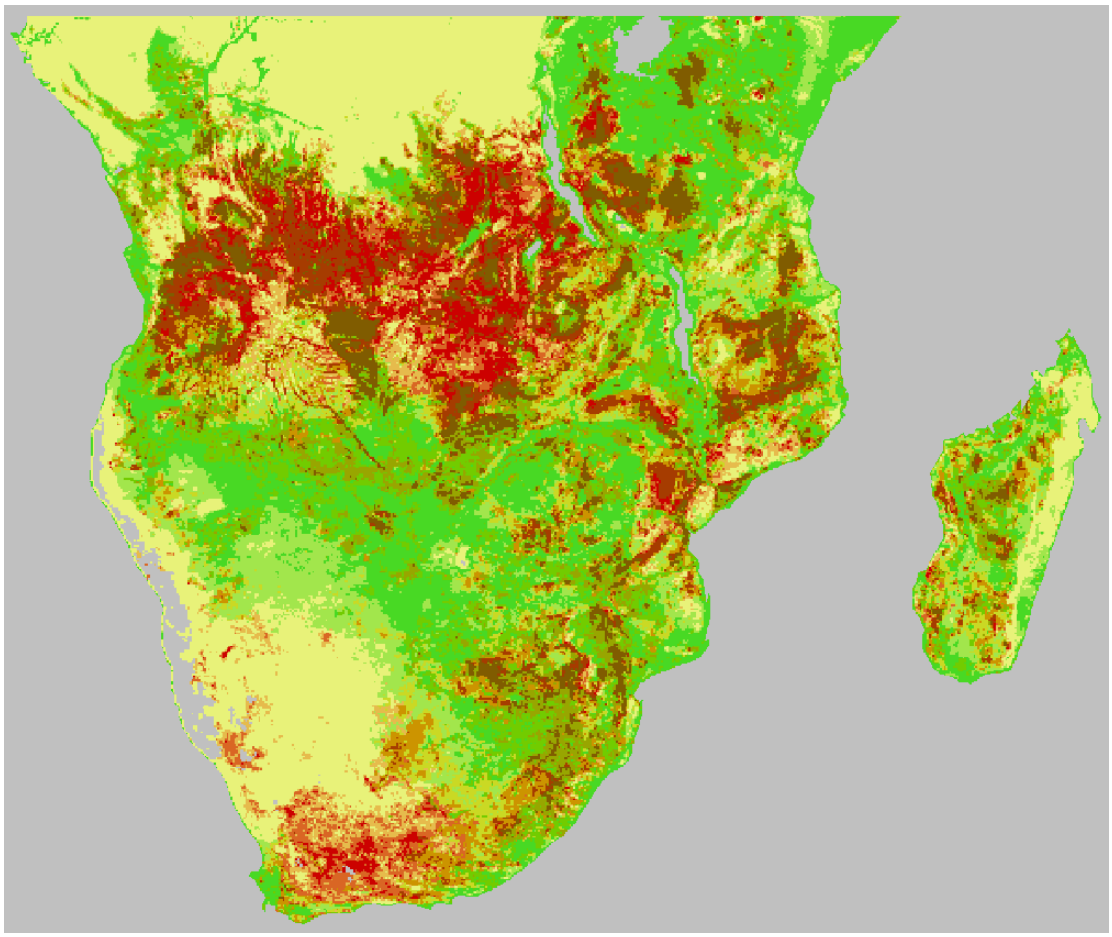
Centre for Geo-Information

Thesis Report GIRS-2009-06

ANALYZING WILDFIRES AND WILDFIRE IGNITION TYPES

in Southern Africa using the L3JRC burned area product

F.W.J. Salet



March 2009

WAGENINGEN UNIVERSITY
WAGENINGEN UR

Cover illustration:

- Cross-classified map visualisation of percentage of burned area per year with herbaceous vegetation cover in Southern Africa

Analyzing wildfires and wildfire ignition types in Southern Africa using the L3JRC burned area product

Frank Salet

Registration number 85 04 01 724 120

Supervisors:

Ron van Lammeren
Frank van Langevelde

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

March 20, 2009
Wageningen, the Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2009-06
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Summary

The research presented in this thesis focuses on the wildfires and wildfire ignition types in Southern Africa. The fire data used in the thesis is derived from the L3JRC product, which consists of detected burnt areas for seven fire years, from 2000 till 2006, with a moderate spatial resolution and a high temporal resolution. Three analyses are performed on the L3JRC product in order to derive spatiotemporal patterns about wildfire in the study area. For the analyses the theory of Exploratory Data Analysis is used. Behaviour of wildfire is made visible by means of map representations. Subsequently there are in each analyses cluster patterns described by investigating the map visualisations. The first analysis aims on the comparison of yearly fire frequency. The second analysis performs a correlation of burned area with two other variables, namely herbaceous vegetation cover and human influence index. The last analysis looks at the variation of monthly fire frequency. One of the results comprehends that fires are mainly occurring in a large cluster between 5° and 15° latitude south, and there are two large clusters that have no or almost no fire, i.e. an area near the equator and a large area in and around the Kalahari. Also we found that the correlation between fire and herbaceous vegetation cover is moderate positive, and the correlation between fire and human influence is negative and small.

Concerning the causes of wildfire we distinguished two ignition types: natural caused fire, which means ignited by lightning, and anthropogenic caused fire, which means ignited by human. In order to assign the two ignition types to wildfires that are represented in the L3JRC product, the so called Ignition Type Model is created. This model aims on assigning ignition types by considering the location in space and time of each wildfire. The Ignition Type Model consists of four steps. In the first step the burned areas of the L3JRC are converted to individual wildfires, and further the starting point of each wildfire is selected. In the second step thunderstorms, which are derived from the Lightning Imaging Sensor product, are selected that might reproduce strikes that ignite wildfire. In the third step of the model these thunderstorms are allocated to wildfires if they comply with certain parameters for distances in space and time. The fourth and last step of the Ignition Type Model implies the allocation of wildfires with human object, like roads and urban areas, based on certain distance parameters.

The Ignition Type Model resulted in 1690546 wildfires. 19.3% of these fires are allocated to either lightning or human objects and therefore received an ignition type. In total 34215 wildfires are assigned as natural caused and 292695 wildfires are assigned as anthropogenic caused. Based on these results the percentage of natural caused and anthropogenic caused wildfires in Southern Africa is calculated as respectively 31.7% and 68.3%.

Two analyses are performed in order to derive spatiotemporal characteristics from the two ignition type datasets that are the result of the Ignition Type Model. The first analysis aims on the monthly fire frequency of both natural caused and anthropogenic caused wildfire. The second analysis performs an investigation of natural caused fire and the land cover type on which these fires occurred. These two analyses revealed that there are differences in spatial clustering of both ignition types. Both the spatial and temporal clustering of natural caused fires is smaller. We also found that 82% of all lightning caused fires are located in woodland areas.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

1.1.1 Context and background

Wild land fire is considered as one of the most important disturbance factors in natural ecosystems. Each year several biotopes are completely lost because of the four million square kilometres wildfire that is consumed each year (Amatulli *et al.*, 2007; Tansey et at., 2004). Burning also leads to emission of vast amounts of gaseous and particulate matter (Andreae & Merlet, 2001). It is estimated that up to 40% of CO_2 production due to human activities may be due to biomass burning (Dwyer *et al.*, 1997). These consequences among others make that the influence of wildfire on environment and climate is large (Guyette *et al.*, 2002). This brings about that the subject of wildfire is investigated many times by researchers.

Africa is often referred to as fire continent or fire centre of our planet, because it has more biomass burning on an annual basis than anywhere else (Sheuyange *et al.*, 2005). African savannah fires may produce as much as a third of the total global emissions from biomass burning (Cahoon *et al.*, 1992). Therefore the implication of African fires on the global climate system is significant (Silva *et al.*, 2005).

Burning vegetation can be caused by humans as well as by natural incidents. Nearly all natural ignitions are caused by lightning strikes. Other possible natural causes are for example volcanic eruptions and spontaneous heating, but they are not treated in this research. From a global perspective, the majority of fires are man-made (Duncan & Schmalzer, 2004; Dwyer *et al.*, 2000). They are used for many reasons such as grassland management, burning of crop residues, elimination of disease bearing insects and snakes, hunting, land clearance and negligence (Andreae & Merlet, 2001; Dwyer *et al.*, 1997). In this thesis the exact reason human ignite wildfire is not of interest. The only two causes of wildfire that we distinct are the natural, which means initiated by lightning, and anthropogenic, which means initiated by human.

Because information about wildfire has been identified as an essential climate variable there are made several attempts to monitor fire activity or burn scars (Carmona-Moreno *et al.*, 2005c). The high demand for long-term, globally consistent records of global fire activity has been satisfied for a short time now. This because the data gap was filled by the L3JRC that came available in the year 2008 (Tansey *et at.*, 2008). The L3JRC is a long-term, moderate spatial resolution, high temporal resolution global burnt area product.

1.1.2 Problem definition

There has been a request from the international science community for information on fire (Dwyer *et al.*, 1997). These include scientists that improve understanding of the role of vegetation fire emissions in atmospheric chemistry, scientists that are involved in environmental policy and management and scientists with land management mandates. Their information requirements include global fire distribution, timing, extent, severity, cause, and return frequency (Amatulli *et al.*, 2007; Dwyer *et al.*, 1997). These descriptors of global wildfire are nevertheless poorly documented (Sheuyange *et al.*, 2005.). The new L3JRC burned area product makes it possible to satisfy to the demands of the scientific community. Therefore in this research we make use of the L3JRC to document a few of the descriptors of wildfire that are not documented yet.

Many research is done on spatiotemporal characteristics of wildfire on local and regional level (e.g. Amatulli *et al.*, 2007; Sheuyange *et al.*, 2005; Forsyth & van Wilgen, 2008; Larjavaara *et al.*, 2005). However, only a few researches described the patterns of wildfire on continental or global scale. Carmona-Moreno *et al.* (2005) and Moreno-Ruiz *et al.* (2006) described the spatiotemporal patterns of global wildfire. Only this is done with a very low spatial and temporal resolution. The new L3JRC product consists of much better spatiotemporal resolutions. Therefore it is necessary to investigate what spatiotemporal patterns can be derived from this product.

Only a few researches are done on the characteristics of wildfires per ignition type. These few researches are working on local scale, and the ignition type are obtaining per fire by consulting registrations or involved parties. Obtaining of ignition type per wildfire, on e.g. continental scale, by involving independent variables, is never tried. In order to get knowledge about causes of wildfire it is necessary to investigate if it is possible to obtain ignition types in this way.

Characteristics of wildfires per ignition type on continental or global scale are unknown in the scientific world. Therefore it is necessary to find out if it is possible to gain knowledge about for example spatiotemporal patterns about both natural and anthropogenic fires.

1.2 Research objectives

Based on the described problems in the previous section the following three research questions are defined:

- What spatiotemporal patterns of wildfires can be derived from the L3JRC burned area product?
- Can the distinction between natural and anthropogenic caused fires be made by independent variables?
- What are the spatiotemporal patterns of natural and anthropogenic caused fires?

Also for the thesis there are three hypotheses defined, which are the starting point of the thesis. The three hypotheses are explained hereafter.

- There is not much known about the characteristics of the content of the new L3JRC product. Therefore the first defined hypothesis is that wildfires are completely

spatiotemporal random, i.e. the amount of burned areas is constant over space and time and the areas are neither clustered nor regularly spaced.

- Previously Buitrago (2008) performed a research in order to describe the statistical correlation between fire frequency and several anthropogenic variables in Southern Africa. The research concluded, because of a low correlation, that the influence of humans on the total fire regime is limited. Therefore this finding is used as the second hypothesis.
- Describing the spatiotemporal patterns of both natural and anthropogenic caused wildfire has never been investigated in science. These patterns and the differences between these patterns are unknown. Therefore the third hypothesis is that both natural caused and anthropogenic caused wildfires contain identical spatiotemporal patterns.

1.3 Thesis outline

The structure of the thesis is as follows.
Chapter 2 explains the methodology of this research, including the definitions of the research area and period, the datasets that are used, and the used methodology concerning the analyses.
Chapter 3 provides an answer on the first research question by investigating the L3JRC burned area product for spatiotemporal patterns by using several visualisation tools for mapping spatiotemporal data.
Chapter 4 presents the new to build Ignition Type Model that attempts to assign the ignition type to wildfires that are derived from the L3JRC burned area product. By means of this model the second research question will be answered.
Chapter 5 covers the converging of the third and fourth chapter in the sense of using the results of the Ignition Type Model and deriving spatiotemporal patterns in the way it is done with the burned areas.
Finally, **Chapter 6** gives an overview of the conclusions of this work and recommendations for future research on this topic.

Chapter 2

Methodology

2.1 Research area and period

For this research the territory of Southern Africa is chosen as research area. This means the African continent on the Southern hemisphere, including Madagascar (see Figure 2.1). The Northern boundary of the study area is the equator and the other boundaries are formed by the coastlines. This study area has a surface area of 9.43×10^6 sq km, which is about 6% of the total earths land surface. Because the whole region is located on one hemisphere we deal in the thesis with only one wildfire cycle (above the equator the fire cycle displaced six months).

Figure 2.1: Topographical map of the study area (adapted from the CIA Factbook).

The time span for which the research is executed starts at April 1, 2000 and ends at March 31, 2007. That means that the research period covers a time span of exactly seven fire years (a fire year starts at April 1st).

2.2 Data management

During the research several data products are used to retrieve the demanded input data in order to perform the analyses. The collected products are ordered in four categories (fire-,

lightning-, anthropogenic- and land cover datasets) and are briefly described per category. In the last paragraph of this section the major remarks about the preparations of the datasets are mentioned.

2.2.1 Description of the datasets

Fire dataset
The most important dataset used during the research is the global burnt surface multi-year product called L3JRC (in Appendix A the explanation of choosing the L3JRC product instead of other burned area and fire products is elaborated). The product consists of all burn scars that are detected worldwide in seven fire years between the fire years of 2000 and 2006, with a resolution of 1 sq km, and a high temporal resolution of daily intervals (Tansey, 2008; JRC-TEM, 2008). The burnt surfaces of the L3JRC are obtained from Spot Vegetation S1 reflectance data, which is available on a global and daily basis and corrected for atmospheric transmissibility. The L3JRC product is released as a binary file with for each burnt pixel the geographic coordinate of the pixel centre and a Julian date of the first detection.

Lightning dataset
The data concerning lightning flashes are from the Lightning Imaging Sensor (LIS) instrument, which is a component of the scientific measurement equipment on board of the Tropical Rainfall Measuring Mission (TRMM). This satellite observes from a nearly circular orbit with an inclination of 35 degrees and an altitude of 350 km (NASA-GHCC, 2005). The LIS is designed for detecting terrestrial lightning activity. The instrument records the time of occurrence, the measured radiant energy, and the location of the lightning events.

Anthropogenic datasets
For the representation of human attendance there are three datasets used. The first anthropogenic dataset is called Human Footprint Map and. This dataset gives a quantitative indication of human influence across the globe (CIESIN, 2008). The second anthropogenic dataset is the Vector Smart Map Level 0 (VMap0) from 2000, that provides us with almost all roads and trails. The VMap0 is currently the best available public domain spatial global road network dataset because it captures between one-quarter and one-third of the global road network (Nelson et al., 2006). The VMap0 also provides us in this research with all human settlements. The third anthropogenic dataset is called GRUMP (CIESIN, 2008), what stands for Global Rural-Urban Mapping Project (CIESIN-SEDAC, 2008). This dataset is used as a representation of all built-up areas.

Land cover datasets
The thesis makes use of the Global Land Cover 2000 (GLC2000). This product consists of approximately 1 km resolution land cover map of the whole globe. The map consists of 25 different classes, with special attention to the forest and savannah biomes (JRC-TEM, 2003). Also a land cover product from the MODIS Vegetation Continuous Fields is used. This product is called the Herbaceous Cover and contains a percent herbaceous cover layer (UMD-GLCF, 2007). The herbaceous vegetation cover is calculated for the period 2000 till 2001.

2.2.2 Preparation of the datasets

The datasets that are mentioned in the previous section are all converted to a shapefile format that is readable by ArcGIS 9.2 and they are stored in a geographical database. In case

of the L3JRC product this implied that a conversion was necessary from text format (that is how it is available, free of charge, at the JRC web portal) to a point feature file. Concerning the LIS product, which is free of charge obtainable at the NASA's Data Pool, this implied that a conversion was necessary from HDF format to text file and subsequently to a point feature file. During the process of making the burned area and lightning data readable in ArcGIS, all attributes that were present in the original data are preserved. In contrast to the fire- and lightning datasets, all other datasets were already available in feature file format. Also these datasets are placed in a geographical database. The lasts action was the process of masking every dataset to the correct spatial extent and transforming the datasets to the same geographical coordinate system (which in this case is WGS84).

2.3 Data analyses

The methods of data analyses are hereafter explained per research question (RQ).

2.3.1 RQ 1: Exploring the L3JRC

The first research question, which is answered in Chapter 3, is aiming on what spatiotemporal patterns of wildfires can be derived from the L3JRC burned area product. In order to answer this question the used method comes from the Exploratory Data Analysis (EDA), which is an approach of getting acquainted with in this case the wildfire data. The EDA approach consists of several techniques for statistical and visually analysing large datasets. However, the technique that is in particular used in this research is the graphical or visual analytics by means of map representations. With the map visuals the goal is to open-mindedly explore the contents of the L3JRC burned area product.

By using the EDA methods we want to answer the first research question and test the first hypothesis. The aspect of the L3JRC product in which we are most interested is therefore fire frequency. The three analyses that are performed in Chapter 3 are all aiming on the patterns of fire frequency in the study area. Recognition of patterns of fire frequency takes place on the basis of natural capabilities that we all possess. The second of the three analyses is performed in order to test the second hypothesis, which is performing a correlation of fire frequency with other variables.

2.3.2 RQ 2: Producing a model for assigning ignition types

The second research question asks if it is possible to use independent variables to make a distinction between the two main ignition types of wildfire. In Chapter 4 we therefore do an attempt to assign to each wildfire that it is ignited by lightning, ignited by human, or that the ignition type is unknown. This is done by the so called Ignition Type Model. The model uses several independent variables that represent lightning or human objects. Based on the spatial and temporal location of a wildfire one of the three categories is assigned to the fire. The verification and discussion sections of Chapter 4 will subsequently answer the second research question.

2.3.3 RQ 3: Exploring the model results

The third research question, which is answered in Chapter 5, is aiming on the spatiotemporal patterns of the two main wildfire ignition types, namely wildfire ignited by lightning and wildfire ignited by human. Therefore it uses the results of the Ignition Type Model. The

method that is used in order to answer this question is for majority the same as the method for answering the first research question. That means we will use the EDA approach. However it is not necessary to explore the structure of the data, because that is already done in Chapter 3. Deriving of patterns will again be done by performing visual analytics by means of map representations. Besides answering the third research question, Chapter 5 also performs the test of the third hypothesis.

Chapter 3

Pattern finding in the fire dataset

This chapter gives answer on the first research question, i.e. the question about what spatiotemporal patterns can be (visually) derived from the L3JRC fire dataset. The investigation that is necessary to answer the research question consists of two parts and both parts are respectively discussed as sections in this chapter. Section 3.1 starts with elaborating on the concerned properties of the fire dataset. When the properties of the data are clear it is necessary to take a couple of decisions before the actual analysis can be performed. These decisions consist of: the way of approaching the dataset, defining the tasks, and choosing proper tools for the analyses. Section 3.1 is based on the theory of Exploratory Data Analysis that is described in Andrienko & Andrienko (2006).

Section 3.2 treats the actual data analyses that are performed on the fire dataset. There are three analyses described and executed and together they give a general indication about how spatiotemporal patterns can be derived from the L3JRC. The analyses are executed by using different visualisation tools to create map images of certain fire behaviour. As a result for each analysis the main patterns from these images are described. Some of these patterns are then in Section 3.3 discussed and related with other researches.

3.1 Theory about analysing spatiotemporal fire data

3.1.1 Perspective on the data

Before the actual analyses can be performed on the *L3JRC fire product*, it is necessary to know a few things about this dataset. The L3JRC fire product consists of one data file for each of the seven fire years. In this research the focus is on all fire years. Therefore, for the sake of convenience, all seven fire years in the L3JRC product are considered as being in one data file. This data file then consists of 6067178 items, which is the number of burned areas that are detected in a period of seven fire years. Because all these records are in one file, all these records have the same components, and each component has the same meaning in each of all the records. The characteristics of these components form the structure of the dataset.

The fire dataset characterizes by a set of properties, which are theme (fires), space (Africa below the equator), and time (from April 2000 till March 2007). Those three properties are the components of the data, which can be divided into *referential components* (indicating the context) and *characteristic components* (representing results or observations). Table 3.1 presents some brief characteristics of the three components from the fire dataset. The location and time are typically treated as referential and the remaining data components are treated as characteristics or attributes referring to location and time. For the fire data the intention is

to study the spatiotemporal patterns, i.e. the variation of the theme over a territory and a time span. Therefore it is appropriate to switch the common division of referential and characteristic components, with a fire occurrence as the context and the location and the moment in time as its attributes.

Table 3.1: Overview of the properties of the three components of the L3JRC fire dataset.

Name:	**FIREnb**	**loc**	**JulianDay**
Description:	Identity of a fire	Location of the fire	Julian day of detection
Component type:	Referential	Characteristic	Characteristic
Ordering	Linear (fully)	Partial	Linear (fully)
Continuity	Discrete	Continuous	Continuous
Element type:	Integer	Referrer	Integer
Element values:	0 till 6067177	References to points	1 till 2556

The first component, FIREnb, takes its values from a fully linearly ordered set of integer elements with each element representing one fire. This set consists of elements starting with zero and subsequently consists of all integer numbers until 6067177. Because all attributes characterising the phenomenon are referring to only one element of this component with a finite set of values, the phenomenon is called a discrete. The underlying set of the location component consists of references to points and the set is therefore continuous. The points represent the centre of the cells of a two dimensional grid overlay, and because the grid has more than one dimension the location set is partially ordered. An item that is referring to a grid cell means that the corresponding fire was detected within the area of the concerned cell. The last component, JulianDay, is the time component with a set of linearly continuously ordering with distances of one day. The set has integer values and the component can choose a value between 1 and 2556, which represents the Gregorian dates from April 1st 2000 till May 31st 2007.

3.1.2 Other relevant principles

The next step of the analysis is formulating questions. The questions that are involved in the data analysis, from here on they are called *tasks*, are for this research merely *synoptic tasks*. That means that the tasks that we use are not referring to single items in the dataset but are referring to more then one item simultaneously as well as the relations between these items.

A synoptic task is involved with one or more reference (sub)sets and the relations between these (sub)sets. A reference (sub)set contains a particular range of values of an attribute, and if this is considered in its entirety, it can be defined as *behaviour*. So it is a task that involves behaviour and the relations between behaviours. A synoptic task deals in this research with the behaviour of fire frequency over space and time.

A *pattern* is something resulting from observation or analysis. It is defined as a construct that reflects essential features of behaviour in a shorter and simpler way, rather then considering and comparing every single feature separately. If essential features of behaviour are detected by means of mental mapping the behaviour is considered as a pattern. While doing analyses with the fire dataset it is important to explicitly distinguish behaviour and pattern.

There are different kinds of Patterns that can be derived from the wildfire dataset concerning fire frequency. To make the different kinds of patterns visible, by means of map visualizations, is mostly not a problem. On the other hand, describing all kinds of patterns is in some cases difficult, time consuming and useless. For example patterns about the course

of changing fire frequency between two locations are often hard to describe. But they are easy to visualize and moreover visualizations in contrast to text do better explain the course of change. Therefore we will not go into detail about these patterns. The kind of patterns we do describe extensively are *clusters*, that are groups of cells that contain roughly the same value (whether high or low) of fire frequency. Clusters will be described by its location and size.

3.1.3 Visualisation tools

Only the human mind actually does the analysis, and the computer tools supply the necessary material. Tools visualise the behaviour and the observer converts the visualised behaviour to patterns. The behaviours are made visible with the help of visualisation tools. That means that the elements of the fire dataset are translated into graphical features. The graphics that are created in this research are various map images. Therefore the tools that construct map images are the main focus for this research.

More specifically, we are using tools that come under the category of display manipulation. Tools within this category are interactive, and support the dynamic modification of the appearance of visual displays. The general purpose of display manipulation is to enhance the image that is created in that way that it is clearer and easier for human to perceive the behaviours. Concerning the display manipulation there are three main definitions that need to be discussed in order to perform the analyses of this research. These definitions are classification, focusing and aggregation; they are shortly explained hereafter.

Classification is an approach that is often used to simplify data by means of generalising data characteristics. The purpose of applying classification to the map is to facilitate the revealing of distinctive features of the spatial distribution of the attribute values and the finding of an appropriate pattern within those features. In cartography the *classified* choropleth is often used as a way to simplify a visualisation. In contrast to an *unclassified* choropleth a considerable amount of information is lost when putting features in classes based on its attribute value. On the other hand the classified choropleth shows more clearly some distinctive features of the spatial distribution of the attribute values than the unclassified map does. The most important notice of classification is the choice about the number of classes and the class breaks.

Data focusing is a way of zooming in the dataset. It involves selection of a data subset rather than viewing all data. The advantage of focusing is that the portrayed data subset can be explored with the maximum possible expressiveness. This technique is used in the first analysis of this chapter (see Paragraph 3.2.1). The advantages of this technique are explained in more detail at the concerned section in this chapter.

Data aggregation tools reduce the amount of data by grouping individual component values into subsets and computing some collective characteristics of each subset. Especially when dealing with very large datasets it is more useful to analyse aggregated data, instead of the original data. Data aggregation is a way of simplification just like classification, with the difference that data aggregation is computing a value out of the characteristics of each subset. This technique is used in all three analyses of Section 3.2 as a kind of pre-processing step.

3.2 Performing analyses

In this section the actual analyses (three in total) are performed. Every paragraph in this section consists of the elaboration of one of the analyses. The paragraphs start with the explanation of the task, subsequently the selection and description of the visualisation tools are given, and finally the main patterns are given as the result of each task. The three tasks together should give a broad overview of what tasks are possible to perform for visually extracting spatiotemporal patterns from the L3JRC. The similarity for the tasks is that they all focus on fire frequency. Therefore each burned area from the L3JRC product is assumed as one fire. So if fire frequency is mentioned then actually the sum of detected burned areas is meant.

3.2.3 Task 1: Comparison of yearly fire frequency

The goal of the first analysis is to investigate the spatial variation between the fire frequencies of different fire years. Thanks to the L3JRC fire dataset we have the disposal of the fire frequency of seven fire years. Therefore the task of the first analysis is formulated as follows: V*isualize and compare the fire frequency of Southern Africa for the fire years of 2000, 2001, 2002, 2003, 2004, 2005 and 2006*. The target of this task is the L3JRC fire dataset from which the fire frequency will be computed. The spatiotemporal constraints of this task are the same as given in Section 2.1.

At first, before the visualisation tool is used, the L3JRC is aggregated to grid segments with a cell size of 200 kilometres and for each cell the fire frequency in a particular week is computed. This aggregated dataset is then visualised by means of *multiple unclassified choropleths*, and the result map can be seen in Figure 3.1. Each map in this figure portrays the fire count of one particular fire year. Some brief statistics of the seven attribute components are given in Table 3.2. The choropleths have a gradual colour scale of decreasing brightness that is matched to the full range of the seven components, with 0 as minimum and 21528 as maximum (occurring in the fire year of 2003).

Table 3.2: Result of univariate statistics on each component of the yearly fire count dataset. The dataset has a cell size of 200 km and for each cell the fire frequency of that particular fire year is counted. The statistics consists of the minimum, maximum, mean and standard deviation of all cell values.

Statistical parameter:	2000	2001	2002	2003	2004	2005	2006
Minimum	0	0	0	0	0	0	0
Maximum	20321	20331	19102	21528	18029	18507	20957
Arithmetic mean	3225	3599	3205	3611	2998	2918	3598
Standard deviation	4281	4053	3667	4345	3766	3695	4370

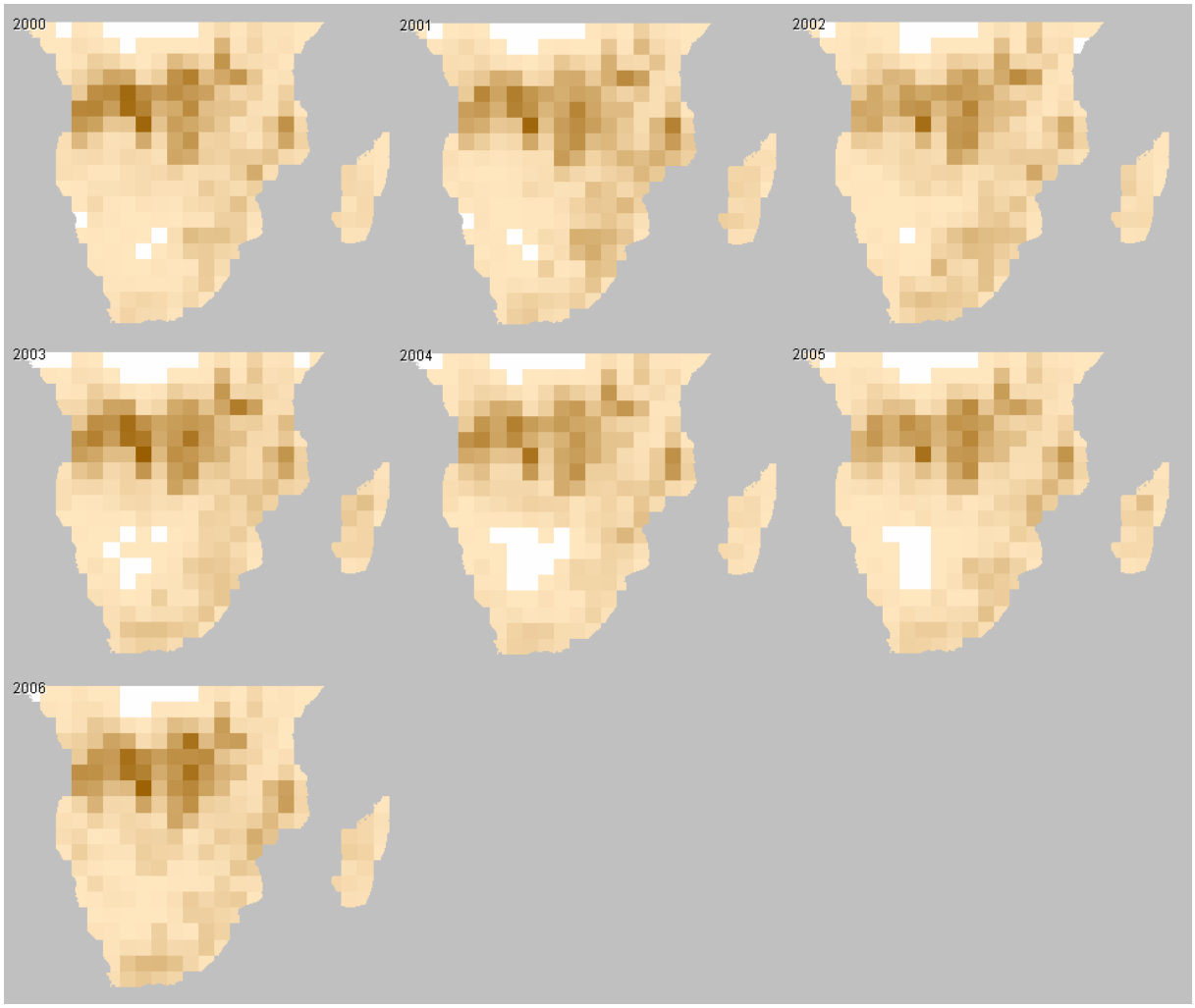


Figure 3.1: Seven maps of Southern Africa with for each map the fire frequency during a fire year for 2000, 2001, 2002, 2003, 2004, 2005 and 2006. The maps have cell sizes of 200 km and are portrayed by means of unclassified choropleths with gradual colour scale of decreasing brightness that is matching the full range of the components minimum and maximum.

The seven portrayed maps of Figure 3.1 show that the behaviour of fire frequency per fire year in Southern Africa. From these behaviours the main patterns will be described hereafter. The first patterns that attract attention are the large clusters in every fire year with high fire frequency in the northern part of the study area, i.e. Zambia, southern part Congo and northern part of Angola. Near the equator there are some areas that do not have fires. Also in the Kalahari Desert there are, at least for 2003, 2004 and 2005 large areas without wildfire. For the rest it seems that each fire year has in general the same distribution of fire frequency. Only the overall shading for the maps of 2004 and 2005 is slightly brighter, which indicates that there are less fires in that fire year. This can also be derived from the yearly statistics that are shown in Table 3.2.

Further it is hard to derive patterns based on these maps because the largest part of each map has almost equal colouring. The few very high values (outliers) cause that all lower values are shown in very light shades and look almost identical. To solve this problem a technique called *focusing* is used. Focusing can be explained as zooming into the data, selecting a data subset and portraying this subset with the maximum possible expressiveness.

The image of Figure 3.2 shows the result of the focusing tool that was used on the identical dataset for Figure 3.1. Again in this image the fire frequency is portrayed by means of choropleth maps for the years 2000 till 2006. The focussing tool resulted in a new colour scale range of 0 – 7500. That means that there is zoomed into the data by excluding the high fire frequencies from the visualisation. The colours of these removed cells are now replaced on the map by triangular symbols. The maximum value of 7500 fires is chosen to remove all cells that are mainly located in the cluster of high fire frequencies. In this way the contrast of shades outside this cluster is maximal.

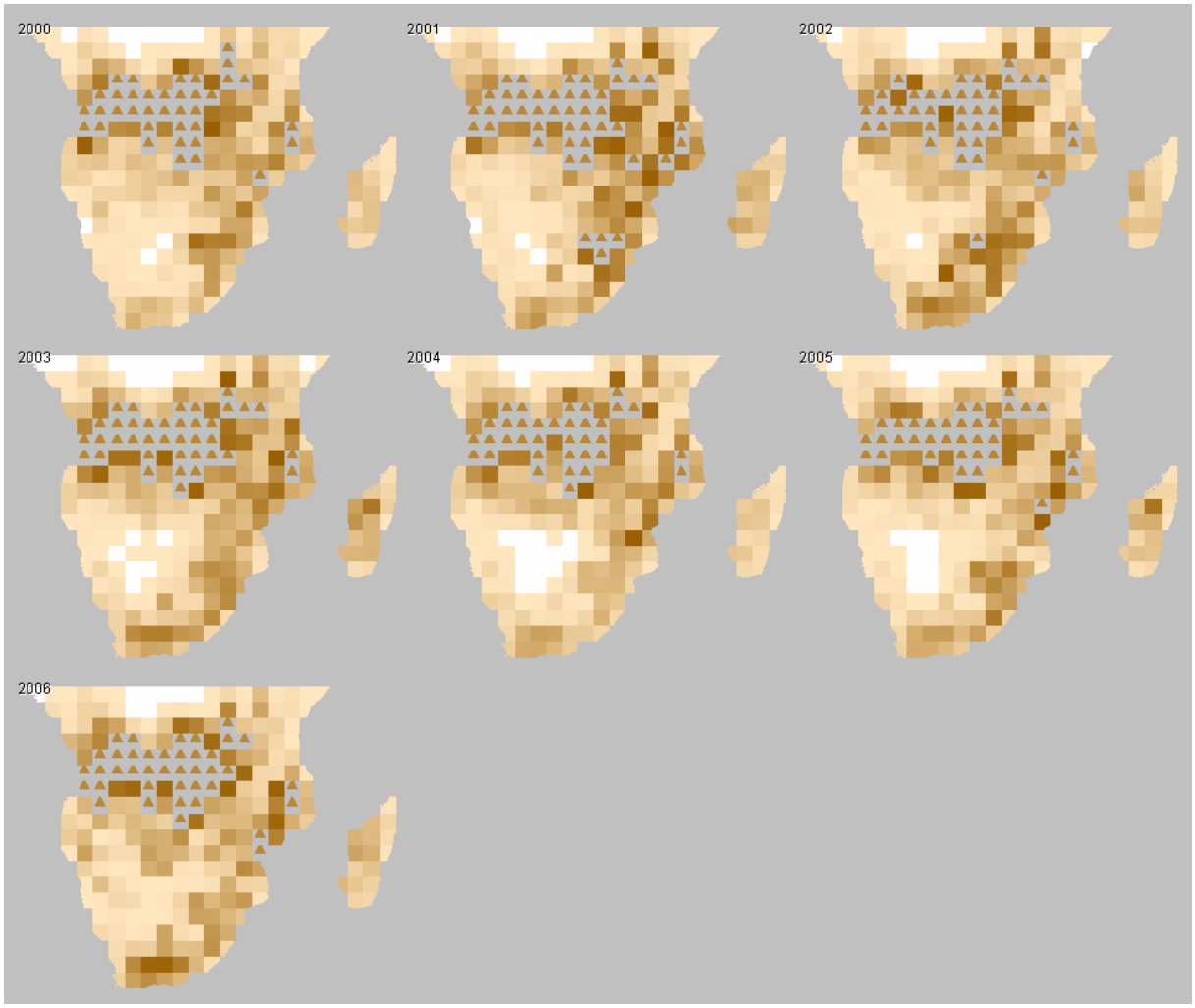


Figure 3.2: Seven maps of Southern Africa with for each map the fire frequency during a fire year for 2000, 2001, 2002, 2003, 2004, 2005 and 2006. The maps have cell sizes of 200 km and are portrayed by means of unclassified choropleths with gradual colour scale of decreasing brightness that is matching the range of 0 till 7500. Cells with a fire frequency of more than 7500 are represented with a triangular symbol.

The maps of Figure 3.2 look much less uniform than the maps of Figure 3.1 for areas outside the cluster with high fire frequencies. Because of the focusing tool the map became for most areas much more expressive. Therefore the spatial variation in the map can be described in much more detail and the temporal variation between the seven maps can better be derived. Now it becomes clear that in and near Mozambique in the year 2001 and 2002 there were many more wildfires than in the other years. Also from the maps can be derived that there is a notable difference between the fire frequencies in South Africa during the years. These patterns could not have been derived from the maps of Figure 3.1.

Another advantage that is involved with performing the focussing tool is that the new maximum value is a sort of threshold that can also be used for comparing. The cells above this threshold are represented with a triangular symbol, and the amount of triangular symbols in each map can be compared with each other. In this way it seems that in general every year the same amount of cells have a fire frequency of more then 7500 wildfires per fire year.

3.2.2 Task 2: Correlation of burned area with other variables

In the research of Buitrago (2008) several environmental, climatic and anthropogenic variables are used to calculate the correlation with fire frequency. As a result, of all variables the herbaceous vegetation has the highest (positive) correlation with fire frequency (founded Pearson correlation coefficient: 0.60). The anthropogenic variables have the lowest correlation with fire frequency (coefficient of population density: 0.003). As a sequel on this finding the following task for the second analysis is formulated: *Visualize an overall view of the correlation of burned area with herbaceous vegetation and burned area with human influence.* The target of this task consists of three different datasets. In both correlation analyses the L3JRC fire data is used from which the average area burned is derived. Besides that, for the one analysis the quantitative herbaceous land cover is used. This data was also used by Buitrago (2008) and has a values ranging from 0, meaning no herbaceous vegetation, till 255, meaning full herbaceous vegetation (UMD-GLCF, 2007). For the second analysis the Human Footprint, a quantitative human influence map, is used. This data is not used by Buitrago (2008) but it is produced with datasets that are used by Buitrago (2008), like population density and land use (CIESIN, 2008). In this dataset the human impact is rated from 0, meaning minimum human influence, till 100, meaning maximum human influence (NASA-GHCC, 2005).

At first, before performing the visualisation tool, an aggregation on all three datasets is performed. The datasets need to have the same segment sizes and for each segment a value is obtained from one of the datasets. A grid overlay with cell sizes of approximately 8 kilometres is used. For each cell the average percentage of burned area per year, the average herb value, and the average human influence value are calculated.

For the two analyses we use classification, as a simplification tool, for generalising quantitative characteristics. Because for each analysis we are dealing with two quantitative characteristics the actual tool that is used is the *cross-classification method*. The idea of this method is to divide the value ranges of each of the two attribute components into subintervals, in the same way as the classification of a single numeric attribute would be performed. If the value range of the first attribute has been divided into M subintervals and the value range of the second attribute into N subintervals, it leads to $M \times N$ different classes. Each class is then a combination of two intervals, one for the first attribute and one for the second attribute. Finally each class is visualised with an own colour.

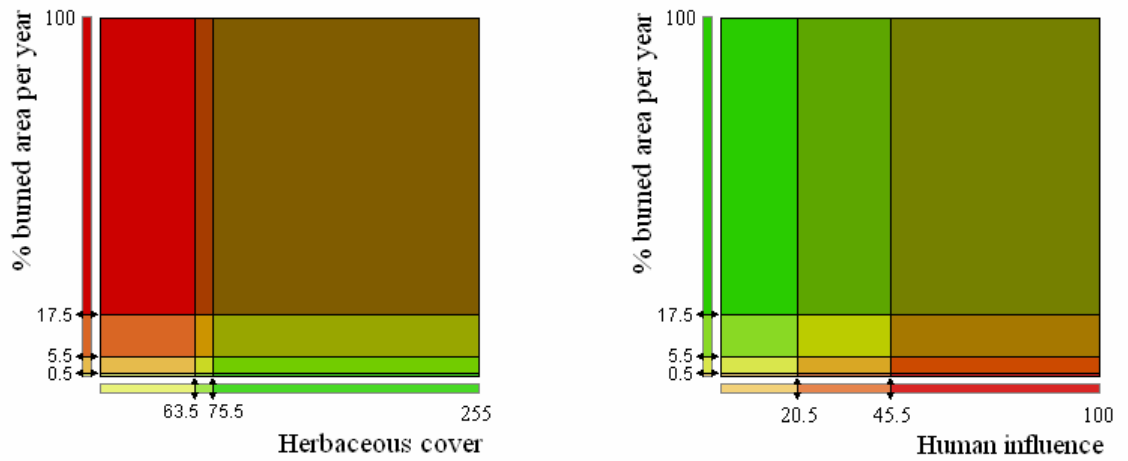


Figure 3.3: Both diagrams show the selection of class breaks and colouring scheme of the cross-classification technique. The left diagram shows the correlation of fire with herbaceous vegetation cover and the right diagram shows the correlation of fire with human influence.

For both cross-classification analyses the percentage of burned area is divided into four subintervals. The division is done on the basis of equal group sizes for each subinterval. The three class breaks, which are visualized in Figure 3.3, of burned area are 0.5, 5.5 and 17.5. The other two variables are divided into three subintervals, and also this is done by dividing into equal group sizes. The class breaks of herbaceous vegetation cover are 63.5 and 75.5; the class breaks of human influence are 20.5 and 45.5. This means that for both cross-classification analyses twelve classes are created. The colours that are given to the classes of both analyses are visualised in Figure 3.3.

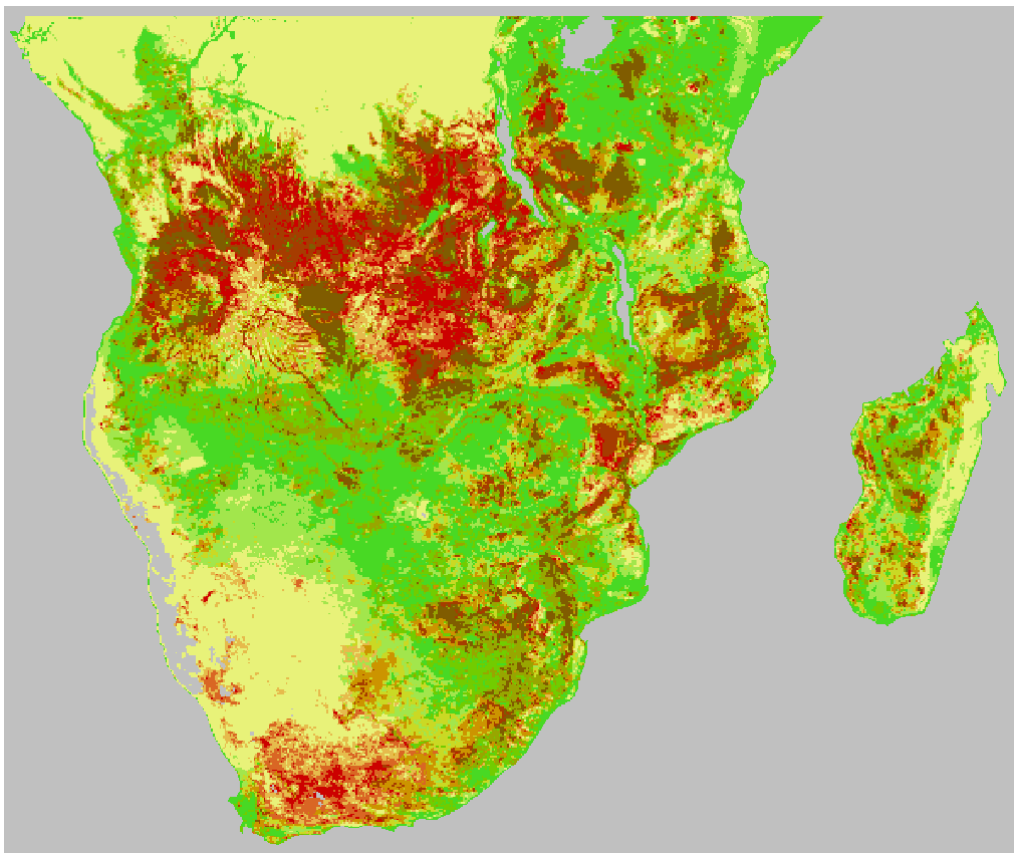


Figure 3.4: Result map of the cross-classification technique of burned area per year with herbaceous vegetation cover. The concerned classification choices are visualised in the left diagram of Figure 3.3.

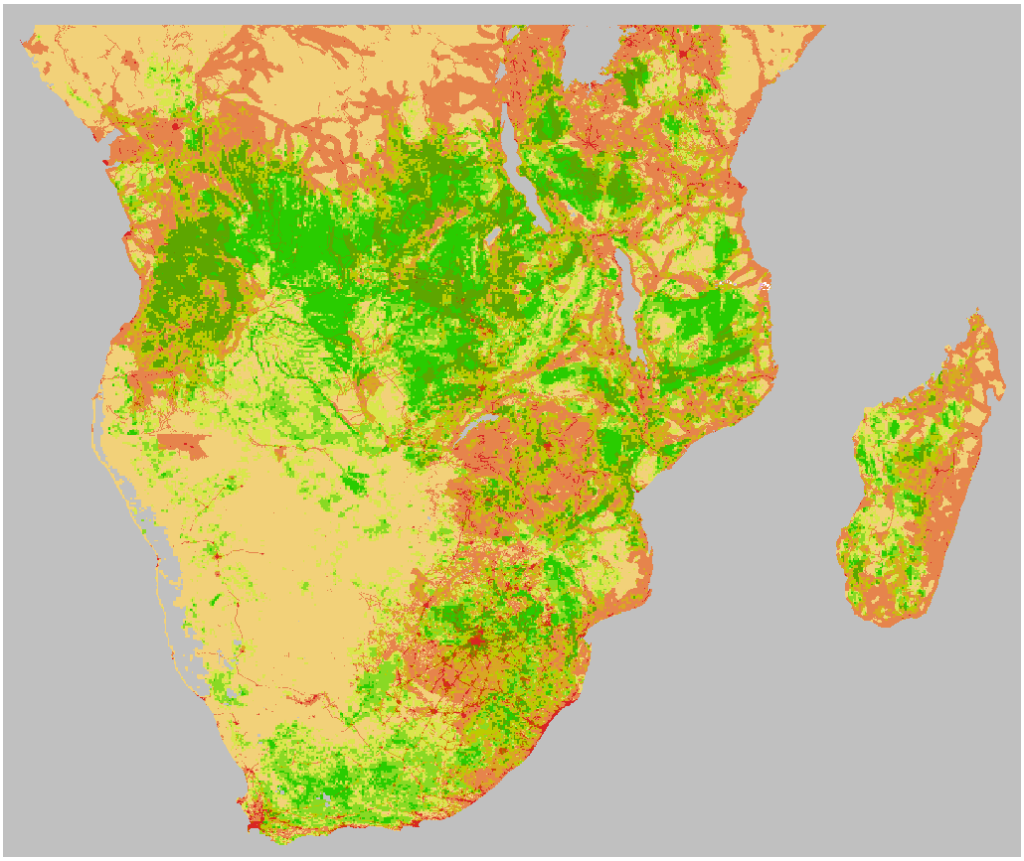


Figure 3.5: Result map of the cross-classification technique of burned area per year with human influence. The concerned classification choices are visualised in the right diagram of Figure 3.3.

The map images of the cross-classification techniques of fire with herbaceous vegetation and with human influence are respectively shown in Figure 3.4 and Figure 3.5. These two maps give the behaviour of the correlated variables. For deriving patterns from these results there are made two simplified representations with only the main patterns of the colours. These can be seen in Figure 3.6.

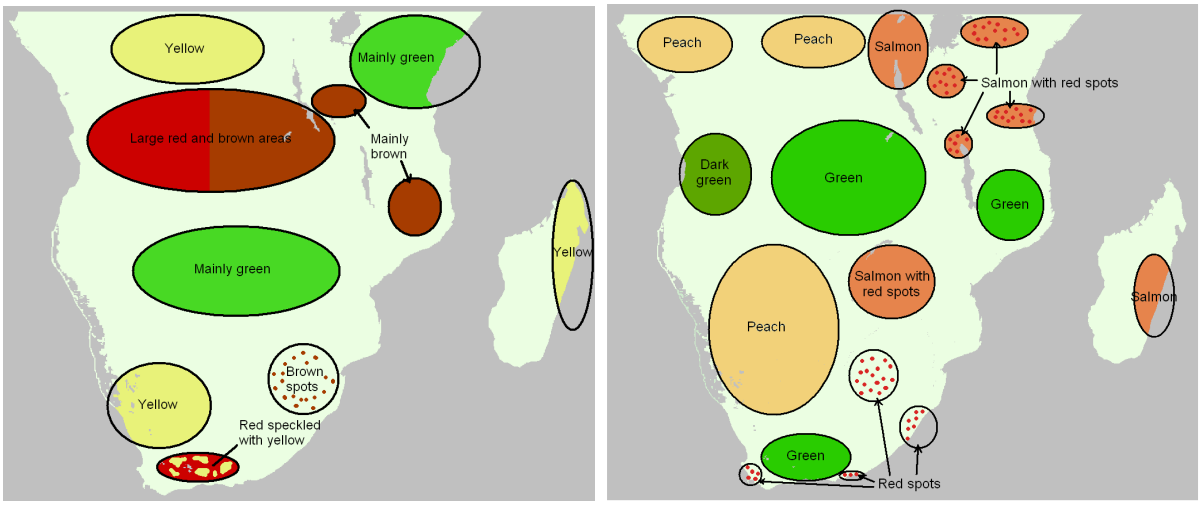


Figure 3.6: A simplified representation of the main patterns of the cross-classification results of fire with herbaceous vegetation on the left side and fire with human influence on the right side.

What can be seen in the left map of Figure 3.6 are two large areas that consist of merely yellow colouring. This means that in these areas, which are located near the equator and around the Kalahari Desert, there is no herbaceous vegetation cover and also no wildfire. Furthermore there are large unbroken planes with mostly green colouring with big concentration in the central part of the study area. These green planes mean that in these areas there is much herb vegetation but there are almost no fires counted. Further there are two areas with many red colouring. Red means that there are a lot of fires but there is no herbaceous vegetation. Apparently it is in these areas something else than herbaceous vegetation that is burning. The brown areas, which are concentrated in small planes in the middle of the study area, indicate that there are many fires and there is high herbaceous cover. The fuel for the fires in the brown areas is probably herbaceous vegetation.

Concerning the cross-classification with human influence we see in the left map of Figure 3.6 three large areas with peach colouring. These indicate that in these areas there is both a low fire frequency and low human influence. The areas are found in the Kalahari Desert and its surroundings and near the equator at the east part of the study area, i.e. in tropical climates. Further in the results some concentrations of green are found with at some locations adjoining small dark green areas. These green colourings represent that there are many fires but the influence of human is very low. Also in the simplified representation we see places with salmon colouring. This colour means the opposite of the green colours, i.e. few fires and large human footprints. The red spots are located on places that have a really high human influence, which mostly represent the urban districts.

By comparing the locations of the main patterns of both cross-classification results we see some similarities. The location of the yellow colouring in the first correlation is mainly represented by peach colouring in the second correlation. This means that locations with a very low fire frequency have no herbaceous vegetations but also no human influence. Red and brown colouring in the result of the first correlation analysis is mainly replaced by green and dark green colouring in the result of the second correlation analysis. This indicates that areas with a high percentage of yearly burned area have no herbaceous vegetation and a very small human footprint. Areas with green colouring in the first analysis are in many situations replaced by peach and salmon colouring in the second analysis. That means that there are many locations that have no fires, high herbaceous vegetation cover, and low or moderate human influence.

The correlation between fire frequency and one of the two variables would be high if the biggest part of the map is coloured by the colours that are near the diagonal of the table in Figure 3.3. That means the colours yellow, orange and brown or dark green. Yellow colouring is representing areas with low fire frequency and a low variable value, and orange and brown or dark green colouring represents areas with moderate to high fire frequency and higher variable value. Concerning the result of the correlation with herbaceous vegetation we see that the map is not dominated by yellow, orange and brown. Therefore it can be concluded that there is indeed a moderate positive correlation. Concerning the result of the correlation with human influence we see that half of the map is coloured peach and salmon, i.e. low fire frequency and low human influence. Also there is very few dark green colouring. That indicates that there are very few spots that have both a high fire frequency and a high human influence. Therefore it can be concluded that the correlation between these variables is probably negative and very low.

3.2.1 Task 3: Monthly fire frequency

For the third and last analysis there is again another task defined that will be performed and also for this task a visualisation tool is used. The task for this analysis is formulated as follows: *Describe the monthly variation of fire frequency in Southern Africa between the fire years of 2000 and 2006.* The target of this task is the monthly fire frequency, which can be derived from the L3JRC fire product.

It is necessary to first use a tool that counts the number of fires for all 84 months (multiplication of seven years and twelve months) for a certain area. Because the task is aiming at a variation over a large area, it would be appropriate to use as segmentation a two-dimensional grid with large cell sizes. Therefore again the cells of 200 kilometres are used. This means that before the actual visual analysing tool is performed, a data aggregation tool is necessary to derive the monthly fire count from the fire dataset. The aggregated data then consists of 260 nearly equal sized cells with for each cell the fire count of all 84 months.

Because we are dealing with a large number of data, to be precise 21840 values (multiplication of 260 cells and 84 months), from which we want to achieve some understanding, it is necessary to use an image simplification technique. The visualisation tool that is used to execute the task is called the *class mosaic mapping technique*. With this technique the image simplification can be achieved by ordering marks, i.e. the elements that became graphical, in a convenient order. In the aggregated data, the fire count is placed in the characteristic component, and these are referring to two referential components. The one referential component is the corresponding grid cell and the other is the corresponding month. That means that each fire count value is referring to a temporal component, and because time is a linearly ordered set, it can serve as the referrer that orders the values. A linearly ordered set can be represented within one-dimension. But time can also be seen as a cyclic ordered set, because a certain month comes back every year. In that case, with time as a cyclic ordered set, time elements can be represented within two-dimensions. In our situation it is indeed more relevant to order fire counts by means of a cyclic ordering, with one dimension used to represent a cycle, i.e. twelve months of a fire year, and the other dimension used to represent the subsequent cycles. The marks of each segment are accordingly placed in a two-dimensional arrangement, which we call mosaics. When these mosaics are placed on the map, a single image (see Figure 3.7) represents the monthly fire count at each segment during seven fire years.

As shown in the map of Figure 3.7, for each of the 260 segment cells there is a corresponding mosaic. Each mosaic consists of 84 small square tiles that correspond to the fire count in that segment during that month. Tiles within a mosaic are organised in seven rows and twelve columns. Each row represents a fire year, and each column represents a certain month (starting with April) of the fire year. A tile is shaded according to the number of fires during that month: the higher the number, the darker red the colour. A green tile indicates that there were no fires during that month. The tiles that do have a fire count of more than zero are classified into four classes, because this number is about the maximum of different colours on one gradual colour scale a human can distinguish (Bertin, 1984). The class breaks are 5, 500 and 5000 fires, according to an almost equal distribution of the number of tiles belonging to each class.

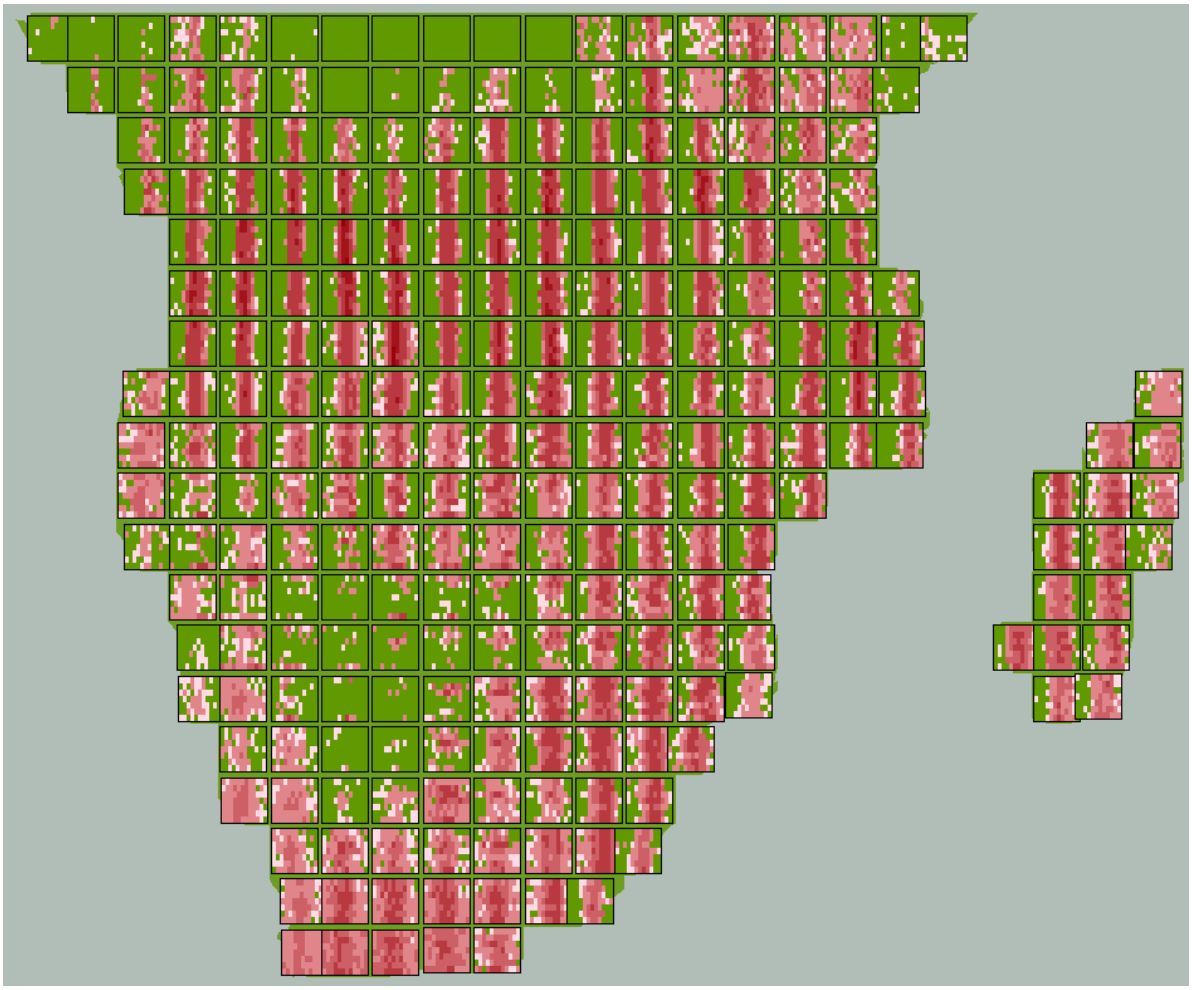


Figure 3.7: Map image of the result of the class mosaic tool performed on the monthly fire frequency of the fire years 2000 till 2006. Each segment cell has a mosaic. The mosaics have 84 tiles that are organised in seven rows, which correspond to the seven fire years, and twelve columns that correspond to the months within a fire year. Each tile has a colour that represents the classified fire frequency. Green means no fire, white means between one and five fires and further the more red the colour the higher the number of fires.

As result, Figure 3.7 shows the behaviour of the monthly fire frequency in Southern Africa during seven fire years from 2000 till 2006. The patterns that can be derived from this image can give a description of the variation of fire frequency. For that reason there are hereafter three major patterns described that are found in the image.

Roughly three types of interpretations of the mosaics can be found, as can be seen in Figure 3.8. Firstly, the mosaics that show no or almost no fires in the seven years time span (indicated with a 1). This type can be found in areas with tropical climate near the equator and in areas with no vegetation, like the Kalahari Desert. Secondly, the mosaics that show only fires during a certain half of the year, but does not show fires during the other half of the year (indicated with a 2). This type is mainly present at the middle part of the study area, between 5° and 15° latitude south. Thirdly, mosaics can be found that show fires throughout the whole year (indicated with a 3). This type of mosaic is particularly present in the most southern part of the study area and in a small area near the equator.

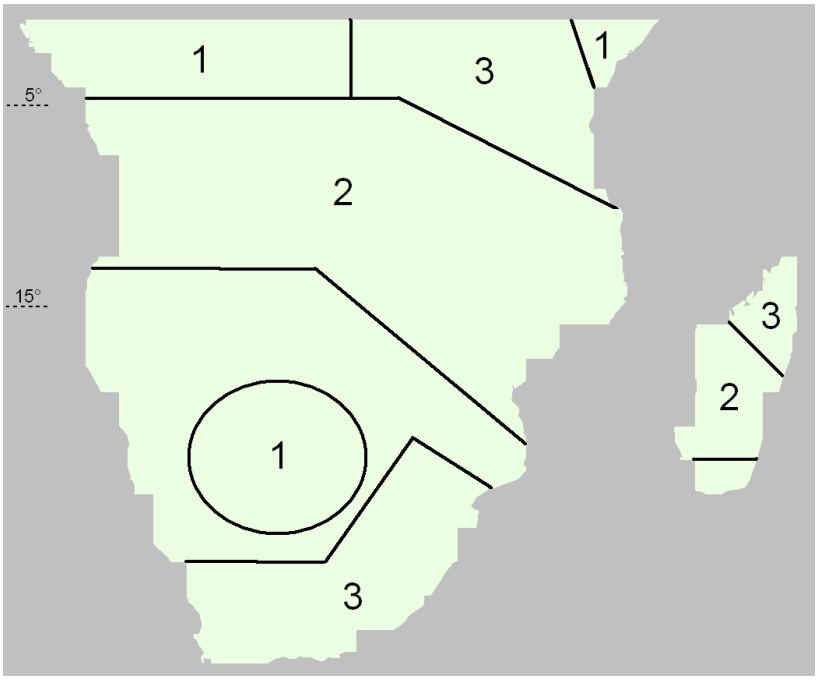


Figure 3.8: Simplified visualisation of the result class mosaic map of Figure 3.7. The fire seasons in the study area are divided in three classes based on the spread of the monthly fire frequency during a fire year. Areas indicated with a 1 have no or almost no fires during the year. Areas indicated with a 2 have mainly their fires during a one half of the year. Areas indicated with a 3 have fires during the whole fire year.

Within a mosaic, the colouring of the seven rows can be compared with each other. By doing this comparison it seems that most patterns of fire frequency throughout the years look almost identical. That means that every fire year within a segment has a fire distribution with mainly the same trend, i.e. a fire season and the fierce of the fire season almost equally every year. Off course these conclusions are dependent on the choices made concerning the classification.

In the image can be seen that there is a difference in the fire season between certain areas. The middle of the fire season is for the whole Southern Africa around September and October. For describing the pattern of fire season we introduce two properties that together describe can roughly characterize fire season. The first characteristic is the temporal length of the fire season. The second characteristic is the fierceness of a fire season, which means the number of fires that burn at the same time. Both properties are strongly related: short fire seasons are fiercer and long fire seasons are less fierce, i.e. less fires at the same time. Southern Africa can roughly be split horizontally in two, with as separator the parallel of 15° latitude south. Above this line most fire seasons are short and fierce (indicated with a 2) and beneath this line the fire seasons are mostly long and moderate (indicated with a 3).

3.3 Discussion

In the last section of this chapter we discuss some patterns about wildfire that we found in the previous sections. At first the found patterns are compared with findings of other researches. The second paragraph considers the fire regime in one specific ecological region, namely the Fynbos Biome.

3.3.1 Patterns of wildfire in other researches

In this paragraph we shortly elaborate on the findings of other researchers about wildfires in Southern Africa in relation to the spatial patterns, the inner-year temporal patterns, the multi-year temporal patterns, and at last, the correlation with other variables.

The spatial patterns of fire frequency in Southern Africa found in this chapter, e.g. the clustering between 5° and 15° latitude and the absence in the Kalahari Desert and near the equator, are also described in several other researches (Moreno-Ruiz *et al.,* 2004/2006; Cahoon *et al.,* 1992; Dwyer *et al.,* 1997/2000; Carmano-Moreno *et al.,* 2005a/2005b/2005c; Riano *et al.,* 2007; Barbosa *et al.,* 1999; Giglio *et al.,* 2006; Justice *et al.,* 2002; Tansey *et al.,* 2004). These researches used burned area or fire products other then L3JRC to describe spatial clusters in (Southern) Africa. But in all cases they used datasets with smaller spatial resolutions. The high spatial resolution of the L3JRC product makes it possible to describe spatial patterns of wildfire in much more detail (Tansey *et al.*, 2008).

Concerning the temporal clustering there are several researches that investigated the inner year fire frequency, and found in general the same patterns (Moreno-Ruiz *et al.,* 2006; Korontzi *et al.,* 2003; Dwyer *et al.,* 1997/2000; Carmano-Moreno *et al.,* 2005a/2005b/2005c; Riano *et al.,* 2007; Barbosa *et al.,* 1999). Almost all of these researches describe the monthly change of fire frequency in a fire year. With the L3JRC product it is however possible to describe the weekly change with a high accuracy. In this research we described fire frequency in much more detail then other researches did. In that way the start and ending of the fire season for a specific region can be described in more detail. Besides that, from these researches only Korontzi *et al.* (2003) describes, just like we did, the multi-year inter-annual variation of fire frequency, which gives much more information about fire frequency.

Differences in fire frequency during multi-year time periods in (Southern) Africa are also investigated in several researches (Moreno-Ruiz *et al.*, 2006; Cahoon *et al.*, 1992; Carmano-Moreno *et al.*, 2005a/2005b/2005c; Riano *et al.*, 2007; Korontzi *et al.*, 2003; Tansey *et al.*, 2004). Compared to the fire datasets that are used in other researches, in order to investigate multi-year fire frequency, the L3JRC covers a relatively short period. Seven years is for example too short for analysing the influence of El Niño on fire regimes.

Relating fire frequency with other variables for the African continent is done by Lehsten *et al.* (2008) and Korontzi *et al.* (2003). Both researches correlated fire frequency in order to estimate wildfire emissions. The researches investigated respectively the yearly and seasonal variation with a low spatial resolution. That means that with the L3JRC the estimations can be estimated with both a higher spatial and temporal accuracy.

3.3.2 Wildfire in the Fynbos Biome

This paragraph elaborates on the reflection of the results that are given in the previous sections by means of knowledge that is found in scientific research. But in this paragraph the focus is on relatively small part of the study area that is called the Cape Floristic Region (located in the far south of the African continent), which is mostly covered with Fynbos. This consists of shrubland or heathland vegetation types and is characterized by its high richness in plant species and its high endemicity (Low & Rebelo, 1996; Forsyth & van Wilgen, 2008). An important issue of the Fybos vegetation is that it must burn every 6 to 45 years of age in order to sustain its plant species. Many species store their fruit in fire-safe cones for release after a fire, and ants are enticed to bury fruit where they are safe from rodents and fire (Low & Rebelo, 1996).

If we investigate the result of the mosaic-mapping technique of Paragraph 3.2.1 we see in the Fynbos region that there are wildfires detected throughout the whole year. There is only a small increase of fire frequency in the fire season, which is in Augustus, September and October. This finding does not correspond with the findings of other researches. Forsyth & van Wilgen (2008) state that 90% of the Fynbos area burnt was due to fires in summer or autumn. The difference in findings is probably caused by the inaccuracy in which we made our conclusion, which is by considering the map of Figure 3.7, i.e. the segment sizes are too large to conclude about the Fynbos Biome. However the mosaic-map shows fires throughout the whole year and that corresponds to the finding of Scott (1993), who stated that Fynbos fires occur in all seasons. Fires in summer and autumn are regarded as optimal for the health of fynbos ecosystems (Van Wilgen *et al.*, 1992). Further research by using visualisations with smaller segment sizes can report whether or not the current fire regime should be a cause for concern.

Chapter 4

The Ignition Type Model

This chapter elaborates on the *Ignition Type Model*. We developed this model to assign the ignition type to wildfires. By that this chapter gives answer to the second research question, which is formulated as: *Can the distinction between lightning ignited and human ignited fires be made by independent variables?*

This chapter consists of four sections. The first section gives the aim of the Ignition Type Model and gives an explanation of the processes that are part of the model. The processes are divided into four distinct parts that are called steps. For each step of the Ignition Type Model in section 4.1 the processes and parameters are explained. Section 4.2 gives the results of all four executed steps and thus the outcome of the model. The third part of this chapter, Section 4.3, elaborates on the verification of the model. Verification is done by analysing the outcome of the model and by comparison of the outcome with the results of the literature review. The last section of Chapter 4 gives some points of discussion about the Ignition Type Model.

4.1 Explanation of the model

4.1.1 Aim of the model

For this research the distinction of ignition types is made by considering all possible causes of wildfires in two groups. The first group contains the *natural-causes* of fire. For this group the primary ignition source is lightning. Other possible natural ignition types, like spontaneous heating and volcanic activity, are not considered in this research. Further, the second group of wildfire causes is the *human-caused* fires, or anthropogenic fires. This includes all kinds of human actions, like land management, arson and human carelessness (Andreae & Merlet, 2001). However, for this research we are not interested in the actual form of human ignition, that a fire is human ignited is enough.

The goal for this chapter is to make a distinction between the two ignition types. This is done by developing the Ignition Type Model. The main input of this model is the L3JRC burned area product, which delivers all wildfires in Southern Africa during seven fire years. For all these fires the model will investigate the ignition type. The processes that are part of the Ignition Type Model can be divided into four steps. Figure 4.1 gives a graphical overview of the processes of the model and the four steps. The first step realizes the extraction of the commencing fire areas from the L3JRC. At first in this step the burned areas from the L3JRC are converted to individual fires and thereafter only the areas of the first detected part of an individual fire is selected. The selection, i.e. the commencing fires, is then used for proximity analysis with both lightning (Step 3) and anthropogenic (Step 4) variables. Based on these

analyses, which are performed independently of each other, the fires that meet with certain conditions are then assigned as lightning or human ignited. Before Step 3 can be carried out, which is the proximity analysis with lightning variables, a selection on the attribute of the lightning dataset is necessary because not every lightning strike causes fire. This selection process is fulfilled in Step 2. The four steps are treated in the next part of this section and for each step the procedure is explained and the chosen parameters are founded.

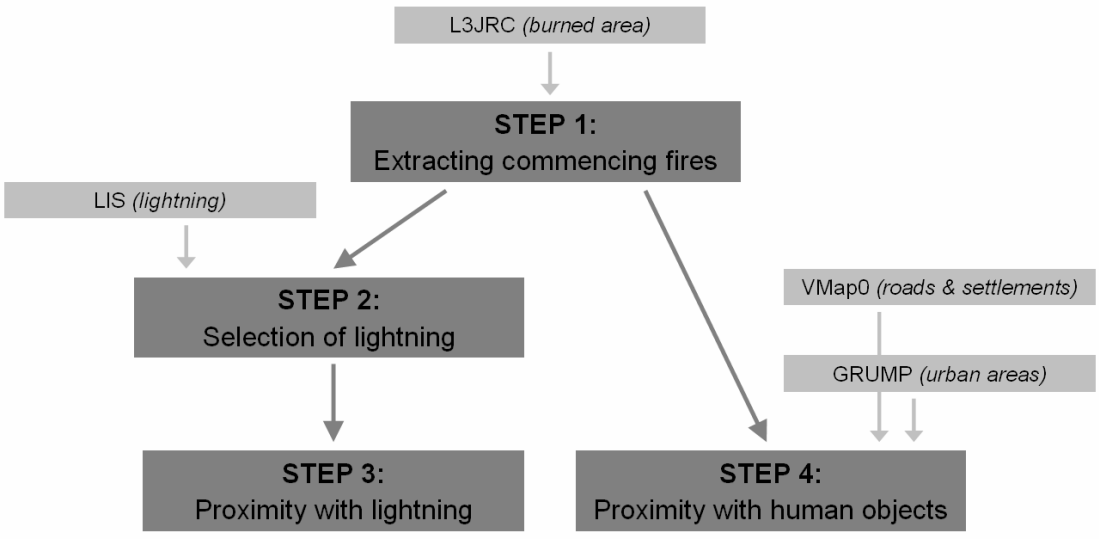


Figure 4.1: Process and dataflow diagram of the Ignition Type Model. The four processing steps are represented in dark grey rectangles and the pale grey rectangles represent the datasets required by the processes. The arrows represent the flow of data

4.1.2 Step 1: Extracting commencing fires from L3JRC

A wildfire originates, regardless of the ignition type, at a certain moment in time and space. From the location where the wildfire originated it will expand and or relocate in the course of time. For this research it is only relevant to know that wildfires of the L3JRC did expand, at least to a certain size so that detection of burned area is possible in the SPOT-VGT data. Because the fires are expanding and relocating it is possible to distinguish in the L3JRC dataset *burned area of commencing fires* from *burned area of continuing fires*. In this research we are only interested in the burned area of commencing fires because these can tell something about the situation in which the fire started. The problem is that the L3JRC does not indicate whether or not adjacent burned areas are caused by the same or by several fires. Therefore, to extract the commencing fires, it is necessary to do assumptions about which burned areas from the L3JRC dataset do and which do not belong to commencing fires. The assumptions are written hereafter and these form the prescriptions of the performed extraction.

The first assumption is defined with the purpose of assigning multiple burned areas as caused by individual fires. This is done by assembling burned areas of adjacent grid cells at a particular time state with each other (see the example of Figure 4.2). The cell sizes have a surface area of 1 sq km, so burned area that is detected at the same time in two adjacent cells, that means within 2 sq km, is assumed to be assigned to the same fire. The detection of adjacent cells is based on the von Neumann neighbourhoods, which means that the four cells that are located one cell away, horizontally or vertically, are taken into consideration (Missoum *et al.*, 2005). Of course there will be several situations in which this assumption is

not true. But for this research the assumption complies because even two fires within 2 sq km have most likely the same properties and consequently the same ignition type.

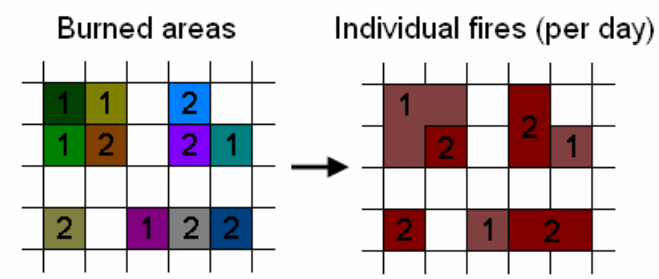


Figure 4.2: The first step of the extraction of commencing fires consists of assembling adjacent burned areas. The day of detection (day one or day two) is given for each cell with burned area. All individual burned areas from the L3JRC fire dataset are assembled, based on the von Neumann neighbourhoods, to individual fires per day.

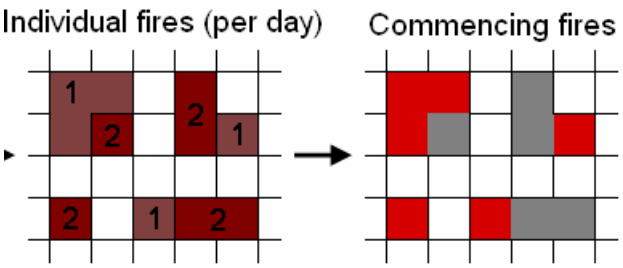


Figure 4.3: In the second step of the extraction of commencing fires the individual fires per day are assembled with adjacent fires of the previous or next day. This results in new multiple day fires that are then split in commencing fires (red) and continuing fires (grey). Only the commencing fires are useful in this research.

The second assumption is that burned areas are caused by the same fire, if they are in different time states, separated by one day maximum, and in addition to this also meet the first assumption. This step is illustrated with an example in Figure 4.3. With the second assumption fires at different time moments are assembled if they are detected in adjacent cells at adjacent days. Subsequently, the burned areas of a fire that are detected at first are assigned as commencing, and are selected for further use in the Ignition Type Model.

4.1.3 Step 2: Attribute selection on LIS

Lightning can be divided based on the target of a strike into cloud-to-ground (CG) and cloud-to-cloud (CC). The latter one occurs six times more often than the former one (Preston-Whyte & Tyson, 1988). Of course the CC lightning can not cause a wildfire. Next, CG strikes can be divided based on its discharge into positive and negative. The lightning strikes with a positive discharge are far more treacherous for igniting wildfire because these strikes last longer and carry ten times more power with them. Of all CG strikes, about ten to twenty percent is a positive strike (NOAA, 2006). The lightning strike that is positively discharged can reach temperatures approaching 30000 °C (UCAR, 2007). There is little known about the form of lightning that ignites wildfire, like for example what percentages of positive CG lightning ignites wildfire.

The LIS sensor detects both CG as CC lightning but does not make a distinction between them. The same is true for positive and negative strikes because they are also both detected but not distinguished. The LIS conversely does register for each item a couple of properties about the lightning strike that can be used to remove at least the items that do not consist of positive CG strikes, and therefore can not cause wildfire. The fact is that small

thunderstorms do not generate positive strikes and as the thunderstorm becomes larger the occurrence of positive CG strikes becomes more frequent (Larjavaara *et al.*, 2005). And a higher density of CG strikes gives a higher chance for ignition (Fuquay *et al.*, 1980). Now the intention is to perform a selection on one of the attributes of the LIS dataset.

For each LIS item the following attributes are present: *Groups*, *Events* and *Radiance*. This corresponds to more familiar physical features such that an item is a thunderstorm, Groups are flashes, and Events are strikes. The other attribute, Radiance, is the measured cloud-top radiance of the thunderstorm. It is the radiance as seen from space, and not radiances directly from the lightning source. The detected value is a result of multiple scatterings of the lightning flash within the cloud and that makes the detected value dependent on the optical depth of the concerned cloud. Therefore this attribute is not valid for relating it to the size of a thunderstorm. By contrast, the other two attributes, Groups and Events, together do give an indication of the size of a thunderstorm. Event is a subdivision of Group, which means that Event is the sum of all strokes that occurred in all groups within a thunderstorm. A single attribute, named Events, is consequently used for deselecting the thunderstorm that does not bring about wildfire.

So for the second step we need to define a threshold value for the selection of LIS items that can cause wildfire. This threshold value defines the amount of events a thunderstorm should consist of to involve the thunderstorm with the continuation of the Ignition Type Model. The frequency distribution of the Event attribute is visualised in Figure 4.4. Further a univariate analysis on the Events attribute of the LIS dataset, containing 1196012 records, is performed and the results are given in Table 4.1. In this table there are a number of parameters that quantify the central tendency of the attribute. The mean value and the median value differ strongly, which means that the univariate dataset consists of outliers. Therefore the mean is not useful as the threshold value. It is better to use the median parameter, i.e. 24 events, for it. The value of the median parameter of the Events attribute will be the threshold that represents the large thunderstorms that can consist of strikes that cause wildfire.

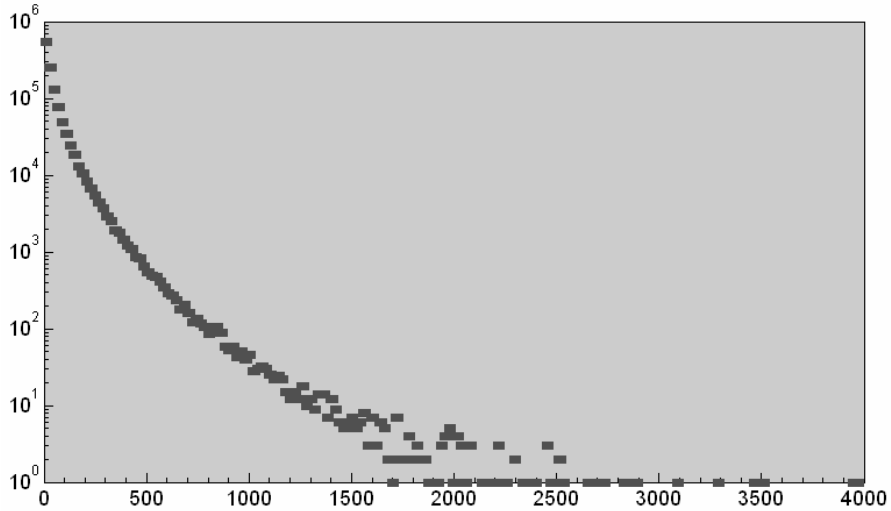


Figure 4.4: Distribution histogram of the Events attribute from the LIS (lightning) dataset. On the X-axis the Event values are given and on the Y-axis with a logarithmic scale the frequency is portrayed.

Table 4.1: The Events attribute in the LIS dataset contains the number of lightning strikes a thunderstorm consists of. The results of the univariate analysis of this attribute are shown in this table.

Statistical parameter:	Result value:
Minimum	1
Median	24
Arithmetic mean	49.51
Standard deviation	78.97
Maximum	3978

4.1.4 Step 3: Proximity of wildfire with lightning

For Step 3, the input datasets are the commencing fires that are extracted from the L3JRC product and the larger thunderstorms that are extracted from the LIS product. The aim of Step 3 is to create spatiotemporal relations between the lightning's and wildfires. Therefore two parameters are of interest. These parameters indicate the threshold value of spatial distance and the threshold value of temporal duration between a thunderstorm and a fire. Having a distance and duration lower than the threshold value, then the fire can be assigned as lightning induced.

The first parameter of interest is the distance threshold between a lightning and a fire. To define the parameter we should know something about the horizontal distance that a lightning stroke can cover between the anvil portion of a storm and the location on the ground where the stroke starts a fire.

To determine this distance we first investigated the agility of the (larger) thunderstorms. The difficulty is that the LIS product observes a single thunderstorm for only 80 seconds (Christian *et al*., 1992). So the LIS does not register where the storm is moving to, with what speed and for how long it last. But besides the agility of a thunderstorm we can instead investigate the reach of a lightning strike. According to NOAA (2006) a lightning can strike the ground 16 to 24 kilometres from the anvil portion of a thunderstorm.

But then, if we use as distance threshold of 16 kilometres, it means that around every single thunderstorm all grid cells within a radius of 16 kilometres can potentially be related. That means potentially that 804 sq km of burned area can be assigned as caused by one thunderstorm. Seeing that this is not realistic, the choice is made to use a smaller threshold, of exactly the half of the investigated distance, i.e. 8 kilometres. That means that lightning strike of a large storm can reach an area of 201 sq km. Because the size of this threshold value we can neglect other factors that could have influence on this parameter, like the accuracy of both datasets.

The second parameter to create a relation between fires and lightning concerns the time span between the two. The question that we should ask to determine a threshold is how long it takes between the ignition of wildfire and the detection of burned area. On the one hand we should consider that the L3JRC detects burned area only when the circumstances are good enough. In many situations it can occur that the burned area is not detected at the very first satellite overpass (NASA-GHCC, 2005). On the other hand the fire needs some time to expand to a certain size at which it can be detected. The speed of expanding is depending on the fuel conditions. In substantial situations it can last a couple of days before a fire becomes visible (Rorig & Ferguson, 1999). The parameter value for the temporal duration between the induction of fire and the detection of burned areas is therefore assumed as maximum three days.

4.1.5 Step 4: Proximity of wildfire with anthropogenic

The fourth step of the Ignition Type Model is the spatial proximity analysis between wildfires and anthropogenic variables. Based on the distance between a fire and a human object we can assign to a fire whether it is ignited by humans. For the representation of anthropogenic variables we use two datasets (VMAP0 and GRUMP) that together give the objects of roads, urban and settlements. The distance threshold for the proximity analysis for each variable is explained hereafter.

The road network of the VMap0 is classified as divided highways, primary roads, path/trails and railroads. After investigating the dataset it appears that the classification is inconsistent across countries. As can be seen in Figure 4.5, there is a sharp interruption of the road network in the dataset between primary roads and path/trails. Highways do not seem to exist in Southern Africa according to the VMap0. That this inconsistency is caused by serious limitations with this dataset is stretched in Nelson *et al.* (2006). These interruptions across tile lines or country boundaries make it impossible to perform further analysis based on the initial classification. Therefore there is defined only one threshold value that is valid for all roads from VMap0.

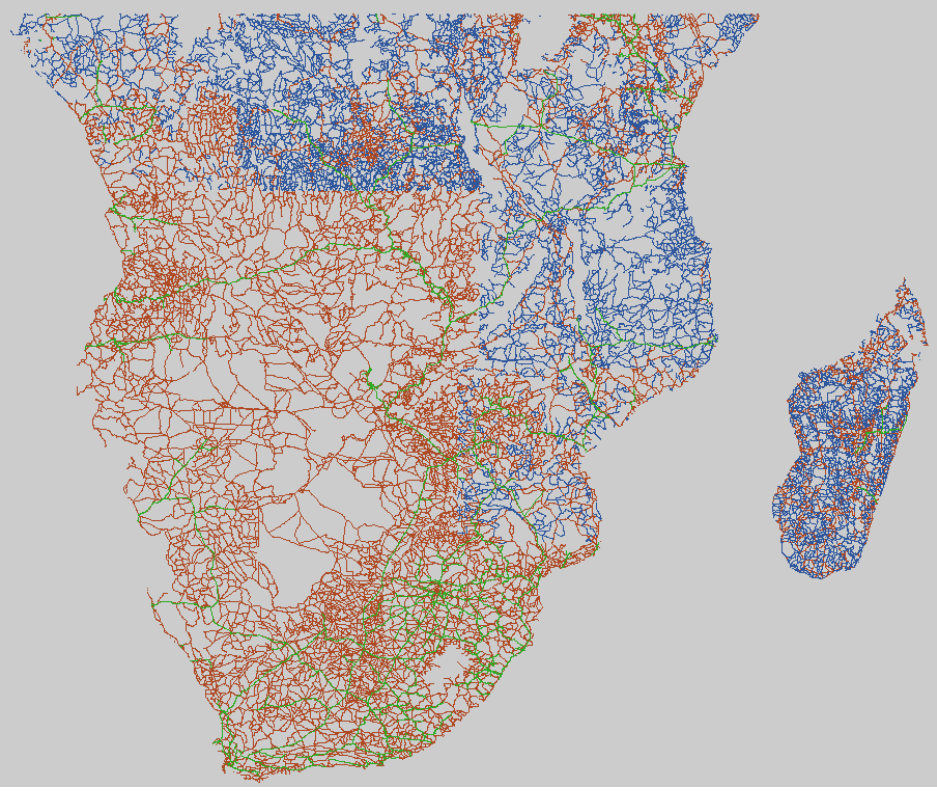


Figure 4.5: Map visualisation of the VMap0 road dataset. The initial classification is used, which is found inconsistent due to the sharp interruptions at tiles and boundaries. Primary roads are red, path/trails are blue, and railroad is represented with green.

The distance threshold value of roads will be defined by performing a univariate statistics analysis on the Euclidean distances of all fires to the nearest road. The nearest distances are graphically displayed in a histogram as can be seen in the illustration of Figure 4.6. The frequency distribution is organized in classes of 0.2 km size. The distances of lower than 0.6 km are excluded in this illustration. This is done because the structure of the fire dataset has a point feature that represents the state of a grid cell that is overlapping a radius of roughly

0.6 km from the point. Fires that are less than 0.6 km from a road are anyhow assigned as human ignited. The histogram shows a distribution with a positive skewness, which indicates that the distribution is spread out more to the right of the mean value. The idea is to define a threshold value based on the illustration of Figure 4.6. This threshold value should point the distinction between the fires that are human ignited and the fires from which the ignition type is unknown. Between the bars of 1.0 km and 1.4 km there is a slightly higher kurtosis than between other bars. This indicates that distances at the left of the beginning of this kurtosis representing fires near roads that are ignited by humans. Therefore 1.2 km is the chosen as threshold value for roads.

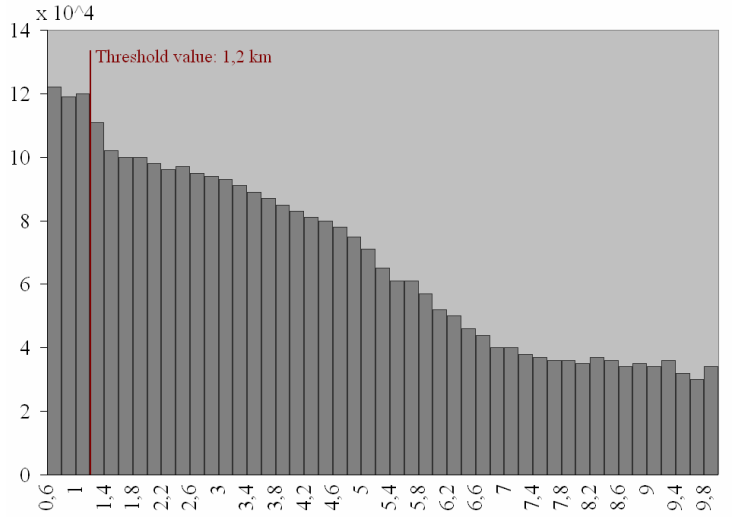


Figure 4.6: Distribution histogram of Euclidean distances between each fire and its nearest road. The histogram is set up with a value range from 0.6 till 10 km and it has class sizes of 0.2 km. Based on this histogram a threshold value of 1.2 km is defined for the proximity analysis with roads.

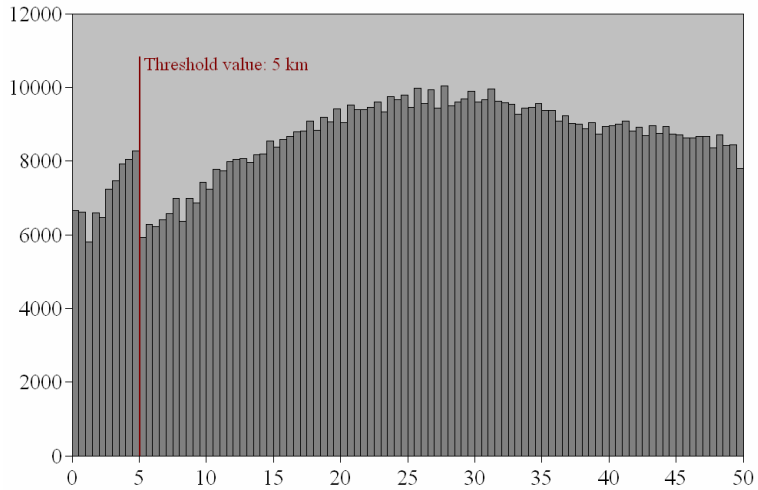


Figure 4.7: Distribution of Euclidean distances between fires and the nearest urban areas. The histogram has a value range from 0 till 50 km and classes of 0.5 km. Based on this illustration the defined threshold for the urban proximity analysis is 5 km.

The second part of the proximity processing of fires with anthropogenic variables is done with the means of the GRUMP dataset (see Section 2.1 for a description). This dataset consist of all urbanized areas in Southern Africa. The fires that occur within urban extents are assigned as human caused fires. Subsequently the same univariate analysis on the Euclidean distance between fires and the closest urban area is performed. The result, again a

distribution histogram, is shown in Figure 4.7. The frequency distribution is organized in classes of 0.5 km and the distances ranges from 0 till 50 km. The shape of the histogram shows a bimodal distribution, i.e. the histogram has two modes. The first mode has a strong negative skewness that peaks at 5 km and has a large drop in frequency thereafter. Therefore the first mode is used as the threshold value for the distance between a fire and urban areas for assigning the human ignition type.

The third and last part of this proximity process of anthropogenic variables concerns the nearest distance of fires to settlements. The settlements are point features that are gained from the VMap0 dataset. Feature elements that are overlapping with urban areas of GRUMP are removed. This resulted in total 13761 settlements. Also on this dataset the same analysis is performed as with the previous variables, and the result distribution histogram is shown in Figure 4.8. The histogram is in the same way organized as the one with the urban areas, so with a range from 0 till 50 and class sizes of 0.5 km. In the illustration can be seen that the distribution of distances from fire to settlement is totally different as the one with urban areas. Settlements are not attracting fires with the same force as urban areas do. Based on the illustration it seems that most fires are between 5 km and 10 km distance from a settlement. But when a random spread of 13761 points in the study area is used to measure the distance of each fire to the nearest random point it also results in a similar distribution. This can be seen in the frequency histogram that is shown in Figure 4.9. This figure shows the same distribution as Figure 4.8, with the difference that the nearest distance to settlements has a mode around 7 km and the random points around 12 km. That means that the settlements are noticeably attracting wildfire. But this also means that settlements are not useful for further use in the Ignition Type Model, because it is not possible to define a threshold value that represents wildfires that are ignited by humans.

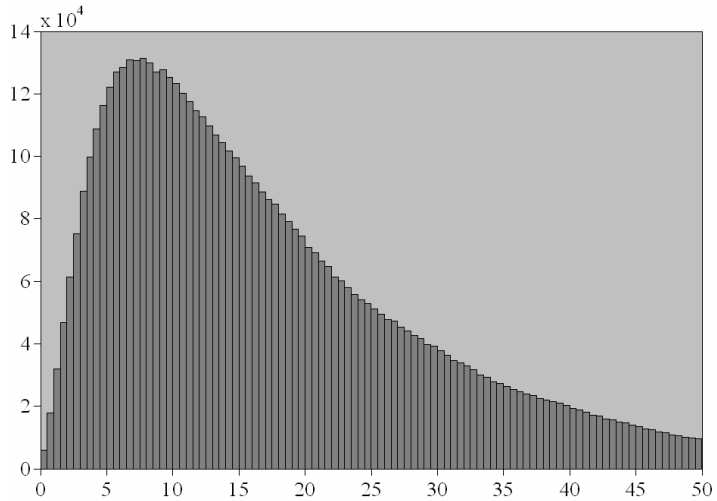


Figure 4.8: The histogram gives the distribution the Euclidean distances from each fire to its nearest settlement. The histogram has a value range from 0 till 50 km and classes of 0.5 km.

Figure 4.9: Distribution of Euclidean distances between each fire and one of the 13761 random points (which is the same amount as the amount of settlements). The histogram further has the same properties as the histogram of Figure 4.8.

This means that for Step 4 two anthropogenic variables are used, namely roads and urban areas, to assign the human ignition type to wildfires. For both variables a threshold value is defined that declares whether or not the human ignition type can be assigned to a fire.

4.2 Results

4.2.1 Step 1: Commencing fires

The 6067178 burned areas of the L3JRC dataset are transformed to individual fires based on the distances in space and time between burned areas. This resulted in a total of 1690546 individual fires in Southern Africa in seven fire years (with an average of 3.6 assembled burned areas per fire). From these fires the locations of the firstly detected burned areas are selected for further use in the ignition type model. Table 4.2 shows the results of the extraction in numbers for each fire year.

Table 4.2: For each fire year: (1) the amount of burned areas in the L3JRC dataset, (2) the amount of fires that are the result of the assembling of the burned areas, (3) the amount of burned areas that are assigned as commencing, and (4) the amount of burned areas that are assigned continuous.

Fire year	# of burned areas in L3JRC	# of fires	# of commencing burned areas	# of continuous burned areas
2000	843650	236480	730768	112882
2001	945623	243844	832427	113196
2002	839260	233777	743802	95458
2003	946246	268791	833265	112981
2004	783875	222108	693802	90073
2005	765439	223133	686257	79182
2006	943085	262413	823130	119955
Total:	6067178	1690546	5343451	723727

4.2.2 Step 2: LIS attribute selection

As threshold for the attribute selection, in order to deselect the thunderstorms from LIS that does not cause wildfire, the median value is chosen. Therefore exactly half of the original 1196012 items remain, i.e. 598006 items, for further use in the ignition type model.

4.2.3 Step 3: Lightning proximity

For the processes of Step 3 we used the 1690546 (commencing) fires of Step 1 and the 598006 thunderstorms of Step 2. The two datasets are related, and next, the fires are selected that meet the two threshold values of maximum spatial distance and maximum temporal duration. This process resulted in 35509 fires that are therefore assigned as ignited by lightning. A yearly overview of the results is shown in Table 4.3.

Table 4.3: The results of Step 3 per year. First is given the number of fires, then the number of thunderstorms and last the number of fires that are assigned as ignited by lightning.

Fire year	# of fires	# of thunderstorms	# of lightning fires
2000	236480	70538	5229
2001	243844	97830	5701
2002	233777	66614	3967
2003	268791	104228	5364
2004	222108	83778	4966
2005	223133	73605	4016
2006	262413	101413	6266
Total:	1690546	598006	35509

4.2.4 Step 4: Anthropogenic proximity

The results of Step 4 are described in two parts, starting with the proximity analysis with the road variable and thereupon the proximity analysis with the urban area variable. In the first part the proximity analysis is performed to indicate which of the 1690546 fires are within a distance of 1.2 km from the road network with a total length of 713097 kilometres. It turned out that in seven fire years a total of 263551 fires start within this road buffer. This result counts for 15.6% of all fires. The results per fire year are given in the third column of Table 4.4. The total buffer surface area around the roads in Southern Africa is 1007238 sq km. So the buffer area consists of 9.6% of the total study surface area (which is 10510527 sq km). That means that 15.6% of the fires are located in 9.6% of the area. Therefore we can conclude that roads do attract fire and that it is more certain that the fires we assigned as anthropogenic ignited are actually anthropogenic ignited.

The second proximity analysis is done by means of the urban areas. From a total of 769 urban areas spread over Southern Africa the fires are selected that overlap the objects or are within a buffer distance of 5 km from the objects. In this way a total of 44010 fires are selected and assigned to as human ignited. The results per year are given in the fourth column of Table 4.4.

Table 4.4: Given per fire year: (1) the number of wildfires, (2) the amount of fires that are the result of the road proximity analysis, (3) the amount of fires that are the result of the urban areas proximity analysis, and last (4) the total result of both proximity analyses.

Fire year	# of fires	# of 'road fires'	# of 'urban fires'	# of anthrop. fires
2000	236480	37882	5673	41741
2001	243844	39833	7723	45133
2002	233777	37783	6934	42512
2003	268791	40560	6447	45044
2004	222108	33452	4691	36804
2005	223133	33597	5747	37608
2006	262413	40444	6795	45147
Total:	1690546	263551	44010	293989

The last column of Table 4.4 shows the total number of fires that are assigned as human-caused as a result of the proximity analyses with both variables. For the seven fire years period there are 293989 wildfires assigned as human ignited, which is 17.4% of the total number of wildfires. Note that the values of the last column are not the sum column (1) and (2). This because lots of fires are both near roads and near urban areas and therefore double assigned as human-caused.

4.3 Verification

In this section we verify the Ignition Type Model based on the results as they are shown in Section 4.2. Part of the verification is already performed in the previous sections of this chapter. For example the verification of some parameters is done while choosing the parameters by considering its distribution. Also in Section 4.2, in which the results of the human proximity analysis are given, there is verification performed. This concerns the comparison of the buffer surface area ratio with the ratio of selected fires that received the anthropogenic ignition type. This section contains five different parts and for each part a manner of verification is performed by considering the results of the Ignition Type Model. All amounts are still representing the seven fire year's period.

4.3.1 Sensitivity of the parameters

For each step in the Ignition Type Model one or more parameters are defined and used. These parameters form the most important aspects of the model because these contain the criteria for which a fire receives an ignition type or not. In Section 4.1 where the parameters are chosen there is a justification included. In case of the lightning selection and proximity analysis this is done by performing multivariate statistics and consulting other research. In case of the proximity analysis with anthropogenic objects the parameters are chosen with the help of distribution analyses of Euclidean distances. For additional justification it is necessary to investigate the sensitivity of the parameters. This can be done by performing the model with other parameter values and then by comparing the model outputs the sensitivity of each parameter can be described. The problem herein is that executing the model is time consuming because of the complexity of the model and the extension of the input datasets. Therefore it is for this research not possible to further describe the sensitivity of the parameter values that are used in the Ignition Type Model.

4.3.2 Double assigned ignition types

All fires that are produced in the first step of the Ignition Type Model make up the input for Step 3, which is the proximity analysis with thunderstorms. The identical dataset with fires is also used for Step 4, which is the proximity analysis with human objects. Theoretically this means that it is possible that a fire is assigned as both natural and anthropogenic, i.e. the fire is related with a thunderstorm and also near to a road or urban area. The first way of verifying the Ignition Type Model concerns the investigation of double assigned ignition types. None or little double assignments signify that the steps are processing the two ignition types separately, which will result in reliable results. Many double assignments signify that there is something wrong in the model, like for example that the parameters are not correct. Investigation of the double assignments in the results as shown in Section 4.3 shows that there are 1294 fires that received two different ignition types (see also Figure 4.10). Concerning the lightning fires this implies 3.6% and concerning the anthropogenic fires this implies 0.4%. This means that the assignments are reliable, and that the anthropogenic assignment is little more reliable then the lightning assignments. For the record, we remove the double assigned fires from the results of the Ignition Type Model. The results of the total lightning and anthropogenic fires are altered to respectively 34215 and 292695 fires.

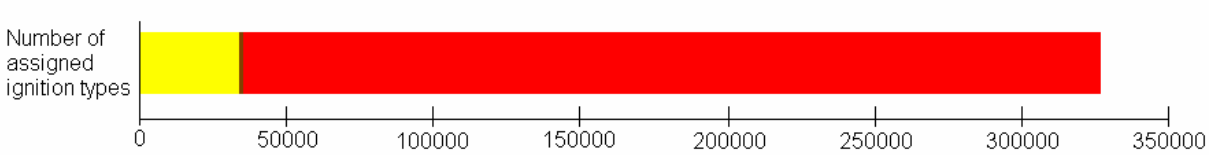


Figure 4.10: The total amount of wildfires for which an ignition type is assigned by the Ignition Type Model (i.e. 328204). The yellow bar represents the 34215 lightning ignited fires and the red bar represents the 292695 anthropogenic ignited fires. A total of 1294 fires (represented by the narrow brown bar, in between the yellow and red bar) received both the lightning as the anthropogenic ignition type.

4.3.3 Performing the model with and without Step 1

The first step of the Ignition Type Model performs the assembling of burned area and the selection of the commencing areas. The model can also be executed without performing Step 1, so that each single burned area is treated as one single fire. By comparing the results of an executed model with and without performing Step 1 might give a verification of the first step. Therefore in this paragraph the comparison is carried out. The results of the Ignition Type Model with and without Step 1 are shown in Table 4.5.

Table 4.5: Results of Ignition Type Model with and without performing of Step 1, i.e. the assembling of burned area and the selection of the commencing areas. The percentages in parentheses for lightning fires and anthropogenic fires give the proportions in relation to the total number of fires. The percentages in parentheses for double assigned fires give the proportions in relation to the sum of lightning fires and anthropogenic fires.

	Step 1 included	**Step 1 excluded**
# fires	1690546	6067178
# lightning fires	34215 (2.0 %)	127777 (2.1 %)
# anthropogenic fires	292695 (17.3 %)	785049 (13.0 %)
# double assigned fires	1294 (0.4 %)	3801 (0.4 %)

When comparing the values of the executed model 'Step 1 included' and 'Step 1 excluded' (see Table 4.5) the first thing to notice is that there is a large difference between the total number of fires that form the input of the Ignition Type Model. For the lightning proximity

part of the model this difference does not give a significant discrepancy in the results. With or without performing Step 1 gives the same proportion of lightning fires that are output in relation to the number of fires that are input. What was more likely to expect is that the proportion of lightning fires in 'Step 1 excluded' is much larger. This because lightning fires are in general larger (Podur *et al.*, 2003), so cause more burned area, and therefore it was expected that a single lightning fire in 'Step 1 included' exists of multiple lightning fires in 'Step 1 excluded'. In other words, against our expectations, the input of 'Step 1 included' consists on average of 3.6 burned areas and the lightning-output of 'Step 1 included' consists on average of almost the same, i.e. 3.7, burned areas.

In contrast to the lightning fires, the anthropogenic fires have a difference in proportion between the input and output. Of the input of 'Step 1 included' 17.3% is anthropogenic and of the input of 'Step 1 excluded' 13.0% is anthropogenic. This means that the fires that are assigned as human ignited in 'Step 1 included' are small fires, i.e. fires with relatively few burned area cells. This finding corresponds with earlier research that concluded that anthropogenic fires are typically smaller is size (Pyne, 1994; Larjavaara *et al.*, 2005).

The double assigned fires form in both ways of performing the model the same proportion in relation to the input of the models. These are therefore most likely the same group of fires, that are both assigned as lightning and human ignited. To conclude about this verification process, we can say that assembling of burned area, i.e. Step 1, is not crucial.

4.3.4 Fire frequency versus thunderstorm size

In the Ignition Type Model a selection is performed on the LIS dataset (containing 1196012 thunderstorms) because not all thunderstorms do bring forth lightning strikes that can cause wildfire (see Paragraph 4.1.3). For the selection a parameter is defined, which we verify in this paragraph, that represents a minimum number of strikes a thunderstorm should consist of. The thunderstorms that belong to this selection, that are 598006 thunderstorms, are then used for the proximity analysis with wildfires. This resulted in 34215 thunderstorms that produced lightning that ignited a wildfire.

So there are in the model three lightning datasets: the initial LIS thunderstorms, the selected thunderstorms and the thunderstorms that ignited wildfire. These three datasets all consist of the Event attribute, which signifies the amount of lightning strikes coming from each thunderstorm. From these three datasets the correlation between Events attribute and fire frequency is calculated and visualised. The goal of this analysis is investigating if there actually is a positive relation between the size of a thunderstorm and natural caused wildfires, as been stated in Larjavaara *et al.* (2005). If this is indeed true then we correctly included Step 2 in the Ignition Type Model.

In these three correlation analyses one variable that is used in all three cases is the fire frequency. Actually we are only interested in the natural caused wildfires, but we do not have a dataset with only natural wildfires. We can however exclude the anthropogenic results from the whole fire dataset. That means that from the total 1690546 wildfires in the study area the anthropogenic model results, i.e. 292695 fires, are removed. This results in a new fire dataset with 38.4% lightning fires and 61.6% anthropogenic fires, which is then the best available fire dataset for the analyses. From this fire dataset we make an aggregated set with average number of fires per year per segment of 40000 sq km. An unclassified choropleth map of this variable is shown in figure 4.11.

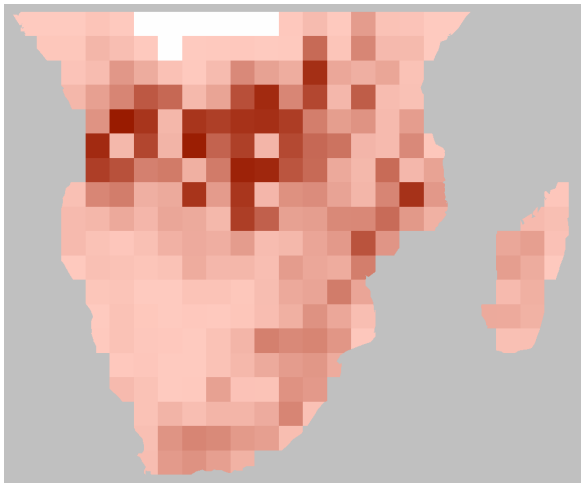


Figure 4.11: Unclassified choropleth map of the average yearly fire frequency per segments of 200 by 200 kilometres. Seven segments have on average of zero fires a year. The other 253 segments have an average fire frequency per year ranging from 1 to 15000.

Also for the three lightning datasets an aggregation is performed. The segment size for these aggregations is also 40000 sq km and for each segment the average number of lightning strikes per thunderstorm is calculated. The three aggregated datasets, i.e. respectively (1) all LIS data, (2) the results of the selection (Step 2 of the Ignition Type Model), (3) and the results of the proximity analysis (Step 3 of the Ignition Type Model), are visualised in Figure 4.12 by means of identical choropleth mapping techniques. By investigating the three lightning strike maps in comparison with the fire frequency map it becomes clear that visual analysing is very hard because of the indistinct patterns.

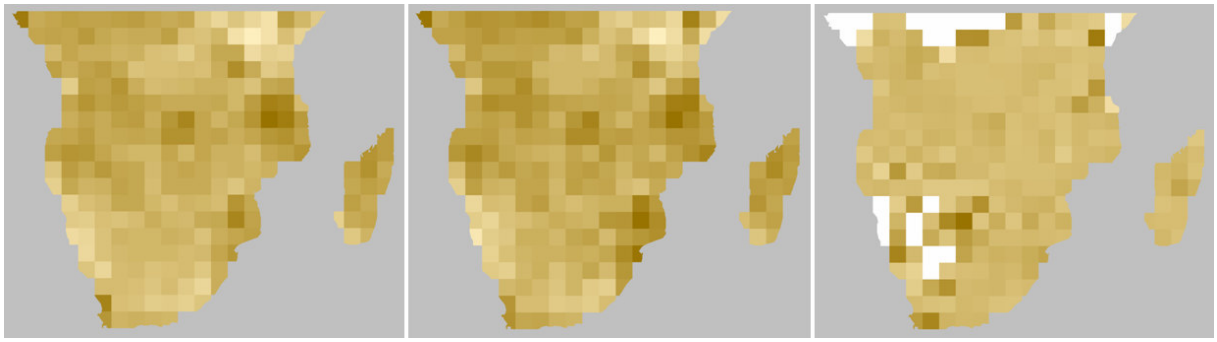


Figure 4.12: The average Event attribute values (number of strikes in a thunderstorm) visualised by means of choropleth maps of respectively all LIS data, the result of the attribute selection (Step 2), and the result of the proximity analysis (Step3).

By just considering the four maps it is impossible to conclude about the relation between the sizes of thunderstorms and fire frequency. Therefore an additional technique is used, which are the three correlation graphs that are shown in Figure 4.13. The graphs contain a linear regressed trend line, which in this case visualises to what extent the two variables are related. By investigating the three trend lines it becomes clear that the third graph has a steeper positive line. This indicates that the sizes of thunderstorms are positively related to fire frequency for thunderstorms that are already allocated with fires. The sizes of thunderstorms that are not allocated have only a very small positive relation with fire frequency, as can be seen in the first two graphs of Figure 4.13.

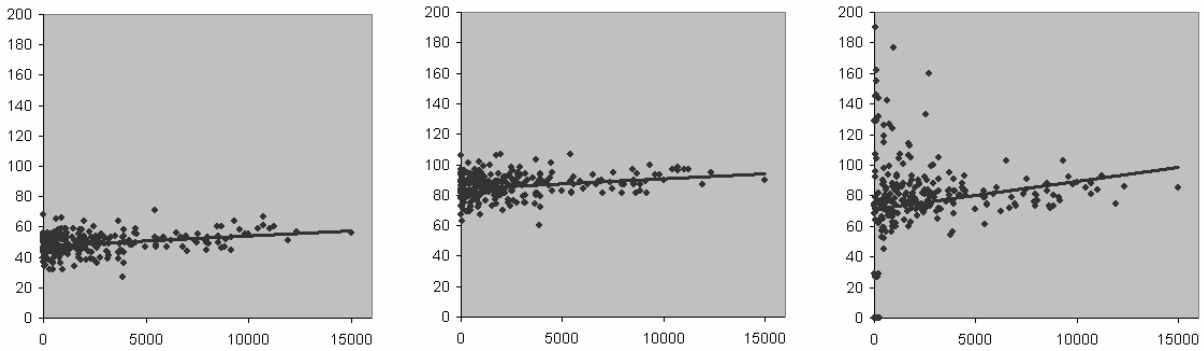


Figure 4.13: Correlation graphs of fire frequency with the thunderstorm size of respectively all LIS data, the result of the attribute selection, and the result of the proximity analysis. Each graph is set up with the average number of lightning events per thunderstorm per segment cell on the Y-axis and the average yearly number of fires per segment cell on the X-axis. Additionally a linear regressed trend line is added to each graph.

Based on this analysis it can be concluded that the assertion of Larjavaara *et al.* (2005) is partly true. The size of a thunderstorm is positively related with fire frequency, but this prevails only for larger thunderstorms that are allocated with wildfire. Furthermore we can also conclude that the use of lightning attribute selection is correctly used in the Ignition Type Model

4.3.5 Effectiveness of the model

In Paragraph 4.4.2 the calculations are given that results in the knowledge that for 19.3% of the input fires the ignition type is defined. Also in that Paragraph the total ratio of both ignition types is calculated and with that we can calculate the effectiveness of the Ignition Type Model per ignition type. The total number of lightning fires is 535897 (see Paragraph 4.4.2 for the calculation) but the outcome of the model contains 34215 lightning fires, so the effectiveness of the model concerning lightning fires is 6.4%. This low outcome for lightning fires is mainly caused by the fact that the LIS dataset, having 59240018 lightning strikes, contains only 6.4% of the total amount of lightning in the study area, which are almost one billion lightning strikes (see Paragraph 4.5.3 for more information hereabout). The same percentage can also be calculated for human fires. The amount of human fires is the difference between the total amount of fires and the total amount of lightning fire, which are then 1154649 fires. The outcome of the model contains 292695 human fires. That means that the effectiveness of the model concerning human fires is 25.4%. Now we know that one quarter of all human fires are located near roads and urban areas. The results that are given in this paragraph are summarized in Table 4.5.

Table 4.5: Overview of the results of the model and the total amount in the study area per ignition type and the sum of the two ignition types. The last column gives the effectiveness of the model in percentages.

	Model results	**Total in study area**	**Model effectiveness**
Lightning ignited fire	34215	535897	6.4%
Anthropogenic ignited fire	292695	1154649	25.4%
Total	326910	1690546	19.3%

4.4 Discussion

This section discusses three different aspects of the Ignition Type Model. This includes (§4.5.1) the percentage of lightning that ignites wildfire, (§4.5.2) the ratio of all wildfires in the study area that is lightning- or human ignited, and (§4.5.3) the assessment of the results of the Ignition Type Model.

4.4.1 Lightning that ignites wildfire

This paragraph elaborates on the comparison of lightning that ignites fire as found in literature and as calculated from the results of the Ignition Type Model. A lightning can be classified as either cloud-to-ground or cloud-to-cloud, where the latter outnumbers the former by a ratio of 6.3:1 near the equator and 4.2:1 at latitude of 30° (Collier *et al.*, 2006). Based on this the average ratio is calculated for the whole study area by linearly interpolating and averaging by means of surface area. This resulted in a ratio of 5.17:1, which means that 16.2% of all lightning is CG. Next, the ratio of CG lightning that ignites wildfire is found in literature. By studying the ignition type of 17 000 naturally ignited wildfires in the United States, Hall (2007) found that 0.35% of all recorded CG lightning strikes are associated with these fires. By considering the results of the literature study (i.e. 6.2% of all lightning is CG and 0.35% of all CG lightning ignites wildfire), the ratio of all lightning that ignites wildfire is 0.0567%.

The lightning dataset used as input for the Ignition Type Model contains a total of 59240018 lightning strikes. The model assigned to 34215 fires that they are lightning-ignited, what means that the same amount of lightning strikes is responsible for the ignition of these fires. The percentage of all lightning that ignited a wildfire is therefore 0.0578%. This ratio does slightly differ from the ratio found in literature, i.e. 0.0567. Based on this the lightning part of the Ignition Type Model can be seen as reliable.

4.4.2 Total ratio of natural-/anthropogenic fire

The Ignition Type Model assigned an ignition type to a total of 326910 fires, which is 19.3% of all fires (i.e. 1690546) that form the input of the model. That means that based on the final results it is impossible to say something about the overall ignition type ratio, i.e. what percentage of all fires is lightning caused and what percentage is human caused. But it is possible to estimate this ratio by considering the percentage of lightning that ignites a wildfire. With this percentage and the number of lightning strikes the total amount of lightning fires can be calculated. The amount of lightning strikes in Southern Africa is according to Collier *et al.* (2006) 4.2 strikes per second, which is for a seven fire years period almost one billion lightning strikes. The calculation, using the percentage of lightning that ignites wildfire and the total amount of lightning, resulted in an occurrence of 535897 lightning fires in seven fire years in the study area. Now we can finally calculate the ratio of the natural and anthropogenic ignition types. From the total number of fires (i.e. 1690546) the number of lightning fires is calculated as 535897, so the percentage of lightning fires in the project area is 31.7% (see Table 4.6). This percentage is not significantly different from the estimation of Van Wilgen *et al.* (1990), who calculated that 30 percent of all wildfires in Southern Africa are caused by lightning.

Table 4.6: The ratio of lightning ignited and human ignited wildfire in the study area based on the Ignition Type Model and based on literature review is given in this table.

	According to Ignition Type Model	**According to Van Wilgen (1990)**
Lightning ignited fire	31.7%	30%
Anthropogenic ignited fire	68.3%	70%

4.4.3 Assessment of the model results

The very last part of Chapter 4 aims on describing the usefulness of the final results of the Ignition Type Model. While doing this we will focus on the intentions of the third research question that is treated in Chapter 5. That research question has the goal to investigate spatiotemporal patterns of the two ignition types. Therefore the usefulness of the model results for this approach is described hereinafter.

Concerning the lightning ignited fires that are the result of the Ignition Type Model it is possible to find patterns in space and time. However, the model resulted in only 6.4% of the total number of lightning fires. That means that for the pattern analyses roughly 5000 lightning fires are available for an area of 10510527 sq km and a time span of one year. On the one hand it is therefore not possible to describe patterns in great detail and on the other hand the researcher should consider if the available lightning fires are a reliable sample from the actual number of lightning fires. Based on the characteristics of the LIS dataset and the verification of the Ignition Type Model we can ascertain that the result of the model is a reliable random sampling of all lightning fires. However one should be conscious of the fact that the sampling is relatively small. To conclude about the lightning fires we can say that it is possible to perform spatiotemporal analyses, while keeping in mind that describing patterns in great detail is not possible and that the sample is relatively small.

The second dataset for which the usefulness needs to be defined is the anthropogenic part of the results of the Ignition Type Model. In contrast to the lightning part, the anthropogenic part has a large sampling of more than one quarter of all human fires. Therefore it is reasonable to make conclusions out of the anthropogenic fires dataset. The dataset consists of roughly 42000 human fires in the study area for each year, and therefore the detail in which the analyses can be performed is larger. However there is one important difficulty in performing spatial pattern finding that is caused by the methods that are used in the model. This implies the proximity analysis with human objects and therefore all anthropogenic fires in the dataset are located near these objects. Fires located near roads for example are assigned as anthropogenic caused and when doing spatial analysis the analyst will find that most human fires are located near roads. It is therefore not reasonable to make conclusions about the spatial patterns about anthropogenic fires based on the results of the Ignition Type Model. On the other hand there are no hindrances in finding temporal patterns in this dataset.

Chapter 5

Patterns of natural-/anthropogenic fires

This chapter performs analyses in order to find patterns about natural caused fires and anthropogenic caused fires. Therefore it uses the results of the Ignition Type Model that are explained in Chapter 4. The characteristics of the two result datasets are explained and assessed extensively in Chapter 4. Concerning the analyses that are executed in this chapter, the same theories and choices as given in Chapter 3 are applicable. That means, among others, that again a task is defined, techniques for map visualisation are chosen, spatiotemporal patterns are visualized and spatiotemporal clusters are described. In Section 5.1 there is dealt with two tasks that together give the patterns that are possible to derive from the two ignition type datasets. The second section of this chapter contains the discussion of the results of the analyses.

5.1 Performing analyses

5.1.1 Task 1: Monthly fire frequency of all fires and per ignition type

For the first analysis concerning the wildfires per ignition type the following task is defined: *Visualize and compare the average monthly fire frequency of all wildfires, the natural caused wildfires, and the anthropogenic caused wildfires.* The target of this task aims at three variables. These are: (1) all wildfires, i.e. the assembled burned areas from Step 1 of the Ignition Type Model, (2) the lightning caused fires, and (3) the human caused fires, which both are the result of the Ignition Type Model. The constraint is that the targets are analysed by means of the average monthly fire frequency that is calculated by considering seven fire years.

At first it is necessary to perform an aggregation tool on all three variables in order to calculate the average monthly fire frequency for each segment cell. As segmentation we use a grid overlay with cell sizes of 200 kilometres. Secondly we choose a tool to visualize the variables. The most suitable tool for this task would be a *bar charts map*, so that each variable is portrayed in one map. The result maps of this tool are shown in Figure 5.1 (all fires), Figure 5.2 (lightning fires) and Figure 5.3 (anthropogenic fires). In these maps a two-dimensional space is used to portray the geographical referrer, i.e. the 260 segment cells. Each segment cell has a chart, whose space is outlined by a frame. The values of the twelve attributes are represented by bar-shaped marks within these frames. The marks of one cell are arranged horizontally in an appropriate way, which in this case means that the marks of the twelve months are arranged in chronological order, starting with April. The vertical spatial dimension of the bar chart is used to represent the attribute values. The height of the bars is dependent on the maximum value of all twelve attributes, which fills the total vertical

space of the frame. Therefore it is not possible to compare the values of the three variables by means of the maps since the sizes of the bars are not proportional to the variable values.

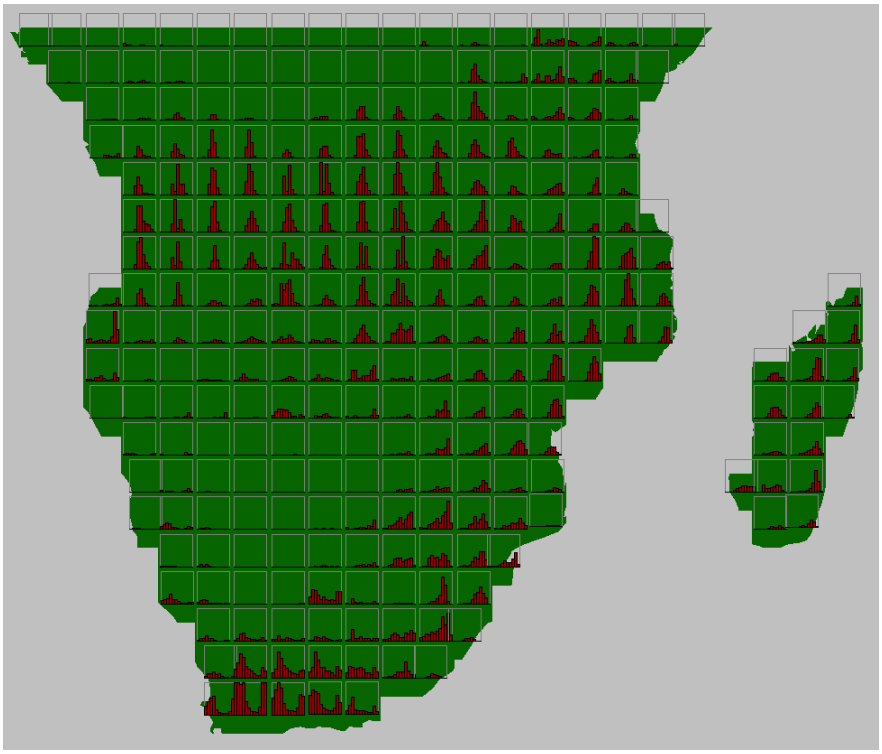


Figure 5.1: Bar chart map of the monthly fire frequency of *all fires* portrayed per segment with 200 km cell sizes. The charts for each segment are portrayed in frames with as horizontal dimension the twelve months and as vertical dimension the attribute values with as maximum value 9504 fires.

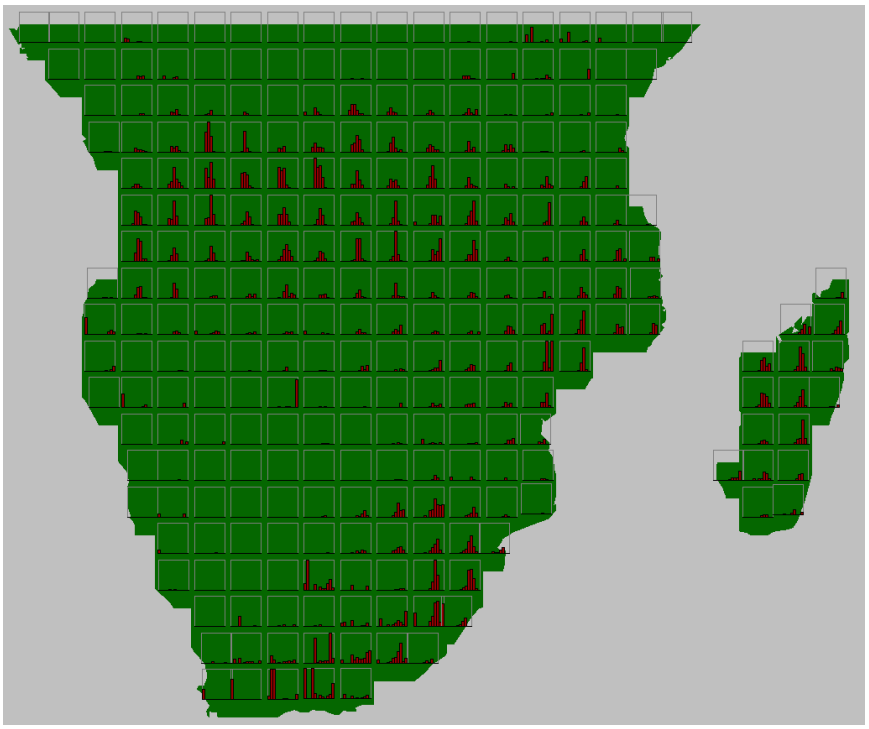


Figure 5.2: Bar chart map of the monthly fire frequency of *lightning fires* portrayed per segment with 200 km cell sizes. The charts for each segment are portrayed in frames with as horizontal dimension the twelve months and as vertical dimension the attribute values with as maximum value 527 fires.

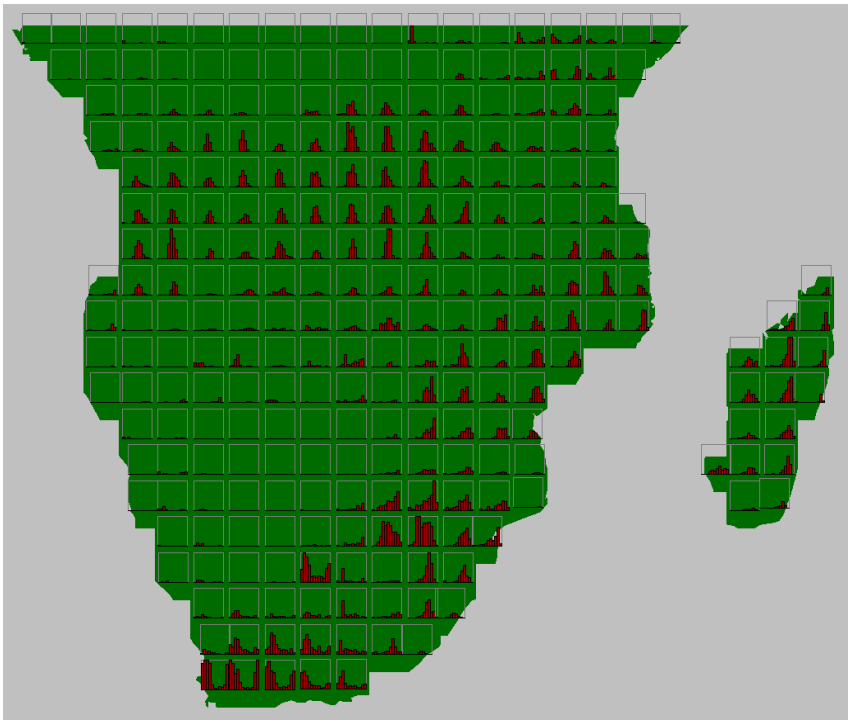


Figure 5.3: Bar chart map of the monthly fire frequency of *anthropogenic fires* portrayed per segment with 200 km cell sizes. The charts for each segment are portrayed in frames with as horizontal dimension the twelve months and as vertical dimension the attribute values with as maximum value 2855 fires.

The three result maps portray a large amount of information, which makes it hard to interpret the behaviour. Therefore, before we describe the main patterns of the three variables, we perform a focussing tool, i.e. we zoom into the data by selecting only that information that is necessary to visualize the main patterns. In this case we remove all frames that are not relevant. As threshold for the focussing tool we take 10 percent of the total frame area that is filled with bar area. That means for the bar chart map of Figure 5.1 that only the frames that contain in total of more then 11405 fires (10% of 12 times the frame height value, which is then for this variable 9501 fires) are selected by the focussing tool. For Figure 5.2 this is 632 fires and for Figure 5.3 this is 3426 fires. The results of the three focussing processes are respectively shown in Figure 5.4, Figure 5.5 and Figure 5.6.

Another advantage of the focussing tool, besides eliminating the information overkill, is that visual comparison is based on a certain amount of yearly fires. Only the segments that meet this amount of yearly fires are provided with a chart. By considering which segments have a chart one can immediately know where the majority of the fires are located.

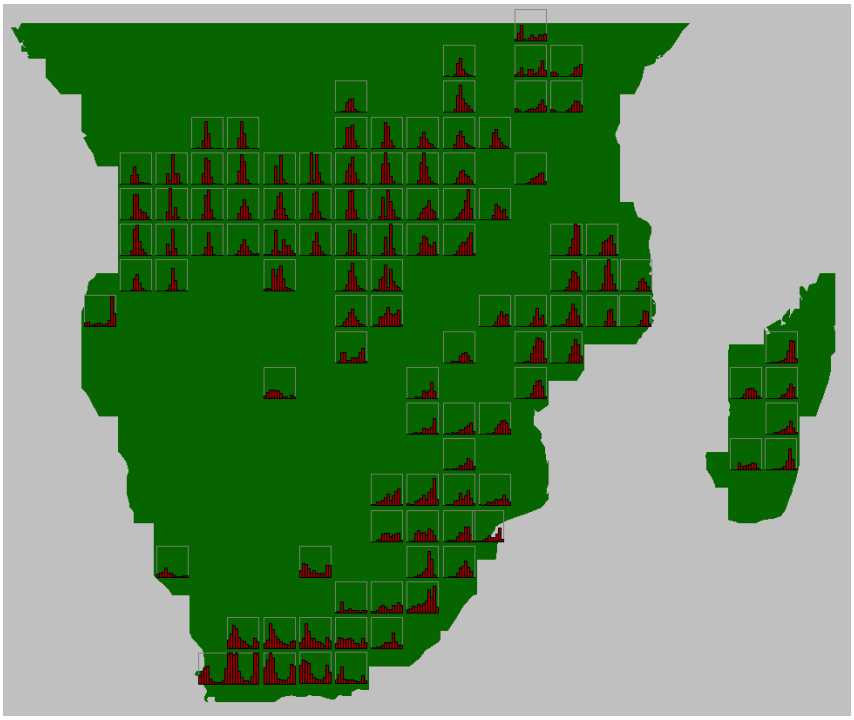


Figure 5.4: Bar chart map of the monthly fire frequency of *all fires*, with only the frames of Figure 5.1 that have more then 10%, which is more then 11405 fires per year, of the total frame area filled with bars.

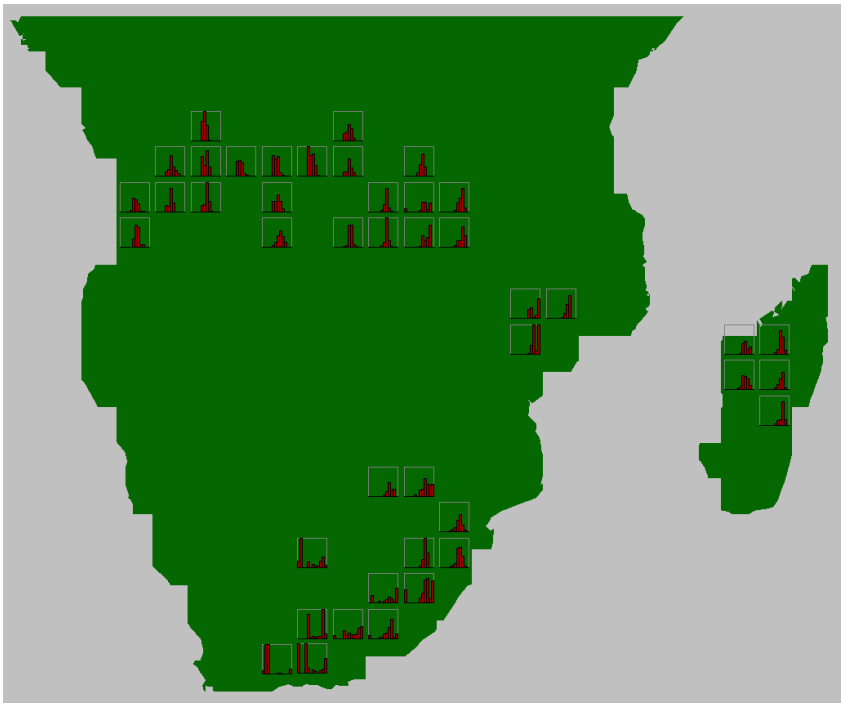


Figure 5.5: Bar chart map of the monthly fire frequency of *lightning fires*, with only the frames of Figure 5.2 that have more then 10%, which is more then 632 fires per year, of the total frame area filled with bars.

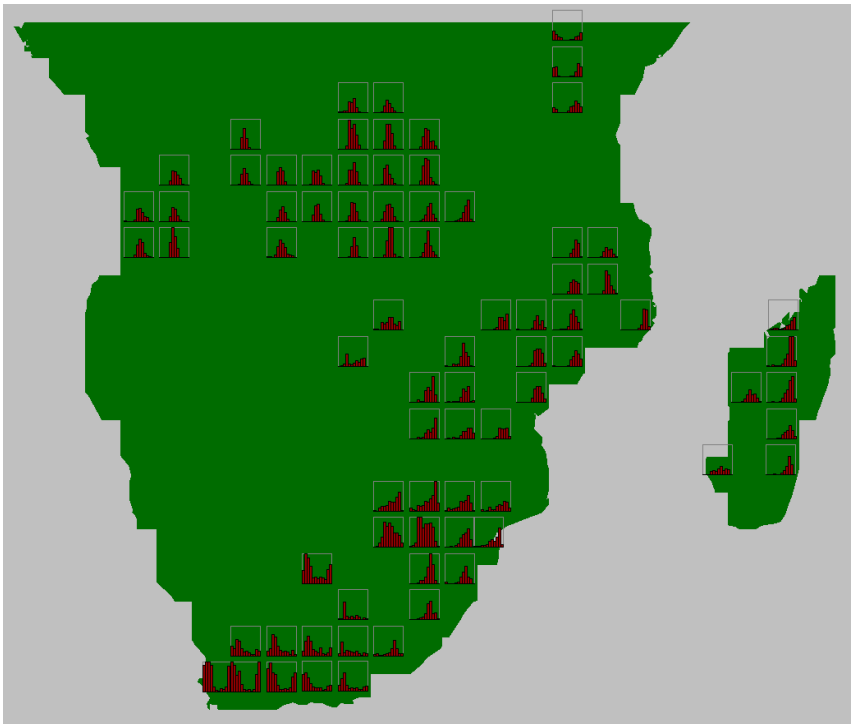


Figure 5.6: Bar chart map of the monthly fire frequency of *anthropogenic fires*, with only the frames of Figure 5.3 that have more then 10%, which is more then 3426 fires per year, of the total frame area filled with bars.

At first we investigate the result maps of Figure 5.1 and Figure 5.4, i.e. the bar chart maps of all wildfires that are produced by assembling adjoining burned areas by means of Step 1 of the Ignition Type Model. So the dataset used for this analysis is another dataset then the one used for the analysis in Chapter 3. Due to Step 1 there are less fires used in this analysis, and especially during periods in which lots of burned area was detected the number of assembled burned area per wildfires increased. This can be seen for example in some segment cells that have a dip during the fire season. This is probably caused because the numerous burned areas are assembled to one wildfire. This is obviously a defect in this analysis, but, as can be seen in the result maps, the general distribution during a fire year is kept intact. Moreover, the distribution within a fire year is much clearer in comparison with the mosaic-analysis, because of the averaging of fire frequency that is performed.

When we investigate only Figure 5.4 we see that a little less then half of all segments comply with the 10% norm. That means that the total amounts of wildfire within all segments are relatively close to each other. Only the segments that have no or almost no fires and segments that have an unclear distribution of fire frequency are removed from this map by means of the 10% norm. This resulted in a map with two main clusters on the African mainland. The first cluster is located between 5° and 15° latitude south, in Congo and Angola. The second cluster is located vertically parallel to the east coast ranging from South Africa to Tanzania.

Next, we consider the results of the lightning bar chart maps, which are shown in Figure 5.2 and Figure 5.5. For this analysis a very small amount of fires is used in comparison with the other two analyses. Therefore the distribution of fire frequency might not be as clear as for the other two. This might be the reason that there are several lonely peaks in the charts. On the other hand, we used the average of lightning fire frequency of seven years, which would normally lead to a smoothed and solid distribution. Therefore the high isolated peaks might indicate that the lightning caused wildfires only occur during a small period in a year.

These periods do in most cases correspond to the centre of the fire season as been recognized in the analysis with all wildfires.

After the segment selection based on the 10% norm we see that only a few, in total 43, segments remain (see Figure 5.5). This indicates that the total number of lightning fires per segment is disproportional, i.e. most segments have no or few lightning fires and some segments have lots of lightning fires. So the lightning fires, as been assigned by the Ignition Type Model, are located on only a few locations but within two main clusters. These clusters do have the same location as the two main clusters that have been described for all wildfires. Only the lightning clusters are much smaller. The last remark is about Madagascar, which seems to have relatively to surface area much more lightning fires then the mainland.

When investigating the anthropogenic fires we should cautious conclude about the spatial patterns of this dataset. Because of the way of assigning human caused fires by the Ignition Type Model (see Paragraph 4.4.4). However we see in Figure 5.6 roughly the same cluster shapes and cluster locations as for analysis with all wildfires. This may indicate that the proportion of fire amount between segments is correct. However, we do not further conclude about spatial patterns of anthropogenic fires. Although we assume that the spatial patterns of anthropogenic wildfire is about the same as the spatial patterns of all fires. The two reasons for this are: (1) the results of the analysis (Figure 5.1 and 5.3) show roughly the same patterns, and (2) anthropogenic fires contribute a large portion (i.e. 68.3%) of all wildfires.

The second aspect that we investigate concerning the anthropogenic fires is the monthly distribution of the average fire frequency. It seems that these do highly correspond to the distributions of all wildfires. One remarkable difference is that the monthly distributions of anthropogenic fire do not have the dips that we observed in the distribution of all wildfires. The reason for this is not clear. In contrast to the lightning fires the distribution of anthropogenic fires has no peaks. Its distribution is smoother, just like all wildfire result. Through that the season of anthropogenic fires is longer (in some places much longer) then the season of lightning fires. However, both seasons of lightning and human caused fires have their centre in the same month.

5.1.2 Task 2: Lightning fires per land cover type

The second task that is defined in order to investigate the spatiotemporal patterns of wildfire ignition types is aiming exclusively on the lightning caused fires. The task is formulated as: Visualize in which land cover types lightning caused wildfire occurs the most. This task has as target only the lightning ignited fires that were the result of the Ignition Type Model. In addition to this an independent dataset will be used, i.e. is the GLC2000 that delivers us of a map with the land covers that are present in the study area. The goal of this task is to investigate if there are certain characteristics in the relation between the lightning fires and the land cover type on which the lightning fire is burning. By that, we try to gain knowledge about notable behaviour lightning fires in relation with land cover. It is by this probably necessary to compare this finding with the characteristics of all wildfires in relation to land cover, in order to indicate whether the finding is indeed a characteristic of lightning fire. Therefore also the wildfire dataset that was the result of Step 1 of the Ignition Type Model is used during this research.

The GLC2000 dataset contains for the study area 24 different land cover types. In Figure 5.7 a cut out of this dataset concerning the study area is shown. The next step of this analysis is adding to each wildfire on which land cover type it is located. Because of the high resolution of the GLC2000 (i.e. 1 kilometre) the assigning of land cover type is performed very accurately. As a result, for both fire datasets, which are all wildfires and lightning caused wildfires, a 100% stacked histogram is produced (see Figure 5.8). For this histogram only eleven land cover types remain that are for that reason the land cover types on which most fires are occurring (i.e. contributing for more then 1%).

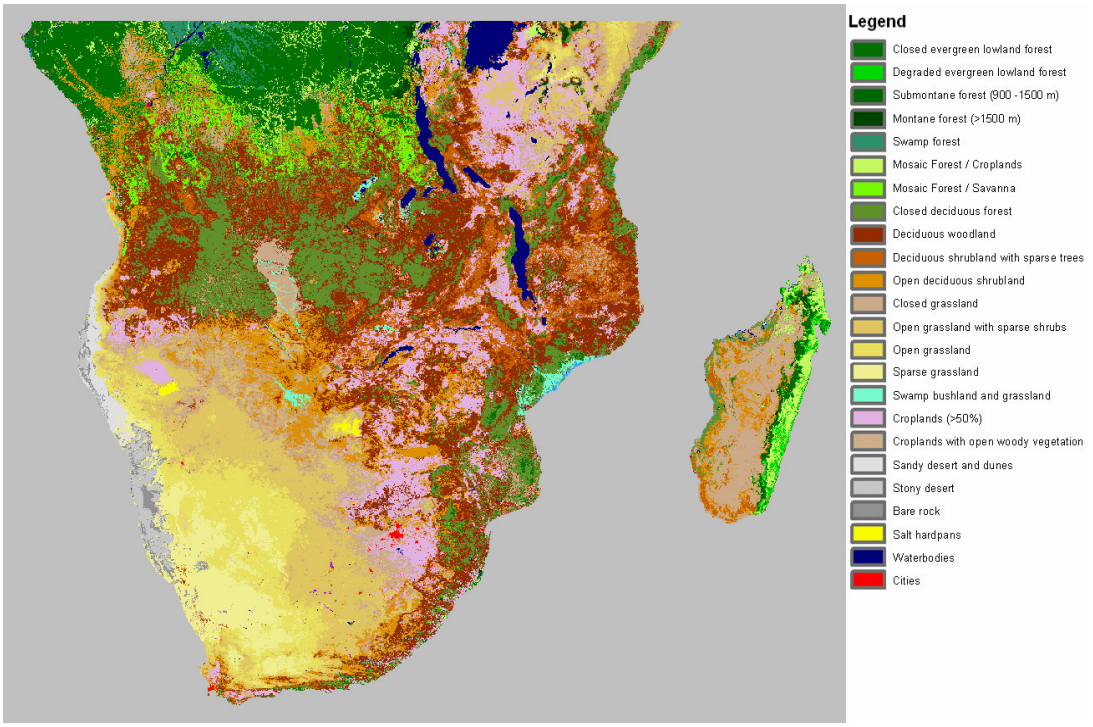


Figure 5.7: Clip of the Global Land Cover 2000 product with its original classes and colouring scheme.

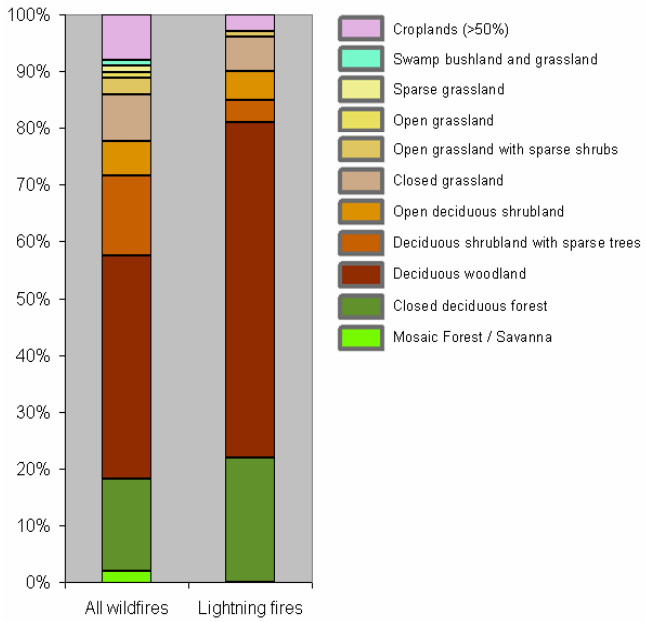


Figure 5.8: 100% stacked histogram of the land cover types on which a wildfire occurs in the study area. The left bar represents all wildfires and the right bar represents lightning caused wildfires that are the result of the Ignition Type Model.

By investigate this histogram we first of all notice a difference in the amount of land cover types on which the fires of the two datasets occur. All wildfires are located mostly on eleven land cover types and lightning fires are located on only seven land cover types. That means that lightning fires occur on only a few specific land cover types in comparison to all wildfires (which also include lightning fires). The second aspect we notice is that lightning fires mainly occur on only two specific land cover types. These are closed deciduous forest (22%) and deciduous woodland (59). That means that a total of 82% of all lightning fires occur in woodland areas, in contrast to a total of 55% of all wildfires that occur in woodland areas. This difference would even be larger if instead of all wildfires we would consider only anthropogenic fires (the anthropogenic results of the Ignition Type Modal are not sufficient for this analysis).

In the next step we try to perform the same analysis but this time by visualizing the behaviour of lightning wildfire with respect to land cover types. Let us consider the classified choropleth map in Figure 5.9, which gives the average yearly fire frequency of all lightning fires over seven fire years. What we could do is visually comparing this map with the land cover map of Figure 5.7 in order to find out on which land cover types these fires do and do not occur. Although this is quite difficult to do, we can however derive a pattern from this relation. There are several land cover types that have no or almost no lightning fire, e.g. the desert land covers and certain forest types. However, most lightning fires seem to occur in deciduous forest types.

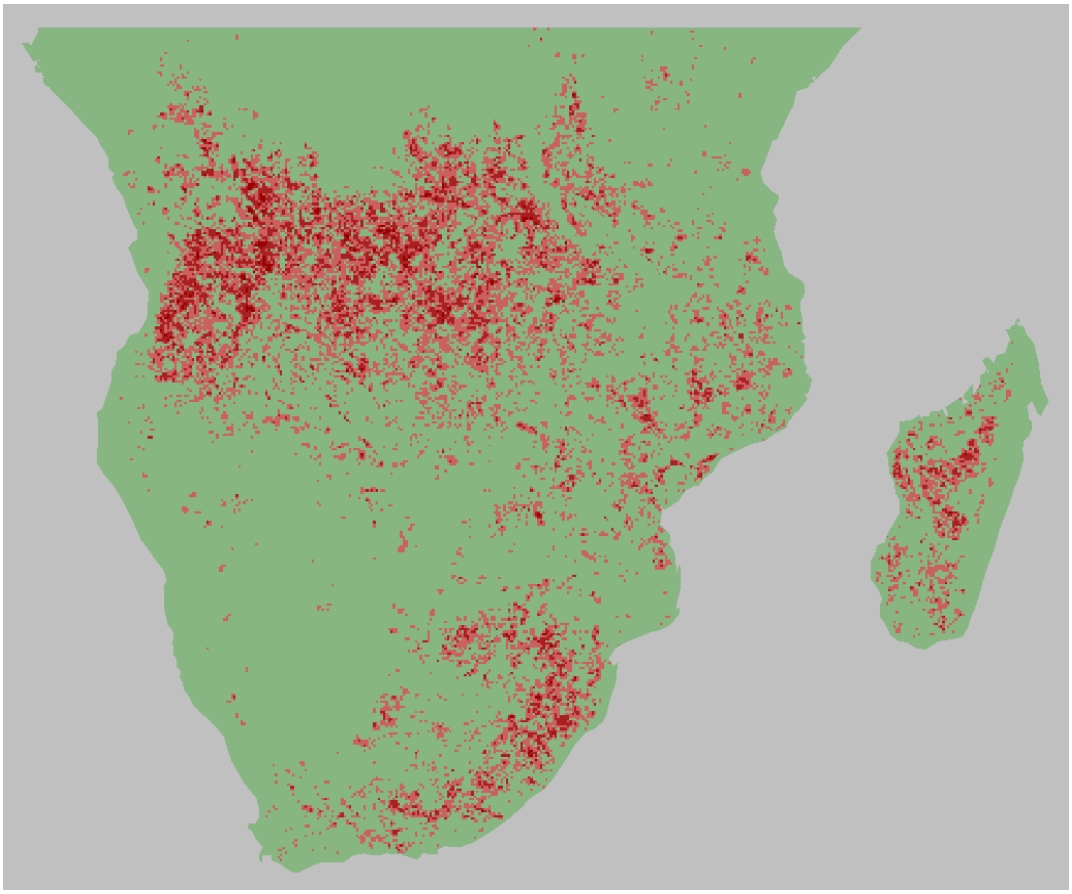


Figure 5.9: Classified choropleth map of the average yearly frequency of lightning caused wildfires that are the result of the Ignition Type Model (in total 34215 fires). Classification is based on three equal sized classes.

For better visual presentation of the contribution of lightning fires on the deciduous forest types we perform an *attribute selection tool*. In this case we select by means of the land cover type attribute, in order to represent in a map illustration only the elements of a certain attribute value. To perform this tool we consider the visualisation of Figure 5.9. To that we perform the attribute selection with the purpose to show merely the lightning fires that are located in deciduous forests (they are called deciduous woodland and closed deciduous forest). Also we perform this tool with the purpose to show the lightning fires that are not located in deciduous forests. The two result maps are shown in Figure 5.10.

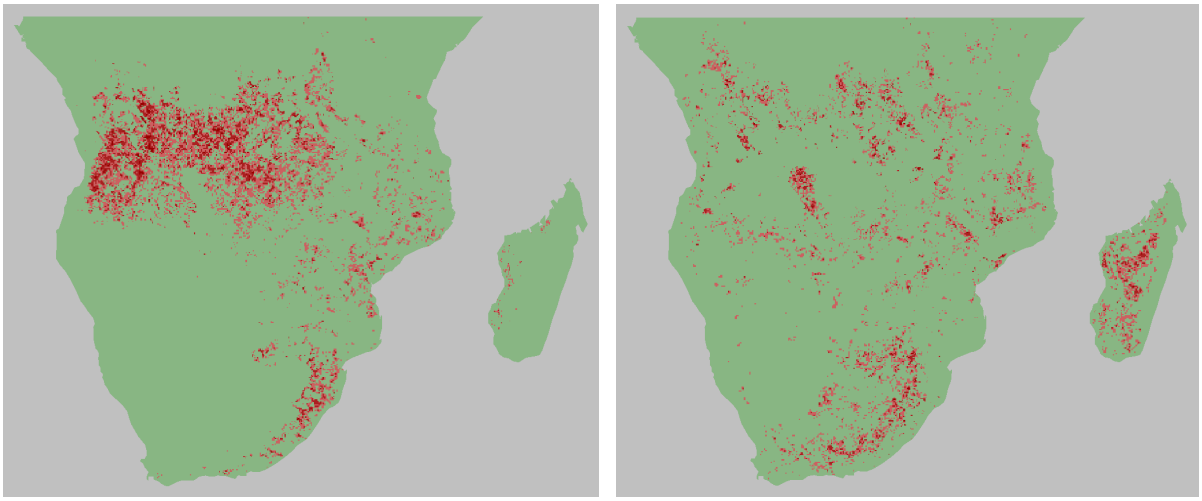


Figure 5.10: Classified choropleth maps of the average yearly frequency of lightning caused wildfires. The left map shows only the fires that occur in deciduous woodland and closed deciduous forest. The right map shows all other fires. Classification is based on same selection as the map of Figure 5.9.

With the two illustrations of Figure 5.10 we can easily recognize that most lightning caused wildfires do occur in woodland areas. The large spatial cluster of lightning fires between 5° and 15° latitude south that we already described in the previous task is again visible. Now we know that this cluster is almost totally build up by lightning fires occurring in woodland areas. Also within this cluster an area with closed grassland that has a high fire frequency is clearly visible. Also on Madagascar it is this single land cover type that is burning the most. Further we see that near the south of the study area there are some fire concentrations that are not located in woodland areas but on grasslands.

5.2 Discussion

In the last section of this chapter we discus the findings with results of other researches. The first chapter elaborates on the characteristics of lightning caused fires as found in the Section 5.1 in relation with patterns of lightning frequency that is found by Collier *et al.* (2006). The second paragraph shortly discusses the finding of lightning caused wildfires occurring in wooded areas.

5.2.1 Comparison of lightning fire patterns and lighting patterns

Collier *et al.* (2006) performed a research on the spatial and seasonal variation of lightning activity over Southern Africa. There are some strong similarities between his findings and the patterns of lightning caused wildfires found in this chapter. These similarities indicate that the flash rate is positively related to frequency of lightning ignited fire. Collier *et al.*

(2006) found that lightning activity is concentrated predominantly in the northern regions of the study area. We found large amounts of lightning caused fires in the northern regions, except from the tropical areas near the equator, which are obviously too wet to burn. Collier *et al.* (2006) also found large contributions of lightning frequency in the western part of South Africa and in Madagascar. Precisely in these areas we found more lightning caused fires then in other regions.

Concerning the seasonal variation of lightning Collier *et al.* (2006) concluded that the maximum activity occurs in South Africa on the east and west in different seasons. The central and south part of South Africa has its maximum lightning activity in the autumn and the east part of South Africa has its maximum lightning activity in the spring and summer. Both conclusions do correspond with the frequency of lightning caused fires that we found. Especially in the south part of South Africa we saw peaks, which are occurring in the autumn. We also saw that on the east coast of South Africa the lightning caused fires are occurring in spring and summer.

5.2.2 Lightning caused fires in wooded areas

In this chapter we found that lightning caused fires in Southern Africa are mainly occurring on deciduous woodland and closed deciduous forest. It seems that there is little known in literature about our finding. So our finding can be seen as new scientific knowledge. Lightning caused fires have been investigating extensively for boreal forest areas, e.g. in Canada, Russia and Scandinavia (Larjavaara *et al.*, 2005). Because of the differences in forest type, climate, etcetera, it is impossible to compare their findings with our results lightning ignited wildfire.

Chapter 6

Conclusions and recommendations

This chapter completes the thesis by giving the final conclusions concerning the presented work. In the first paragraph of the first section we return to the starting points of this thesis, which are the three hypotheses. Further in Section 6.1 the answers on the research questions and all conclusions are given. In the second section there are stated two recommendations for further research on the wildfire topic.

6.1 Conclusions

6.1.1 Hypotheses

Hereafter the three hypotheses, formulated at the beginning of this thesis, are finalized.

The *first hypothesis* presupposed that wildfires in the study area are completely spatiotemporal random. However, in all visualisations of the L3JRC product that are made in the thesis, we clearly see that this statement is false. The variation of fire frequency in the study area is large, with areas with no fires and areas with many fires. Concerning the spatial patterns we clearly observe for instance that there are two large areas with almost no fire and two large areas with lots of fire in every visualised map. Within these clusters there is a difference in total amount of fire but a similarity in temporal patterns. Also regarding the temporal patterns the seasonality of fire regimes are observed. Most areas have a clear seasonality but some small areas have a constant fire frequency throughout the year. Based on these findings the first hypothesis is declared as false.

The *second hypothesis* stated that the influence of humans on the total fire regime is limited. Based on the findings of this thesis we can adapt this proposition because we found that the influence of humans is large. Buitrago (2008) concluded that on places where the fire frequency is high the population density is very low. In addition, this research found that the correlation between wildfire and human influence is very low. But that does not mean that all wildfires, that are located on places where almost no human lives are, not caused by humans. It only needs one human to set a wildfire in an remote area. The results of the Ignition Type Model showed that of all wildfires in the study area almost two third is human caused. Furthermore the model found that three quarters of all anthropogenic fires do not start near roads or urban areas. But most important of all, we found that roads, urban areas, and human settlements do attract wildfire. All these findings indicate that the hypothesis is false, i.e. the influence of humans on fire regime is large, and in fact is the most dominant ignition type for Southern Africa.

The *third hypothesis* described that natural and anthropogenic fires have the same spatiotemporal patterns. Because of the methods used for the Ignition Type Model there were some difficulties in deriving spatiotemporal patterns from both ignition type datasets. Nevertheless, we did found some differences in their behaviour. One of the analysis found that anthropogenic fires occur in larger spatial clusters in comparison with lightning fires. Lightning fires are mainly located in woodlands and forests. Concerning the temporal aspect we found that the fire season of lightning ignited fires are short in comparison with the fire season of anthropogenic fires that at some locations even occur during the whole year.

6.1.2 Research questions

In the introduction three research questions have been defined:

- What spatiotemporal patterns of wildfires can be derived from the L3JRC burned area product?
- Can the distinction between natural and anthropogenic caused fires be made by independent variables?
- What are the spatiotemporal patterns of natural and anthropogenic caused fires?

In the following the conclusions and answers on these research questions are presented.

Concerning the first research question the L3JRC burned area product is investigated in order to describe spatiotemporal patterns of wildfire in Southern Africa. The patterns that are described are derived by means of analysing the behaviour in map visualisations. Concerning the spatial and temporal patterns these included fire clustering, fire absence, and change of fire frequency. New findings that are found by answering this question concerns the comparison of the spatial and temporal patterning of wildfire. Spatially the wildfires occur in clusters. Within these clusters the spatial fire amount differs, but the temporal distribution (seasoning) stays about the same.

By means of the Ignition Type Model we tried in Chapter 4 to distinguish between natural and anthropogenic wildfire by considering the location of wildfire ignition in space and time. Therefore the model used the L3JRC burned area product as its main input and several independent variables to allocate fire to lightning ignited, human ignited or unknown ignition. This thesis showed that it is possible to distinguish in this way wildfires by ignition type to a certain extent. The wildfires that are allocated to thunderstorms (and therefore assigned as lightning ignited) form a reliable model output. The verification of the lightning fires show similar results as results found in literature, like the percentage of lightning that ignites wildfire and the ratio lightning versus human caused fires. On the other hand, concerning the anthropogenic fires, the Ignition Type Model performs mediocre. It seems hard to validate the human caused fires. Although the distribution of Euclidean distances between wildfire and human objects clearly show that these attract each other. Most of the fires that are close to human objects are most likely anthropogenic, but it is suggestive to state that all these fires are anthropogenic. The chosen threshold are still arbitrary and need more verification. Nevertheless, the human caused fires that are the result of the Ignition Type Model are, even as the lightning caused fires, appropriate for deriving spatiotemporal patterns.

In Chapter 5 the spatiotemporal patterns of natural and anthropogenic fires are investigated. As a result of the analyses we can state that lightning fires in comparison with anthropogenic fires hold smaller spatial clusters and have shorter duration of fire seasoning.

However the location of the cluster and season is the same for both ignition types. At last we found that lightning fires in contrast to human caused fires occur in a great extent in woody areas.

6.1.3 Main conclusions

Based on the answers on the research questions three main conclusions of the thesis are defined. These conclusions are new findings that might be useful for further scientific research on wildfires. These three main conclusions are:

- Although the spatial variance of fire frequency in Southern Africa is large, they mainly occur in a few large clusters. Within these clusters the total fire amount differs strongly per area. But the temporal distribution within the clusters is about the same, i.e. they have identical seasonal patterns. For further research on spatiotemporal fire patterning this conclusion should be the new hypothesis.
- As a result of the thesis we can announce that it is possible to use independent variables to assign ignition types to some wildfires. With that we can supply the demand of the scientific community with a solution to obtain one of the wildfire descriptors.
- For further research on lightning caused wildfire some findings that are found in this thesis are useful as new hypothesis. The most important finding that was not known yet is that lightning caused fires have a very short but fierce fire seasoning in comparison with anthropogenic wildfires.

6.1.4 Other results

Other results that can be drawn from the thesis are:

- We found in Southern Africa one main spatial cluster in which most wildfires occur every fire year again. This is a large inland area that is roughly located between 5° and 15° latitude south, i.e. extended out over the countries of Angola, Congo and Zambia. There are also two large areas found in which no or almost no fires occur. The first one is located near the equator in Congo; the second one is located in the Kalahari Desert in Namibia and Botswana.
- The correlation between fire frequency and herbaceous land cover is moderate positive. There are large areas that have herbaceous vegetation and a high fire frequency, but there are also large areas that have something else burning then herbaceous vegetation. Further, there are very few spots that have both a high human influence and a high fire frequency. The correlation between these two is negative and very small. Most wildfires occur in areas with a small human footprint.
- Concerning the fire frequency within the years we can distinguish three categories. The first one has almost no fires throughout the year. This category can mainly be found near the equator and in the desert. Secondly there are areas that have all their wildfires only in a small period in the year, which is the fire season. These areas are located between 5° and 15° latitude south. The third and last category consists of areas with wildfires throughout the whole year, without having a clear fire season. This category can be found in the southern part of the study area and near the equator in Kenya and a part of Tanzania.

- Human objects like roads, urban areas and settlements attract wildfire. A total of 15.6% of all wildfires are located within 1.2 km from a road (that implies a surface area of 9.6% of the study area). Also 2.6% of the wildfires are located in or near the few urban areas of Southern Africa. Concerning the human settlements we found that they also attract fire and that on average a wildfire is located 7 km away from a settlement.
- There is a positive relation between lightning fire frequency and sizes of thunderstorms. It seems that thunderstorms that produce positive lightning strikes are on average larger at areas that have a high lightning fire frequency. That means that large thunderstorms that produce positive lightning strikes are most likely to ignite wildfire.
- The percentage of all lightning strikes (including cloud-to-cloud and negative strikes) that ignites a wildfire (that is large enough to be detected) is calculated as 0.0578%.
- The percentage of lightning caused and human caused wildfire in the study area is calculated as respectively 31.7% and 68.3%.
- About 82% of all lightning caused fires are located on woodland areas (that are to be precise the deciduous woodland and closed deciduous forest classes of the GLC2000 product).

6.2 Recommendations

In particular there are two issues that have been encountered that need attention in future research on this topic. These two recommendations are discussed hereinafter.

The performed analyses on fire data in this thesis are focussed on using EDA techniques for visual analytics by means of two-dimensional maps. For every analysis there is one or more map visualisation produced and with that the patterns of fire behaviour are visualized and subsequently described. However, there are more EDA techniques, apart from map visualisations, that can be used for deriving patterns from fire data. For example the computational EDA techniques, that include both simple basic statistics and more advanced multivariate exploratory techniques. By that the patterns can be quantified. Also several other visualisation techniques are part of EDA, such as all kinds of graphs and plots. These techniques are sparsely used in the thesis because the choice was made to concentrate on visual analytics by means of maps. For further wildfire data exploring it is recommended to use the EDA techniques already used in the thesis in combination with the other abovementioned EDA techniques. By that it should be possible to improve the description of patterns and even find patterns that are not found during the thesis.

The Ignition Type Model of this thesis assigns ignition types to wildfires by considering its location in space and time with respect to independent variables. There are several ways thinkable to extent and improve this model. The most appropriate way of doing this would be to include expert knowledge. This includes knowledge about behaviour of several fire aspects, i.e. behaviour of wildfire (e.g. in relation with precipitation), behaviour of humans (e.g. information about land clearance) in relation to wildfire, and behaviour of (positive cloud-to-ground) lightning strikes. All these knowledge can contribute in extending the model by improving the parameters, adding more criteria's, and implementing better ways of validation. However, their will arise some difficulties while dealing with knowledge about behaviour of these aspects because of the large extent of the area. A good solution for this problem is to first use segmentation to explore the behaviour on regional scale. The

segmentation can be made by land cover and by country, because these have the most influence on the behaviour of the abovementioned aspects. By doing first the exploring on regional scale and then performing the continental processing the Ignition Type Model will improve in reliability and effectiveness.

References

Amatulli, G., F. Perez-Cabello, et al. (2007). "Mapping lightning/human-caused wildfires occurrence under ignition point location uncertainty." Ecological Modelling **200**(3-4): 321-333.

Andreae, M. O. and P. Merlet (2001). "Emission of trace gases and aerosols from biomass burning." Global Biogeochemical Cycles **15**(4): 955-966.

Andrienko, N. and G. Andrienko (2006). Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach, Birkhäuser.

Barbosa, P. M., J. M. Gregoire, et al. (1999). "An algorithm for extracting burned areas from time series of AVHRR GAC data applied at a continental scale." Remote Sensing of Environment **69**(3): 253-263.

Bertin, J. (1984). "Semiology of graphics: diagrams, networks, maps."

Buitrago, M.F. (2008). "Predicting the occurrence of fires in Africa."

Cahoon, D. R., B. J. Stocks, et al. (1992). "Seasonal distribution of African savanna fires." Nature **359**(6398): 812-815.

Carmona-Moreno, C., A. Belward, et al. (2005a). "Spatial-temporal Correlation Analyses of Global Burned Surface Time Series from Remote Sensing data (1982-1999)." Global Change Biology.

Carmona-Moreno, C., A. Belward, et al. (2005b). "Global Fire Calendar Probability maps from the Analysis of Global Burned Surfaces Time Series (1982-1999)." Global Change Biology.

Carmona-Moreno, C., A. Belward, et al. (2005c). "Characterizing interannual variations in global fire calendar using data from Earth observing satellites." Global Change Biology **11**(9): 1537-1555.

CIESIN-SEDAC (2008). "Center for International Earth Science Information Network - Socioeconomic Data and Applications Center." **http://sedac.ciesin.columbia.edu/gpw/** (Accessed: January 2009).

CIESIN (2008). "Center for International Earth Science Information Network." **http://microwave.nsstc.nasa.gov:5721/dataset_documents/lis_dataset.html** (Accessed: January 2009).

Collier, A. B., A. R. W. Hughes, et al. (2006). "Seasonal and diurnal variation of lightning activity over southern Africa and correlation with European whistler observations." Annales Geophysicae **24**(2): 529-542.

Duncan, B. W. and P. A. Schmalzer (2004). "Anthropogenic influences on potential fire spread in a pyrogenic ecosystem of Florida, USA." Landscape Ecology **19**(2): 153-165.

Dwyer, E., J. M. Gregoire, et al. (1997). Global distribution and characterisation of vegetation fires using NOAA AVHRR data. Proceedings of SPIE - The International Society for Optical Engineering.

Dwyer, E., J. M. C. Pereira, et al. (2000). "Characterization of the spatio-temporal patterns of global fire activity using satellite imagery for the period April 1992 to March 1993." Journal of Biogeography **27**(1): 57-69.

Forsyth, G. G. and B. W. van Wilgen (2008). "The recent fire history of the table mountain national park and implications for fire management." Koedoe **50**(1): 3-9.

Fuquay, D. M. (1980). "Lightning that ignites forest fires." Proceedings of the Sixth Conference on Fire and Forest Meteorology: 109–112.

Giglio, L., G. R. van der Werf, et al. (2006). "Global estimation of burned area using MODIS active fire observations." Atmospheric Chemistry and Physics **6**(4): 957-974.

Guyette, R. P., R. M. Muzika, et al. (2002). "Dynamics of an anthropogenic fire regime." Ecosystems **5**(5): 472-486.

Hall, B. L. (2007). "Precipitation associated with lightning-ignited wildfires in Arizona and New Mexico." International Journal of Wildland Fire **16**(2): 242-254.

JRC-TEM (2003). "Joint Research Centre - Terrestrial Ecosystem Monitoring - Global Land Cover 2000." **http://www.tem.jrc.it/glc2000** (Accessed: January 2009).

JRC-TEM (2008). "Joint Research Centre - Terrestrial Ecosystem Monitoring - Global Burnt Areas." **http://www-tem.jrc.it/Disturbance_by_fire/products/burnt_areas/GlobalBurntAreas2000-2007.htm** (Accessed: January 2009).

Justice, C. O., L. Giglio, et al. (2002). "The MODIS fire products." Remote Sensing of Environment **83**(1-2): 244-262.

Korontzi, S., C. O. Justice, et al. (2003). "Influence of timing and spatial extent of savanna fires in southern Africa on atmospheric emissions." Journal of Arid Environments **54**(2): 395-404.

Larjavaara, M., T. Kuuluvainen, et al. (2005). "Spatial distribution of lightning-ignited forest fires in Finland." Forest Ecology and Management **208**(1-3): 177-188.

Lehsten, V., K. J. Tansey, et al. (2008). "Estimating carbon emissions from African wildfires." Biogeosciences Discussions **5**(4): 3091-3122.

Low, A. B. and A. G. Rebelo (1996). "Vegetation of South Africa, Lesotho and Swaziland." Pretoria: DEAT **http://www.ngo.grida.no/soesa/nsoer/Data/vegrsa/vegstart.htm** (Accessed: January 2009).

Missoum, S., Z. Gürdal, et al. (2005). "Study of a new local update scheme for cellular automata in structural design." Structural and Multidisciplinary Optimization **29**(2): 103-112.

Moreno-Ruiz, J. A., J. Baron-Martinez, et al. (2006). "Burned data time series for Tropical Africa (1981-2000) with the Daily Tile NOAA/NASA Pathfinder AVHRR 8-km Land dataset." International Journal of Remote Sensing **27**(6): 1119-1134.

Moreno Ruiz, J. A. and M. Canton Garbin (2004). "Estimating burned area for tropical Africa for the year 1990 with the NOAA-NASA pathfinder AVHRR 8 km land dataset." International Journal of Remote Sensing **25**(17): 3389-3410.

NASA-GHCC (2005). "NASA Global Hydrology and Climate Center -Introduction to Lightning Imaging Sensor (LIS)." **http://microwave.nsstc.nasa.gov:5721/dataset_documents/lis_dataset.html** (Accessed: January 2009).

Nelson, A., A. De Sherbinin, et al. (2006). "Towards development of a high quality public domain global roads database." Data Science Journal **5**: 223-265.

NOAA (2006). "National Oceanic and Atmospheric Administration." **http://www.noaa.gov/** (Accessed: January 2009).

Podur, J., D. L. Martell, et al. (2003). "Spatial patterns of lightning-caused forest fires in Ontario, 1976-1998." Ecological Modelling **164**(1): 1-20.

Preston-Whyte, R. A. and P. D. Tyson (1988). Atmosphere And Weather Of Southern Africa, Oxford University Press Southern Africa.

Pyne, S. J. (1994). "An introduction to Antrhopogenic Fire." Elsevier Science Ltd: Chemosphere **29**(5): 889-911.

Riano, D., J. A. Moreno Ruiz, et al. (2007). "Global spatial patterns and temporal trends of burned area between 1981 and 2000 using NOAA-NASA Pathfinder." Global Change Biology **13**(1): 40-50.

Rorig, M. L. and S. A. Ferguson (1999). "Characteristics of lightning and wildland fire ignition in the Pacific Northwest." Journal of Applied Meteorology **38**(11): 1565-1575.

Scott, D. F. (1993). "The hydrological effects of fire in South African mountain catchments." Journal of Hydrology **150**(2-4): 409-432.

Sheuyange, A., G. Oba, et al. (2005). "Effects of anthropogenic fire history on savanna vegetation in northeastern Namibia." Journal of Environmental Management **75**(3): 189-198.

Silva, J. M. N., A. C. L. Sa, et al. (2005). "Comparison of burned area estimates derived from SPOT-VEGETATION and Landsat ETM+ data in Africa: Influence of spatial pattern and vegetation type." Remote Sensing of Environment **96**(2): 188-201.

Tansey, K., J. M. Gregoire, et al. (2008). "L3JRC - A global, multi-year (2000-2007) burnt area product (1 km resolution and daily time steps)."

Tansey, K., J. M. Grégoire, et al. (2004). "Vegetation burning in the year 2000: Global burned area estimates from SPOT VEGETATION data." Journal of Geophysical Research D: Atmospheres **109**(14).

UCAR (2007). "University Corporation for Atmospheric Research." **http://www.ucar.edu/** (Accessed: January 2009).

UMD-GLCF (2007). "University of Maryland - Global Land Cover Facility." **http://glcf.umiacs.umd.edu/** (Accessed: January 2009).

Van Wilgen, B. W., W. J. Bond, et al. (1992). "Ecosystem management." The ecology of fynbos: nutrients, fire and diversity: 345–371.

Van Wilgen, B. W., W. S. W. Trollope, et al. (1990). "Fire management in southern Africa: some examples of current objectives, practices and problems." Goldammer JG: Fire in the tropical biota.

Appendix A

L3JRC and other global fire products

In several research fields there is a high demand on global fire products (Tansey *et al.*, 2008). Like for example for researches in modelling the carbon cycle, in understanding the relationship between fire regime and climate, or in investigating the impact of vegetation burning on land cover change. For those researches, the L3JRC is currently the most recent global fire dataset available. With the release of this dataset at the end of 2007 a data gap was filled. Before L3JRC became available there was a lack of accurate, consistent long-term information on global burnt area (Carmano-Moreno *et al.*, 2005a). In my opinion, the new L3JRC global fire product provides the scientific community with a good useable, reliable, multi-annual product to strengthen the arguments of relationships between fire, vegetation and climate.

Global fire products occur in two types: datasets with detected burn scars and datasets with detected active fires. The big advantage of the burn scars products are that a fire does not need to be detected at the time of burning but can be detected at the next satellite overpass with suitable viewing conditions. The drawback is that this can add a delay to the time of detecting. Examples of burn scars products are, besides the L3JRC, the ESA GlobScar and the GBA2000 (Korontzi *et al.*, 2003). The latter product is developed by the Joint Research Centre and is the predecessors of the L3JRC. Moreover, the Global Burned Areas 2000 has been produced by the same algorithm as the L3JRC. The disadvantage of the GBA2000, as the name already suggests, is that it is limited to the year 2000 and therefore did not fulfil the need of scientist for a multi-year product (Moreno-Ruiz *et al.*, 2004).

The active fire datasets only contains information of fires that are burning at the time of the satellite overpass. The fires are usually presented in the form of bounded fire masks at a coarse spatial resolution. When the detected fires are counted for a time period of one month or the like, the 'fire count product' can familiarize with aspects like the spatial distribution and seasonality of burning. But it is difficult to use them for calculating actual area burned, because of the inadequate temporal sampling. Nevertheless, active fire datasets were often used as a proxy for burned area (Giglio *et al.*, 2006), because of the lack of long-term and accurate fire products, like L3JRC.

An example of an active fire product is the GBS1982-1999, which is produced by the Joint Research Centre, and is the predecessor of the GBA2000 and L3JRC. The Global Burned Surface, generated out of AVHRR (Advanced Very High Resolution Radiometer) daily global observations between 1982 and 1999, is a multi-year, weekly fire product at a resolution of 8 sq km (Moreno-Ruiz *et al.*, 2006). Due to this product, which has also been used by Buitrago (2008), scientists obtained a better understanding of the characteristics of fire activity in both northern an southern hemispheres on the basis of average seasonal cycle and inter-annual variability (Riano *et al.*, 2007). Other examples of active fire products are the Terra MODIS active fires (Giglio *et al.*, 2006), the ATSR andATSR-2 active fire count (Lehsten *et al.*, 2008), and further several other products received from AVHRR active fire data (Barbosa *et al.*, 1999).