomalies. McKee originally calculated the SPI with a moving window of 3-, 6-, 12-, 24,- and 48- months (McKee et al., 1993). The SPI function uses the months within the moving window and historical data and fits it to a gamma distribution (Beguería et al., 2017). The SPI can be interpreted in terms of occurrence probability of a given precipitation amount compared to historical data, which makes it a relative drought index (Brede et al., 2015). Values from -1.5 onwards are considered very dry and values from 1.5 onwards are considered very wet, see table A.1.

For this research a moving window of 6 months was used to calculate the SPI index for the Amazonia. The SPI was calculated for the dry season before occurrences of the fires. We assumed the dry season to be between July and September, a simplification often used by other studies (Samanta et al., 2012). The 10% of the lowest SPI pixel values were used to identify drought stress and the 10% of the highest SPI pixel values were used to identify non-drought stress. The TRMM dataset was used to calculate the SPI and stores the precipitation in millimetres per hour-1. The cumulative monthly precipitation was estimated in mm per month considering a 30-day month for all the dataset (Aragao et al., 2007). A disadvantage of using TRMM datasets to calculate the SPI is the need of long time-series (Hayes, 1999). Guttman (1999) recommends a time-series of at least 50 years of data to analyse drought periods. However other studies used shorter time periods to calculate the SPI: Lingtong et al. (2012) and Naresh Kumar et al. (2009) used respectively 12- and 30-years. Unfortunately, in this research we only had a time-series of 18 years available for the calculations of the SPI.

TABLE A.1: SPI classification (McKee et al., 1993)

SPI value	Category
>2.0	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
-0.99 to 0.99	Near normal
-1.0 to -1.49	Moderately dry
-1.5 to -1.99	Very dry
<-2.0	Extremely dry

Appendix B

Window sizes

B.1 Sensitivity analysis of the window size on temporal EWMs for the local scale

The estimations of the metrics were influenced by the choices made in the application of the metrics. (i) The rolling window size could influence the results and therefore window sizes of 25%, 50% and 75% are used in the calculations. (ii) Regular time series are of importance in the case of indicators such as autocorrelation that estimate the memory in time series. Interpolating the missing values in time series could cause spurious correlations, therefore interpolated time series are often checked against the original time series (Dakos et al., 2008). Unfortunately, the datasets used in this research make it impossible to compare the interpolated time series against the original time series. Due to thick cloud covers it is impossible to select enough complete time series to compare the interpolated data against.

Table B.1 shows the results of the sensitivity analysis per dataset when using a window size of 25%, 50% and 75% to calculate the temporal EWMs. The sensitivity analysis is performed on 10% of the total data per dataset (5% fire areas and 5% non-fire areas).The performance of the LST day and night dataset increases when using a bigger window size. The NBR dataset performs best when using a window size of 25% and the predictive performance decreases when using a window size of 50%. The NDVI dataset is relatively stable in its predictive performance when changing the window sizes.

B.2 Sensitivity analysis of the window size on temporal EWMs for the regional scale

Table B.2 show the results of the sensitivity analysis per dataset when using a window size of 25%, 50% and 75% to calculate the temporal EWMs. The sensitivity analysis is performed on 10% of the total data per dataset (5% fire areas and 5% non-fire areas). The performance of the LST day dataset is best at a window size of 25% and decreases when using a bigger window size. The performance of the LST day dataset increases when using a bigger window size. The NBR dataset performs best when using a window size of 50% and performs worst of all datasets when using a window size of 75%. The performance of the NDVI dataset is relatively stable when changing the window sizes.

B.3 Recommendations for further research

A recommendation for further research is to experiment with different window sizes. Each dataset has a different temporal resolution: NDVI (16- day product), NBR (8-day product), LST (8-day product) and VOD (daily product. Using datasets with different temporal resolutions could mean that using one temporal window size could be beneficial for one dataset and not for the other.

TABLE B.1: Sensitivity analysis for the local scale on window sizes

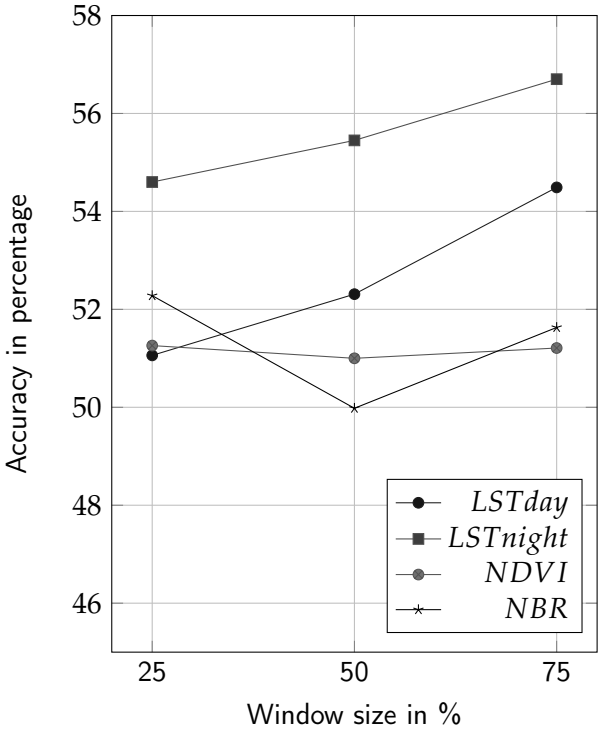
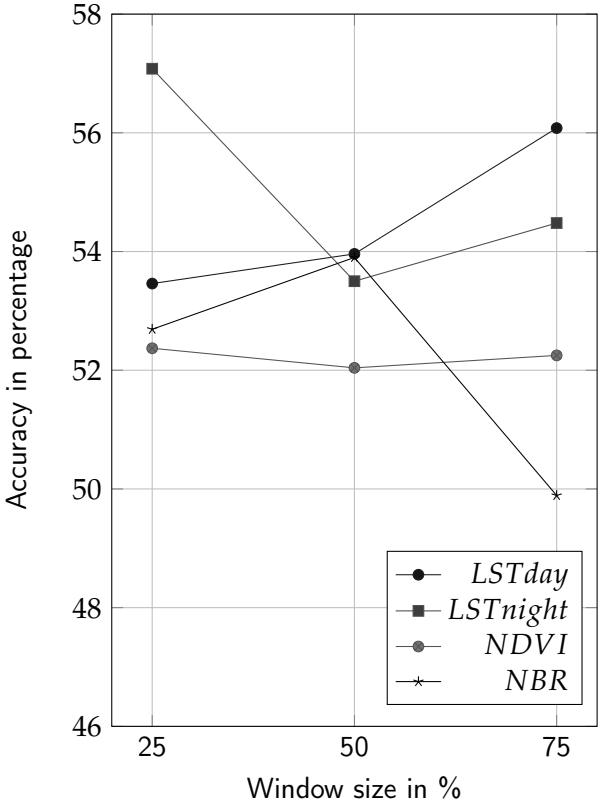


TABLE B.2: Sensitivity analysis for the regional scale on window sizes



Appendix C

Mean values per metric local scale

TABLE C.1: Significance test for fire and non-fire pixels

Metric per index	μ fire	μ non-fire	*p*-value
NDVI	8545.942	8531.996	<0.001
NDVI contemporary spatial correlation	0.57802	0.45196	<0.001
NDVI contemporary spatial skewness	-0.42826	-0.37139	<0.001
NDVI contemporary spatial variance	5014.644	3453.776	<0.001
NDVI spatial correlation	-0.00747	-0.00671	0.06202
NDVI spatial skewness	-0.0062	-0.004979	<0.001
NDVI spatial variance	0.01803	0.01493	<0.001
NDVI temp variance	0.00072	-0.00743	<0.001
NDVI temp skewness	-0.00355	-0.01248	<0.001
NDVI temp kurtosis	-0.00439	-0.02042	<0.001
NDVI temp conditional hetroscedasticity	-0.00664	-0.001434	0.4698
NDVI temp autocorrelation	-0.02015	-0.00183	<0.001
NDVI temp return rate	0.02015	0.00180	<0.001
LST day	300.3892	0.300.6382	<0.001
LST day contemporary spatial correlation	-0.00313	-0.08410	<0.001
LST day contemporary spatial skewness	0.05372	0.08011	<0.001
LST day contemporary spatial variance	0.00212	0.00280	<0.001
LST day spatial correlation	-0.01785	-0.01564	<0.001
LST day spatial skewness	0.00822	0.00681	<0.001
LST day spatial variance	-0.00146	0.00900	<0.001
LST day temp variance	0.03536	0.02958	<0.001
LST day temp skewness	-0.04363	-0.01922	<0.001
LST day temp kurtosis	-0.07265	-0.02394	<0.001
LST day temp conditional hetroscedasticity	-0.02190	-0.02292	0.6602
LST day temp autocorrelation	-0.00313	0.08410	<0.001
LST day temp return rate	0.00313	-0.08410	<0.001
LST night	295.9961	295.8813	<0.001
LST night contemporary spatial correlation	0.17643	0.35644	<0.001
LST night contemporary spatial skewness	0.03714	0.03286	0.1556
LST night contemporary spatial variance	0.00083	0.00114	<0.001
LST night spatial correlation	0.01605	-0.00431	<0.001
LST night spatial skewness	-0.00044	-0.00068	0.3912
LST night spatial variance	0.00222	-0.01879	<0.001
LST night temp variance	0.05564	0.06400	<0.001

Continued on next page

Table C.1 – continued from previous page

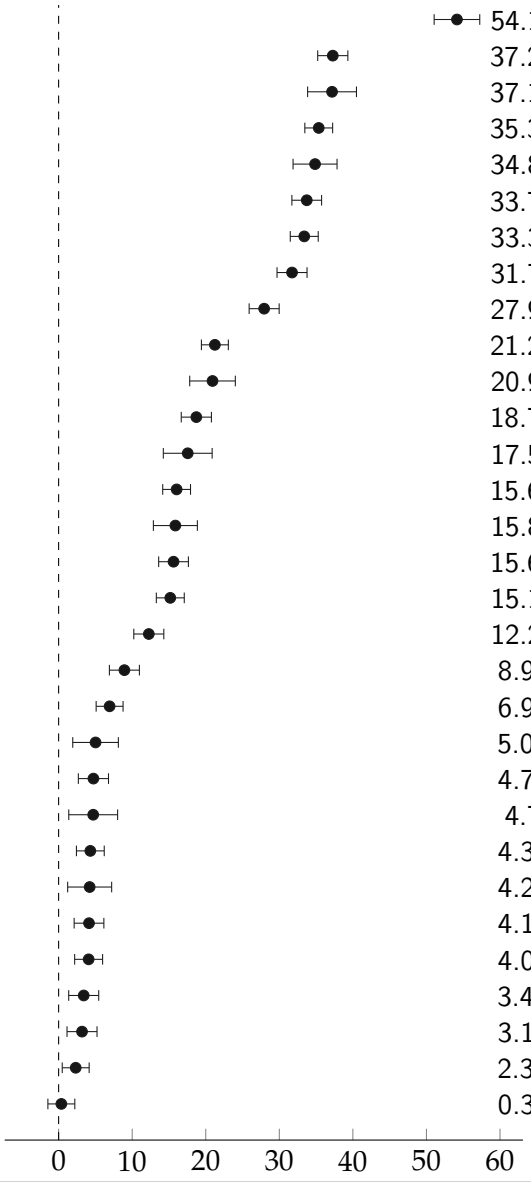
Metric per index	μ fire	μ non-fire	*p*-value
LST night temp skewness	-0.11546	-0.07331	<0.001
LST night temp kurtosis	0.12879	0.08719	<0.001
LST night temp conditional hetroskedasticity	-0.06689	-0.10548	<0.001
LST night temp autocorrelation	-0.30101	-0.20170	<0.001
LST night temp return rate	0.30101	0.20170	<0.001
NBR	0.73306	0.718017	<0.001
NBR contemporary spatial correlation	0.09082	1.10519	<0.001
NBR contemporary spatial skewness	-0.11156	-0.11634	0.1334
NBR contemporary spatial variance	0.000002	0.00002	<0.001
NBR spatial correlation	-0.00265	0.00216	<0.001
NBR spatial skewness	0.00517	0.00759	<0.001
NBR spatial variance	0.02115	0.01915	<0.001
NBR temp variance	-0.01146	-0.02206	<0.001
NBR temp skewness	-0.18965	-0.18928	0.8785
NBR temp kurtosis	-0.13799	-0.15060	<0.001
NBR temp conditional hetroskedasticity	0.0538	0.04678	0.1528
NBR temp autocorrelation	0.15189	0.11884	<0.001
NBR temp return rate	-0.15189	-0.11883	<0.001

Appendix D

Mean importance local scale

TABLE D.1: Continuation of table 4.4

Metric	Mean
LST night spatial moran	54.15
LST night contemporary spatial correlation	37.27
NDVI contemporary spatial variance	37.17
LST day contemporary spatial correlation	35.36
LST night	34.86
LST day contemporary spatial variance	33.73
LST day	33.39
LST night spatial variance	31.73
NDVI	27.95
NDVI contemporary spatial correlation	21.24
LST night temporal skewness	20.92
LST night contemporary spatial variance	18.72
NBR spatial variance	17.55
NBR contemporary spatial correlation	15.62
LST night temporal kurtosis	15.87
NBR contemporary spatial variance	15.62
NBR	15.19
LST night temporal CH	12.26
LST day temporal variance	8.94
LST day temporal kurtosis	6.93
NBR contemporary spatial skewness	5.02
NBR temporal variance	4.73
LST day contemporary spatial skewness	4.7
LST day temporal skewness	4.31
NDVI spatial variance	4.22
NBR temporal autocorrelation	4.13
NBR temporal return rate	4.08
LST night contemporary spatial skewness	3.41
LST day spatial skewness	3.18
NDVI contemporary spatial skewness	2.33
NDVI spatial skewness	0.37



Appendix E

Mean values per metric regional scale

TABLE E.1: Significance test for drought and non-drought pixels

Metric per index	μ drought	μ non-drought	p-value
NDVI	8439.435	8491.193	<0.001
NDVI contemporary spatial correlation	0.11103	0.14655	<0.001
NDVI contemporary spatial skewness	-0.29943	-0.35267	<0.001
NDVI contemporary spatial variance	0.11103	0.14655	<0.001
NDVI spatial correlation	0.03196	0.026911	<0.001
NDVI spatial skewness	-0.00526	-0.00324	<0.001
NDVI spatial variance	0.00204	0.02097	<0.001
NDVI temp variance	-0.01934	-0.01870	<0.001
NDVI temp skewness	0.06434	-0.00267	<0.001
NDVI temp kurtosis	0.03193	-0.02505	<0.001
NDVI temp conditional hetroscedasticity	0.02439	0.03371	<0.001
NDVI temp autocorrelation	-0.08181	-0.02867	<0.001
NDVI temp return rate	0.08164	0.02864	<0.001
LST day	15073.33	15080.37	<0.001
LST day contemporary spatial correlation	0.32905	0.35691	<0.001
LST day contemporary spatial skewness	0.02261	0.07048	<0.001
LST day contemporary spatial variance	3.13909	4.63065	<0.001
LST day spatial correlation	-0.00189	0.00582	<0.001
LST day spatial skewness	0.00422	0.00856	<0.001
LST day spatial variance	-0.00285	0.5797432	<0.001
LST day temp variance	0.04468	0.04359	<0.001
LST day temp skewness	-0.07392	0.00867	<0.001
LST day temp kurtosis	-0.14383	-0.06236	<0.001
LST day temp conditional hetroscedasticity	-0.029618	-0.03241	0.0136
LST day temp autocorrelation	0.04731	-0.01162	<0.001
LST day temp return rate	-0.04731	0.01161	<0.001
LST night	14781.39	14749.30	<0.001
LST night contemporary spatial correlation	0.45063	0.33086	<0.001
LST night contemporary spatial skewness	0.06705	0.10704	<0.001
LST night contemporary spatial variance	3.11676	3.44870	<0.001
LST night spatial correlation	-0.00387	-0.00801	<0.001
LST night spatial skewness	0.00231	-0.00301	<0.001
LST night spatial variance	-0.01013	-0.00931	<0.001
LST night temp variance	0.08999	0.08519	<0.001

Continued on next page

Table E.1 – continued from previous page

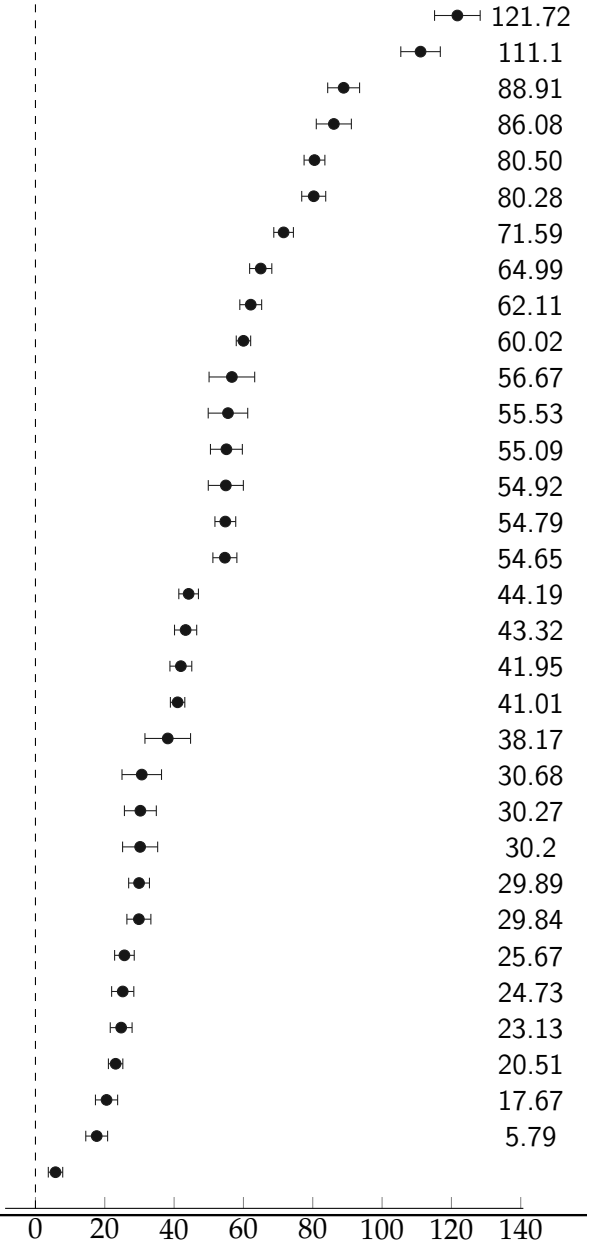
Metric per index	μ drought	μ non-drought	*p*-value
LST night temp skewness	-0.02949	0.08962	<0.001
LST night temp kurtosis	0.10662	0.17224	<0.001
LST night temp conditional hetroscedasticity	0.02439	0.03371	<0.001
LST night temp autocorrelation	-0.08795	-0.05266	<0.001
LST night temp return rate	0.08795	0.05266	<0.001
NBR	0.71487	0.71058	<0.001
NBR contemporary spatial correlation	0.60803	0.45992	<0.001
NBR contemporary spatial skewness	-0.13082	-0.17729	<0.001
NBR contemporary spatial variance	0.000007	0.000006	<0.001
NBR spatial correlation	-0.00624	0.00264	<0.001
NBR spatial skewness	0.00476	0.00612	<0.001
NBR spatial variance	0.01635	0.01296	<0.001
NBR temp variance	-0.10012	-0.06169	<0.001
NBR temp skewness	-0.00608	-0.00949	<0.001
NBR temp kurtosis	-0.07075	-0.02221	<0.001
NBR temp conditional hetroscedasticity	0.09645	0.07937	<0.001
NBR temp autocorrelation	0.07493	0.03809	<0.001
NBR temp return rate	-0.07493	-0.03808	<0.001
VOD	1.05152	0.81513	<0.001
VOD contemporary spatial correlation	0.45062	0.33086	<0.001
VOD contemporary spatial skewness	0.06705	0.10704	<0.001
VOD contemporary spatial variance	3.11676	3.44870	<0.001
VOD spatial correlation	-0.00113	-0.00673	<0.001
VOD spatial skewness	0.01979	0.00644	<0.001
VOD spatial variance	0.02452	0.01068	<0.001
VOD temp variance	0.03201	0.02944	<0.001
VOD temp skewness	0.00842	-0.00691	<0.001
VOD temp kurtosis	0.01797	0.05395	<0.001
VOD temp conditional hetroscedasticity	0.00773	0.05979	<0.001
VOD temp autocorrelation	-0.05526	-0.08574	<0.001
VOD temp return rate	0.05526	0.08574	<0.001

Appendix F

Mean importance regional scale

TABLE F.1: Continuation of table 4.8

Metric	Mean
LST night	121.72
LST day	111.1
LST night spatial autocorrelation	88.91
LST day contemporary spatial variance	86.08
LST night contemporary spatial autocorrelation	80.50
LST day contemporary spatial autocorrelation	80.28
LST day contemporary spatial autocorrelation	71.59
NDVI spatial autocorrelation	64.99
LST night contemporary spatial variance	62.11
NBR contemporary spatial variance	60.02
NBR spatial variance	56.67
NBR	55.53
NDVI	55.09
NDVI contemporary spatial variance	54.92
LST night temporal variance	54.79
NDVI contemporary spatial autocorrelation	54.65
NDVI contemporary spatial skewness	44.19
NDVI spatial skewness	43.32
LST day temporal kurtosis	41.95
LST day temporal variance	41.01
LST day temporal skewness	38.17
NBR temporal autocorrelation	30.68
NBR temporal return rate	30.27
LST day temporal autocorrelation	30.2
LST day temporal return rate	29.89
LST day contemporary spatial skewness	29.84
LST night contemporary spatial skewness	25.67
NDVI temporal kurtosis	24.73
NBR contemporary spatial skewness	23.13
LST day temporal CH	20.51
NBR temporal kurtosis	17.67
NBR temporal CH	5.79



Appendix G

Forest mask

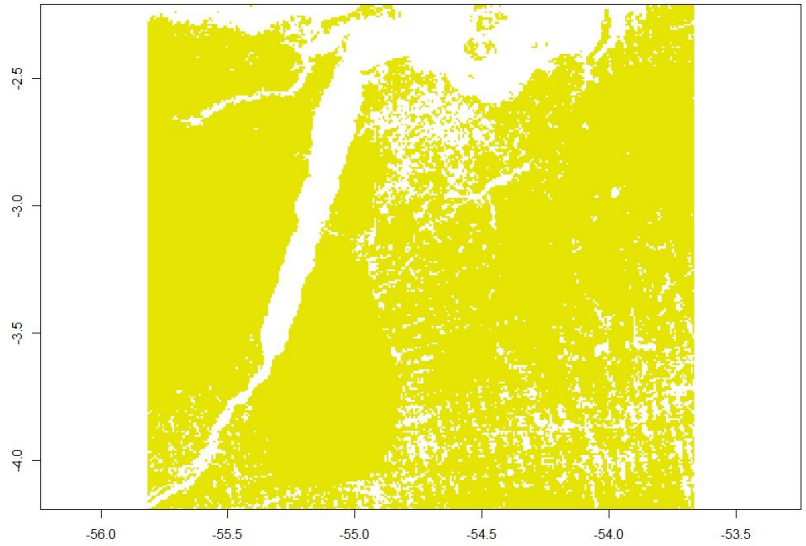


FIGURE G.1: Forest mask local scale

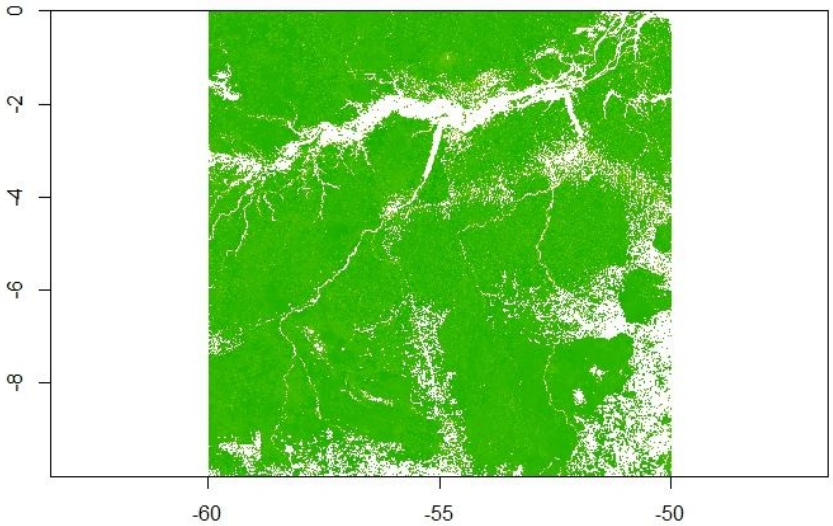


FIGURE G.2: Forest mask regional scale

Appendix H

VOD in relation to SPI

Figure H.1 shows the relationship between SPI values and VOD values before fire occurrence. The blue line shows the fitted regression line with a negative R-square of -0.1, which indicates that there is a weak negative correlation between VOD and SPI.

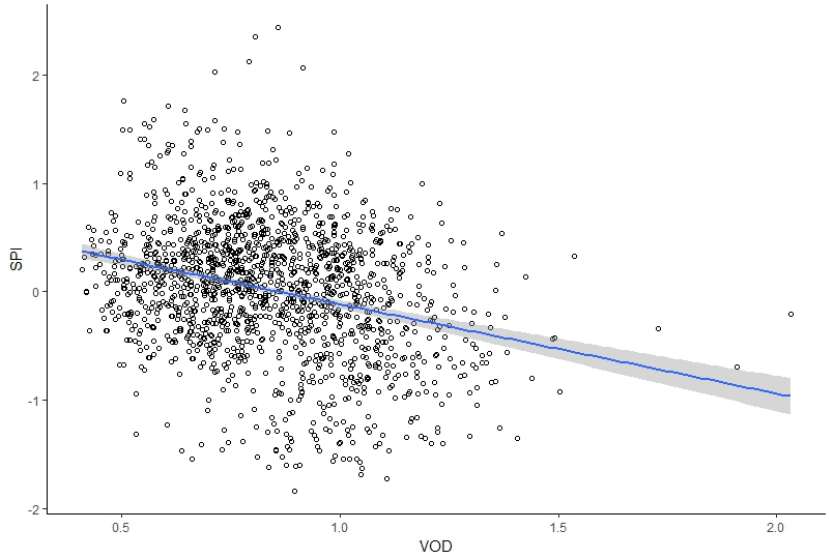


FIGURE H.1: This figure shows the correlation between VOD and SPI.