The Case for Effective Fire Suppression in Onatrio

Dissertation submitted in partial fulfilment of the requirements for the Master of Science Degree in Evironmental Sustainability & Green Technology, Keele University.

Student Number:08010027
(**Supervisor: Dr Peter Thomas**)

31st August 2012

Abstract

A great deal of effort and resources are expended in suppressing wildfires in the boreal forests of Ontario, Canada. Fire managers at Ontario's Ministry of Natural Resources had assumed these efforts to be effective, given that data from Ontario showed fire suppression had significantly reduced the annual area burned by wildfires over recent decades. However, fire–ecologists have brought to light flaws in the fire manager's analysis, making their assumption of effectiveness unsound.

The challenge this research sought to address therefore, was to improve upon these past efforts at measuring the effectiveness of forest fire suppression in Ontario, whilst maintaining a scientific integrity in both the research design and analysis. The study involves comparing the fire size distributions between several areas with contrasting forest fire management strategies, selected to control for other significant causal factors. With this method, it is possible to calculate precisely the extent to which a more aggressive fire suppression strategy has effectively reduced the damage caused by wildfires over recent decades.

Acknowledgments

The work presented here would not have been possible without the invaluable guidance of my dissertation supervisor Dr Peter Thomas and the help of Rob Luik at Ontario's Ministry of Natural Resources for for providing the data used in this research.

Contents

List of Figures

List of Tables

1 Introduction

A great deal of effort and resources are expended in suppressing wildfires in the boreal forests of Ontario, Canada. Fire managers at Ontario's Ministry of Natural Resources had assumed these efforts to be effective, given that data from Ontario showed fire suppression had significantly reduced the annual area burned by wildfires over recent decades. However, fire–ecologists have brought to light flaws in the fire manager's analysis, making their assumption of effectiveness unsound (see *Chapter 2:Literature Review*).

The challenge this research sought to address therefore, was to improve upon these past efforts at measuring the effectiveness of forest fire suppression in Ontario, whilst maintaining a scientific integrity in both the research design and analysis (see *Chapter 3:Research Design*). The study involves comparing the fire size distributions between several areas with contrasting forest fire management strategies, selected to control for other significant causal factors (see *Chapter 4:Methodology*). With this method, it is possible to calculate precisely the extent to which a more aggressive fire suppression strategy has effectively reduced the damage caused by wildfires over recent decades (see *Chapter 5:Results*).

It is hoped this research provides a definitive solution to the Ontario debate. However, the interpretation of these results may also be of interest to all users of quantitative methods, as special consideration was given to the limitations of this methodology (see *Chapter 6:Discussion*). The analysis was developed using the open-source R system for statistical computation, which allows other researchers to replicate this study and refine the methodology in future work. To aid with the dissemination of this research, the R–code of the project has been uploaded to the author's github repository (`https://github.com/craigrshenton/masters_thesis`) and distributed under the creative commons licence.

1.1 Fire Size Distributions

Over the past two decades, researchers have increasingly used fire size distributions in the analysis of forest fire dynamics (Cui and Perera 2008:236). Research on the distribution of forest fires covers both the use of simulated models (see Li *et al.* 1999; Cumming 2000; and Song *et al.* 2001) and the analysis of empirical field data (see Ward and Tithecott 1993; Ward *et al.* 2001; and Cumming 2005).

A fire size distribution is simply a probability distribution of forest fire sizes, over a given study area (Cui and Perera 2008:236). They can be described using the following elements:

- (*i*) A study area with a defined spatial extent, over a specified period of time.
- (*ii*) The rate of occurrence of fires within a specified range of sizes over this predefined area.

An individual fire's size is typically measured in hectares (1 ha = 10,000 square metres), however, fires are also classified into fire–size categories (i.e., 4-40 ha, 40–100 ha and so on). These measurements are taken from the maximum extent of the fire's perimeter, thereby, include any unburned areas, known as fire residuals (Heyerdahl *et al.* 2001).

The size of the study area must be sufficient in magnitude (typically over 1000's of ha) and contain a sufficient sample size (i.e., the total number of fires recorded), for statistical inference to be valid (Cui and Perera 2008:236). This is due to the extreme stochastic nature of fire size distributions ('randomness' in layman's terms), with fire sizes ranging from 0.1 ha to over 10,000 ha, all within the same forest. As such, the observation period must be long enough (typically over decades) to capture this variation, else the distribution will inevitably display a bias towards smaller forest fires.

Much of the literature on fire size distributions comes from research carried out in the boreal forests of Canada. However, the knowledge gained from these studies can be generalised to other forests ecosystems around the world (Cui and Perera 2008:235). For example, one of the most startling facts about forest fire distributions in Canada, is that while the vast majority of fires are relatively small, the few extremely large fires that do occur each year, account for up to 96% of the total area burned (Strauss *et al.* 1989:319). This phenomena has also been found to exist in forests in the United States (Heyerdahl *et al.* 2001); Spain (Moreno *et al.* 1998; Vazquez and Moreno 2001); and Australia (Haydon *et al.* 2000).

1.2 Forest Fire Suppression

In many areas of Ontario, forests are now managed to generate sustainable yields of timber, and these resources are pivotal to Canada's continued economic development (Drushka, 2003). The management strategy employed, relies on the timely and effective suppression of forest wildfires, given that for the most part, burned trees cannot be commercially harvested (Bridge *et al.* 2005:41).

While there is no doubt that fire suppression is a direct factor affecting fire size distributions, forests and forest fires are both highly complex systems (shown in *Figure 1*), therefore, fire suppression's true significance has been at the centre of many academic debates (see *Chapter 2*).

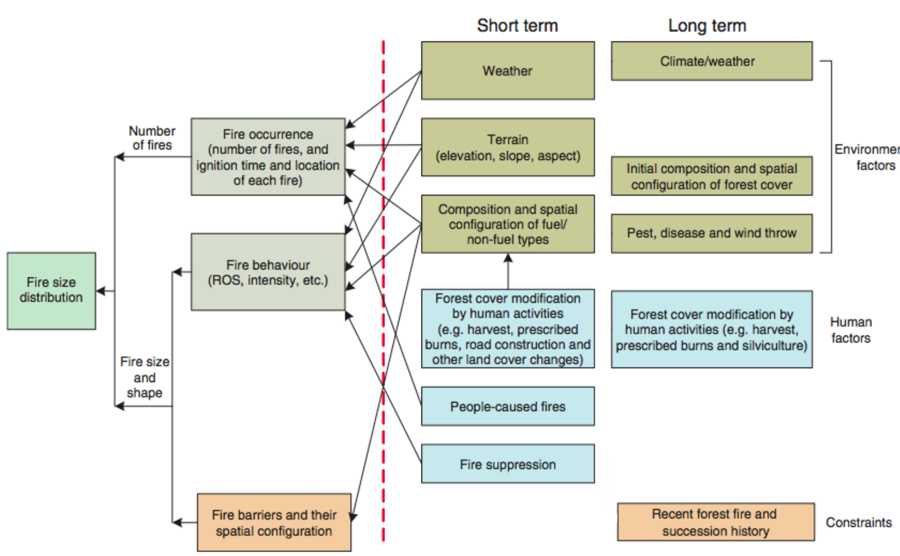

Figure 1: *The major causal factors of forest fire size distributions (adapted from Cui and Perera 2008:239).*

Outside of Ontario, research has shown fire suppression reducing both the number and frequency of large fires, examples include British Columbia (Hawkes *et al.* 1997), Alberta (Cumming 2005), and Italy (Telesca *et al.* 2005). However, other research, again in Alberta (Cumming 2000), but also in Yellowstone National Park (DiBari 2003) shows quite the opposite, in that fire suppression has resulted in fires larger than normal. The explanation offered for these rather counterintuitive results, is that the suppression of individual fires leaves behind fuel that would have normally burned. So that over the course of many years, forests protected from the natural fire cycle, become ever more dense, thereby, becoming prone to catastrophic large fires due to the accumulated build–up of fuel (Flannigan *et al.* 2009).

The extent to which human action influences fire size distributions has therefore, significant implications for the future of commercial timber extraction in Ontario, and for the forest managers and policy makers in Ontario's Ministry of Natural Resources, who are tasked with determining the most sustainable harvest levels (Bridge *et al.* 2005:41). Predictions generated from fire size distribution research have been used before, in both fire management planning, and in the development of future fire suppression policy (Minnich 1983; Minnich and Chou 1997).

1.2.1 Fire Management Planning

Fire managers must contend with considerable uncertainty when attempting to deploy fire suppression resources where they will be most effective (Martell *et al.* 1999:135). However, historical fire size distribution data could be used to develop estimates of future fire behaviour, thereby helping fire managers make more informed decisions (Cui and Perera 2008:240). In addition, fire size distributions, within a specific fire management zone, can be combined with readily available information on the cost of suppressing fires in each class size, to estimate the cost of continued fire suppression (Cumming 2000; Calkin *et al.* 2005; Donovan and Noordijk 2005).

1.2.2 Developing Fire Suppression Policy

Changing fire size distributions can be used as an indicator of how a forest's fire regime may be changing over time (Cui and Perera 2008:241). Correlating these changes with differing fire suppression strategies can be used to evaluate the effectiveness of fire suppression policies (Ward and Tithecott 1993; Ward *et al.* 2001; Bridge *et al.* 2005). Furthermore, historical fire size distribution data can be used to predict the probability of large fire events, thereby helping policy makers develop adequate disaster preparedness plans (Malamud *et al.* 1998; Song *et al.* 2001).

1.3 Limitations of Fire Size Distributions

While it is clear that fire size distributions have many applications in fire suppression research, a certain amount of caution is needed. Fire size distributions are highly specific in both spatial and temporal frames (Cui and Perera 2008:241). As such, they must not be treated as universal constants, and may not be applicable in all areas.

1.3.1 Spatially Explicit

Given that fire size distributions are affected by a highly complex set of causal factors (see *Figure 1*, on page 3), the probability estimates are likely to be highly specific to the particular area from where they were derived. This means that fire size distributions from one forest, climate or fire management zone can not be extrapolated elsewhere, without further calculations (Cui and Perera 2008:241).

1.3.2 Scale Explicit

Furthermore, probability estimates are specific to the particular scale and resolution of the observations from which they were derived. Therefore, estimates derived from only a few months of data for example, can not be used to predict effects happening over decades, and vice versa (Cui and Perera 2008:241).

1.4 Forest Fire Management in Ontario

In Canada, provincial governments are responsible for fire–suppression activities, and in Ontario, these activities are organized by the Aviation Flood and Fire Management Branch (AFFMB) of Ontario's Ministry of Natural Resources (OMNR 1997). Ontario's Ministry of Natural Resources expends a great deal of effort and resources suppressing wildfires, particularly in those areas where commercial timber harvesting is in operation (Cumming 2005:772). These resources include over 200 permanent staff, 600 seasonal firefighters, a fleet of amphibious air–tankers, helicopters and fire detection aircraft. Additional firefighting resources are also contracted in from the private sector during periods of high activity (Martell *et al.* 1999:132).

Fire managers assume these efforts to be effective, given that data from Ontario showed fire suppression had significantly reduced the annual area burned by wildfires over recent decades (Ward and Tithecott 1993). Some calculations suggest that the Ministry's fire suppression activities save up to 35% of the province's annual timber harvest from being burned each year (OMNR 1997).

However, not all forests in Ontario are managed in the same way. Rather, the province is divided into several fire management zones (OMNR 1997), with each zone offering a different level of protection (Bridge *et al.* 2005:41). These areas are classified as the Intensive, Measured, and Extensive fire management zones (see *Figure 2*).

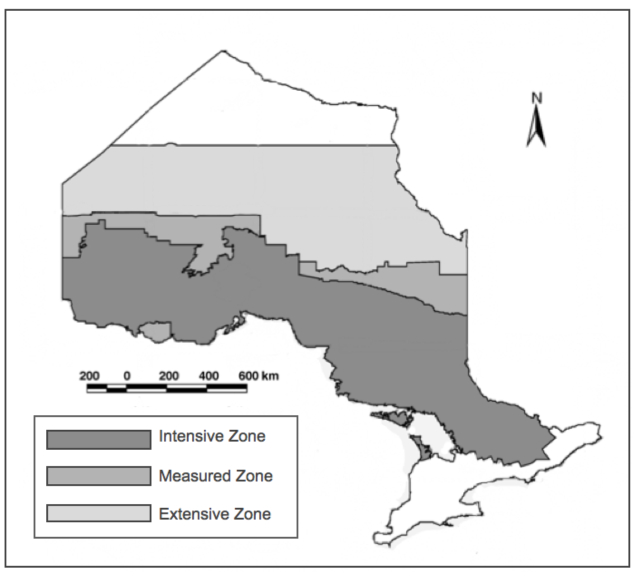

Figure 2: *Approximate distribution of the fire management zones in the province of Ontario (adapted from Bridge* et al. *2005:41).*

1.4.1 Intensive Zone

The intensive protection zone covers areas of significant commercial timber production, as well as those areas in which forest fires threaten residential housing (Bridge *et al.* 2005:42). The strategy employed in the Intensive zone, therefore, is designed to minimise the damage done by forest fires by aggressively attacking all fires in the region until they are fully extinguished. At first, this action, called 'Initial Attack', involves dispatching small crews of firefighters to halt the initial spread of a new fire (Merrill and Alexander 1987). However, if a fire fails to be extinguished while small, the fire is said to have 'escaped' initial attack (Cumming 2005:773). After this point, fire fighters move to a different strategy (hereinafter referred to as *Continued Suppression*), designed to manage and ultimately suppress these larger 'campaign' fires.

1.4.2 Measured Zone

The areas covered by the Measured zone also include commercial timber production, however these areas are considered of lesser value that those in the Intensive zone (Bridge *et al.* 2005:42). For this reason, while all new fires are aggressively attacked, those fires that escape are assessed on a cost–benefit analysis (i.e., the cost of fire–fighting verses the value of the area at risk), as to whether fire suppression should continue (Hirsch *et al.* 1998).

1.4.3 Extensive Zone

The Extensive zone covers the most northern reaches of the provence. Only when lives or property are at risk, is fire suppression employed in this region.

1.4.4 Preventive Measures

In addition to the activities outline above, the Ministry also employs a range of preventive measures (Martell 2001). These include, the use of prescribed burning for silvicultural management (Martell *et al.* 1999:133), and educational programmes in fire safety, such as the Woods Modifications Guide, for those companies operating in forested areas (OMNR 1989).

2 Literature Review

The debate over Ontario's forest fire suppression policy has centered on the credibility of a retrospective study carried out by fire managers at Ontario's Ministry of Natural Resources (Ward and Tithecott 1993). This study has been cited many times in support of the supposed 'effectiveness' of fire suppression, despite being an internal OMNR document, not peer-reviewed science (see Li *et al.* 1996; Welsh and Venier 1996; OMNR 1997; Bourgeau-Chavez *et al.* 2000). The study has faced criticism, most notably from fire ecologists (Johnson *et al.* 2001; Miyanishi and Johnson 2001; Miyanishi *et al.* 2002). While, Ward and Tithecott have sought to defended their original analysis (Ward, Tithecott and Wotton 2001), the case for effective fire suppression in Ontario remains in dispute.

2.1 Ward and Tithecott's Original Thesis

Ward and Tithecott's original report discussed the role played by fire suppression in managing Ontario's forests (Ward, Tithecott and Wotton 2001:1468). Part of their analysis dealt specifically with the impact of fire-suppression activities in regulating fire–size distributions.

The authors began with a simple hypothesis:

> "*A fire suppressed when small, would have otherwise grown to become large in the absence of that suppressive action*" (Ward, Tithecott and Wotton 2001:1469).

From this hypothesis they made a prediction–given two similar forests, one with fire suppression and one without, the forest with suppression would have more small fires, and fewer large fires than in the unsuppressed forest (Ward, Tithecott and Wotton 2001:1469).

To assess the validity of this prediction, Ward and Tithecott (1993) compared the fire distributions between a forested area with some degree of protection, and an area without, over the same 15 year period. The fire–size distribution in areas with fire protection was calculated for Ontario's 'Intensive' and 'Measured' fire management zones (i.e., those areas with fire suppression), using annual fire statistics over the period 1976 to 1990, derived from provincial fire management records in the Ontario Ministry of Natural Resources (OMNR) fire database (Ward, Tithecott and Wotton 2001:1473). Data drawn from the 'Extensive' fire management zone, in which fires are generally (though not exclusively) not suppressed, was used for what Ward and Tithecott (1993) argue is a "*surrogate indicator of the natural distribution of fires by size class in the pre-suppression era.*"

The results were presented as plots of the mean annual numbers of fires by size–class, in the unprotected zone (i.e., Ontario's Extensive fire management zone) and the protected zones (i.e., Ontario's Intensive and Measured fire management zones) (see *Figure 3*), (Ward and Tithecott 1993, cited in Miyanishi and Johnson 2001:1463). These graphs show a relatively flat distribution for the area without fire suppression (*Figure 3a*), and a highly right–skewed distribution for those areas with fire suppression (*Figure 3b*), (Bridge, Miyanishi and Johnson 2005:43).

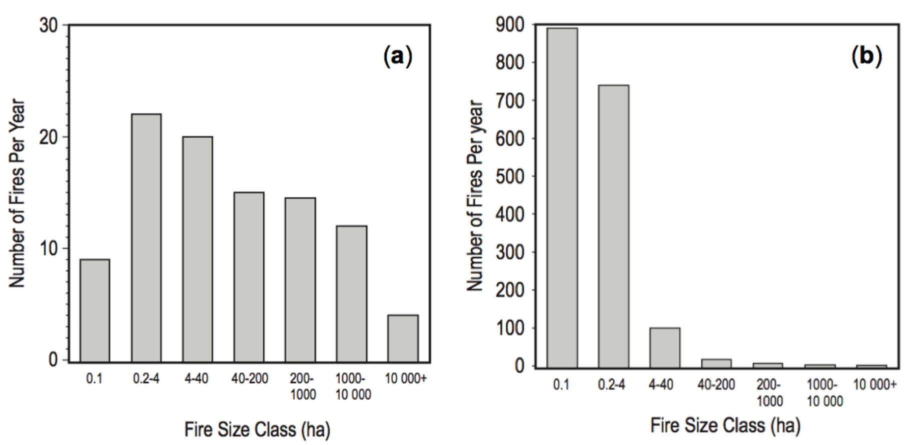

Figure 3: *Distribution of the mean annual fire–size classes (* ha*) in (**a**) Ontario's unprotected Extensive Fire management zone, and (**b**) protected Intensive and Measured Fire management zones (1976–1990), (adapted from Ward and Tithecott 1993, cited in Miyanishi and Johnson 2001:1463).*

Ward and Tithecott (1993) interpreted these results as follows: *Figure 1a* shows a natural fire regime, similar to the pre–suppression era, in which fires would have been naturally distributed across a wide range of size classes (Ward, Tithecott and Wotton 2001:1473). Whereas *Figure 1b* shows very few fires reaching the larger class–sizes. The explanation given for this difference in the shape of the distribution of fire sizes is that, unlike in the unprotected zone, fire suppression in the Intensive and Measured fire management zones prevents many smaller fires from becoming large, compared with what would have been observed without fire suppression (Ward, Tithecott and Wotton 2001:1473). Thereby, fire suppression has skewed the number of fires recorded towards the smaller fire size classes, into a 'J' shaped distribution (Bridge, Miyanishi and Johnson 2005:43).

The same dataset was also presented in Ward and Mawdsley (2000). However, here the protected and unprotected fire management zones were combined into a single graph showing the annual number of fires as a percentage (*Figure 4*). The same conclusion was drawn however, in that fire suppression in the protected zone is creating far fewer moderate– and large–sized fires than in the unprotected zone (Ward and Mawdsley 2000).

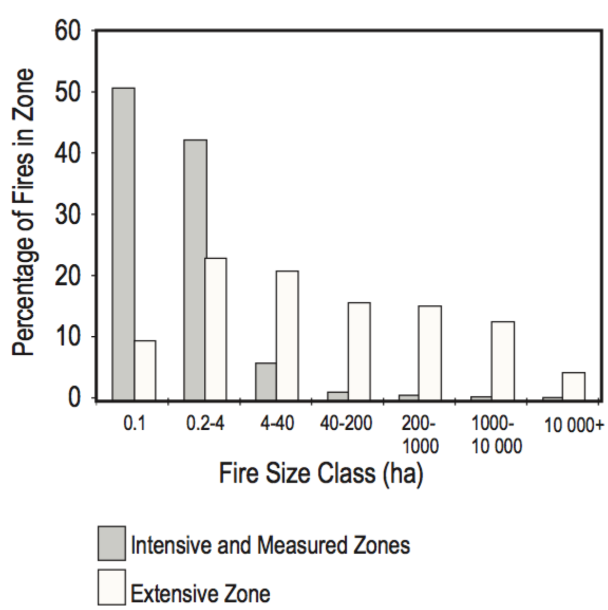

Figure 4: *Distribution of percentage of fires by size class in Ontario (1976–1990) for the unprotected and protected fire management zones (adapted from Ward and Mawdsley 2000:73).*

Ward and Tithecott argue that these results show fire suppression has indeed "*reduced the overall area burned in the protected forest, compared with what would have occurred in the absence of fire suppression*" (Ward, Tithecott and Wotton 2001:1479), and therefore, prove the effectiveness of fire suppression in Ontario (Ward and Tithecott 1993; Ward, Tithecott and Wotton 2001; and Ward and Mawdsley 2000).

2.2 Fire–Ecologist's Critiques

Fire–ecologists have argued however, that Ward and Tithecott's analysis contains several fatal flaws (Miyanishi and Johnson 2001; Johnson *et al.* 2001; and Miyaishi *et al.* 2002).

2.2.1 Difference in Fire Detection Resolution

In the first case, Miyanishi and Johnson (2001:1463) are suspicious of the data presented in *Figure 3*, on page 9 (Ward and Tithecott 1993, cited in Miyanishi and Johnson 2001:1463), which suggests a great disparity in the annual number of fires between the protected Intensive and Measured zones ($\sim$2000 fires per year), and the unprotected Extensive zone (only $\sim$100 fires per year). Furthermore, the number of small fires ($<4ha$) in the Extensive zone (*Figure 3a*) seems extremely small (at only $\sim$30 fires per year), especially considering the vast size of the area in question.

Ward and Tithecott (1993) noted that they may have underestimated the number of small fires in the Extensive zone, as the detection of fires in this zone is thought to be biased towards larger fires. This is common error in estimating fire occurrences (Ricotta *et al.* 1999; Holmes *et al.* 2004; Telesca *et al.* 2005), especially in areas where small fires are not considered important enough to record (Ward *et al.* 2001; Bridge *et al.* 2005). It is therefore possible that these discrepancies are due to the fact that in the sparsely populated Extensive zone, many small fires are simply never detected. Miyanishi and Johnson (2001:1464) conclude that "*any comparison of the numbers of small fires between the two zones cannot be valid due to these differences in detection resolution.*"

However, Ward and Tithecott believe that their original analysis is still valid, as the area burned by small undetected fires is relatively insignificant compared to the overall area burned each year, and therefore, recording their full impact is not strictly required to understand the role of fire suppression in Ontario (Ward and Tithecott and Wotton 2001:1473).

2.2.2 Misrepresenting the Distribution of Larger Fires

Miyanishi and Johnson's (2001:1464) second critique relates to the way in which the comparison between the number of fires in the protected and unprotected zones has been presented. They note that the x–axis on the two graphs in *Figure 3*, on page **??** (i.e., The number of fires per year) are plotted over two different scales (0–30 in *Figure 3a*, and 0–900 in *Figure 3b*). This gives the impression that there are fewer large fires in the protected zone than in the unprotected zone. However, when the Ontario dataset is plotted with the x–axis on the same scale, as Miyanishi and Johnson (2001:1464) have shown below (*Figure 5*), there is little difference between the two zones in the number of moderate– and large–sized fires (i.e., fires >40 *ha*).

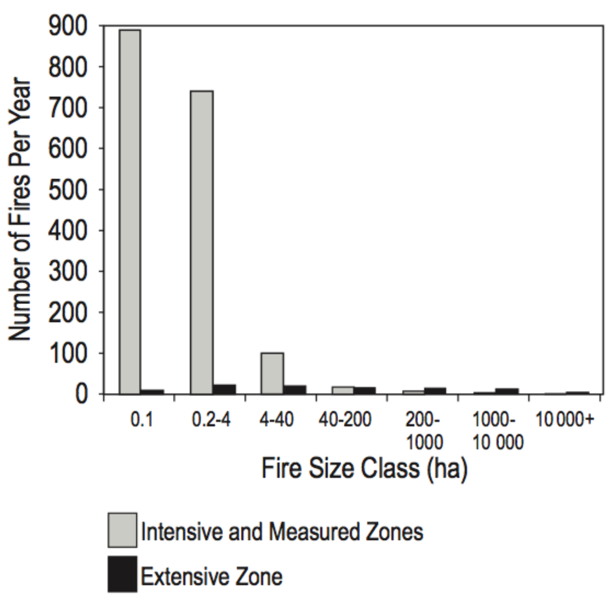

Figure 5: *Distribution of mean annual fire sizes (*ha*) within Ontario's unprotected Extensive Fire management zone and protected Intensive and Measured Fire management zones (1976–1990) using the same scale for the vertical axis. Redrawn from data presented in Ward and Tithecott (1993), (adapted from Miyanishi and Johnson 2001:1464).*

While this mistake was corrected in Ward and Mawdsley (2000:73), the x–axis was plotted as a percentage of the number of fires in each zone (see *Figure 4*, on page 10). Miyanishi and Johnson (2001:1464) argue that this representation of the data can also give a mistaken impression that there are fewer large fires in the protected zones than in the unprotected zone. As previously noted, the data on the number of small fires occurring in the unprotected zone is likely to be censored, therefore, those missing small fires would in effect, skew the distribution towards larger fire classes. Miyanishi and Johnson (2001:1464) conclude, contrary to Ward and Tithecott (1993), that the Ontario dataset shows fire suppression to have not significantly changed the observed distribution of medium– and large–forest fires.

2.2.3 Lack of Consideration of Environmental Variables

As noted in Miyanishi *et al.* (2002), the analysis presented in Ward and Tithecott (1993) and Ward et al. (2001) does not adequately consider the differences that exist between the various fire management zones. It is implicitly assumed in the research design of these studies, that the fire management zones differ only in their fire–suppression policies, and therefore, inferred that the differing levels of fire suppression must account for the differences seen in the distribution of large forest fires (i.e., that the distribution of forest fires is dependent on the fire suppression strategy being employed). The problem, however, is that Ontario's fire–protection zones conform to a general north–south gradation in several environmental variables associated with forest fire dynamics, such as climate, topography, tree species, ignition sources and human activity (Miyanishi *et al.* 2002:1177), as well as longitudinal gradients in humidity and fire frequency (Beverly and Martell 2005; Hills 1959; Suffling 1995). As such, the spatial variation of these variables could also have a causal link to the distribution of fires and, under ideal scientific conditions, should have been controlled for in the original research design. Unfortunately, in this case quantitative disturbance records from Ontario's pre–suppression era do not exist (Ward, Tithecott and Wotton 2001:1470), making a comparison of fire size distributions in a single fire management zone, before and after a substantial period of fire suppression unworkable.

2.2.4 Stochastic Nature of Forest Fires

The final problem outlined in Miyanishi *et al.* (2002), relates to a more fundamental issue with retrospective statistical research. They argue that the extreme size, and stochastic nature of forest fires in boreal forests (see Weir *et al.* 2000), leads to a very high annual variation in area burned. The occurrence (or omission) of a single large fire can therefore, have a significant effect on the results of any statistical analysis (Miyanishi *et al.* 2002:1178). This is especially true of studies from which the data covers only a relatively short time period. Miyanishi *et al.* (2002:1178) cite an example by Bridge (2001), in which the estimates of average annual burn rates were compared, based on 24 years (1972–1995) versus 75 years (1921–1995) of fire data in Ontario. Given that the analysis was using partially the same data, one would expect the estimates to be fairly similar. Bridge (2001) however, found these estimates could differ by several orders of magnitude. Miyanishi *et al.* (2002:1178) also carried out their own analysis of average annual burn rates in Ontario, this time based on 13 years (1976–1988) versus 41 years (1955–1995) of fire data. They found that due to the very large fires in 1956 and 1961, which where not included in the shorter dataset, the estimated average annual burn rates differed wildly based on the omission of just these two years data.

2.3 Defining of Fire Suppression Effectiveness

While both Ward and Tithecott (1993) and Miyanishi *et al.* (2001) agree that the objective of fire suppression is to reduce the area of a forest burned, Cumming (2005:781) argues that neither have been able to developed a workable definition to measure this effect.

Fire ecologists are typically concerned with how the suppression of wildfires effects the structure of boreal forest. This is usually measured by the change in mean annual burn rate (σ), with fire suppression expected to reduce this statistic (Van Wagner 1978). While even small changes in σ can have enormous ecological significance, this effect can only be estimated over very long periods of time, as some causal factors are related to changes in climate over the last 200 years (Bergeron *et al.* 2004).

Unfortunately, while national fire statistics have been recorded since 1918 in Canada, large swathes of the remote northern regions of Ontario were not documented prior to the 1950s (Murphy *et al.* 2000; Stocks *et al.* 2002). Therefore, one must conclude that the empirical data available in Ontario is simply inadequate for determining the effect of fire suppression using this method (Miyanishi *et al.* 2002:1178).

Cumming (2005:773) argues that a definition of 'effectiveness' only needs to be relative to its objectives. In this case, the objective of fire managers in Ontario is to reduce the proportion of 'large' fires (i.e., those fires on which suppression has failed) in fire management zones with aggressive fire suppression policies (Martell 2001, cited in Cumming 2005:781). Therefore, one could argue that an effective fire suppression policy requires that the observed proportion of large fires in areas with aggressive suppression only be lower than in those areas without.

Cumming (2005) has demonstrated this method by comparing the proportions of large fires between areas with contrasting fire management strategies in Alberta, while controlling for other factors (both spatial and temporal) likely to affect this distribution. He believes this effect is best measured as a distribution of the the odds that a randomly chosen fire will be suppressed (Cumming 2005:781).

2.4 Considerations for Future Research

While the case for effective fire suppression in Ontario has not been settled, several important observations can be made:

(i) Quantitative disturbance records from Ontario's pre–suppression era do not exist, making a direct comparison of a single area, with and without fire suppression unworkable. (see *Chapter 2.1*).

(ii) The difference in fire detection resolution between the unprotected and protected fire management zones also makes a comparison unworkable (see *Chapter 2.2.1*).

(iii) The results of any statistical analysis need to be presented in a way that does not distort the data (see *Chapter 2.2.2*).

(iv) Other differences between the fire management zones must be controlled for in any statistical analysis (see *Chapter 2.2.3*).

(v) Larger datasets will give more accurate results (see *Chapter 2.2.4*).

(vi) The definition of fire suppression effectiveness needs to be related to its objectives (see *Chapter 2.3*)

In considering further research on the impact of fire suppression, each of these points should be taken into account at the research design stage of the analysis.

3 Research Design

To overcome the problems found in previous analyses (*Chapter 2.4*), this study will improve on Ward and Tithecott's (1993) research design in several ways.

Miyanishi and Johnson (2001:1464) argued that any comparison between the numbers of fires between the Extensive and Intensive fire management zones cannot be valid due to a difference in fire detection resolution (see *Chapter 2.2.1*). However, no such difference exists between the Intensive and Measured zones, due to the fact that all fires in both the Intensive and Measured zones are actioned and, therefore, recorded whether suppressed or not.

For a comparison to be made however, each zone needs to employ a different fire management strategy. In this case the fire management strategy employed after Initial Attack has failed (i.e., fires that grow >3 *ha*) does in fact differ to a significant degree. In the Intensive zone, all fires that escape Initial Attack, continue to be aggressively attacked until extinguished. Whereas, in the Measured zone, escaped fires are assessed on a cost–benefit analysis, as to whether fire suppression should continue (Hirsch *et al.* 1998). It can be said therefore, that the 'aggressiveness' of continued fire suppression employed to fight escaped fires in the Intensive zone is, to some unknown degree, likely to be greater than that being employed in the Measured zone. Given this simple proposition, one could hypothesis that, on average, more escaped fires would be suppressed in the Intensive zone, than in the Measured zone. However, if the opposite is found to be true, the effectiveness of continued suppression must be in doubt.

As such, testing this hypothesis should provide evidence as to whether continued suppression has been effective (using Cumming's definition of 'effectiveness', see *Chapter 2.3*).

3.1 Hypotheses

Given that our definition of effective fire suppression requires that the observed proportion of large fires in areas with aggressive suppression be lower than in those areas without (*Chapter 2.3*), the following hypotheses were developed:

$H_0 \Leftarrow$ Null hypothesis: that the proportion of large fires is independent of the fire suppression strategy being employed.

$H_\mathrm{S} \Leftarrow$ Strategy hypothesis: that the proportion of large fires is dependent on the fire suppression strategy being employed$*$.

$*$ With the proportion of large fires expected to be lower in the Intensive zone than the Measured zone, if fire suppression is effective.

Operationalising these hypotheses will require a value to signify when continued suppression can said to have failed. For this, I decided to classify fires that burn more than 200 hectares as having 'escaped' continued suppression (i.e., have failed to be suppressed). While somewhat arbitrary, Cumming (2005:2) suggests that it is at this point where such fires become highly distructive events, as fires this large are thought to account for the vast majority (up to 96%) of the area burned by forest fires every year (Strauss *et al.* 1989:319; Johnson *et al.* 2001). Preventing fires from crossing this threshold, therefore, is a primary concern for fire managers in Ontario and was also the main focus of Ward and Tithecott's (1993) original analysis.

3.2 Materials and Sources

The analysis will use annual fire statistics over the period 1989–2004, derived from provincial fire management records in the Ontario Ministry of Natural Resources (OMNR) forest fire database (retrieved `18 Feb 2012`). This database is an archive of all fires detected and reported to the provincial aviation and forest fire management centre. Over the study period, the OMNR received 22461 such reports. Each report includes the following variables: The year of each fire (`FIRE_YEAR`); The fire management zone in which the fire was found (`FIRE_MGT_ZONE`); If known, the cause of the fire, i.e., lightening or anthropogenic (`GENERAL_CAUSE`); The fire's size, measured in hectares (`FINAL_SIZE`);

and both the Latitude and Longitude of the fire's location (`LATITUDE, LONGITUDE`).

3.3 Experimental Conditions

As discussed in *Chapter 2.3.3*, variation in a multitude of environmental factors could be competing causal elements in the distribution of forest fires between each fire management zone (in addition to the fire management strategy employed). As such, these environmental factors must be controlled for under scientific experimental conditions.

3.3.1 Spatial Variation

The delineation required to control for spatial variation in forest fire distributions will be achieved by sub–dividing the dataset into several 'ecoregions'. These ecoregion classifications are commonly used by fire managers at the Ministry of Natural Resources as a way of categorising areas in Ontario based on ecological factors (OMNR 2007: Perera *et al.* 2009). The system used to create these classifications is called the Ecological Land Classification (ELC) system, which is based on Angus Hills' (1959) original classifications, that were first developed in the 1950s. Since that time, the Ministry of Natural Resources has refined Ontario's ecological regions to be compatible with both national and international classification systems (OMNR 2007).

Each ecoregion is defined by a characteristic range and pattern in the following variables (OMNR 2007:4):

- Geology
- Climate (temperature, precipitation, humidity)
- Physiography (soils, slope, aspect)
- Vegetation (species, biodiversity)

While Ontario is subdivided into a total of 14 of these ecoregions, only 3 have been chosen to be studied (3E, 3W, and 3S), due to the availability of relevant data (see *Figure 6*, on page 20). Given that ecoregion 2 is covered solely by the Extensive zone and ecoregions 4 and 5 covered solely by the Intensive fire

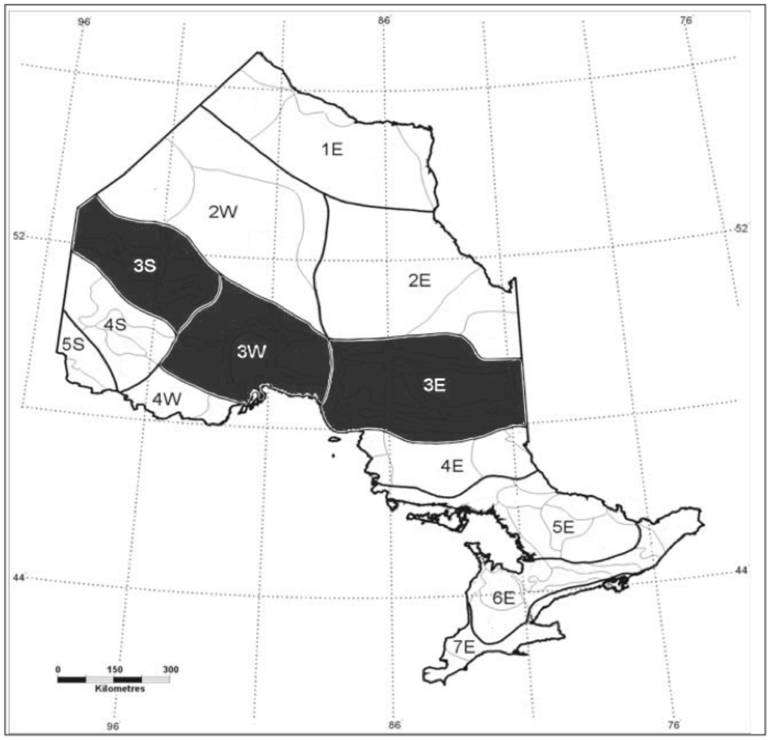

Figure 6: *Ontario's ecoregions 3E, 3W, and 3S (Adapted from Perera* et al. *1998:12).*

management zone, a comparison between fires in the Measured and Intensive zones is not possible within these regions, and therefore, were excluded from this study. In addition, ecoregions 1, 6 and 7 are not subject to forest fire management at all, and therefore, were not applicable to this research.

Despite these limitations, the study area is still highly representative of the boreal forests under active fire management in Ontario (covering approximately 50% of the forested area). The majority of the study area is Crown Land (i.e., government owned) and is managed extensively for timber (Perera *et al.* 2009:2). The relative forest covers are as follows:

- **3E** – 3,541,098 ha (53.5% of total area)
- **3W** – 7,417,259 ha (83.6% of total area)
- **3S** – 12,926,327 ha (94.5% of total area)

3.3.2 Humidity and Fire Frequency

As mentioned in Chapter 2, Ontario's forests exhibit longitudinal gradients in environmental variables associated with forest fire dynamics. Fortunately, sub–setting the dataset into the three ecoregions will also allow for the analysis to account for these variations, given that the 3 ecoregions selected cover the whole province from East to West (48 °– 52 °N and 79 °– 95 °W), (Beverly and Martell 2005; Suffling 1995). While ecoregion 4S has the highest rate of forest fires in Ontario, it is only covered by the Intensive zone (as one would expect), therefore, could not be included in this analysis (Ashiq 2011:13).

The 3 ecoregions also represent a full range of humidity in the region (Ashiq 2011:13). According to Hills' (1959) classifications, the area under study conforms largely to three distinct regions of humidity as follows:

- **3E** – Medium Humid
- **3W** – Dry Humid
- **3S** – Sub Humid

These classification are however, general approximations. According to Ashiq (2011:13), the very Southern part of ecoregion 3S for example, is thought to have a similar climate to 3W (the driest and most fire prone region under study) and therefore experiences similar fire patterns. Analysing such fine spatial variations is however, beyond the scope of this study. What is known, however, is that ecoregion 3W experiences more fires than 3S, due to differences in humidity.

3.3.3 Temporal Variation

As mentioned in *Chapter 2.2.4*, the Boreal forests of Canada are known to exhibit a high annual variation in the area burned by forest fires. Cumming (2005) has argued that this inter–annual variation can potentially confound fire suppression efforts. This is because fire managers do not have an infinite amount of fire fighting resources with which to attack new forest fires. Therefore, on days with high numbers of forest fires, the effectiveness of suppression is bound to be limited (given that fire fighters can only attack so many fires simultaneously). Cumming (2005:776) found that the annual number of fires (annual arrival counts), were strongly correlated with the daily load (i.e., the number of fires, fire fighters have to attack simultaneously), and thus, he suggests these statistics can be used as a proxy in determining any effect the variation in fire load may have on the effectiveness of fire suppression. As such, the effect of temporal variation can be controlled for in the statistical analysis (detailed in *Chapter 4:Methodology*).

3.3.4 Data Error

Podur et al. (2003) and Turner (2009) have both found data points in the OMNR dataset to be erroneous. Their analysis of the dataset shows some fires having started over lakes and rivers, which is clearly impossible. The authors concluded that the source of such error is likely to caused by the relative coarseness of the records. Often, individual fire locations are only recorded to nearest minute. From which, rounding errors can displace a fire's original location by up to 1 km (Turner 2009:203). Fortunately, given the time and distance scale at which this study is operating these types of error will not have a significant affect on the results.

3.4 Reproducibility

As Vitek and Kalibera (2011:36) point out, the confirmation of results by independent researchers is "at the core of the scientific method". However, additional research confirming the results of scientific investigations is usually based only on the information presented in the publication itself. Errors in analysis can often be traced back to the point at which data presented in a paper becomes detached from the operation used to produce that data in the first place. Therefore, while reproducibility is an enviable minimum standard, full replication (i.e., that the original data and code used for analysis is available) is now considered the gold standard in Science (see *Figure 7*).

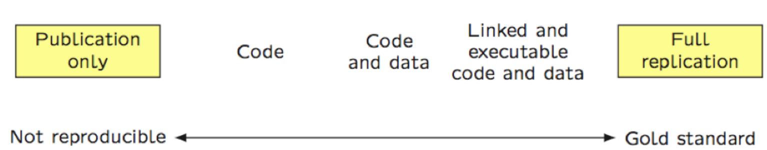

Figure 7: *The Reproducibility Spectrum (adapted from Peng 2011).*

Full replication is better than simply being reproducible, as having the exact code with which the experiment was performed, allows other researchers to easily uncover errors (or fraud). To achieve this gold standard, researchers must fulfil the following requirements (adapted from Schwab *et al.* 2000):

- **Data:** All datasets used in the analysis are made available.
- **Methods:** All computer code underlying the results, figures and tables is made available, and ideally the program required to execute that code is also freely available and open source.
- **Documentation:** The computer code must be documented to a sufficient standard that others are able to repeat the analysis.
- **Distribution:** All of the above must be distributed as openly as possible.

This research will employ all the recommendations outlined above. To do so, all analysis will be carryed out using statistical programming, rather than a propriety statistics application (such as SPSS, STATA etc...) and simply reporting the results in the text. The R system for statistical computation (R Development Core Team 2004) and the R-Studio Integrated Development Environment (R–Studio Development Team 2011) were used throughout. In addition, Sweave (Leisch 2002) will be used to include all of the R code necessary to replicate the analysis within the .Rnw version of this file.

The results presented in this paper therefore, are not simply reported, but calculated inline. This means that there is complete transparency in how the results were analysed, leaving no room to manipulate the figures in support of one particular hypothesis over another. This method of analysis is an example of what is known as 'literate programming' which simply means that the all the data and code necessary for producing the results are contained within the document itself (Gentleman and Lang 2007).

To aid with the dissemination of this resource, the R code of the project will be uploaded to the author's github repository (`https://github.com/craigrshenton/masters_thesis`) and distributed under the creative commons licence.

4 Methodology

Note: The full R code used in the analysis is available both in print (*Appendix A*) and executable format– within the sweave (`.Rnw`) version of this file (`https://github.com/craigrshenton/masters_thesis`).

4.1 Operationalising the Hypotheses

Operationalising the hypotheses described in *Chapter 3.1*, will employ a method similar to Cumming's (2005) approach. This method is relatively simple and well known to researchers performing clinical trails, to whom it is know as a Bernoulli trial (Papoulis 1984).

A Bernoulli trial consists of an experiment in which the outcome is a random variable with one of two possible values– '`success`' or '`failure`' (known as a binomial random variable). Convention has it that these values are encoded as follows:

```
0 = failure
1 = success
```

In each experiment, the probability of a randomly chosen outcome being successful is denoted as (p) and the probability of failure (q). The relationship between (q) and (p) is given by:

$$q = 1 - p$$

The collective results of a number of statistically independent Bernoulli trials are known as a binomial experiment, with the output being a binomial distribution of probabilities or probability ratios.

Binomial experiments are often used to determine the effectiveness of a new drug in a clinical setting. One of the first uses of binomial experiments was in biological assay (bioassay) research (Finney, 1973). This is where researchers administer poisonous substances to animals in order to determine the lethality of its effect. One would expect that the greater the dose, the greater the number of animals killed. A binomial experiment would be used here to show that the probability of an animal being killed by the poison is significantly greater when given a high dosage, than when given a low dosage (in other words, that the probability of death is dependent on the dosage of poison administered). Thereby, proving that the effectiveness, is in fact dependent on the dosage.

For this study, the 'dosage' can be considered equivalent to the 'amount' of fire suppression being deployed, which is expected to increase the number of fires successfully suppressed (i.e., increase the effectiveness). Specifically, we are interested in two different strategies (or 'dosages') of fire suppression. In the Intensive zone, a strategy of high levels of suppression is employed, and in the Measured zone, a strategy of low(er) levels of suppression is employed (see *Chapter 3*). Of course, in this study, these 'trails' are retrospective, having already happened in the past, and therefore, can only be considered quasi-experimental.

As in the Bernoulli trial, the outcome of each fire being attacked with differing levels of suppression, is a random variable with one of two possible values– '`success`' or '`failure`'. As previously discussed (see *Chapter 3.1*), I decided that fires burning more than 200 hectares would signify a failure of continued suppression. Therefore, fires with a final burn area greater that 3 hectares, but less that 200 hectares, are considered as having been successfully suppressed. The following logic operation was used to determine the number of fires both suppressed and unsuppressed (see *Appendix A: 1*):

`Suppressed (`p`) = Final Size` $\geq$ `3 ha & Final Size` $\leq$ `200 ha`

`Unsuppressed (`q`) = Final Size` $\leq$ `200 ha`

Another logic operator was used to sub-divide the dataset into fires in the Intensive (`INT`) and Measured (`MEA`) fire management zones (see *Appendix A: 1.1*):

```
Fire Management Zone = "MEA" & Fire Management Zone = "INT"
```

To sub-divide the dataset into the three ecoregions (3E, 3W and 3S) required for the analysis, another logic operator was used that delineates fires along both latitude and longitude as follows:

```
3E = (Lat ≥ 48°N & ≤ 51°N) + (Long ≤ -80°W & ≥ -86°W)
3W = (Lat ≥ 48°N & ≤ 51°N) + (Long ≤ -88°W & ≥ -92°W)
3S = (Lat ≥ 51°N & ≤ 53°N) + (Long ≤ -88°W & ≥ -96°W)
```

From these operators, the following statistics were calculated for each ecoregion and each fire management zone (see *Appendix A: 1.2–1.3*):

```
S = the number of successfully suppressed fires.
E = the number of escaped (i.e., unsuppressed) fires.
N = the total number of fires (S + E).
```

The results of which were saved as a new dataframe `omnr.glm` (see *Table 1*).

zone	ecoregion	suppressed	unsuppressed
mea	E3	60	20
mea	W3	194	67
mea	S3	111	19
int	E3	152	20
int	W3	91	43
int	S3	44	10

Table 1: *Summary of the omnr.glm dataframe.*

As Cumming (2005:775) suggests, the distribution of fires in each area will be an estimate of the probability (p), that a randomly chosen fire will be suppressed. This binomial random variable will have an expected value of $N \times p$ and a variance of $N \times p(1-p)$.

The hypotheses H_0 and H_S can now be tested by regressing the observed suppression probabilities (p) against the fire management strategy being employed, while controlling for the environmental factors outlined in *Chapter 3.3*.

4.2 Assumptions

Several assumption have been made in this analysis:

(i) It is assumed that the fire detection resolution in both the Intensive and Measured fire management zones to have been both similar and constant over the study period (1989–2004). This is evidenced by reports in Ward and Tithecott (1993); Ward, Tithecott and Wotton (2001); and Cumming (2005).

(ii) It is assumed that all new fires in both fire management zones are actioned in a consistent manner over the study period. While there are no reasons to doubt this assumption, further analysis could detect any differences by looking into the deployment records (i.e., number of fire fighters, spotters, helicopters, and so on). However, performing this analysis was beyond the scope of this study.

(iii) It is assumed that each new forest fire is an independent event. That is to say that each fire is not a causal factor in the distribution of subsequent fires. The veracity of this assumption is discussed in *Chapter 6.1*

4.3 Generalised Liner Models

Cumming (2005:774) argues that the standard linear regression methods used for hypothesis testing are not suitable for binomial experiments. This is because the suppression probability (p) is a binomial variable (i.e., is bounded on both ends, with values in the range $[0, 1]$). This means that (p) is an odds ratio and, therefore requires a log–odds model. In addition, the variance in p is heteroscedastic (i.e., not all variables have the same variance), which can invalidate statistical tests of significance if not accounted for in the research design (Dobson 2002: Chapter 7.2).

Instead, the hypothesis tests will be formulated as generalised linear models (not to be confused with *general* liner models), which were formulated by John Nelder and Robert Wedderburn (1972). The generalisation of linear regression allows the use of a response variable that does not have a normal (or Gaussian) distribution. This is achieved by a link function. While the choice of link function is somewhat arbitrary, McCullagh and Nelder (1983: Chapter 4.1.5) suggest that the logit link is most appropriate for retrospective binomial experiments, as it is the natural log of odds (log odds).

The logit link outlines the relationship between the linear predictor and the non-normal distribution (in this case, a logistic distribution). The linear predictor η (the Greek letter "eta") shows how the suppression probability (p) is related to the parameters β, and their coefficients x by the equation:

$$\eta = \beta x$$

The logit link used for this analysis can therefore, be expressed as (Dobson 2002: Chapter 7.3):

$$\beta x = \log p(1 - p)$$

4.4 Testing the Hypotheses

The null hypothesis (H_0) implies that the suppression probabilities p should be similar over the whole study area. This will be estimated using `model 0` (null) as follows:

$$\texttt{logit}\ p = \beta_0$$

If the difference in strategy has improved the effectiveness of fire suppression (H_{S}), the suppression probability (p) should be larger in the intensive zone than in the measured zone. Which will be estimated using `model` S (strategy):

$$\texttt{logit}\ p = \beta_0 + \beta_1\ \texttt{zone}.$$

To control for environmental factors, another interaction term is needed that accounts for the differences between the ecoregions `3E, 3W` and `3S`. The full `model` SE (strategy + ecoregion) therefore, is as follows:

`logit` $p = \beta_0 + \beta_1$ `zone` $+\beta_2$ `ecoregion`.

It should be noted that the response variable is usually binary coded (i.e., 0 = failure and 1 = success), however, the variables (or 'vectors' in R) '`suppressed`' and '`unsuppressed`' are not binary by default. Rather than convert these variables by hand, it is possible to simply bind them into a matrix using the `cbind()` command in R, which achieves the same result.

Using this method leads to the following equation:

```
glm.eq = "cbind(suppressed, unsuppressed)  zone + ecoregion"
```

The final generalised liner model with logit link function is, therefore, as follows:

```
glm(glm.eq, family=binomial(logit), data=omnr.glm)
```

Note: the data frame onmr.glm was created from the results of previous calculations (see *Chapter 4.2*).

4.5 Controlling for Temporal Variation

As outlined in `Chapter 3.3.3`, the inter–annual variation in forest fires can potentially confound fire suppression efforts. Therefore, to test the effect of temporal variation has on the dataset, a further generalised liner model was developed.

Much like in the previous analysis, the following statistics were calculated for each year over the period 1989–2004 (see *Appendix A: 1.5*)

S_t `= the annual number of successfully suppressed fires.`
E_t `= the annual number of escaped (i.e., unsuppressed) fires.`
N_t `= the total number of fires each year.`

The results of which are saved as a new data frame `temp.glm` (see *Table 2*).

However, the crucial difference in this case, is that N_t is a measure of all fires each year, including fires less that 3 ha. As Cumming (2005:776) suggests, this

zone	year	suppressed	unsuppressed	load
mea	1989	49	7	2429
mea	1990	10	0	1614
mea	1991	52	14	2559
mea	1992	28	2	960
mea	1993	5	2	743
mea	1994	34	4	1053
mea	1995	58	26	2122
mea	1996	33	27	1245
mea	1997	21	6	1634
mea	1998	23	4	2278
mea	1999	12	6	1017
mea	2000	6	0	644
mea	2001	6	2	1561
mea	2002	11	6	1138
mea	2003	10	6	1036
mea	2004	1	0	428
int	1989	137	7	2429
int	1990	83	5	1614
int	1991	12	8	2559
int	1992	79	7	960
int	1993	42	1	743
int	1994	93	1	1053
int	1995	199	27	2122
int	1996	89	54	1245
int	1997	73	5	1634
int	1998	150	15	2278
int	1999	83	12	1017
int	2000	57	1	644
int	2001	101	0	1561
int	2002	38	10	1138
int	2003	77	10	1036
int	2004	16	1	428

Table 2: *Summary of the temp.glm dataframe.*

variable is strongly correlated with the daily load (i.e., the number of fires, fire fighters have to attack simultaneously), and thus, can be used as a proxy in determining how temporal variation in fire load may impact on the effectiveness of fire suppression (see *Chapter 3.3.3*).

The same generalised liner model was used as before (i.e., a binomial, logit link model). However, to determine the significance of temporal variation in fire load, another interaction term was added, giving `model` SL (strategy + load) as follows:

$$\texttt{logit}\ p = \beta_0 + \beta_1\ \texttt{zone} + \beta_2\ \texttt{load}.$$

Which was coded as;

```
glm.eq = "cbind(suppressed, unsuppressed)  zone + load"
```

and;

```
glm(glm.eq, family=binomial(logit), data=temp.glm)
```

It should be noted that originally, the intent was to analyse both the spatial and temporal variance in one single GLM, that included all 3 interaction terms. However, it was found that the variance (standard deviation) at such a fine scale (i.e., the number of fires suppressed; per zone, per ecoregion, per year) was so large as to make a statistical test of any significance, all but impossible. This is largely due to the extreme stochastic nature of forest fires, and shows the limits of statistical analysis (see *Chapter 6.2.2* for discussion).

5 Results

From the method outlined in *Chapter 4*, the following statistics were derived for the study area (i.e., fires in ecoregions 3E, 3W and 3S, over the period 1989–2004).

5.1 Total Number of Forest Fires

The number of new forest fires–per year (N_{t}), *Table 3*, was calculated by aggregating the total number of data points (in this case, the final size of each fire), in each year (see *Appendix B: 1.4* for calculations);

Year	Fires
1989	478
1990	323
1991	736
1992	219
1993	113
1994	303
1995	771
1996	508
1997	506
1998	755
1999	293
2000	194
2001	430
2002	361
2003	395
2004	144

Table 3: *The number of forest fires–per year (Nt)*

The total number of forest fires (N) was then calculated by aggregating (N_{t});

N = 6525 fires.

Therefore, we can say that in the study area there were a total of 6,529 fires, over the period 1989–2004. Of those fires, the vast majority, 5607 (86%), were suppressed (via Initial Attack) before becoming >3 ha in size. 744 (11%) fires were put out by continued suppression (i.e., fires >3ha but <200 ha), and only 178 (3%) escaped continued suppression, (i.e., burning >200 ha of forest).

5.2 Area Burned by Forest Fires

The area burned by forest fires, per year (A_t), *Table 4*, was calculated by aggregating the final size of every fire, per year (see *Appendix B: 1.4* for calculations):

Year	Area (ha)
1989	17615
1990	3584
1991	19876
1992	10244
1993	18483
1994	2336
1995	158269
1996	236434
1997	22459
1998	45144
1999	138822
2000	632
2001	1268
2002	14759
2003	103703
2004	172

Table 4: *The area burned by forest fires, per year (At)*

The total area burned by forest fires (A) over the study period is therefore, the summation of (A_t):

A = 793797 ha

These result can also be interpreted graphically by taking the mean area burned per year, as shown in *Figure 8*, (see *Appendix B: 1.5* for calculations).

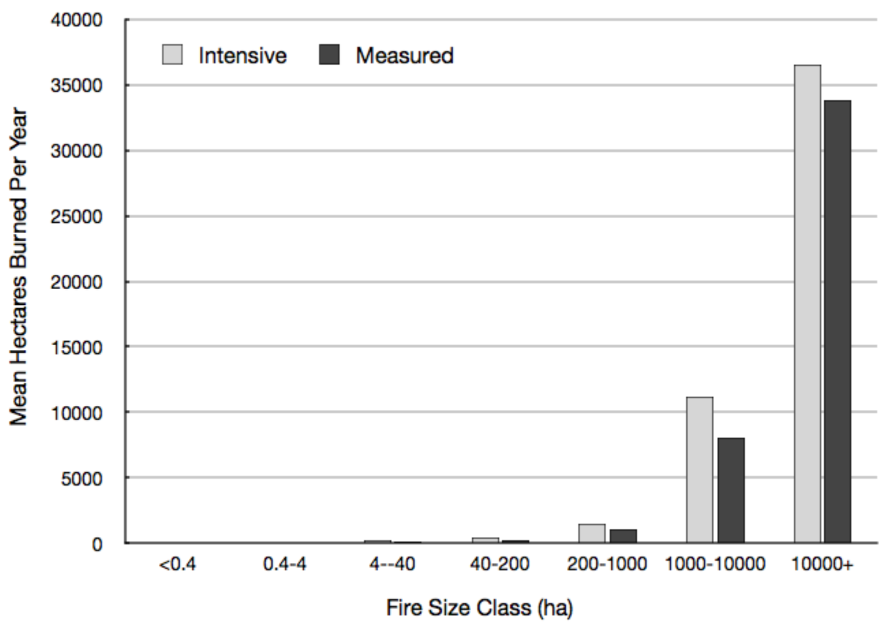

Figure 8: *Mean hectares burned by size class (1989–2004) in the Intensive and Measured fire management zones.*

These results are consistent with Cumming's (2005:2) observation (see *Chapter 3.1*) that fires greater that 200 ha, account for the vast majority of the area burned by forest fires every year (98% in this case). Hence, why preventing fires from becoming this large is a primary concern for fire managers in Ontario.

5.3 Fire Size Distributions

In Ward and Tithecott's (1993) original analysis (see *Chapter 2.1*), the authors used the relative fire size distributions in different fire management zones as an indicator of the impact of fire suppression in Ontario. They predicted that an area with aggressive fire suppression would have fewer large fires, than an area without.

To see how this prediction has fared over the period 1989–2004, the fire size distributions (measured as a percentage of fires in that zone) were calculated (see *Appendix B 1.1-1.3* for calculations). The results of those calculations are shown in *Figure 9*, on page 35.

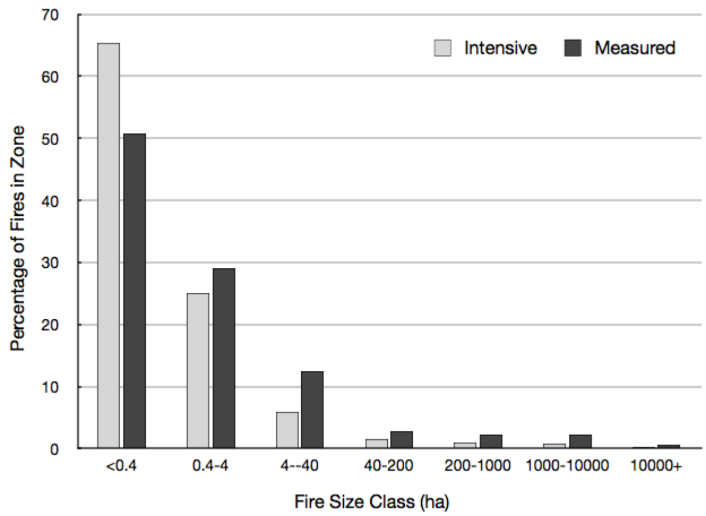

Figure 9: *Percentage of fires by size class (1989–2004) in the Intensive and Measured fire management zones.*

While it is not readily apparent, there is indeed a significant difference in the fire size distributions between the Intensive and Measured fire management zones. If we focus in on fires at the large end of the size class spectrum (see *Figure 10*), it can clearly be shown that there is approximately twice the number of moderate– and large–sized fires (i.e., fires >40 ha) in the Measured zone compared to the Intensive zone.

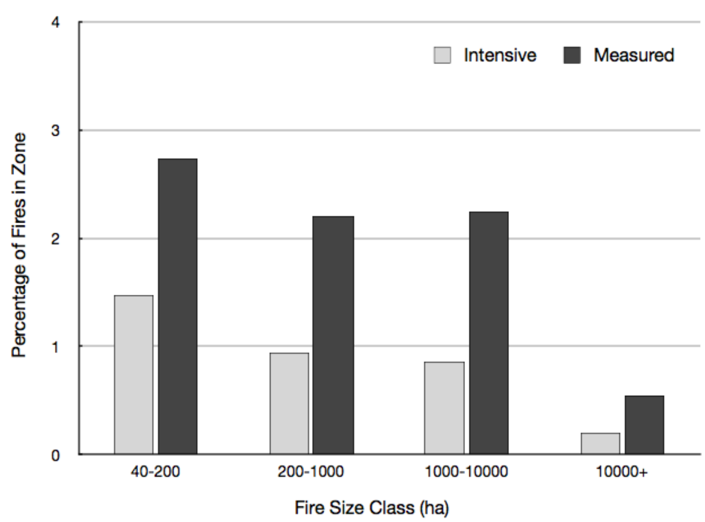

Figure 10: *Percentage of fires for size classes 40 ha to 10000+ ha (1989–2004) in the Intensive and Measured fire management zones.*

From the analysis presented above, one might be tempted to draw the same conclusion as Ward and Tithecott (1993), that aggressive fire suppression in the Intensive zone has lead to far fewer moderate– and large–sized fires. However, this would be based on the counter factual– that without fire suppression the fire size distribution in the Intensive zone would have been the same as that of the Measured zone. Therefore, while these results are certainly indicative of the impact fire suppression has had, it can not alone, prove the effectiveness of fire suppression in Ontario.

5.4 The Generalised Liner Model

To measure the effectiveness of fire suppression in a more scientifically robust manner, a generalised liner model (GLM) was developed (see *Chapter 4.3*), that would not only estimate the suppression probabilities between the fire management zone, but also control for other environmental factors that may be influencing the distribution of forest fires in Ontario.

The suppression probability (p) in both the Intensive and Measured zones were estimated using the generalised liner model (GLM) as follows:

```
glm.eq = ”cbind(suppressed, unsuppressed)  zone + ecoregion”

glm(glm.eq, family=binomial(logit), data=omnr.glm)
```

The results of which, can then be displayed using the `summary()` function:

```
summary(glm.out)
```

The results of which are detailed in *Table 5.*

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.6691	0.1819	9.18	0.0000
zonemea	-0.0043	0.1822	-0.02	0.9813
ecoregionS3	0.0101	0.2751	0.04	0.9708
ecoregionW3	-0.7142	0.2150	-3.32	0.0009

Table 5: *Coefficients for the Generalised Liner Model (zone + ecoregion)*

5.5 Interpreting the GLM

The way in which the `glm()` command works is to estimate the suppression probability (p) in the (`intensive zone / ecoregion 3E`) as default and then add the other interaction terms sequentially– as follows; (`measured zone / ecoregion 3E`), (`intensive zone / ecoregion 3S`), (`measured zone / ecoregion 3S`), (`intensive zone / ecoregion 3W`), (`measured zone / ecoregion 3W`).

The last coefficient (`ecoregion 3W`) in the summary, therefore shows the estimate, standard error and significance value for the complete model.

The estimate is reported as -0.7142, however this is a log of odds. Therefore, to find the suppression probability (p) (as an odds ratio), it must first be interpreted as the antilog of the model estimate:

`exp`(-0.7142) $= p$

$p = 0.4896$

Therefore;

All else being equal (i.e., when interaction with the ecoregion variable is controlled), the odds of a fire being suppressed in the Measured zone is less than half (0.5) that of the Intensive zone, giving the Intensive zone a significant advantage in terms of fire suppression effectiveness.

5.6 Measuring the Performance of the Model

The performance of the model can be evaluated in two distinct ways, via ANOVA and AIC.

5.6.1 ANOVA

ANOVA (analysis of variance) is often used to analyse the results of experiments by testing for statistical significance. The results outlined above, may be called statistically significant if they are unlikely to have occurred by chance alone. A calculated probability is used to determine if this is true. If the probability is less than a pre–determined threshold, called a significance level (usually 0.05, i.e., 5% chance), we can safely reject the null hypothesis. By testing hypotheses in this way, we can significantly limit Type `I` errors (i.e., false positives).

In this case, the null hypothesis H_0, is that there is no difference between the fire suppression probabilities of the Intensive and Measured fire management zones (in statistics this usually means that all the results are samples of the same population). This implies that the different fire management strategies have no discernible effect on the resulting fire suppression probabilities. Ergo, rejecting the null hypothesis would imply that the different fire management strategies *have* in fact altered fire suppression probabilities.

Testing the results will be operationalised as a chi square test, which will show how well the interaction terms (zone and ecoregion) explain the deviation between the model and the data.

The advantage of statistical programming is that the output of the generalised liner models is saved as a single vector (glm.out), which can then easily be used in further analysis. Therefore, while a chi square test can often be complicated to implement, it can be programmed in R simply using the following command:

```
anova(glm.out, test="Chisq")
```

The results of which are detailed in *Table 6*.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			5	26.92	
zone	1	0.60	4	26.32	0.4381
ecoregion	2	17.19	2	9.12	0.0002

Table 6: *Analysis of variance for the Generalised Liner Model (zone + ecoregion)*

The results of the chi square test show the null deviance to be 26.92 points on 5 degree of freedom. This number represents how well the response variable (number of fires suppressed) is predicted by the model with only the intercept. Adding in the interaction terms (`zone` and `ecoregion`) decreased this deviance by 17.2 points on 2 degree of freedom. When interpreted as a chi square value

(0.00018), it indicates a highly significant decrease in deviance (where <0.05 is considered to be statistically significant). This gives us a $>99\%$ confidence interval in that these results are not due to random chance alone. While these values are computed automatically in R, it is still possible to use F tables in the Engineering Statistics Handbook (Croarkin and Tobias 2003) and look up these values by hand.

One should note however, that this test rests upon one key assumption about the data. That is in the independence of each observation (Snedecor and Cochran 1967:321). In this case, independence means that each forest fire needs to be an independent event. However, fire ecologists may point out that forest fires (and the suppression of those fires) may change the distribution of subsequent fires (i.e., that suppressing a fire today, simply adds more fuel for fires to burn in the future). While this is undoubtably the case, modelling this behaviour is far beyond the scope of this project and, therefore, for the sake of simplicity, independence was assumed.

From the generalised liner model we have found the suppression probabilities (p) in the Intensive and Measured zones to be different by a ratio of 2 to 1. By analysing of variance of this model, we now know that this deviation can be explained by the different suppression strategies employed in each fire management zone. Therefore, the null hypothesis H_0 was rejected.

5.6.2 AIC

The Akaike Information Criterion (AIC) was developed by Hirotsugu Akaike (1974), and is used to test a goodness of fit of a statistical model. The AIC also takes into account the tradeoff between the accuracy and complexity of the model. This is done by penalising a model for increasing the number of interaction terms, thereby, discouraging over–fitting (i.e., adding spurious covariates simply to increase the goodness of fit). However, the AIC value only provides a relative measure of fit, and can therefore, only be used as a method for choosing between competing models (rather than hypothesis testing).

The AIC value is derived from the following equation;

$\texttt{AIC} = 2k - 2\texttt{ln}(L)$

Where;

k is the number of interaction terms used

L is the maximum value of the likelihood of the model

Therefore, given several different models that may explain the data, the model with the lowest AIC value is preferred. In *Chapter 4.4*, the different models used were outlined as model S (strategy) and model SE (strategy + ecoregion), (*Note model 0 (null) has already been rejected).

The AIC values for these models are as follows;

model $S = 59.14$

model $SE = 45.95$

As such, model SE (i.e., the full model used in the ANOVA test above) best fits the data.

5.7 The Temporal Model

To test the effect temporal variation has on the dataset, a further generalised liner model was developed, that estimates the suppression probability (p) in both the Intensive and Measured zones, on an annual basis. The GLM was coded as follows:

```
glm.eq = "cbind(suppressed, unsuppressed)  zone + load"

glm(glm.eq, family=binomial(logit), data=omnr.glm)
```

The results of which are detailed in *Table 7*, on page 41.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.9210	0.1895	10.14	0.0000
zonemea	-0.9464	0.1377	-6.87	0.0000
load	0.0001	0.0001	1.00	0.3192

Table 7: *Coefficients for the Generalised Liner Model (zone + load)*

A chi square test was aslo performed to see if the load interaction term explained the deviation (*Table 8*).

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			31	238.46	
zone	1	44.03	30	194.43	0.0000
load	1	0.99	29	193.44	0.3188

Table 8: *Analysis of variance for the Generalised Liner Model (zone + load)*

The results of the chi square test show the null deviance to be 238 points on 31 degree of freedom. This is a very large deviance. Adding in the interaction term (`zone`) decreased this deviance by 44 points on 30 degree of freedom. Further more, adding the interaction term (`load`), only reduced this deviance by 1 point on 29 degree of freedom. When interpreted as a chi square value (0.32), it indicates a very low decrease in deviance (where <0.05 is considered to be statistically significant). This gives us no confidence in the results of this model. In addition, the AIC value of model $SL = 295.8$, which is very high, suggesting that the model is not a good fit.

This analysis suggest that fire suppression is not confounded, to any significant degree, by the annual number of fires in each zone. Therefore, the only parsimonious conclusion from the evidence presented here, is that the difference in fire management strategies between the intensive and measured zone has been primarily responsible for the observed differences in fire distributions over the period 1989–2004 and, as such, the aggressive fire suppression strategy employed by the OMNR can be considered effective at reducing the number of large forest fires in Ontario.

6 Discussion

Based on Cumming's (2005:773) definition of effectiveness (i.e., that the observed proportion of large fires in an area with aggressive suppression is lower than in areas without), it can now be said that the fire suppression strategy employed by the OMNR has indeed been effective over the study period 1989–2004. This study vindicates those fire managers at the Ministry of Natural Resources, who argued that fire suppression significantly reduces the annual area burned by large wildfires. Proving that the OMNR's fire management strategy has delivered in its objectives will help protect those services against calls to reallocate both funding and resources.

The new data on fire size distributions also helps fire managers plan new strategies over the short– to medium–term. The data can be used to develop estimates of how the forests in other fire management zones would react to an increase in fire suppression activities, thereby helping fire managers make more informed decisions about the delineation of fire management zones. Measuring the impact of aggressive fire suppression (i.e., the total area saved from forest fires due to fire management) could also help quantify the value of those resources protected by existing fire management strategies, details which are highly pertinent to the ongoing commercial timber operations in Ontario. Calculating this impact, however, requires the use of counterfactual modelling, and was, therefore, beyond the scope of this study.

6.1 Epistemological Limitations

While the results presented in this paper support the OMNR's strategy, there are still concerns over the long term sustainability of forest management in Ontario. These concerns arise from the epistemological limitations of the methods used in fire size distribution research itself.

The dynamics of forests and forest fire are complex and poorly understood phenomena (see *Chapter 1.2*). The resultant fire size distributions are determined by many causal factors. These factors may interact with fire suppression, to either enhance or neutralise its effectiveness (Cui and Perera 2008:241). For example, while in this study, each forest fire was assumed to be an independent event (see *Chapter 5.6.1*), we know that this in not true in reality. Forest fires can change the distribution of subsequent fires, which in turn, can determine the composition of future forests and their susceptibility to wildfire. Even research based on computer modelling of fire size distributions is not immune from oversimplification (Song *et al.* 2001). Long term factors, such as climate patterns, pest epidemics and demographic changes could all significantly challenge the continued effectiveness of fire suppression (Cui and Perera 2008:238). Overcoming these limitations would require further studies, using GIS to explore the relationships between successive fires, on a much finer spatial scale than what could be achieved here.

Holling and Meffe (1996) have warned against overlooking the complexities of forest management. They argue that the aggressive fire suppression strategies, designed to increase the short–term predictability of forest ecosystems, may be leading to long–term instability, by reducing the forest's natural reliance to disturbances. In Yellowstone national park in the United States for example, fire suppression efforts were at first very successful. However, the consequence of those actions only became apparent years later, when the accumulation of unburned fuel eventually lead to "*fires of an intensity, extent, and human cost never before encountered*" (Kilgore 1976; Christensen *et al.* 1989). In the clearing created by the first large fire, huge swathes of new forest grew together, laying the foundations of the next fire, creating what was described as an "*unending cycle of monster fires and blackened landscapes*" (Bonnicksen 2002:2). Holling and Meffe (1996:330) believe that the conflagrations seen in Yellowstone are an inevitable consequence of the same approach to forest fire management taken by the OMNR in Ontario.

The proposed solution to these concerns (which is also supported by Donovan and Brown 2007) requires forest managers to accept the natural range of variation within the forest ecosystem, rather than trying to change or control them (Holling and Meffe 1996:334). With the aim to restore forest ecosystems so that they are more representative of the prehistoric landscape (Bonnicksen 2002:5). These policies would require the OMNR to abandon large swathes of the provence to wilderness and halt further development. Considering the forests commercial significance, however, these ideas are highly unlikely ever to be implemented.

6.2 Statistical Limitations

While statistical analysis makes scientific studies possible, the methods contains within are not without their problems. As Nassim Nicolas Taleb (2008) states:

> "*Statistics is the core of knowledge; the logic of science; the tools of epistemology. However, statistics can fool you.*"

Most of these problems originate from the necessary abstractions and assumptions made in the research design (see *Chapter 3*). However, there are other, deeper issues, originating from studying the philosophy of science. While none prove fatal to the evidence presented here, for the sake of completeness and a more open science, they are worthy attention.

6.2.1 Probabilistic Modus Tollens fallacy

Hypothesis testing (as used in *Chapter 4*) rests on a logical argument known as modus tollens (also known as 'denying the consequent') (Gill 1999:7). The logic can be explained as follows:

> If P implies Q, and Q is found to be false;
> Then one must concluded that P must also be false.

In scientific research, this method involves verifying a theoretical assumption with observations of real–world events. For example, if the null hypothesis H_0 is true, then the data should follow an expected pattern (usually $\theta = 0$). If the data does not follow this pattern therefore, one must conclude that H_0 must be false.

However, when we replace certainties with probabilities (as is the case with the method employed in this study), the logic of modus tollens no longer holds true (Gill 1999:7). For example, it would be a fallacy to assert that; if the null hypothesis H_0 is true, then the data is *likely* to follow an expected pattern. And if the data does not follow this pattern, one must concluded that H_0 is therefore *unlikely*. Simply obtaining a sample of data that is atypical (such as that found in this study), in no way proves H_0 to be false. Merely that, over the given study period, H_0 has not yet been observed.

To put this in more concrete terms; while it can be said that over the study period 1989–2004, the aggressive fire suppression strategy has been effective at reducing the number of large forest fires, we cannot say with any certainty that this effect will continue indefinitely.

6.2.2 Predicting Small Probability Events

If we did wish to know whether the effect of fire suppression will continue in the future, more observations would be needed. However, given that the probabilities involved with large forest fires are very small, the number of observations needed grows exponentially. For example, to predict 1 year ahead, at least 10 or more years of data would be required. However, to predict 10 years ahead, more than 1000 years of data would be required to give us any remotely acceptable accuracy (Makridakis and Taleb 2009:2) In addition, estimates derived from these observations tend to come with greater error than for more frequent events. To solve this problem statisticians assume that these events follow a probability distribution (For example, in this study a logistic distribution was assumed, see *Chapter 3*). This is done so that we can extrapolate that the pattern observed at the centre, out towards the extreme ends or 'tails' of the distribution. However, these distributions are often chosen in a rather tautological fashion. On one hand, we need the data to show us which probability distribution will allow us to extrapolate our existing knowledge, however, we also need to choose a probability distribution, *a priori*, in order to gauge weather the data conforms to that particular distribution (Taleb 2007:2). These problems presents somewhat of a philosophical quandary, from which there is no obvious solution.

7 Conclusion

The challenge this research sought to address was to improve upon previous studies aimed at measuring the effectiveness of forest fire suppression in Ontario. Epistemological limitations aside, the only parsimonious conclusion to be made from this study, is that that the aggressive fire suppression strategy employed by Ontario's Ministry of Natural Resources–designed to limit the number of large forest fires, has indeed been effective over the period 1989–2004. This research has therefore, vindicated previous studies, which also found fire suppression to have significantly reduced the annual area burned by large wildfires, and is consistent with findings outside of Ontario, that showed fire suppression to reduce both the number and frequency of large fires in the forests of British Columbia and Alberta.

The study involved comparing the fire size distributions between several forested areas with contrasting fire management strategies, selected to control for other significant causal factors. Generalised Liner Models were used to calculate precisely the extent to which a more aggressive fire suppression strategy has effectively reduced the damage caused by wildfires over recent decades.

This new data on fire size distributions helps fire managers estimate how the forests in other fire management zones may react to changing fire suppression activities, thereby helping fire managers make more informed decisions about the province's forest management strategy. It will now also be possible to quantify the value of those resources protected by the OMNR's fire management strategy. Given that the OMNR expends a great deal of resources suppressing these wildfires, proving the fire management strategy employed, has delivered in its objectives, will go a long way to protecting those services against calls to reallocate funding and resources.

The methods used to derive these results may also be of interest to researchers using quantitative methods, as special consideration was given to ensuring the reproducibility of all calculations. The analysis was developed using the open-source R system for statistical computation and the code used in the analysis is include in the digital version of this file. These steps will therefore, allow for full replication, which is considered the gold standard in open Science. To aid with the dissemination of this resource, the R code of the project will be uploaded to the author's github repository (`https://github.com/craigrshenton/masters_thesis`) and distributed under the creative commons licence.

8 References

Armstrong, G.W. 1999. A stochastic characterisation of the natural disturbance regime of the boreal mixed–wood forest with implications for sustainable forest management. *Canadian Journal of Forest Research*, 29, 424–433.

Allen, C.D., Savage, M., Falk, D.A., Suckling, K.F., Swetnam, T.W., Schulke, T., Stacey, P.B., Morgan, P., Hoffman, M., and Klinglel, J.T. 2002. Ecological restoration of southwestern ponderosa pine ecosystems: A broad perspective. *Ecological Applications*, 12(5), 1418–1433.

Allen, C.D. 2007. Interactions Across Spatial Scales among Forest Dieback, Fire, and Erosion in Northern New Mexico Landscapes. *Ecosystems*, 10(5), 797–808.

Ashiq, M.W. 2011. Ontario boreal fire regimes in the context of lightning–caused ignition point spatial patterns. M.Sc. Thesis. Geography, University of Waterloo, Waterloo, Canada. [Online] Available at: http://uwspace.uwaterloo.ca/handle/10012 /6257. [Accessed 12 July 2012].

Bergeron, Y., Flannigan, M., Gauthier, S., Leduc, A., and Lefort, P. 2004. Past, current and future fire frequency in the Canadian boreal forest: implications for sustainable forest management. *Ambio*, 33, 356–360.

Beverly, J. L., and Martell, D. L. (2005). Characterizing extreme fire and weather events in the boreal shield ecozone of Ontario. *Agricultural and Forest Meteorology*, 133, 5–16.

Bonnicksen, T.M. 2002. Oversight Hearing on Wildfire on the National Forests. Committee on Forests and Forest Health, United States House of Representatives. [Online] Available at: http://naturalresources.house. gov/uploadedfiles/ thomas–bonnicksen–testimony–7.11.02.pdf. [Accessed 21 July 2012].

Bridge, S.R.J., Miyanishi, K. and Johnson, E.A. 2005. A Critical Evaluation of Fire Suppression Effects in the Boreal Forest of Ontario. *Forest Science*, 51(1), 41–50.

Calkin, D.E., Gebert, K.M., Jones, J.C., and Neilson, R.P. 2005. Forest Service large fire area burned and suppression expenditure trends, 1970–2002. *Journal of Forestry*, 103, 179–183.

Christensen, N.L., Agee, J.K., Brussard, P.F., Hughes, J., Knight, D.H., Minshall, G.W., Peek, J.M., Pyne, S.J., Swanson, F.J., Thomas, J.W., Wells, S., Williams S.E., and Wright, H.A. 1989. Interpreting the Yellowstone fires of 1988. *Bioscience*, 39, 678–685.

Collett, D. 1991. *Modelling binary data*. London: Chapman and Hall.

Croarkin, C. and Tobias, P. 2003. Engineering Statistics Handbook: Chapter 1.3.6.7.3 Upper Critical Values of the F Distribution. [Online] Available at: http://w ww.itl.nist.gov/div898/handbook/eda/section3/eda3673.htmONE–01–1–10. [Accessed 21 June 2012].

Cui, W. and Perera, A.H. 2008. What do we know about forest fire size distribution, and why is this knowledge useful for forest management? *International Journal of Wildland Fire*, 17, 234–244.

Cumming, S.G. 2001. A parametric model of the fire–size distribution. *Canadian Journal of Forest Research*, 31, 1297–1303.

Cumming, S.G. 2005. Effective fire suppression in boreal forests. *Canadian Journal of Forest Research*, 35(4), 772–786.

Cunningham, A.A., and Martell, D.L. 1976. The use of subjective probability assessments to predict forest fire occurrence. *Canadian Journal of Forest Research*, 6, 348–356.

DiBari, J.N. 2003. Scaling exponents and rank–size distributions as indicators of landscape character and change. *Ecological Indicators*, 3, 275–284.

Dobson, A.J. 2002. *An Introduction to Generalized Liner Models.* London: Chapman and Hall.

Donovan, G.H. and Noordijk, P. 2005. Assessing the accuracy of wildland fire situation analysis (WFSA) fire size and suppression cost estimates. *Journal of Forestry*, 105, 10–13.

Donovan, G.H., and Brown, T.C. 2007. Be careful what you wish for: The legacy of Smokey Bear. *Frontiers in Ecology and the Environment*, 5(2), 73–79.

Drushka, K. 2003. *Canada's forests: a history.* Montreal: McGill Queen's University Press.

Finney, D.J. 1973. *Statistical Methods in Bioassay.* New York: Hafner.

Flannigan, M.D., Krawchuk, M.A., de Groot, W.J., Wotton, B.M. and Gowman, L.M. 2009. Implications for changing climate for global wildland fire. *International Journal of Wildland Fire*, 18, 483–507.

Franklin, J.F. 1993. Fueling controversy and providing direction: lessons from old growth. *Journal of Forestry*, 91, 11–13.

Friedrich Leisch, F. 2002. *Dynamic generation of statistical reports using literate data analysis.* In: "Compstat 2002 – Proceedings in Computational Statistics". (Eds) Hardle, W. and Ronz, B. Heidelberg: Physica Verlag, 575–580.

Gauthier, S., Valliancourt, M.A., Kneeshaw, D., Drapeau, P., Grandpre, L.D., Claveau, Y. and Pare, D. 2009. *Forest ecosystem management: origins and foundations.* In Gauthier, S., Valliancourt, M.A., Leduc, A., Grandpre, L.D., Kneeshaw, D., Morin, H., Drapeau, P. and Bergeron, Y. (eds.), "Ecosystem management in the boreal forests". Quebec: Presses de l'Universite du Quebec.

Gentleman, R. and Lang, D.T. 2007. Statistical Analyses and Reproducible Research. *Journal of Computational and Graphical Statistics*, 16(1), 1–23.

Gill, J. 1999. The Insignificance of Null Hypothesis Significance Testing. *Political Research Quarterly*, 9, 1–32.

Hawkes, B., Vasbinder, W., and DeLong, G. 1997. Retrospective Fire Study–Fire Regimes in the SBSvk & ESSFwk2/wc3 Biogeoclimatic Units of North–Eastern British Columbia. McGregor Model Forest Network: Prince George, BC.

Haydon, D.T., Friar, J.K., Pianka, E.R. 2000. Fire–driven dynamic mosaics in the Great Victoria Desert, Australia. *Landscape Ecology,* 15, 373–382.

Heyerdahl, E.K., Brubaker, L.B., and Agee, J.K. 2001. Spatial controls of historical fire regimes: a multiscale example from the interior west, USA. *Ecology*, 82, 660–678.

Hills, G. A. (1959). A ready reference to the description of the land of Ontario and its productivity. Toronto, ON: Ontario Department of Land and Forests.

Hills, G.A. (1969). A physiographic approach to the classification of terrestrial ecosystems with respect to representative biological communities. Toronto, ON: Ontario Department of Land and Forests.

Hirsch, K.G., Corey, P.N., and Martell, D.L. 1998. Using expert judgment to model initial attack fire crew effectiveness. *Forest Science*, 44, 539–549.

Hirotugu, A. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Holling, C.S. and Meffe, G.K. 1996. Command and Control and the Pathology of Natural Resource Management. *Conservation Biology,* 10(2), 328–337.

Johnson, E.A., Miyanishi, K. and Bridge, S.R.J., 2001. Wildfire Regime in the Boreal Forest and the Idea of Suppression and Fuel Buildup. *Conservation Biology*, 15(6), 1554–1557.

Kilgore, B.M. 1976. Fire management in the natural parks: an overview. *Proceedings of Tall Timbers fire ecology.* Tallahassee: Florida State University.

Latremouille, C., Parker, W. C., McPherson, S., Pinto, F., Fox, B., and McKinnon, L. (2008). Ecology and management of eastern white pine in the lake abitibi (3E) and lake temagami (4E) ecoregions of Ontario. Science Development and Transfer Series No. 004. Sault Ste Marie, Ontario: Ontario Forest Research Institute, Ontario Ministry of Natural Resources.

Leisch, F. 2002. Sweave: Dynamic generation of statistical reports using literate data analysis. In Hardle, W. and Ronz, B. (eds), *Compstat 2002 – Proceedings in Computational Statistics*, 575–580.

Makridakis, S. and Taleb, N. 2009. Decision making and planning under low levels of predictability. *International Journal of Forecasting*, 5, 1–18.

Malamud, B.D., Morein, G., and Turcotte, D.L. 1998. Forest fires: an example of self–organized critical behavior. *Science*, 281, 1840–1842.

Martell, D.L. 1994. The impact of fire on timber supply in Ontario. *The Forestry Chronicle*, 70, 164–173.

Martell, D.L., Kourtz, P.H., Tithecott, A., and Ward, P.C. 1999. The Development and Implementation of Forest Fire Management Decision Support Systems in Ontario, Canada. USDA Forest Service General Technical Report PSW–GTR–173, 131–142.

Martell, D.L. 2001. 'Forest fire management'. In *Forest Fires: behavior and ecological effects.* E.A. Johnson and K. Miyanishi (Eds). San Diego, California: Academic Press, 572–583.

Martell, D.L. 2002. Wildfire regime in the boreal forest. *Conservation Biology*, 16(5), 1177–1178.

McCullagh, P., and Nelder, J.A. 1983. *Generalised linear models.* London: Chapman and Hall.

Merrill, D.F., and Alexander, M.E. 1987. *Glossary of forest fire management terms.* 4th Edition. Canadian Committee on Forest Fire Management, National Research Council of Canada, Ottawa.

Minnich, R.A. 1983. Fire mosaics in Southern California and Northern Baja California. *Science*, 219, 1287–1294.

Minnich, R.A., Chou, Y.H. 1997. Wildland fire patch dynamics in the chaparral of southern California and northern Baja California. International Journal of Wildland Fire, 7, 221–248.

Miyanishi, K., and Johnson, E.A. 2001. A re–examination of the effects of fire suppression in the boreal forest. *Canadian Journal of Forest Research*, 31, 1462–1466.

Miyanishi, K., Bridge, S.R.J., and Johnson, E.A. 2002. Wildfire regime in the boreal forest. *Conservation Biology*, 16(5), 1177–1178.

Moreno, J.M., Vazquez, A., and Velez, R. 1998. *Recent history of forest fires in Spain.* In 'Large Forest Fires'. Leiden: Backhuys Publishers.

Murphy, P.J., Mudd, B.J. Stocks, E.S., Kasischke, D., Barry, M.E., Alexander, N.F.H., and French, N. 2000. *Historical fire records in the North American boreal forest.* In "Fire, Climate Change, and Carbon Cycling in the Boreal Forest", *Ecological Studies*, 138, 274–288.

Nelder, J. and Wedderburn, R. 1972. *Generalized Linear Models. Journal of the Royal Statistical Society.* Series A (General), 135(3), 370–384.

Ontario Ministry of Natural Resources / Ministere des Richesses Naturelles (OMNR). 1989. *Guidelines for modifying woods operations.* Aviation, Flood, and Fire Management Branch, Ontario Ministry of Natural Resources. Toronto, Canada.

Ontario Ministry of Natural Resources / Ministere des Richesses Naturelles (OMNR). 1997. *Ontario's fire management strategies– Background, current state, and future challenges.* Toronto, Canada.

Ontario Ministry of Natural Resources / Ministere des Richesses Naturelles (OMNR). 2001. *Ontario's forests fact sheet: The boreal forests.* [Online] Available at: http://www.mnr.gov.on.ca/stdprodconsume/groups/lr/@mnr/@forests/documents/document/241209.pdf. [Accessed 12 Apr 2012].

Ontario Ministry of Natural Resources / Ministere des Richesses Naturelles (OMNR). 2007. *Ecological Land Classification Primer.* Toronto, Canada. [Online] Available at: http://www.mnr.gov.on.ca/en/Business/LUEPS/Publication/264779.html. [Accessed 12 July 2012].

Papoulis, A. 1984. "Bernoulli Trials." Chapter 3.2. In *Probability, Random Variables, and Stochastic Processes.* New York: McGraw-Hill.

Peng, R.D. 2011. Reproducible research in computational science. *Science*, 334, 1226–1227.

Perera, A. H., Baldwin, D. J. B., Schnekenburger, F., Osborne, J. E., and Bae, R. E. (1998). Forest fires in Ontario: a spatio–temporal perspective. *Forest Research Report No. 147.* Sault Ste Marie, ON. Ontario Forest Research Institute, Ontario Ministry of Natural Resources.

Perera, A. H., Cui, W., and Ouellette, M. R. (2009). Size class distribution and spatial proximity of fires in a simulated boreal forest fire regime in relation to Ontario's policy directions for emulating natural disturbance. *Forest Research Report No. 170.* Sault Ste Marie, ON: Ontario Forest Research Institute, Ontario Ministry of Natural Resources.

Podur, J., Martell, D. L., and Csillag, F. (2003). Spatial patterns of lightning–caused forest fires in ontario, 1976–1998. *Ecological Modelling*, 164, 1–20.

R Development Core Team. 2004. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

R–Studio Development Team. 2011. R–Studio Integrated Development Environment (Version 2.13.0). [Computer program] Available at: http://www.rstudio .org/. [Accessed 09 Apr 2012].

Racey, G. D., Wiltshire, R. O., and Archibald, D. J. (2000). Ecoregional forest composition analysis for northwestern Ontario: Present forest condition. *NWST Technical Report TR–123*. Thunder Bay, ON: Northwest Science and Technology Ontario Ministry of Natural Resources.

Schwab, M., Karrenbach, M. and Claerbout, J. 2000. Making scientific computations reproducible. *Computing in Science and Engineering*, 2.6, 61–67.

Snedecor, G.W. and Cochran, W.G. 1967. *Statistical Methods* (6th ed.). Iowa: Iowa State University Press.

Song. W., Fan. W., Wang. B., and Zhou. J. 2001. Self–organized criticality of forest fires in China. *Ecological Modeling*, 145, 61–68.

Stocks, B.J. 1991. "The extent and impact of forest fires in northern circumpolar countries." In, *Global biomass burning: atmospheric, climatic and biospheric implications*. Levine, J.S. (ed.). Cambridge, MA: MIT Press.

Stocks, B.J., Mason, J.A., Todd, J.B., Bosch, E.M., Wotton, B.M., and Amiro B.D. 2002. Large forest fires. *Canadian Journal of Geophysysical Research*, 107, 8149.

Strauss, D., Bednar, L., and Mees, R. 1989. Do one percent of forest fires cause ninety–nine percent of the damage? *Forest Science*, 35(2), 319–328.

Suffling, R. (1995). Can disturbance determine vegetation distribution during climate warming? A boreal test. *Journal of Biogeography*, 22, 501–508.

Taleb, N.N. 2007. Black Swans and the Domains of Statistics. *The American Statistician*, 61(3), 1–3.

Taleb, N.N. 2008. The Fourth Quadrant: A Map of the Limits of Statistics. *Edge*. [Online] Available at: http://www.edge.org/3rd–culture/taleb08/taleb08–index.html. [Accessed 22 July 2012].

Telesca, L., Amatulli, G., Lasaponara, R., Lovallo, M., and Santulli, A 2005. Time–scaling properties in forest–fire sequences observed in Gargano area (southern Italy). *Ecological Modeling*, 185, 531–544.

Turner, R. (2009). Point patterns of forest fire locations. *Environmental and Ecological Statistics*, 16, 197–223.

Van Wagner, C.E. 1978. Age–class distribution and the forest fire cycle. *Canadian Journal of Forest Research*, 8, 220–227.

Vazquez, A. and Moreno, J.M. 2001. Spatial distribution of forest fires in Sierra de Gredos (Central Spain). *Forest Ecology and Management*, 147, 55–65.

Vitek, J. and Kalibera, T. 2011. Repeatability, Reproducibility and Rigor in Systems Research. *EMSOFT '11*, 33–38.

Ward, P.C., and Tithecott, A.G. 1993. *Impact of fire management on the boreal landscape of Ontario.* Ministry of Natural Resources, Aviation, Flood and Fire Management Branch, Sault Sainte, Marie, Ontario.

Ward, P.C., Tithecott, A.G., and Wotton, B.M. 2001. Reply–A re–examination of the effects of fire suppression in the boreal forest. *Canadian Journal of Forest Research*, 31, 1467–1480.

Weber, M.G., and Stocks, B.J. 1998. *Forest fires in the boreal forests of Canada.* In, "Large forest fires." Leiden, Netherlands: Blackhuys Publishers.

Weir, J.M.H., Johnson, E.A., and Miyanishi, K. 2000. Fire frequency and the spatial age mosaic of the mixed–wood boreal forest in western Canada. *Ecological Applications*, 10, 1162–1177.

9 Appendix A

9.1 Programming Operations

With the OMNR dataset .csv file downloaded to the working directory, the file is loaded into R.

```
omnr <- read.csv("~/Dropbox/Masters/m_thesis/OMNR/OMNR_data.csv")
```

A simple operation sub–divides the dataset into fires in the Intensive (INT) and Measured (MEA) fire management zones.

```
omnr.clean <- subset(omnr, subset = omnr$FIRE_MGT_ZONE == "MEA" |
omnr$FIRE_MGT_ZONE == "INT")
```

9.1.1 Sub–setting Ecoregions

The dataset is again sub–divided into the three ecoregions (3E, 3W and 3S) using logic operations that delineate along both latitude and longitude.

```
omnr.3E <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=48 &
omnr.clean$LATITUDE<=51) & (omnr.clean$LONGITUDE<=-80 & omnr.clean$
LONGITUDE>=-86))

omnr.3W <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=48 &
omnr.clean$LATITUDE<=51) & (omnr.clean$LONGITUDE<=-88 & omnr.clean$
LONGITUDE>=-92))

omnr.3S <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=51 &
omnr.clean$LATITUDE<=53) & (omnr.clean$LONGITUDE<=-88 & omnr.clean$
LONGITUDE>=-96))
```

9.1.2 Calculating Fire Suppression Ratios

Calculating the number of suppressed and unsuppressed fires in both Measured and Intensive zones, per ecoregion.

Ecoregion 3W

Suppressed fires in ecoregion 3W, Measured zone.

```
omnr.3W.mea.s <- subset(omnr.3W, subset = (omnr.3W$FIRE_MGT_ZONE==
"MEA" & omnr.3W$FINAL_SIZE>3 & omnr.3W$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3W, Measured zone.

```
omnr.3W.mea.e <- subset(omnr.3W, subset = (omnr.3W$FIRE_MGT_ZONE==
"MEA" & omnr.3W$FINAL_SIZE>200))
```

Suppressed fires in ecoregion 3W, Intensive zone.

```
omnr.3W.int.s <- subset(omnr.3W, subset = (omnr.3W$FIRE_MGT_ZONE==
"INT" & omnr.3W$FINAL_SIZE>3 & omnr.3W$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3W, Intensive zone.

```
omnr.3W.int.e <- subset(omnr.3W, subset = (omnr.3W$FIRE_MGT_ZONE==
"INT" & omnr.3W$FINAL_SIZE>200))
```

Ecoregion 3E

Suppressed fires in ecoregion 3E, Measured zone.

```
omnr.3E.mea.s <- subset(omnr.3E, subset = (omnr.3E$FIRE_MGT_ZONE==
"MEA" & omnr.3E$FINAL_SIZE>3 & omnr.3E$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3E, Measured zone.

```
omnr.3E.mea.e <- subset(omnr.3E, subset = (omnr.3E$FIRE_MGT_ZONE==
"MEA" & omnr.3E$FINAL_SIZE>200))
```

Suppressed fires in ecoregion 3E, Intensive zone.

```
omnr.3E.int.s <- subset(omnr.3E, subset = (omnr.3E$FIRE_MGT_ZONE==
"INT" & omnr.3E$FINAL_SIZE>3 & omnr.3E$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3E, Intensive zone.

```
omnr.3E.int.e <- subset(omnr.3E, subset = (omnr.3E$FIRE_MGT_ZONE==
"INT" & omnr.3E$FINAL_SIZE>200))
```

Ecoregion 3S

Suppressed fires in ecoregion 3S, Measured zone.

```
omnr.3S.mea.s <- subset(omnr.3S, subset = (omnr.3S$FIRE_MGT_ZONE==
"MEA" & omnr.3S$FINAL_SIZE>3 & omnr.3S$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3S, Measured zone.

```
omnr.3S.mea.e <- subset(omnr.3S, subset = (omnr.3S$FIRE_MGT_ZONE==
"MEA" & omnr.3S$FINAL_SIZE>200))
```

Suppressed fires in ecoregion 3S, Intensive zone.

```
omnr.3S.int.s <- subset(omnr.3S, subset = (omnr.3S$FIRE_MGT_ZONE==
"INT" & omnr.3S$FINAL_SIZE>3 & omnr.3E$FINAL_SIZE<200))
```

Unsuppressed fires in ecoregion 3E, Intensive zone.

```
omnr.3S.int.e <- subset(omnr.3S, subset = (omnr.3S$FIRE_MGT_ZONE==
"INT" & omnr.3S$FINAL_SIZE>200))
```

9.1.3 Aggregating Suppressed and Unsuppressed Fires

Aggregating the number of suppressed and unsuppressed fires in both Measured and Intensive zones, per ecoregion.

Ecoregion 3W

```
omnr.3W.mea.Ne <- aggregate(omnr.3W.mea.e$FINAL_SIZE ~ omnr.3W.mea.
e$FIRE_YEAR, data=omnr.3W.mea.e, FUN=length)

omnr.3W.mea.Ns <- aggregate(omnr.3W.mea.s$FINAL_SIZE ~ omnr.3W.mea.
s$FIRE_YEAR, data=omnr.3W.mea.s, FUN=length)

omnr.3W.int.Ne <- aggregate(omnr.3W.int.e$FINAL_SIZE ~ omnr.3W.int.
e$FIRE_YEAR, data=omnr.3W.int.e, FUN=length)

omnr.3W.int.Ns <- aggregate(omnr.3W.int.s$FINAL_SIZE ~ omnr.3W.int.
s$FIRE_YEAR, data=omnr.3W.int.s, FUN=length)
```

Ecoregion 3E

```
omnr.3E.mea.Ne <- aggregate(omnr.3E.mea.e$FINAL_SIZE ~ omnr.3E.mea.
e$FIRE_YEAR, data=omnr.3E.mea.e, FUN=length)

omnr.3E.mea.Ns <- aggregate(omnr.3E.mea.s$FINAL_SIZE ~ omnr.3E.mea.
s$FIRE_YEAR, data=omnr.3E.mea.s, FUN=length)

omnr.3E.int.Ne <- aggregate(omnr.3E.int.e$FINAL_SIZE ~ omnr.3E.int.
e$FIRE_YEAR, data=omnr.3E.int.e, FUN=length)

omnr.3E.int.Ns <- aggregate(omnr.3E.int.s$FINAL_SIZE ~ omnr.3E.int.
s$FIRE_YEAR, data=omnr.3E.int.s, FUN=length)
```

Ecoregion 3S

```
omnr.3S.mea.Ne <- aggregate(omnr.3S.mea.e$FINAL_SIZE ~ omnr.3S.mea.
e$FIRE_YEAR, data=omnr.3S.mea.e, FUN=length)

omnr.3S.mea.Ns <- aggregate(omnr.3S.mea.s$FINAL_SIZE ~ omnr.3S.mea.
s$FIRE_YEAR, data=omnr.3S.mea.s, FUN=length)

omnr.3S.int.Ne <- aggregate(omnr.3S.int.e$FINAL_SIZE ~ omnr.3S.int.
e$FIRE_YEAR, data=omnr.3S.int.e, FUN=length)

omnr.3S.int.Ns <- aggregate(omnr.3S.int.s$FINAL_SIZE ~ omnr.3S.int.
s$FIRE_YEAR, data=omnr.3S.int.s, FUN=length)
```

9.2 Generalised Liner Models

Creating a new data frame of the number of suppressed unsuppressed fires in each ecoregion/fire management zone (data taken from the results of previous calculations).

```
omnr.glm <- data.frame(zone=rep(c("mea","int"),c(3,3)), ecoregion=
rep(c("E3","W3","S3")),suppressed=c(60,194,111,152,91,44),
unsuppressed=c(20,67,19,20,43,10))
```

Equation to be used in the generalised liner model:

```
glm.eq <- "cbind(suppressed, unsuppressed) ~ zone + ecoregion"
```

Generalised liner model with logit link function:

```
glm.out <- glm(glm.eq, family=binomial(logit), data=omnr.glm)

summary(glm.out)
```

9.3 Temporal Model

Calculating the annual load count:

```
load <- aggregate(omnr$FINAL_SIZE, by=list(omnr$FIRE_YEAR),length)

names(load) <- c("Year", "Nt")
```

Next the annual number of suppressed and unsuppressed fires, in both the Intensive and Measured zones was calculated.

```
mea.ns <- aggregate(omnr.mea.s$FINAL_SIZE, by=list(omnr.mea.s$FIRE_
YEAR),length)

mea.ne <- aggregate(omnr.mea.e$FINAL_SIZE, by=list(omnr.mea.e$FIRE_
YEAR),length)

int.ns <- aggregate(omnr.int.s$FINAL_SIZE, by=list(omnr.int.s$FIRE_
YEAR),length)

int.ne <- aggregate(omnr.int.e$FINAL_SIZE, by=list(omnr.int.e$FIRE_
YEAR),length)
```

A new data frame of the annual number of suppressed unsuppressed fires in each fire management zone, along with the annual load counts was created from the results of the previous calculations.

```
temp.glm <- data.frame(zone=rep(c("mea","int"),c(16,16)), year=rep
(1989:2004,2),suppressed=c(49,10,52,28,5,34,58,33,21,23,12,6,6,11,
10,1,137,83,12,79,42,93,199,89,73,150,83,57,101,38,77,16),
unsuppressed=c(7,0,14,2,2,4,26,27,6,4,6,0,2,6,6,0,7,5,8,7,1,1,27,
54,5,15,12,1,0,10,10,1),load=c(2429,1614,2559,960,743,1053,2122,
1245,1634,2278,1017,644,1561,1138,1036,428,2429,1614,2559,960,743,
1053,2122,1245,1634,2278,1017,644,1561,1138,1036,428))
```

Equation to be used in the generalised liner model

```
glm.eq <- "cbind(suppressed, unsuppressed) ~ zone + load"
```

Generalised liner model with logit link function

```
glm.out <- glm(glm.eq, family=binomial(logit), data=temp.glm)
```

```
summary(glm.out)
```

9.4 ANOVA test

Equation used for the Analysis of Variance test:

```
anova(glm.out, test="Chisq")
```

9.5 Results

Antilog of the zone/mea coefficient:

```
exp(-0.930)
```

```
1/exp(-1.0521)
```

10 Appendix B

10.1 Programming Operations

With the OMNR dataset .csv file downloaded to the working directory, the file is loaded into R.

```
omnr <- read.csv("~/Dropbox/Masters/m_thesis/OMNR/OMNR_data.csv")
```

Sub–divide the dataset into the Intensive (INT) and Measured (MEA) fire management zones.

```
omnr.clean <- subset(omnr, subset = omnr$FIRE_MGT_ZONE == "MEA" |
omnr$FIRE_MGT_ZONE == "INT")
```

10.1.1 Sub–setting Ecoregions

Sub–divide into the three ecoregions (3E, 3W and 3S).

```
omnr.3E <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=48 &
omnr.clean$LATITUDE<=51) & (omnr.clean$LONGITUDE<=-80 & omnr.clean$
LONGITUDE>=-86))

omnr.3W <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=48 &
omnr.clean$LATITUDE<=51) & (omnr.clean$LONGITUDE<=-88 & omnr.clean$
LONGITUDE>=-92))

omnr.3S <- subset(omnr.clean, subset = (omnr.clean$LATITUDE>=51 &
omnr.clean$LATITUDE<=53) & (omnr.clean$LONGITUDE<=-88 & omnr.clean$
LONGITUDE>=-96))

omnr.total <- rbind(omnr.3E, omnr.3W, omnr.3S)
```

10.1.2 Fire–size Distributions (a-g) in the Intensive zone

Total fires in INT.

```
int <- 4886
```

Fire size class (a).

```
a <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 0.4)

burn <- aggregate(a$FINAL_SIZE, by=list(a$FIRE_YEAR),sum)

burn <- mean(burn$x)

a <- aggregate(a$FINAL_SIZE, by=list(a$FIRE_YEAR),length)

names(a) <- c("Year", "Fires")

a <- sum(a$Fires)

ap <- (a/int)*100
```

Fire size class (b).

```
b <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 4)

burn <- aggregate(b$FINAL_SIZE, by=list(b$FIRE_YEAR),sum)

burn <- mean(burn$x)

b <- aggregate(b$FINAL_SIZE, by=list(b$FIRE_YEAR),length)

names(b) <- c("Year", "Fires")

b <- sum(b$Fires)
```

```
b <- b - a

bp <- (b/int)*100
```

Fire size class (c).

```
c <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 40)

burn <- aggregate(c$FINAL_SIZE, by=list(c$FIRE_YEAR),sum)

burn <- mean(burn$x)

c <- aggregate(c$FINAL_SIZE, by=list(c$FIRE_YEAR),length)

names(c) <- c("Year", "Fires")

c <- sum(c$Fires)

c <- c - b - a

cp <- (c/int)*100
```

Fire size class (d).

```
d <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 200)

burn <- aggregate(d$FINAL_SIZE, by=list(d$FIRE_YEAR),sum)

burn <- mean(burn$x)

d <- aggregate(d$FINAL_SIZE, by=list(d$FIRE_YEAR),length)

names(d) <- c("Year", "Fires")

d <- sum(d$Fires)
```

```
d <- d - c - b - a

dp <- (d/int)*100
```

Fire size class (e).

```
e <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 1000)

burn <- aggregate(e$FINAL_SIZE, by=list(e$FIRE_YEAR),sum)

burn <- mean(burn$x)

e <- aggregate(e$FINAL_SIZE, by=list(e$FIRE_YEAR),length)

names(e) <- c("Year", "Fires")

e <- sum(e$Fires)

e <- e - d - c - b - a

ep <- (e/int)*100
```

Fire size class (f).

```
f <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 10000)

burn <- aggregate(f$FINAL_SIZE, by=list(f$FIRE_YEAR),sum)

burn <- mean(burn$x)

f <- aggregate(f$FINAL_SIZE, by=list(f$FIRE_YEAR),length)

names(f) <- c("Year", "Fires")

f <- sum(f$Fires)
```

```
f <- f - e - d - c - b - a

fp <- (f/int)*100
```

Fire size class (g).

```
g <- subset(omnr.total, subset = omnr.total$FINAL_SIZE > 10000)

burn <- aggregate(g$FINAL_SIZE, by=list(g$FIRE_YEAR),sum)

burn <- mean(burn$x)

g <- aggregate(g$FINAL_SIZE, by=list(g$FIRE_YEAR),length)

names(g) <- c("Year", "Fires")

g <- sum(g$Fires)

gp <- (g/int)*100
```

10.1.3 Fire–size Distributions (a-g) in the Measured zone

Total fires in MEA.

```
mea <- 1643
```

Fire size class (a).

```
a <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 0.4)

burn <- aggregate(a$FINAL_SIZE, by=list(a$FIRE_YEAR),sum)

burn <- mean(burn$x)

a <- aggregate(a$FINAL_SIZE, by=list(a$FIRE_YEAR),length)
```

```
names(a) <- c("Year", "Fires")

a <- sum(a$Fires)

ap <- (a/mea)*100
```

Fire size class (b).

```
b <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 4)

burn <- aggregate(b$FINAL_SIZE, by=list(b$FIRE_YEAR),sum)

burn <- mean(burn$x)

b <- aggregate(b$FINAL_SIZE, by=list(b$FIRE_YEAR),length)

names(b) <- c("Year", "Fires")

b <- sum(b$Fires)

b <- b - a

bp <- (b/mea)*100
```

Fire size class (c).

```
c <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 40)

burn <- aggregate(c$FINAL_SIZE, by=list(c$FIRE_YEAR),sum)

burn <- mean(burn$x)

c <- aggregate(c$FINAL_SIZE, by=list(c$FIRE_YEAR),length)

names(c) <- c("Year", "Fires")
```

```
c <- sum(c$Fires)

c <- c - b - a

cp <- (c/mea)*100
```

Fire size class (d).

```
d <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 200)

burn <- aggregate(d$FINAL_SIZE, by=list(d$FIRE_YEAR),sum)

burn <- mean(burn$x)

d <- aggregate(d$FINAL_SIZE, by=list(d$FIRE_YEAR),length)

names(d) <- c("Year", "Fires")

d <- sum(d$Fires)

d <- d - c - b - a

dp <- (d/mea)*100
```

Fire size class (e).

```
e <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 1000)

burn <- aggregate(e$FINAL_SIZE, by=list(e$FIRE_YEAR),sum)

burn <- mean(burn$x)

e <- aggregate(e$FINAL_SIZE, by=list(e$FIRE_YEAR),length)

names(e) <- c("Year", "Fires")
```

```
e <- sum(e$Fires)

e <- e - d - c - b - a

ep <- (e/mea)*100
```

Fire size class (f).

```
f <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 10000)

burn <- aggregate(f$FINAL_SIZE, by=list(f$FIRE_YEAR),sum)

burn <- mean(burn$x)

f <- aggregate(f$FINAL_SIZE, by=list(f$FIRE_YEAR),length)

names(f) <- c("Year", "Fires")

f <- sum(f$Fires)

f <- f - e - d - c - b - a

fp <- (f/mea)*100
```

Fire size class (g).

```
g <- subset(omnr.total, subset = omnr.total$FINAL_SIZE > 10000)

burn <- aggregate(g$FINAL_SIZE, by=list(g$FIRE_YEAR),sum)

burn <- mean(burn$x)

g <- aggregate(g$FINAL_SIZE, by=list(g$FIRE_YEAR),length)

names(g) <- c("Year", "Fires")
```

```
g <- sum(g$Fires)

gp <- (g/mea)*100
```

10.2 Forest Fire Statistics

The number of suppressed and unsuppressed fires in both Measured and Intensive zones, per ecoregion.

Calculate the number of arrivals per year.

```
Nt <- aggregate(omnr.total$FINAL_SIZE, by=list(omnr.total$FIRE_YEAR),length)

names(Nt) <- c("Year", "Fires")
```

Calculate total number of fires.

```
N <- sum(Nt$Fires)
```

Calculate number of small (<3 ha) fires.

```
omnr.small <- subset(omnr.total, subset = omnr.total$FINAL_SIZE < 3)

small <- aggregate(omnr.small$FINAL_SIZE, by=list(omnr.small$FIRE_YEAR),length)

names(small) <- c("Year", "Fires")

small <- sum(small$Fires)

percent.small <- (small/N)*100
```

Calculate number of large (>200 ha) fires.

```
omnr.big <- subset(omnr.total, subset = omnr.total$FINAL_SIZE > 200)
```

```
big <- aggregate(omnr.big$FINAL_SIZE, by=list(omnr.big$FIRE_YEAR),length)

names(big) <- c("Year", "Fires")

big <- sum(big$Fires)

percent.big <- (big/N)*100
```

Calculate number of suppressed fires.

```
suppressed <- N-(big + small)

percent.suppressed <- (suppressed/N)*100
```

Details of the Machine used for this analysis:

```
> print(sessionInfo(), locale=FALSE)

R version 2.15.0 (2012-03-30)
Platform: x86_64-apple-darwin9.8.0/x86_64 (64-bit)

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base

loaded via a namespace (and not attached):
[1] tools_2.15.0
```