Modelling the effects of mitigation methods on forest fire spread in Ontario

GROUP 6
Belle Tuen 20779888
Gagandeep Singh 20780377
Noah Cameron 20790140
Samuel Giesbrecht 20857231

1. **Introduction**

Forest fires are an intricate part of ecosystems. A rudimentary process in which forest 'fuel' is ignited by a causation, forest fires are both a cleansing force and a dangerous one. In moderation forest fires help stabilize carbon intensive resources in natural systems, (Pettit et al., 2010) but in greater frequency can be a threat such as heightening the greenhouse effect. (Jaffe et al., 2011) On average, 2.5 million hectares of land are burned yearly in Canada because of forest fire activities since 1990 as reported by Natural Resources Canada. One study also showed that within Canada, forests that contain coniferous dense trees or forests that are within 50—150km from populated areas are more susceptible to fires. (Gralewicz et al., 2012) While there are countless literatures detailing mitigative practices and susceptibility, many do not take modeling these mitigation techniques into consideration.

As such the mitigation of forest fires has been studied in an attempt to find the most effective method of dealing with forest fires. (Glasa et al., 2006 & Encinas et al., 2006 & Zheng et al., 2016) Forest fire mitigation is any method set in place by humans aiming to reduce the total amount of area that a forest fire will spread. The spatial implications have specifically led to modelling resources that accentuate the importance of fire susceptibility mapping. In the continued search for a solution to this wildfire problem, the following research paper presents a marriage between the commonly used alleviating forest fire practices with their practical impacts. Utilizing a variety of geospatial variable criteria modelling, the intention is to derive measurable mitigation effectiveness better quantitatively for forest fires under various conditions. Moreover, the culmination of such research should be a supporting branch of forest fire mitigation, a developmental and analytical approach tackling an unbound spatial phenomenon occurring globally at increasingly greater temporal frequencies (Gheshlaghi, Hassan Abedi., 2019).

Despite the evident importance of forest fire mitigation, there is still much room for further investigation on the topic. Currently Canada has separated different forest fire prone areas into 'zones' to test and study the effectiveness in many different forest fire management practices to indicate what prevention plan the province should take for each danger zone. (Tymstra et al., 2020) We have selected to focus on one mitigation technique, fire compartments/fire breaks, and compare its effectiveness against a fire without any mitigation.

1.1 Fire Compartments and Firebreaks

Firebreaks is a term relating to a break in the forest that could potentially stop the spread of a fire, this could be rivers, roads, or deforested gaps. Fire compartments is a term relating to a section of the forest that is sectioned off entirely by firebreaks. For example, rivers and roads may create a square in the forest, if a fire were to occur in this square (fire compartment) it would not be able to spread to the rest of the forest without passing a road or river (firebreak).

Price et al. (2007), studied 623 events of wildfires in Australia where the fire met a permanent firebreak and compiled the results of these occurrences. The results were as follows. "Cliffs were more effective than streams at stopping fires, which were more effective than roads. Larger streams were more effective than small ones. The largest streams stop 75% of early dry season fires, but there are no firebreak types with more than 50% likelihood of stopping a late dry season fire." (Price et al., 2007). Plana et al. (2005), expand on this, mentioning both the benefits, and challenges associated with firebreaks, stating both how "fire prevention infrastructure (firebreaks, water points, etc.) have considerably increased the success in fighting wildfires." But also, Unfortunately, for many reasons, firebreaks may either not be wide enough to be effective under high winds, but also, can act as a corridor for wind to travel through and increase fire spread and intensity. (Plana et al., 2005). This once again shows us that this one mitigation method will not be the universal solution to forest fires everywhere. If an area is too windy, firebreaks and compartments may have the adverse of the desired effect, this will be tested in our model, and we will determine whether or not this particular mitigation method is a reliable, and viable option. Seidl et al. (2011) state "Susceptibility to fire depends on the properties of living and dead vegetation as fuel". (Seidl et al., 2011). This cements the notion that fire breaks are one of the most effective strategies for mitigation forest fire risk. The elimination of these fuels act as a barrier which the fire cannot pass.

1.2 Current Forest Fire Spread Models

Forest fire spread models are models that simulate how fire may spread throughout a forest given a starting point. From reviewing current literature on fire spread models, we found it to be true that models choose to base their experiment on historical forest fire data, and not experimental data. This data is used to determine how impactful a factor, such as wind speed, is at predicting the extent of fire spread. This could be done in a variety of ways; one method uses Random Forests (a machine learning technique) to create a predictive forest to predict the number of "bins" that fire will spread to. Then a "feature importance method" can be used to determine from the forest what the impact of each variable was (Moustakas & Davlias, 2021). This method of using "bins" to represent how far a fire will spread is common among simple models when simulating a forest fire. This idea can be found in use regardless of location, for example a stochastic fire spread model in North Patagonia (Morales et al., 2015) also uses this method. The basis of the bin concept is that fire spreads from bin to bin across an area, but this will be discussed later in the Methods section.

1.3 Research Question and Objectives

There is a large gap in the literature relating to forest fires. While there has been extensive research done on how effective mitigation methods may be, and lots of research into forest fire spread models, we could not find any models that consider how the total area burned by forest fires may differ if a mitigation technique were in place. This large gap is what we were aiming to address and potentially help fill with our research question. What is the effectiveness of firebreaks/ fire compartments at controlling the spread of forest fires in Ontario? To answer this question, we set out to [1] Create a multiple linear regression to determine the impact that numerical variables have on spread of forest fire. [2] Create susceptibility tessellations using the qualitative data acquired […] [3] Create a model that will estimate the effectiveness that firebreaks have on eliminating the spread of forest fires given an initial fire point and select geographical conditions, then comparing it to a fire given the same conditions but with no mitigation techniques applied.

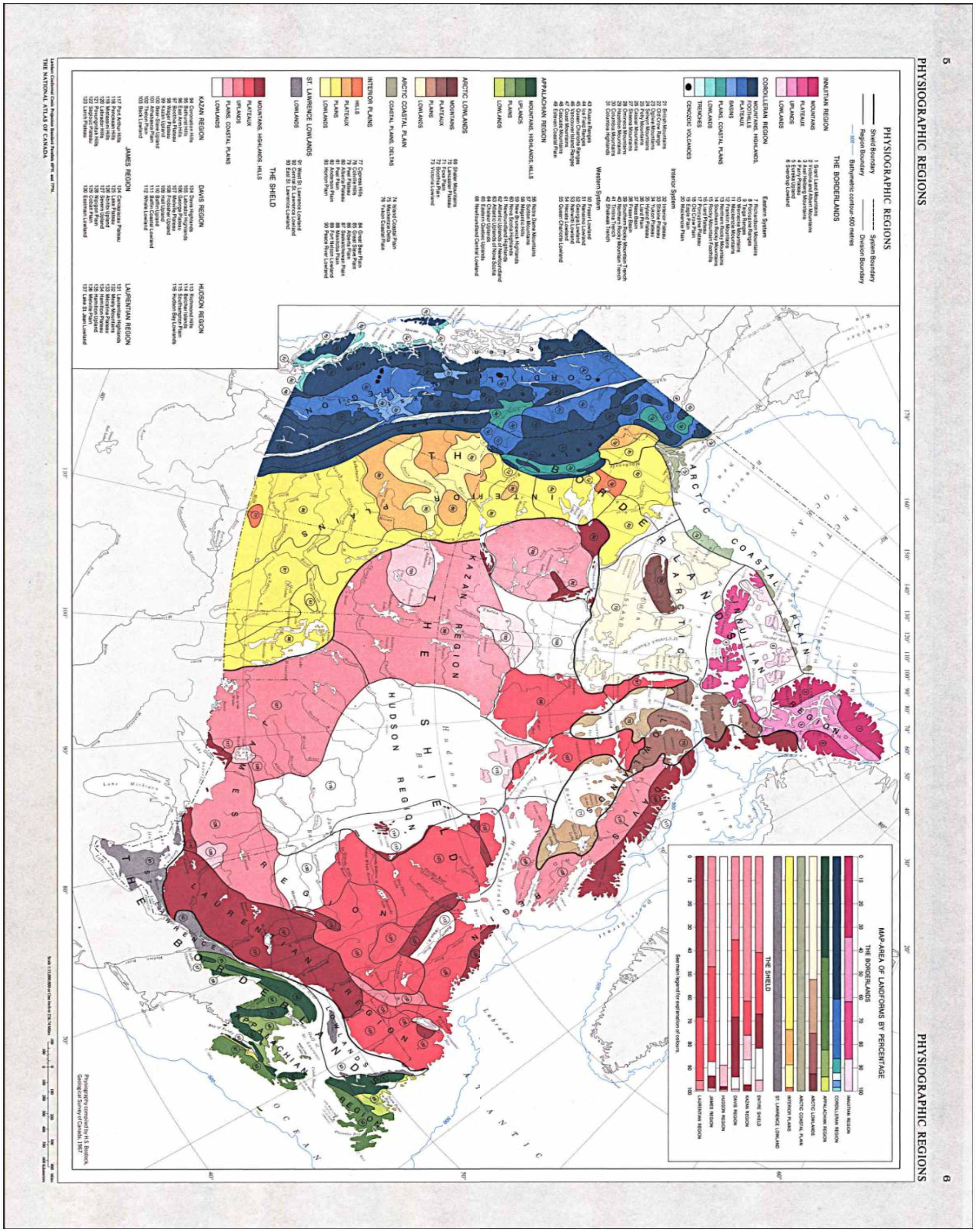

Figure 1 – Geography of Canada – Data showing the different regions of Canada, relevant to forest fire susceptibility as discussed below.

2. Methods

2.1 Study Area

The study area of choice is the province of Ontario, Canada, which comprises of diverse landscapes and environmental conditions as well as available spatial data that can be utilized for modelling forest fires. For the applicable research the limiting temporal constraints are that of 1986 to 2020. This has been selected because the metric of ‘burned area’ is recorded by the ‘Canadian National Fire Database’ for the time period stated herein. Selecting this study area allows us to start on a smaller, more refined scale in comparison to the entire country of Canada which would require further research and data collection would is always a possibility for the future. Once this model is working well with this selected area, it would be possible to outline different variables and scenarios which would allow us to expand the study area into much larger areas, like potentially across Canada.

While there is a great possibility that valuable data exists prior to 1986, that research cannot verify the quality and relevant standard to which Natural Resources Canada provides for the data referenced. Highlighted to a greater degree in the limitations, this research stands to carry relevant relational fire phenomenon over the specified time and location respectfully.

Canada has a diverse climate, with fifteen different terrestrial ecozones, all of which with very different features and characteristics (Wiken et al., 1996). much of which is very prone to forest fires. Most specifically the Mid-western areas such as eastern Alberta and Saskatchewan. which are very flat and dry. On top of this, this region also experiences some of the highest potential evapotranspiration in Canada, which is a direct correlation to dryness. (Atlas of Canada, 1974) Dryness is the primary cause, as dry conditions lead to dead plants and other fuel being more likely to ignite. The flatness also promotes high winds which cause fire to spread. Although other provinces may not be as dry, they still contain forests large enough to not only experience forest fires due to their excess fuel, and high winds, but to also have suitable data to compare with other, similarly sized forests. Although it might not be as common for the ignition to occur, it does happen naturally as an essential part of the forest’s development.

2.2 Variables

The variables we chose to use for our model are temperature, precipitation, elevation, tree cover type, and soil type cover. These were selected since they were found used by other existing fire spread models, which is visualized in figure 2. Although not all variables are used by all models, we deemed that if one model uses a variable it may be of use to analyze its impact in MLR. Secondly, we reduced our variable list to match variables for which we could find available data layers. Other variables such as soil moisture, may have been impactful on fire spread, but we do not have data for them, so they are excluded.

Table 1 Shows the variables which are used in other existing fire spread models. This was used in selecting our variables. Green squares represent that a variable is used in that model, and red squares mean that the variable was not used by the model.

Model	Elevation	Precipitation	Temperature	Tree Cover Type	Soil Cover Type	
FARSITE (Pinto et al., 2016)						
Patagonia Model (Morales et al., 2015)						
Canadian Fire Behaviour Prediction Model (Wotton et al., 2009)						
LANDIS 4.0 (Sturtevant et al., 2009), (He et al., 2005)						

2.3 Overview of Fire Modelling

Overall, our three objectives work together to create a forest fire spread model that predicts how far a fire will spread from a given starting point. Objective one, we analyze the numerical variables using MLR. This will tell us how likely fire is to spread based on variables like temperature and precipitation. In objective two, variables that were not numerical are analyzed using a weighted overlay analysis. This creates a tessellation showing the susceptibility of areas. Lastly, in objective three our model weighs both factors together and calculates the probability of spread between neighbours. This calculation is used to create an ArcGIS model.

All these objectives, while separate processes, all come together to one outcome, as shown below in figure 1. The blue outlines represent objective one, orange objective two, and green objective three. The beige-filled squares represent data.

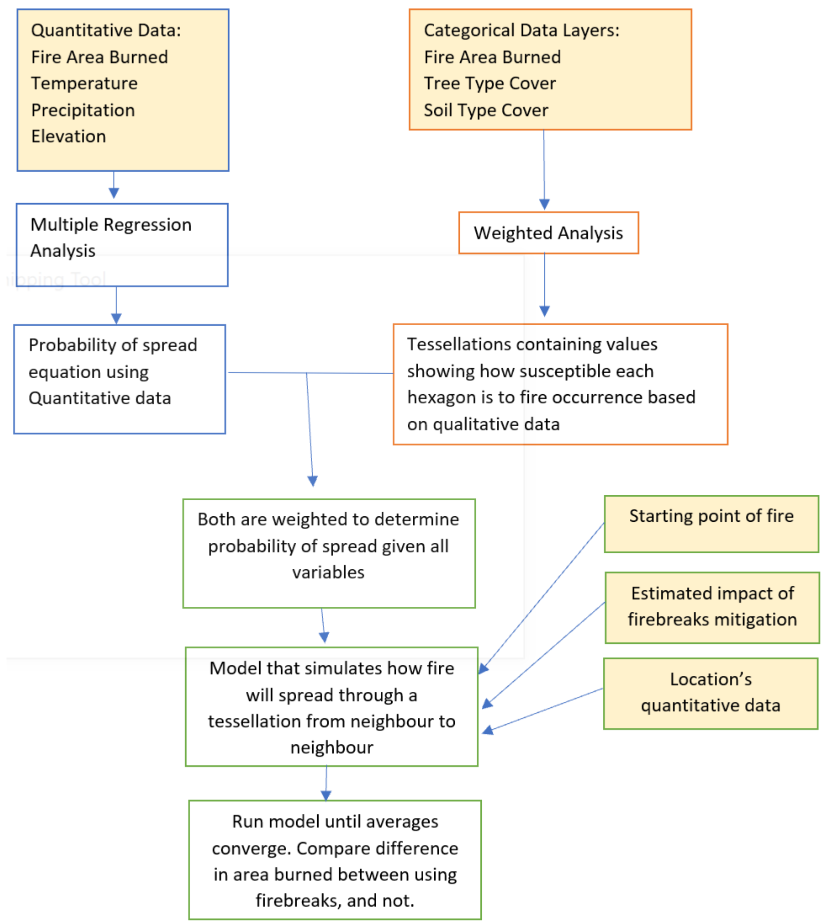

Figure 2 – Objectives Flow - Shows how the three objectives and datasets fit together. The blue outlines represent objective one, orange objective two, and green objective three.

2.4 Multiple Regression

Before creating a model of forest fire spread, we first need to determine which variables are most impactful in causing spread of fire and use these variables to create an equation outputting a value between 0 and 1 which will show the probability of fire spreading. We use MLR (multiple regression) to complete these tasks, this is the first objective.

Before running MLR, we must format the data into a table which we can run MLR on. In this table each row entry will show the data for one fire, and have columns for: area burned, elevation, precipitation, and temperature. Firstly, the forest area burned data are from Geohub Ontario and come in polygon format. We will use the calculate field tool in ArcGIS to find the amount of area burned in Km^2 of each fire. For our table we will first normalize the values so that they are between 0 and 1 (justified later). The values are normalized by dividing each area by the largest fire area recorded in the data. Secondly, the elevation data is also from Geohub Ontario and comes in the form of a DEM raster. The average elevation for each fire is calculated by running the GIS function zonal statistics as table on the fire area polygons and the elevation raster. This adds a column to the fire area polygons showing the average elevation (metres) of each fire area. Lastly, both temperature and precipitation data are collected from Environment and Climate Change Canada (ECCC) station data. These values are approximated by going to the nearest ECCC station, downloading the data between the fire start and end date (given from fire area polygons) and then calculating the average daily precipitation (mm) and temperature (Celsius).

Before performing MLR on the data we should first plot all the variables against each other and look at the correlation between them. This serves the purpose of ensuring our data is good, removing outliers, and determining if two independent variables are too correlated. For example, if we had temperature in Celsius and Fahrenheit as two variables, we would see that they are the same measure, and could remove one.

Finally, we can create an R script which will call the linear model function on our data which creates an MLR model. The main output from this will be an equation predicting the area burned by a fire (normalized) in the format:

Area Burned (between 0-1) = Intercept + x*Elevation + y*Precipitation +z*Temperature

This equation predicts the area burned by a fire, given the three independent variables. However, this equation will be used to predict the probability that fire will spread from cell to cell in our model. The reason we can use the two interchangeably is because they are both measuring the probability of fire spread, just on different scales. For example, if a fire is unlikely to spread from cell to cell, then this fire is likely to die out resulting in a smaller area burned by fire. Since they are also both measures between 0-1, we can use the same equation to predict likelihood of spread between cells in our model. It is important to note that the two values may not be the same, however, this equation makes for a strong estimate.

This use of the equation, for probability of spread, is also why we chose to normalize the values. This way the resulting value can be easily used for probability since it is already in decimal format. We believe this measure could have been more accurate if we could run the MLR to predict the percentage of forest that was burned by a fire (area burned / total forest area). However, we were not able to find forest polygon data to calculate total forest area, therefore we chose to use normalized values as an alternative.

It is also important to examine the accuracy of our MLR, we will focus on two measures. The strongest measure of accuracy is the R^2 value which will show how accurately the model is able to predict area burned from the independent variables. We will also look at the p-values for each of our independent variables. These values show the p-value for the test that a variable has a strong impact in predicting the area of a forest fire. These values range from 0-1, if this value is very low, the independent variable has a large impact on the MLR model, whereas if this value is high, it means we could remove that variable from the model without making much impact on the model’s accuracy.

2.5 Susceptibility Tessellations

Following the completion of the MLR (Multiple Linear Regression) section we can determine the impact of the decided variables on fire spread. In doing so, there are still some outstanding variables that are qualitative in nature that cannot be implemented numerically. As such, these variables are attributed to tessellations that will be used in the following Fire Model.

A tessellation in our case is a geometric lattice covering a spatial extent with a set encompassing area for each feature. As we are looking at Ontario, we build a hexagonal tessellation cover Ontario, with the limitation being the computational power of the computer generating such a grid. After testing, sizes of 100, 50 and 25 kilometer squared the device being used was able to safely generate tessellations encompassing ten squared kilometers of area for Ontario. Such a tessellation still contains roughly 90,000 polygons given the size of Ontario. An area of discussion would be how the size of a tessellated polygon and or the number of polygons impact the spread. If there are smaller polygons, meaning a greater of polygons then the influence each polygon is smaller. Larger polygons compensate more for their area and have a larger influence on the surroundings. This gives merit to deducing an ideal sized polygon that covers a reasonable area of Ontario, but also behaves similarly in terms of influence to a forest fire.

The choice of tessellation was made to be able to include multiple qualitative parameters beyond what a raster grid could provide. A raster grid of Ontario would have one defined pixel value for the pixel attributed to any spatial location in Ontario, making it difficult for the Fire Model to differentiate without the usage of normalization for the values. A polygon variant such as the hexagonal tessellation used can have many 'spatially joined' attribute fields that can be read by the Fire Model Script and factored along with the MLR equation to look at spread ability.

To normalize the values, we first have to look at the characteristics of the qualitative variables chosen. The three variables were soil type, forest / tree cover type and fire disturbance history. For soil type a characteristic that impacts forest fire spread is 'soil heat capacity', often referenced in terms of agriculture this parameter gauges flux of heat transfer between soils. The next variable's characteristic observed is the forest / tree cover type index location. This value is provided through the Ontario Geohub's dataset that coagulates forest environments together with their respective proximity to infrastructure. From that we have environments and their types with the criticality of spatial importance. Lastly, fire disturbance history and their size combine roughly 40-50 years of Canadian fire history, providing a good indicator of fire likeliness. Factoring in where fires are more likely to occur and their respective growth size from previous historical fires that have covered the same spatial extent through time.

Factoring these variables together in a 'weighted overlay analysis' creates a susceptibility index field that can be attributed to every polygon in the tessellation. Ranging from 0 – 1 the susceptibility index is 'spatially joined' by location to the polygons. Lastly, there are some parts of the tessellations that 'cannot' catch fire, or in our case are deemed not part of the burnable area. These are the water bodies and infrastructure in Ontario. To remove this area, we perform a pairwise polygon erase with the datasets available from NRCAN. Thus, we have a final tessellation environment that contains variables not factored in the MLR and can spread fire visually.

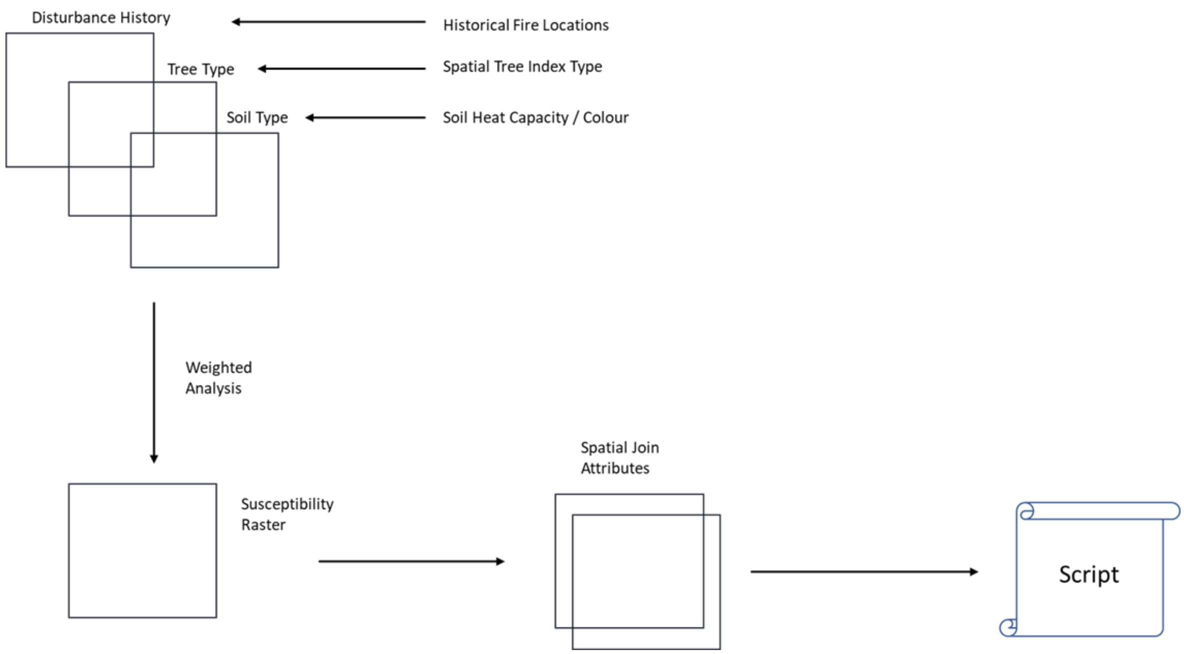

Figure 3 – Shows the Geoprocessing steps involved in preparing the tessellation for the script

For the purposes of the fire model runs, we have also broken the Ontario tessellations into smaller forest sections based on the inventory polygons also provided from the Ontario Geohub. This means we can test fire spread on largely forested areas, sections with high amount of water bodies and eventually for the mitigation, sections that contain firebreaks (road polyline attribution).

2.6 Forest Fire Model GIS Scripting

The base of this model uses the previous two objectives to calculate the chance of a burnt area to spread to the next. Using ArcGIS and Arcpy, we created a script tool that takes in temperature, precipitation, elevation, initiation point and burn time so that different variations can be used when running the tool. The first three variable input can then be put into the MLR formula to create a base percentage. The script will then create a copy of the susceptibility tessellation and select the starting hexagon from the tessellation. From there we start running a nearest neighbour approach where the script runs through all the neighbours that area touching a burnt hexagon. Each hexagon has a susceptibility value from objective two, so it factors in the susceptibility percentage with the base percentage from the MLR. This gives the likelihood of the hexagon burning. The script will then generate a random number and if the number is smaller or equal to the chance of it burning, it will set the hexagon to burning and recurse through its neighbours again.

In order to factor in mitigation techniques, the susceptibility tessellation factored in all hexagons that had roads that ran across it. This represented road breaks when the fire is spreading. Finally, to compare the results of the two models, the tool will have to run multiple times to get a definite result due to the randomness of the model. Using the same variables and same starting location, it is possible to obtain different result every time. It is decided that the tool would run until the variation between the runs are no more than five km^2 in area burnt.

3. Results

3.1 Multiple Regression

After collecting the forest fire data and organizing it for MLR we had a table of twenty-one fires, we visualize our data in figure 3. This figure shows the final data we used, the variables are listed on the diagonal with their graphs on the bottom, and correlation coefficients (r-value) on the top. Graphs/r-values comparing two variables can be found by following one variable along the row, and the other down the column and finding their intersection. The data does not have any problems visually: the correlation between independent variables is not too high, there are not any clear outlying points, and there does appear to be correlation between AreaNorm and the other variables. However, this correlation between AreaNorm and the other variables is not as high as we had expected, no |r| value exceeds 0.31. In previous tests of our data this visualization of data helped to identify and remove outliers, which are not shown here as they have been removed.

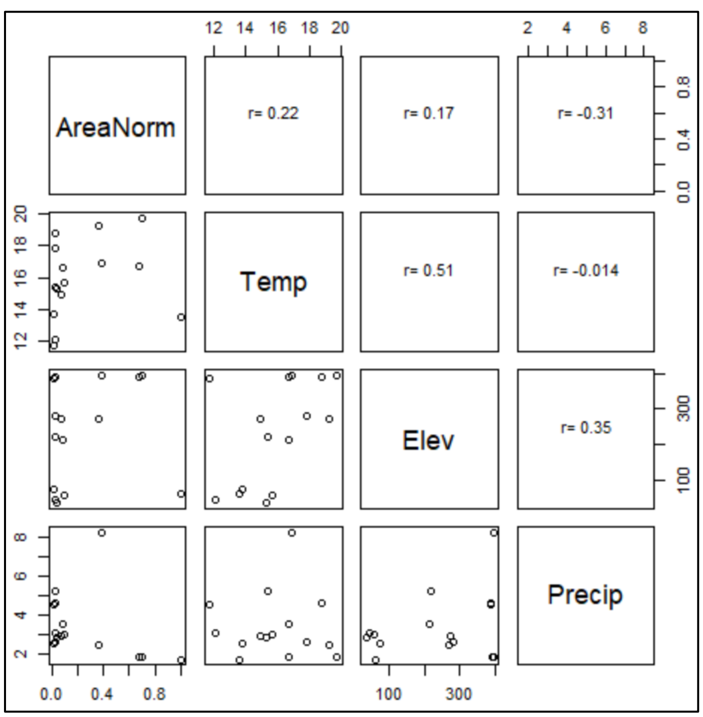

Figure 4 – Plots showing all the variables graphed against each other, as well as their correlation coefficient values.

```
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.197072   0.654628   0.301    0.769
Temp         0.010101   0.042240   0.239    0.815
Elev         0.000592   0.000766   0.773    0.456
Precip      -0.075466   0.056097  -1.345    0.206
```

Figure 5 – Summary statistics from the MLR model.

After fitting our MLR model to our data we get the results shown in figure 4, giving us the equation:

Area Burned (0-1) = 0.197 + 0.010*Temperature + 0.001*Elevation – 0.075*Precipitation

This equation is what we expected, higher temperatures, and less precipitation promote more fire spread. We also got an R^2 value of 0.186 and an adjusted R^2 of -0.036. Although this R^2 is not as high as we hoped, it is not terrible. We will go over how we could have increased this value in the discussion section. The negative R^2 value is very concerning at first, however from looking into it, this is a normal case that implies the sample size of the data is not large enough, relative to the number of independent variables, or when some of the model's variables are not helping to predict the response variable (Gouda & El-Hoshy, 2020).

Another crucial factor to analyze is the Pr(>|t|) column in figure 4, these are the p-values discussed in the methods section. From these values, we can see that precipitation is the most important variable with a value of 0.206, whereas temperature is the least important with a value of 0.815.

3.2 Forest Fire GIS Model Results

Both models with no mitigation technique and road break technique was run multiple times until the variation between the mean of the last five runs were below 0.5 hexagons, which is equivalent to 5 km^2.

Table 2 First 5 results of the no mitigation technique model

RUN NUMBER	**1**	**2**	**3**	**4**	**5**
	2	2	2	2	2
		26	26	26	26
			1	1	1
				2	2
					152
AVERAGE	2	14	9.666667	7.75	36.6
STANDARD DEVIATION	0	6	4.961581	4.31099	11.93818

Table 2 above shows the first five runs of the model with no mitigation techniques applied with the average of each run being the mean of all the runs up till the current one. The standard deviation here is the standard deviation of the current run mean and the last four run means to show the variation between the means. For example, by run five the standard deviation between the means is 11.93818. This number is very high, and it means that the averages are still very spread apart and has not converged to a similar number yet. The model is then continuously run and table 3 below is the last five runs of the model with no mitigation.

Table 3 Last 5 results of the no mitigation technique model

RUN NUMBER	34	35	36	37	38
AVERAGE	45.05882	45.34286	45.27778	44.91892	44.05263
STANDARD DEVIATION	0.818124	0.649302	0.570904	0.507326	0.464305

By run thirty-eight, the standard deviation fell below 0.5 hexagons meaning the last 5 means are very similar and the variability between them do not go beyond 5 km^2 since each hexagon is 10 km^2. In figure 6 below, the chart is a visual representation of the means after each run on the model.

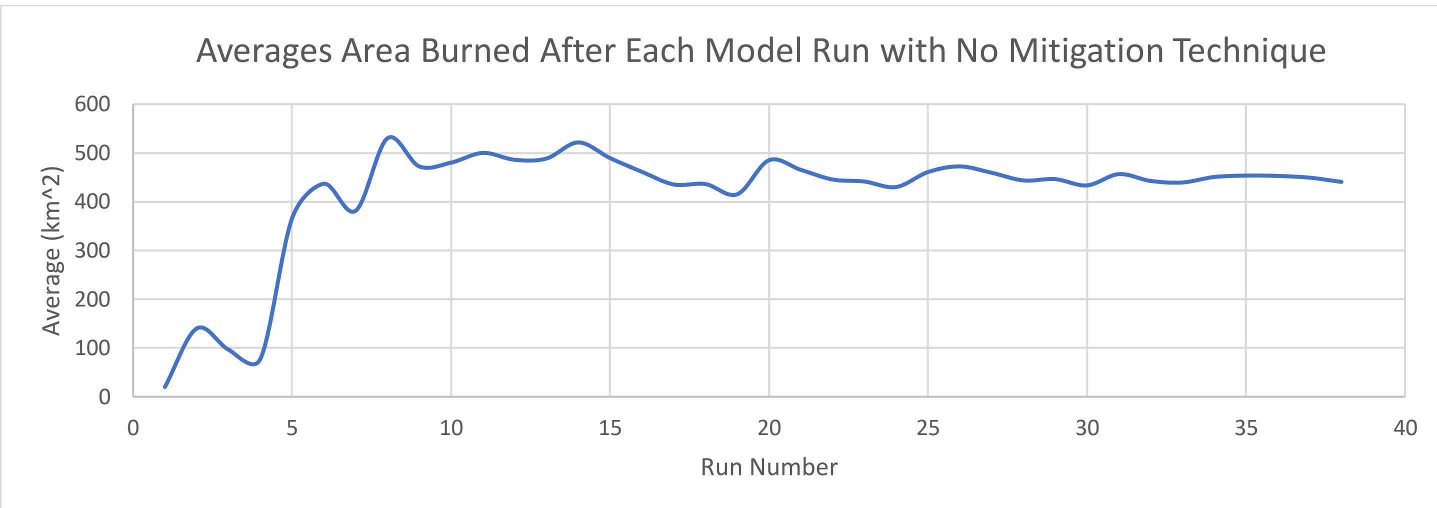

Figure 6 – Line Graph of the Average Area burnt in relation to the run number.

It is visible in figure 6 above that in the beginning there was a lot of variability between the numbers since the mean after each run ranged from 5 km^2 to 520 km^2. After the 32nd run the line starts to become straighter and has converged into a smaller range of numbers. This shows the deviation between the means at that point is not very big and by run 38 the means are so close that it is less than 5km^2 between the averages and the mean of the averages.

With the no mitigation model completed, we need a road breaks model to compare the results to. Using the same concept, we ran the model until the standard deviation of the last five averages is less than 0.5 hexagons.

Table 4 First 5 results of the road breaks model

RUN NUMBER	11	12	13	14	15
AVERAGE	9.363636	8.666667	8.769231	9	8.933333
STANDARD DEVIATION	0.910651	0.663358	0.75618	0.57982	0.267754

From Table 3 and Table 4 we can acquire the average of the runs when the standard deviation is lower than 0.5. This would be run 38 for no mitigation and run 15 for the model with mitigation. To determine the effectiveness of the road breaks mitigation technique, we first need to convert the average hexagon data to average area burnt. Each hexagon is 10 km^2 meaning the average area burnt is the average hexagon x 10. From there we can find the relative size of the smaller area in terms of the bigger area by dividing the smaller area by the bigger area which is 8.93333/44.05263 = 0.202788. By doing 1 – 0.202788, we can get the percentage of area that the mitigation technique reduced the burnt area by which is around 80%.

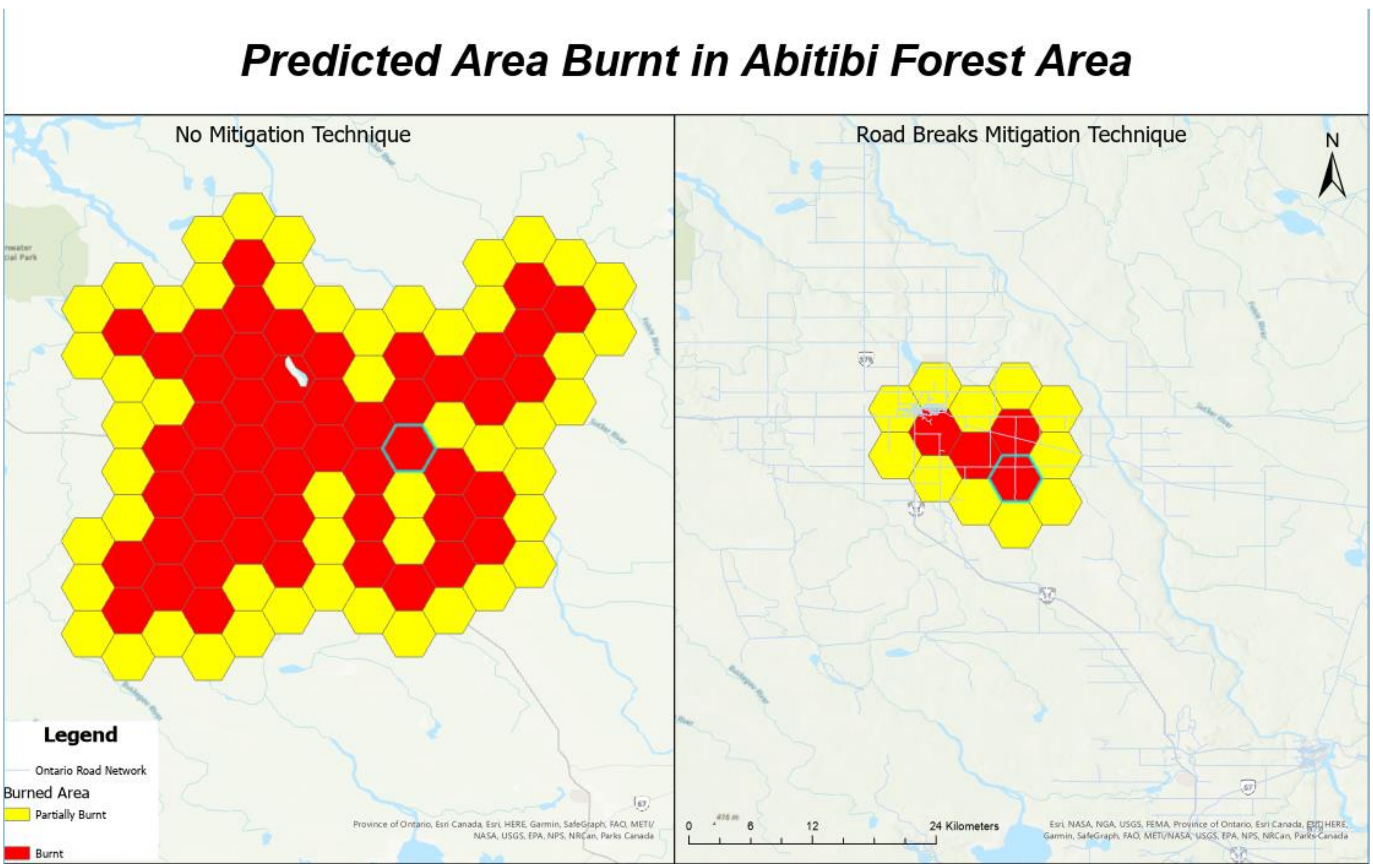

Figure 7 – Map of fire spread with no mitigation technique compared to road break technique

The maps above, figure 7, show one of the runs that were taken when running the scripting tool on the road breaks and no mitigation models. The area is the Abitibi Forest in Ontario, and the tool ran with the same initiation point signified by the highlighted blue polygon. It is visually obvious that the map with the road breaks do significantly better than the map with no mitigation technique applied

4. Discussion

4.1 Broader Implications

The broader implications of our research are to start small in Ontario and leave the possibility available to expand and to create a tool that can be used throughout the vast, and exceedingly broad landscapes that come with that. The influence of the tool we are creating could be vaster than we are hoping for. Of course, we all know that wildfires are not a Canada only occurrence. The southern United States experiences some of the worst wildfires in the world, with approximately 3 million hectares burned annually from 2012 to 2021 (Hoover, K. Hanson, L. 2022). Due to the geographical similarities with the study area and the United States, simply due to the proximity, there is potential to implement the finished tool in the United States, and eventually all around the world. This tool will be useful as there are not any models that use the same principles as ours out there. We believe that using the mitigation strategies and applying it to different forest fires with vastly differing compositions, soil types, and landscapes, we can create a tool which can accurately estimate a forest fire's spread, and then determine how effective the use of our researched mitigation method is. Whether it be firebreaks, or potentially comparing the effectiveness of multiple mitigation methods in the future, the potential of this research is far greater than what we were able to do.

4.2 Improving Multiple Regression Results

In the results of MLR the R^2 of 0.186 was extremely low, and there are varying ways in which we could go about increasing this value. The most obvious would be to increase the sample size from our small 21 fire sample. The negative adjusted R^2 value also points at this being the largest issue. The data is also available making it very possible to increase the sample size, however this process was very slow as we manually were collecting the precipitation and temperature values from the ECCC station data website. One solution to this could be to mass download (although it is very large) select stations data for all years, and then use a script to filter the data to the closest station to the fire, the start and end dates of the fire, and calculate the average precipitation and temperature.

A second way we could have increased the R^2 value was by increasing the diversity of the points we did select. From examining figure 3, we can see from the left column graphs, that many of our fires are about the same size in area. It is difficult to point out a trend predicting area when all the areas are similar. The same can be said about other factors; most of the data points in our sample came from 2016 and beyond because this time frame had temperature and precipitation values for almost all ECCC stations. Also, most of our data came from the middle of summer because this is when most fires occur. Had we searched harder to find a diverse set of data over many years, more diverse months, more diverse size, it is possible we would have seen a more defined trend and increased R^2, however a decreased R^2 is also possible.

Lastly, we could have included more independent variables. Increasing the number of variables will never decrease the R^2 value. This is because in the worst case-scenario if a variable you add is useless, it will be multiplied by zero in the MLR equation, leaving you with the same equation you already had. Because of this adding more variables will almost always increase your R^2 value. (Cobb, C., 2018) There were other variables that other models considered which we excluded due to not finding data for them. This is not to say that the data does not exist however, for example, humidity was another variable which we had anticipated useful due to its use in other models, such as FARSITE (Pinto et al., 2016). Humidity was also available in the ECCC data. However, very few stations recorded humidity data, so we dropped it from our variable list. If we had focused on finding fires which also had humidity measurements, we would have had a higher R^2.

It is also important to note that increasing R^2 is not the ‘at all costs’ goal. We could hand pick points to make our R^2 very strong but this doesn’t mean the model is any better. For example, removing one outlier from our data decreased our R^2 from 0.26 to 0.18. This is not because the model got worse, its because that one outlier was so far away from all the other points (over fifteen times the size of the second largest fire), that its impact on the model was too strong (also due to the small sample size). Although the R^2 decreased, the model is more accurate.

One other aspect of the MLR results which is interesting is the impact of the variables. Our p-value for precipitation indicated that it was important to the spread of fires. This makes sense as dry conditions are known to promote fire spread (Morales, 2015). However, high temperatures have little impact on our model. This is interesting because lab studies show that fires are directly spread by particle heating (Cohen, 2010) which intuitively would be easier under hotter conditions, and other models such as FARSITE use temperature as a variable (Pinto et al., 2016). So why would temperature not appear important in our model? One reason could be that the temperature is more impactful in the ignition of a fire and less impactful to the spread. This would agree with our data as during cold months fires are not occurring, thus why we do not have data for cold weather. The second reason could be that it is our lack of data causing this effect. Like the R^2, maybe if we had a larger collection of fires or just a larger variation in temperature, we would have seen a clearer trend and a higher impact from temperature.

4.3 Improving the GIS Tessellation Results

When constructing the GIS tessellations there were various elements that could be adjusted to receive better justification of results. Of which the size of the tessellations was a contentious topic, with each polygon representing an area of 10 kilometers squared. Forest fires in terms of size hardly spread to the size that is generated in the Fire Model. As such the polygon size maybe too large or too influential on the neighbouring ‘cells’. If a polygon is considered burnt it’s total 10 kilometers squared area is added to the burned area, leaving little room for error. If the polygons were representing a smaller area, in the grand area covered by Ontario, 1 km squared or even less would have a minimal statistical deviation. With this consideration the room for error in terms of burned polygons above expected is the difference of 10-20 squared kilometers of area. Very substantial compared to a plus/minus of 1-2 squared kilometers in an ideal environment (or less given the statistical significance of the maximum allowable confidence interval used).

Another facet that can be explored is the analysis type used to calculate the susceptibility index field for the tessellated polygons. As there were three qualitative variables, the significance in weighted analysis was distributed equally. This was done under the assumption these variables would have identical impacts when paired together but was not tested compared to the MLR equation. As such, the true impact that these variables have on the susceptibility is not fully explored and would be an avenue to increase the realistic nature of fire spread. Since the variables are qualitative in nature, different metrics of the variables would first have to explored to find the most impactful. These characteristics would then have to be studied in relation to each other before a weighted analysis can be performed, so that they are not improperly represented.

The last improvement would be the application of the mitigation method onto each polygon of the tessellation. Currently, if a road polyline (natural break method) intersects a polygon than an identifier is attached that there is a mitigation method applied. Moreover, when the modelling script is run it checks if there is mitigation signifier, then reduces the likelihood of spread by half. This is another assumption that is made, based on the effectiveness of a natural break. Though, the purpose of the paper is to deduce effectiveness of mitigation in reducing burned area, the impact of a mitigation is not fully understood. Moreover, the frequency of intersections in our model does not impact the reduced spread-ability. An example would be a polygon with a small road nicking its boundary still has the reduced signifier applied compared to a polygon with various roads running through. Due to this inconsistency a more detailed approach needs to be taken to apply the reduction in spread-ability signifier to polygons in the tessellation.

Cumulatively, the GIS tessellation is largely constructed from assumptions that were made about the qualitative variables as well as the spatial relationship between polygons. Various elements for influenceability, mitigative implementation and statistical confidence were overlooked to create a working spread environment. Of which, the variables and actual behaviour may not constitute accurately to how a real-life fire may spread. As such, the GIS tessellation can benefit from greater research into its mechanics as well as a greater understanding of the quantitative relationships at play during the process.

4.4 Improving the Forest Fire GIS Scripting Results

In our proposal, we planned to include a few more variables in our research. These variables being wind speed and slope. The reason these data were left out was that these variables require more work to incorporate due to their non-linear nature. Our other data were numerical; however, these two variables require direction as well. If we were to incorporate these variables, this is an outline on how it would be done. First, we would take the centroid of the currently affected polygon and then find the centroid of an adjacent polygon relative to the original polygon. Then we would calculate the angle between them (some value 0-360), which we plug into the equation (Cos((w-n))+1)/2 (Morales et al., 2015), in which w represents the direction the wind is travelling (also 0-360), and n represents the direction between the two polygons. This gives us a value between 0 and 1 which can be weighted into our model (along with the windspeed) as another factor similar to our MLR equation and susceptibility raster. An example of this calculation works is visualized in figure 8 below.

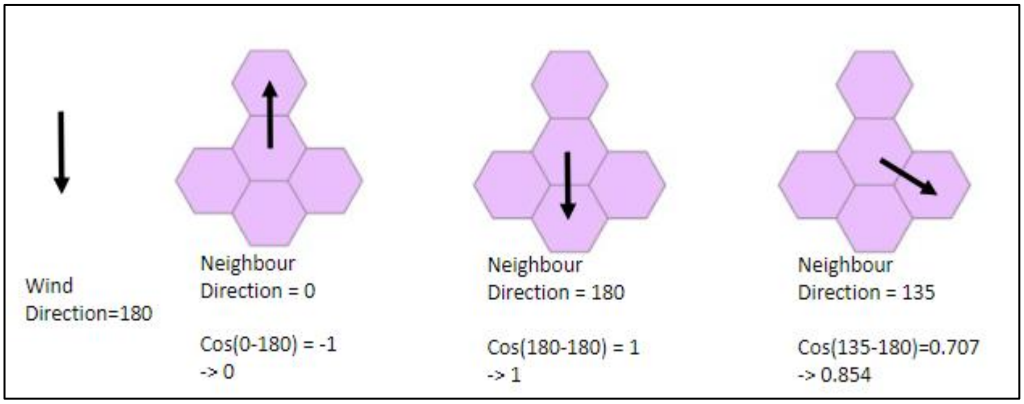

Figure 8 – Example of how the equation would convert to a value between 0-1 for windspeed. When Windspeed is opposite neighbour direction, the result is 0, when it is the same the result is 1, and for in between values the result is relative to the angle between the neighbour direction and wind direction.

Another option to improve the script is to compare the accuracy of the model to real life examples. This can be done using the invariable-variable method of comparing each polygon with a real-life case and running the model many times. (Brown et al., 2005) By using this method and calculating the ratio of correct and incorrect polygons, it is possible to generate a value that dictates whether the model is accurate or precise. (Brown et al., 2005)

4.5 Limitations

Firstly, our model will make assumptions based on lack of data. Our current model is running with an unjustified estimate for how effective firebreaks are. This is because although we did wish to run the MLR on fires using firebreaks, and fires not using firebreaks, this data is difficult to find. The fire datasets we did investigate using, such as Fire Disturbance Areas, and Large Fire Database both had a column representing the mitigation efforts used. However, these columns are very vague. They have listings such as full mitigation effort, monitoring, or partial mitigation. These were not in detail enough to tell which mitigation techniques were applied. Additionally, there is research existing related to fire breaks, but from our literature review we found that this was primarily based on small fires (Agueda, 2008) which is not suitable for what our model wishes to achieve. Our model currently leaves the door open that given a future study that shows the effectiveness of firebreaks, this value could be plugged into our model.

We will also have to make other assumptions to account for the lack of data. For example, many models assume that fuel on the forest floor is evenly distributed (Cohen, 2010), however, this is often not the case. We are also assuming in our model that the quantitative values are inputs for the entire fire. This means that the temperature, elevation, and precipitation are all constant across the potential fire area.

Also, our study's focus was on only how effective our one mitigation method, firebreaks, was at limiting the spread of forest fires. There are other variables that may have been of interest that we chose to ignore. These include the cost of mitigation techniques, difficulty to implement mitigation techniques, and environmental impacts. Something else that we had originally planned to investigate but subsequently chose to ignore were the different mitigation methods for which there is an abundance of research on.

The results from our model are also limited by the fact that the model is currently being run without its accuracy compared to other fires. Initially, the goal was to complete the model and run it at the exact starting location of a historical fire. Then we could compare our fire to the actual fire and tune our variables to be more accurate. However, this objective was not realistic in our given timeframe. For future use of this model, this accuracy should be analyzed.

5. Conclusions

The potential for expanding this work to even greater extents is a very real possibility. Through the work that was completed, we created a basis of approach to determine the effectiveness of mitigation on forest fires. In doing so, we have allowed ourselves the opportunity to continue this research further. This is possible through the broadening of our research area to apply to much greater areas. This could be done by slowly increasing the size, to allow for this research to be done in new places. If we were to decide to expand the study area to be all of Canada for instance, this would entail some more work. By expanding across Canada, we are introducing a lot of new variables that were not present in Ontario. Different soil types, tree types, and many others would impact the accuracy of our results were they not accounted for. Another direction this could be taken in is to not expand the area studied, but to study our original area in more detail.

Overall, he results from MLR show that there is correlation between fire spread and temperature, elevation, and precipitation. The MLR can predict the area burned, however the accuracy could be greatly increased with more data, a larger variation in data, and more variables. The fire spread model weighs both our MLR, and susceptibility analysis to predict forest area burned by a fire. The results from running the model demonstrate how the inclusion of fire breaks can reduce the area burned, with our results showing that fire breaks decrease the area burned by 80%. The model provides a strong structure for a fire spread model that accounts for fire breaks. However, moving forward the model could be strengthened through more research into the effectiveness of fire breaks, the inclusion of directional variables such as windspeed and slope, and smaller cell size of tessellations used by the model.

As such the narrative on such a combustible conclusion is derived by the successful nature of implementation. Perhaps overly rhetoric and simplified, the mitigation explored provides tangible effects in the combat against spatial spread. The potential avenue for discussed elements highlights the critical nature for the fight against frequently increasing forest fires.

Appendices

Appendix A: Group Contributions

Table A1: Description of group member contributions to final paper and project

Section	Group Member	Contribution	Specific Tasks
Title			
	Noah Cameron	50%	Proposed title ideas
	Belle Tuen	50%	Proposed title ideas
Intro			
	Gagandeep Singh	20%	Wrote pre-face for forest fires, mitigation, and intentions.
	Noah Cameron	35%	Transition to research question. Contributed to creation of research question and objectives. Added definitions of terms recommended in peer review.
	Sam Giesbrecht	35%	Wrote majority of intro paragraphs, edited the rest from proposal to fit our final objectives and results
	Belle Tuen	10%	Contributed to creation of research question and objectives.
Methods			
	Noah Cameron	33%	Multiple Regression, variables
	Belle Tuen	33%	Scripting
	Gagandeep Singh	33%	Susceptibility tessellations
Results			
	Belle Tuen	50%	Ran model for no mitigation, wrote script results and final mitigation percentage result
	Noah Cameron	50%	MLR results, running model for fire breaks
Discussion			
	Noah Cameron	25%	Improving MLR results, limitations of data availability, limitations of model accuracy, part of wind/slope
	Belle Tuen	25%	Wrote the improvement of the script model
	Sam Giesbrecht	25%	Wrote broader implications, limitations, part of wind/slope
	Gagandeep Singh	25%	
Conclusions			
	Sam Giesbrecht	65%	Wrote conclusion paragraph detailing future direction

	Noah Cameron	35%	Summary of findings
Appendices			
	Noah Cameron	5%	Set up table
	Everyone	95%	8 references each

References

Àgueda, A., Pastor, E., & Planas, E. (2008). Different scales for studying the effectiveness of long-term forest fire retardants. Progress in Energy and Combustion Science, 34(6), 782–796. https://doi.org/10.1016/j.pecs.2008.06.001

Brown, D. G., Page, S., Riolo, R., Zellner, M., & Rand, W. (2005). Path dependence and the validation of agent-based spatial models of Land Use. *International Journal of Geographical Information Science*, *19*(2), 153–174. https://doi.org/10.1080/13658810410001713399

Cobb, C. (2018). Re: Why does R squared increase with the inclusion of an interaction term?. Retrieved from: https://www.researchgate.net/post/Why_does_R_squared_increase_with_the_inclusion_of_an_interaction_term/5bd9fc8ec7d8ab96c859e6e4/citation/download.

Cohen, J. D., & Finney, M. A. (2010, January 1). An examination of fuel particle heating during fire spread. In: Viegas, D. X., ed. Proceedings of the VI International Conference on Forest Fire Research; 15-2018 November 2010; Coimbra, Portugal. Coimbra, Portugal: University of Coimbra. 13 p. Retrieved May 31, 2022, from https://srs.fs.usda.gov/pubs/39350

Encinas L. Hernández, White S. Hoya, del Rey A. Martín, Sánchez G. Rodríguez. (2007) Modelling Forest fire spread using hexagonal cellular automata. 31(6) 1213-1227. https://www.sciencedirect.com/science/article/pii/S0307904X06000916

Evans, A., Rodriguez, M., & Krasilovsky, E. (2017). Controlled Burning on Private Land in New Mexico. Forest Stewards Guild. Retrieved May 31, 2022, from https://www.foreststewardsguild.org/wp-content/uploads/2019/05/ControlledBurningNMweb.pdf.

Gheshlaghi, Hassan Abedi. (2019) Using GIS to Develop a Model for Forest Fire Risk Mapping 47 1173-1185. https://link.springer.com/article/10.1007/s12524-019-00981-z

Glasa, J., & Halada, L. (2008). On elliptical model for forest fire spread modeling and Simulation. Mathematics and Computers in Simulation, 78(1), 76–88. https://doi.org/10.1016/j.matcom.2007.06.001

Gouda, O. E., & El-Hoshy, S. H. (2020). Diagnostic technique for analysing the internal faults within power transformers based on sweep frequency response using adjusted R-square methodology. IET Science, Measurement & Technology, 14(10), 1057–1068. https://doi.org/10.1049/iet-smt.2020.0048

Gralewicz, N. J., Nelson, T. A., & Wulder, M. A. (2012). Factors influencing national scale wildfire susceptibility in Canada. Forest Ecology and Management, 265, 20–29. https://doi.org/10.1016/j.foreco.2011.10.031

He, H. S., Sturtevant, B. R., Yang, J., Shang, B. Z., Gustafson, E. J., & Mladenoff, D. J. (n.d.). Landis 4.0 Users Guide - US Forest Service. Retrieved July 26, 2022, from https://www.nrs.fs.fed.us/pubs/gtr/gtr_nc263.pdf

Hoover, K., & Hanson, L. A. (2022). Wildfire Statistics. Congressional Research Service. Retrieved June 1, 2022, from https://sgp.fas.org/crs/misc/IF10244.pdf.

Jaffe, D. A., & Wigder, N. L. Ozone production from wildfires: A critical review, Atmospheric Environment, Volume 51, 2012, Pages 1-10, ISSN 1352-2310, https://doi.org/10.1016/j.atmosenv.2011.11.063. (https://www.sciencedirect.com/science/article/pii/S1352231011012507)

Kalabokidis K., Athanasis N., Gagliardi F., Karayiannis F., Palaiologou P., Parastatidis S., Vasilakos C. (2013) Virtual Fire: A web-based GIS platform for forest fire control. 16 62-69. https://www.sciencedirect.com/science/article/pii/S1574954113000332?casa_token=Hzli1GZBm-wAAAAA:7egKW4NJG4VHJw_RPYYLWbu3J5sHUYNuCtiv4jKxTkYEQgMDyxQqNFyesUYIQ55HvMQCNbTPM68

Morales, J. M., Mermoz, M., Gowda, J. H., & Kitzberger, T. (2015). A stochastic fire spread model for North Patagonia based on fire occurrence maps. Ecological Modelling, 300, 73–80. https://doi.org/10.1016/j.ecolmodel.2015.01.004

Moustakas, A., & Davlias, O. (2021). Minimal effect of prescribed burning on fire spread rate and intensity in savanna ecosystems. Stochastic Environmental Research and Risk Assessment, 35(4), 849–860. https://doi.org/10.1007/s00477-021-01977-3

Pettit, N. E., Bayliss, P., Davies, P. M., Hamilton, S. K., Warfe, D. M., Bunn, S. E., & Douglas, M. M. (2010). Seasonal contrasts in carbon resources and ecological processes on a tropical floodplain. Freshwater Biology, 56(6), 1047–1064. https://doi.org/10.1111/j.1365-2427.2010.02544.x

Pinto, R. M. S., Benali, A., Sá, A. C., Fernandes, P. M., Soares, P. M., Cardoso, R. M., Trigo, R. M., & Pereira, J. M. (2016). Probabilistic fire spread forecast as a management tool in an operational setting. SpringerPlus, 5(1). https://doi.org/10.1186/s40064-016-2842-9

Plana, E., Cerdan, R., & Castellnou, M. (2005). Developing firebreaks. In Forest Restoration in Landscapes (pp. 269-273). Springer, New York, NY

Price Owen F., Edwards Andrew C., Russell-Smith Jeremy (2007) Efficacy of permanent firebreaks and aerial prescribed burning in western Arnhem Land, Northern Territory, Australia. International Journal of Wildland Fire 16, 295-307

Savage, David W., et al. "Evaluation of two risk mitigation strategies for dealing with fire-related uncertainty in timber supply modelling." Canadian Journal of Forest Research, vol. 40, no. 6, June 2010, pp. 1136+. Gale Academic OneFile, link.gale.com/apps/doc/A231419173/AONE?u=uniwater&sid=bookmark-AONE&xid=6b9f9044. Accessed 3 June 2022.

Seidl, R., Fernandes, P. M., Fonseca, T. F., Gillet, F., Jönsson, A. M., Merganičová, K., Netherer, S., Arpaci, A., Bontemps, J.-D., Bugmann, H., González-Olabarria, J. R., Lasch, P., Meredieu, C., Moreira, F., Schelhaas, M.-J., & Mohren, F. (2011). Modelling natural disturbances in Forest Ecosystems: A Review. Ecological Modelling, 222(4), 903–924. https://doi.org/10.1016/j.ecolmodel.2010.09.040

Sturtevant, B.R., Miranda, B.R., Yang, J. et al. Studying Fire Mitigation Strategies in Multi-Ownership Landscapes: Balancing the Management of Fire-Dependent Ecosystems and Fire Risk. Ecosystems 12, 445 (2009). https://doi.org/10.1007/s10021-009-9234-8

Suryabhagavan K. V., Alemu Misrak & Balakrishnan (2016) GIS-based multi-criteria decision analysis for forest fire susceptibility mapping: a case study in Harenna forest, southwestern Ethiopia. 57(1) 33-43. http://216.10.241.130/pdf/open/PDF_57_1/4%20Alemu,%20Suryabhagavan%20&%20Balakrishinan_F_.pdf

Trending.com. (2017, April 1). The road that stopped the fire. Trending. Retrieved June 3, 2022, from https://trending.com/posts/2017-04-01/the-road-that-stopped-the-fire

Tymstra, C., Stocks, B. J., Cai, X., & Flannigan, M. D. (2020). Wildfire Management in Canada: Review, Challenges and Opportunities. Progress in Disaster Science, 5, 100045. https://doi.org/10.1016/j.pdisas.2019.100045

Wotton, B. M., Alexander, M. E., & Taylor, S. W., Updates and revisions to the 1992 Canadian forest fire behavior prediction system (2009). Sault Ste. Marie, Ont.; Great Lakes Forestry Centre.

Wiken, E. B., Marshall, I., Lawton, K., & Hirvonen, H. (1996). A perspective on Canada's ecosystems: An overview of the terrestrial and marine ecozones. Canadian Council on Ecological Areas.

Yassemi, S., Dragicevic S., Schmidt M. (2008) Design and implementation of an integrated GIS-based cellular automata model to characterize forest fire behaviour. 210 (1&2) 71-84. https://www.sciencedirect.com/science/article/abs/pii/S0304380007003626?casa_token=MKcW2btLBLgAAAAA:JbHkhnvWFUpcvflWvI7Er9UNb0gNljvU9rCAvHQT0IwamhEpj9bzsKeoR5znYbiXb3B0E0h92a0

Zheng, Z., Huang, W., Li, S., & Zeng, Y. (2017). Forest fire spread simulating model using cellular automaton with Extreme Learning Machine. Ecological Modelling, 348, 33–43. https://doi.org/10.1016/j.ecolmodel.2016.12.022