DETAILS OF ASSIGNMENT			
STUDENT NAME	**Alex Eisner**	**SWINBURNE ID NUMBER**	**100600170**
EMAIL ADDRESS	**100600170@student.swin.edu.au**	**PHONE CONTACT**	**0477990302**
UNIT CODE & NAME	**STA80010 Industrial/Research Project**		
ASSESSMENT TITLE	**Analysis of factors impacting the likelihood of fatal and serious injury crashes in Victoria**		
TUTOR'S NAME:		**DATE OF SUBMISSION:**	**22/11/2020**

.

DECLARATION
I declare that (the first four boxes must be completed for the assignment to be accepted): ☒ This assignment does not contain any material that has previously been submitted for assessment at this or any other university. ☒ This is an original piece of work and no part has been completed by any other student than signed below. ☒ I have read and understood the avoiding plagiarism guidelines at http://www.swinburne.edu.au/ltas/plagiarism/students.htm and no part of this work has been copied or paraphrased **from any other source** except where this has been clearly acknowledged in the body of the assignment and included in the reference list. ☒ I have retained a copy of this assignment in the event of it becoming lost or damaged. ☒ (optional) I agree to a copy of the assignment being retained as an exemplar for future students (subject to identifying details being removed).

Student acknowledgement (by typing your name you agree to the above):	ALEX EISNER	**Date:**	20/11/2020

DETAILS OF FEEDBACK

Abstract

Background: This study examined fatal and serious injury crashes (FSI crashes) in non-rural Victoria. Variables representing three domains were examined. The three domains are atmospheric factors, such as light condition and weather. Vehicle factors, such as vehicle type and movement. Road segment factors, such as road geometry and speed zone. The aim of the current study is to identify if any variables are associated with an FSI crash being more likely to occur than a minor injury crash.

Methods: The data used in the study was obtained from the CrashStats website. The sample included Victorian road crashes which occurred between 2009 and 2019. Crashes occurring in regional areas, involving motorcycles, or resulting in no injury were excluded. The current study involved a descriptive analysis and variable selection. Random forest, gradient boosting, LASSO regression and stepwise logistic regression were used for variable selection.

Results: For the atmospheric factors crashes which occurred in strong wind were the most likely to be associated with an FSI crash, when compared to crashes occurring with no wind. For the road segment factors crashes occurring in 90km/h speed zones were the most likely to be associated with an FSI crash, when compared to crashes occurring in 60km/h speed zones. For the vehicle factors cars parked illegally were the most likely to be associated with an FSI crash, when compared to cars travelling straight ahead.

Conclusion: Variables which increase the probability of an FSI crash occurring were identified for all three domains. These results can be used by future policy makers to improve road safety across Victoria.

The peak of Australian road fatalities was in the early 1970s, in which there were approximately 3500 fatalities on Australian roads each year. Since then the number of road fatalities has been decreasing (Langford & Newstead, 2008). The reduction of road fatalities is due to research which has identified numerous factors that may cause fatal or serious injury crashes, otherwise known as an FSI crash. In Victoria, a fatal crash is one in which at least one person involved in the crash is killed or dies within 30 days of the accident. A crash is classified as a serious injury crash when, at least one person involved in the crash is sent to hospital. Minor injury crashes may also occur. A crash is classified as a minor injury crash when, at least one person involved in the crash required medical treatment at the scene (CrashStats, 2013). Past research has examined many factors which may contribute towards an FSI crash. This research has focused on many areas. The first area of interest in the current study is atmospheric factors, such as light condition and weather. The second area of interest is vehicle factors, such as vehicle type and movement. The third area of interest is road segment factors, such as road geometry and speed zone.

Road segment factors have been suggested to be important when it comes to the rates of FSI crashes. When looking into the past research of road segment factors, intersections have been suggested to have an impact on the severity of a crash. It has been suggested that multiple vehicle crashes are more severe when they occur at a cross intersection, as opposed to a T intersection (Tay & Rifaat, 2007). It has also been suggested that when a cyclist is involved in a crash, it is more likely to be an FSI crash if it occurs at a T intersection (Boufous, de Rome, Senserrick & Ivers, 2012). Road segment speed is another road segment factor, which may impact if a crash outcome is an FSI. It has been suggested by Brodie, Lyndal and Elias (2009) that fatal crashes involving heavy vehicles occurs about 83% of the time when the speed limit of the road is between 100km/h and 110km/h. A similar trend has been identified for crashes involving younger drivers. This trend is that more fatalities are likely to occur when the speed limit is higher (Chen et al., 2010). However, the opposite has been suggested for older drivers (Thompson, Baldock & Dutschke, 2018) and bicycle riders (O'Hern & Oxley, 2018). For these two groups it has been suggested that more severe crashes are likely to occur when the speed limit is less than 60km/h. Another important road segment factor that has been identified, which may increase the severity of a crash is straight sections of road. It has been indicated that heavy vehicle fatal crashes occur one third of the time on straight sections of road (Brodie et al., 2009). It has also been suggested that drivers aged over 75 years old are more likely to crash in an urban environment, when the road section is straight (Thompson et al., 2018).

Atmospheric factors have also been suggested to influence FSI crash occurrence. Light condition is a factor which may contribute to the severity of a crash. A study examining the severity of bicycle-motor vehicle crashes suggested the following. When compared to day light, crashes occurring during dawn and at night on roads that were either lit or unlit were more severe (Asgarzadeh, Verma, Chrisiani, Fisher & Courtney, 2018). It was also suggested in another study that the potential of an FSI crash increases when it is dark with streetlights on (Jung, Qin & Noyce, 2010). This result was found for vehicle crashes which occurred in clear weather. Windspeed is another factor, which may increase the severity of a crash. It has been suggested that injuries resulting from truck crashes increase in severity when recorded windspeed at the time and location of the crash increases (Naik, Tung, Zhao & Khattak, 2016). The results of this study also suggest that truck crashes are more severe when it is raining. Looking further into crashes occurring in rain. It has been suggested that moderate rainfall intensity increases the likelihood of severe injuries, resulting from multiple vehicle crashes (Jung, jang, Yoon & Kang, 2014).

Vehicle factors are another domain of interest, when examining road crash severity. Results of a study examining single vehicle accidents suggest that the probability of severe injury crashes are related to two types of vehicle movement. These movements are when a vehicle is travelling straight ahead or overtaking another vehicle (Dissanayake & Roy, 2014). Wrong way driving is another type of vehicle movement, which has been identified as a contributor to the severity of road crashes. These results suggest that an accident where a car involved was driving the wrong way is likely to have severe injuries. (Pour-Rouholamin, & Zhou, 2016). When comparing single vehicle and multiple vehicle crashes, it has been suggested that trucks are less likely to be involved in single vehicle crashes when compared to cars (Martensen & Dupont, 2013). It has been identified that approximately 26% of fatal bicycle crashes between 2006 and 2015 in Australia occurred in Victoria. Of these crashes approximately two thirds involved a collision between a cyclist and a car or truck (O'Hern & Oxley, 2018). This indicates that when looking at bicycle crash fatalities it is important to study the vehicle types involved in the crashes.

Road policies and awareness campaigns targeting changes in road infrastructure, vehicle technology and driver behaviour have made vast use from the results of the past research of factors which may contribute to road fatalities. One example of this is the arrive alive campaign, which targeted road safety in Victoria. The aim of this campaign was to reduce the number of FSI crashes. It did so by implementing policy changes and campaigns to try and reduce the number of road fatalities. Examples of this are changes to the drink driving legislation, implementation of random drug testing and the wipe off 5 campaign (Transport Accident Commission). For these campaigns and policies to be successful the factors which contribute to FSI crashes need to be identified, so that they can be targeted. As these factors are constantly evolving research needs to continue to identity them. This leads to the purpose of the current study.

The aim of the current study is to identify if any variables are associated with an FSI crash being more likely to occur than a minor injury crash. To address this, the following research questions have been developed. What effect does atmospheric factors have on the probability of FSI crashes occurring, when compared to minor injury crashes? What effect does road segment factors have on the probability of an FSI crashes occurring, when compared to minor injury crashes? What effect does vehicle factors have on the probability of an FSI crashes occurring, when compared to minor injury crashes?

Method

Data

The data used in the study was obtained from the CrashStats website. CrashStats is part of the Victorian governments open data platform. There are 12 downloadable datasets on the CrashStats website. These datasets contain Victorian crash data from 2006 to 2019. The datasets used in the current study were accident, accident event, accident location, atmospheric condition, node, person, road_surface_cond and vehicle.

The outcome variable used in the current study was a dichotomous variable called “FSI or minor injury crash”. The severity variable, which can be found in the accident dataset was used to combine fatal and serious injury crashes into an FSI group. For the outcome variable this was labelled 1. The severity level other injury was used to create the minor injury group. This was labelled 0 for the outcome variable. All the predictor variables used for each analysis can be found in Table 1 of the Appendix. In total there was 148 predictor variables used at the start of the study.

Design

The current study involved a descriptive analysis, which consisted of obtaining summary statistics for each of the predictor variables of interest. These summary statistics were obtained separately for the FSI crashes and the minor injury crashes. Next variable selection was performed to obtain a final model. The variable selection process involved the use of random forest, gradient boosting, LASSO regression and stepwise logistic regression.

Important variables from the random forest analysis were identified using the MeanDecreaseGini index. For the gradient boosting the adaptive algorithm was used. Stepwise logistic regression was performed using the variables identified as important by the random forest, gradient boosting and LASSO regression. This analysis was performed to remove any remaining variables that were not important. It was also performed to identify the variables which significantly increased or decreased the probability of an FSI crash occurring, when compared to a minor injury crash. In the stepwise logistic regression, the reference category was minor injury crashes. All statistical analyses were undertaken using R software. Rattle, an add on to R, was used for the random forest and gradient boosting.

Subjects

Each individual crash from the CrashStats data was a subject in the descriptive study. There were some exclusion criteria for the sampling of these subjects. All crashes which occurred before 01/01/2009 and after 31/12/2019 were excluded. Crashes which occurred in a region labelled as rural Victoria were excluded. All crashes where a motorcyclist was involved were excluded. Lasty any crash classified as non-injury was excluded, as the outcome variable is only interested in FSI and minor injury crashes.

For the random forest and boosting analyses a random sample of the minor injury crashes was taken, so that the dataset would be balanced in terms of the number of crashes in the two groups (FSI or Minor Injury). The random sample of the minor injury crashes was checked to confirm that it had the same distribution of cases as the full sample for each of the predictor variables. The data was then transformed so that each subject in the following analysis were the drivers of the vehicles involved in the crashes. This was done, as it is common to use the drivers as subjects in vehicle crash studies. This is because they are considered to be at fault when it comes to the cause of the crash. For the LASSO regression and stepwise regression, only the crashes which involved two vehicles and two drivers were included.

Results

A descriptive analysis was undertaken to provide a profile of the crashes included in the current study. The results of the descriptive analysis suggest the following. Between 01/01/2009 and 31/12/2019 There was 31772 FSI crashes and 70615 minor injury crashes. The largest proportion of these crashes occurred in the region of Melbourne urban. The number of crashes occurring in Melbourne urban was suggested to be 80.80% of the FSI crashes and 81.01% of the minor injury crashes. 90.91% of the FSI crashes and 91.4% of the minor injury crashes involved up to two drivers. 60.25% of the drivers involved in FSI crashes were male and 36.17% were female. For the minor injury crashes 55.16% of the drivers were male and 40.09% were female.

Looking at vehicle factors the descriptive results suggest the following. For both the FSI and minor injury crashes the largest proportion of the accident type was a collision with another vehicle. This was 60.19% of the FSI crashes and 77.31% of the minor injury crashes. Furthermore, 50% of the FSI and 66% of the minor injury crashes involved two vehicles. This was the largest proportion for both groups. Of the vehicles involved in FSI crashes 86.76% were cars, 7.29% were bicycles, 2.82% were trucks and 3.12% were other or unknown. Of the vehicles involved in minor injury crashes 86.95% were cars, 7.85% were bicycles, 2.04% were trucks and 3.16% were other or unknown.

The results of the descriptive analysis suggested that for road segment factors 97.42% of FSI crashes and 97.20% of minor injury crashes occurred on paved roads. 39.50% of the FSI crashes and 40.61% of the minor injury crashes occurred in a 60km/h speed zone. This was the largest proportion for both groups. The largest proportion of crashes for both groups occurred when the surface of the road was dry. This was 80.64% of the FSI crashes and 77.21% for minor injury crashes.

The results indicated for atmospheric factors that the majority of FSI and minor injury crashes occurred when it was clear. This was 82.85% and 77.96% respectively. It was also indicated that the largest proportion of crashes for both groups occurred during the day. This was 60.91% of the FSI crashes and 67.65% of the minor injury crashes.

The results of the random forest suggested that the 10 most important variables for classifying FSI crashes are, from most to least important: Vehicle_colour_1.A, Age.A, Postcode.No, Vehicel_colour_1.B, Day_of_week, Vehicle_movement_A, Age.B, Speed_zone, Vehicle_year_manufacture_B and Driver_intent.A. The AUC for the test data was 0.61 for the random forest model. The results from the gradient boosting model suggested that the 10 most important variables are, from most to least important: Driver_intent_A, Accident type, Vehicle_movement.A, Traffic_control.A, Vehicle_type.A_recoded, DCA_CAT, Postcode.No, Light_Condition, Surface_snowy, and deg.urban.name. The AUC for the test data for this model was 0.67. Both the random forest and the boosting analyses results suggested variables that related to drivers and vehicles C through to M were not important when classifying the crashes. The AUC for the random forest and gradient boosting were both below 0.70. This means that the study results will not allow for predictions of FSI crashes to be made. Although the model cannot be used to make predictions, it can still be used to identify the most important factors associated with FSI crashes. The results from the LASSO regression suggested that the variables road_surface_type.A and Vehicle_year_manuf.A were not important for classifying FSI crashes.

The results from the stepwise logistic regression suggest that there are 63 variables which significantly increase or decrease the likelihood of an FSI crash occurring, when compared to minor injury crashes. 24 of these are related to atmospheric, road segment and vehicle factors and suggest that the odds of a FSI crash are more likely when compared to minor injury crashes. Table 2 of the Appendix shows all the significant predictor variables in the stepwise logistic regression model.

The following results were identified for weather conditions. Compared to no wind, crashes occurring in strong wind are the most likely to be associated with FSI crashes. Compared to not raining, crashes occurring when it is raining are the next most likely to be associated with FSI crashes. Crashes occurring in clear weather, when compared to not clear weather is the next weather condition associated with FSI crashes. Lastly, compared to no fog, crashes occurring in the fog are associated with FSI crashes. The following results were found for light condition. Compared to crashes occurring in daylight, crashes occurring in the dark with no streetlights on are more likely to be associated with FSI crashes, followed by crashes occurring in the dark with streetlights on. These results can be found in Table 1.

Table 1: *Atmospheric factors, which significantly increase the likelihood of FSI crash*

Reference Category	Variable Name	Variable Description	Odds Ratio	95%CI	P-Value
Light condition Day	LIGHT_CONDITION3	Dark streetlights on	1.33	1.25 and 1.42	<.001
Light condition Day	LIGHT_CONDITION5	Dark no streetlights	1.50	1.24 and 1.83	<.001
Weather not clear	ATMOSPHERIC_CLEAR	The weather is clear	1.68	1.49 and 1.90	<.001
No rain	ATMOSPHERIC_RAINING	The weather is raining	1.82	1.54 and 2.15	<.001
No fog	ATMOSPHERIC_FOG	The weather is foggy	1.49	1.02 and 2.17	0.038
No wind	ATMOSPHERIC_STRONG.WIND	There are strong winds	2.19	1.61 and 2.98	<.001

The following results were identified for road segment factors. Compared to 60km/h speed zones, crashes occurring in 90km/h speed zones are the most likely to be associated with FSI crashes, followed by 100km/h speed zones, 80km/h speed zones and 70km/h speed zones. Compared to cross intersections, crashes occurring not at intersections were the most likely to be associated with FSI crashes, followed by crashes occurring at T-intersections. Crashes occurring when the road surface was wet are more likely to be an FSI crashes, then crashes occurring when the road surface is not dry. Lastly, crashes occurring on paved roads are more likely to be an FSI crash, when compared to crashes occurring on paved roads. These results can be found in table 2.

Table 2: *Road segment factors, which significantly increase the likelihood of FSI crash*

Reference Category	Variable Name	Variable Description	Odds Ratio	95%CI	P-Value
60km/h speed zone	SPEED_ZONE70	70km/h speed zone	1.23	1.14 and 1.32	<.001
60km/h speed zone	SPEED_ZONE80	80km/h speed zone	1.50	1.40 and 1.60	<.001
60km/h speed zone	SPEED_ZONE90	90km/h speed zone	2.11	1.37 and 3.29	<.001
60km/h speed zone	SPEED_ZONE100	100km/h speed zone	1.68	1.51 and 1.88	<.001
Cross intersection	ROAD_GEOMETRY2	T intersection	1.08	1.02 and 1.15	0.014
Cross intersection	ROAD_GEOMETRY5	Not at intersection	1.14	1.05 and 1.22	<.001
Surface not dry	SURFACE_DRY	Surface condition dry	1.27	1.09 and 1.48	0.002
Road surface paved	ROAD_SURFACE_TYPE.3	Road surface gravel	1.27	1.06 and 1.51	0.008

The crashes examined in the final model were two vehicle crashes. The vehicles in these crashes are labelled A or B. There is no specific ordering as to which vehicle is labelled A or B. So, when it comes to the interpretation of these results there is no need to distinguish between them. The vehicle factor results suggest the following. Compared to vehicles traveling straight ahead, vehicles parked illegally are the most likely to be associated with FSI crashes, followed by vehicles traveling the wrong way, vehicles traveling out of control and stationary vehicles. Compared to cars, bicycles are the most likely vehicle to be involved in an FSI crash, followed by other vehicle (such as agriculture machinery or train) and trucks. Lastly, green vehicles are more likely to be involved in an FSI crash compared to black vehicles. These results can be found in table 3.

Table 3: *Vehicle factors, , which significantly increase the likelihood of FSI crash*

Reference Category	Variable Name	Variable Description	Odds Ratio	95%CI	P-Value
Vehicle A travelling straight	VEHICLE_MOVEMENT.A12	Vehicle A parked illegally	2.24	1.03 and 5.09	0.046
Vehicle A travelling straight	VEHICLE_MOVEMENT.A18	Vehicle A moving out of control	1.71	1.51 and 1.94	<.001
Vehicle A travelling straight	VEHICLE_MOVEMENT.A19	Vehicle A traveling wrong way	2.14	1.41 and 3.35	<.001
Vehicle B travelling straight	VEHICLE_MOVEMENT.B15	Vehicle B movement is stationary (other)	1.19	1.08 and 1.31	<.001
Vehicle B travelling straight	VEHICLE_MOVEMENT.B18	Vehicle B moving out of control	1.31	1.09 and 1.57	0.004
Vehicle A colour black	VEHICLE_COLOUR_1.AGRN	Vehicle A primary colour is green	1.25	1.12 and 1.41	<.001
Vehicle A type car	VEHICLE_TYPE.A_RECODED2	Vehicle A type truck	1.46	1.26 and 1.69	<.001
Vehicle B type car	VEHICLE_TYPE.B_RECODED2	Vehicle B type truck	1.26	1.10 and 1.44	<.001
Vehicle B type car	VEHICLE_TYPE.B_RECODED4	Vehicle B type bicycle	1.59	1.44 and 1.75	<.001
Vehicle B type car	VEHICLE_TYPE.B_RECODED5	Vehicle B type other	1.51	1.06 and 2.15	0.024

Note: vehicle type other is parked trailer, train, agriculture machinery etc.

Discussion

The aim of the current study was to identify if any variables are associated with an FSI crash being more likely to occur than a minor injury crash. The results suggested that for all three domains there are variables which significantly increase the likelihood of an FSI crash occurring, when compared to the minor injury crashes.

One of the domains explored in the current study was atmospheric factors. The current results suggested that an FSI crash is more likely to occur when the crash occurs in the dark with streetlights either on or off. This result supports the findings of a study conducted by Asgarzadeh, Verma, Chrisiani, Fisher and Courtney (2018). The results from this study suggested that that crashes are more severe at night or dawn, when compared to crashes occurring during the day. The results of a study conducted by Tay and Rifaat (2007) were also supported by the current results. This study suggested that crashes are more severe at night. Asgrazadeh et al. interpreted such findings as supporting the theory that limited visibility resulting from the night-time environment results in more crashes and crashes that are more severe. The results of the current study may further support this theory. The results indicated that FSI crashes are more likely to occur in the rain. This does not support the findings of Edwards (1998) study. This study suggested that the number of FSI crashes was lower for crashes occurring in the rain. It also does not support the result that there is a negative correlation between rainfall and the number of injury crashes (Bergel-Hayat, Debbarh, Antoniou & Yannis, 2013). It does however support the result that multiple vehicle crashes are more severe when there is moderate rainfall at the time of the crash (Jung, jang, Yoon & Kang, 2014). The result that truck crashes are more severe in the rain may also be supported (Naik, Tung, Zhao & Khattak, 2016). As the results of the current study related to rain supports and opposes the results of past research, it may be necessary to further study this topic. The current result that there is an increased likelihood of an FSI crash when there are strong winds supports the findings that truck crashes are more severe in strong winds (Naik et al., 2016). This may suggest a need to target strong wind environments in Victoria, to reduce FSI crashes.

The second domain explored was road segment factors. The results indicated that when compared to a speed zone of 60km/h, speed zones of 70km/h, 80 km/h, 90km/h and 100km/h had a higher probability of an FSI crash occurring. This result supports Brodie, Lyndal and Elias (2009) study which suggested that fatal crashes involving heavy vehicles are most likely to occur in high speed zones. The current results also support the suggestion that fatal crashes involving younger drivers are more likely to occur when the speed limit is higher (Chen et al., 2010). From their results Chen et al. suggested that there should be a review into the high posted speed zones, aiming to reduce fatal crashes. The current results support this need to review high posted speed zones. When intersections, or road geometry is considered the results of the current study suggested that an FSI crash is more likely to occur when a crash occurs not at an intersection. This was compared to crashes occurring at cross intersections. It was also suggested by the results that an FSI crash is more likely to occur when the crash occurs at a T intersection. This result does not support the findings of Tay and Rifaat (2007) study which suggest that crashes are more severe at cross-intersection compared to T intersections. Looking into the impact that road condition has on the likelihood of an FSI crash occurring the results suggested that an FSI crash is more likely to occur when the road is dry. This result may support previous findings which suggests that the severity of run off road crashes increases when the crash occurs on a road that is dry (Dissanayake & Roy, 2014). To expand on this further it has been suggested that considerably more runoff road crashes are likely to occur when the road is dry (Palamara, Broughton & Fraser, 2009). This may indicate a need to further study the reason for the increase in severity of dry road crashes indicated by the current results. The current results also indicated that an FSI crashes is more likely if a car involved was travelling on a gravel road. Liu and Dissanayake (2009) suggested factors which may contribute to the likelihood of FSI crashes occurring on gravel roads. Two important factors that were identified are higher speed limits on gravel roads and failing to give way at intersections on gravel roads. These may suggest areas that should be targeted to reduce gravel road FSI crashes in Victoria.

The third domain explored in the current study involved vehicle factors. The results suggested that when a vehicle is travelling the wrong way, an FSI crash is more likely to occur. This result supports previous research which suggests that wrong way crashes are likely to result in more severe injuries (Pour-Rouholamin & Zhou, 2016). This research suggested that there are many factors which contribute to wrong way driving. Some of these factors are driving under the influence, driving at night and non-appropriate lighting conditions. These authors identified some changes to road infrastructure which would aim to reduce the number of wrong way driving incidents. Some of which were the implementation of better lighting on intersections where a driver may end up in the wrong lane. The implementation of red reflector strips on the road to indicate when a driver is on the wrong side, was another recommendation by the authors. As the results suggest that wrong way driving is a problem in Victoria, policy makers should consider such changes. The probability of an FSI crash occurring based on vehicle type was also examined in the current study. The results suggested that bicycles are more likely to be involved in an FSI crash, when compared to cars. O'Hern and Oxley (2018) have suggested that Victoria has a high rate of bicycle related FSI road accidents, when compared to other states. O'Hern and Oxley interpreted this finding as a need to improve cycling infrastructure in Victoria. As the current study suggests that cyclists are more likely to be involved in an FSI crash, this need to improve cycling infrastructure is further supported.

There were some limitations to the study. Firstly, no behavioural factors, such as mobile phone use or alcohol consumption was included in the analysis. This was because such information is not included in the public access data. There may have been interactions between some of the factors in the final model and driver behavioural factors, which could not be tested for in this study. Future studies focusing on Victoria could benefit from examining how a driver's behavioural factors may relate to a vehicle's movement, such as the wrong way driving. This is because it has been suggested that intoxication is a factor which influences the severity of wrong way driving crashes (Pour-Rouholamin & Zhou, 2016). Also, the study did not examine vehicle types separately. This means that the results must be generalised to all vehicles, excluding motorcycles as they were not included in the analysis. This may impact the use of the results by polices which aim to target a specific road user group. Lastly it was identified that there is a lag between the FSI crashes being correctly classified and being updated in the CrashStats data. This means that the data from the final two years included in the analysis, may not accurately reflect the true values of the FSI crashes. Future research using the CrashStats data should account for this lag and potentially exclude the most recent two years of data.

This current study has demonstrated that for crashes occurring in non-regional Victoria there are factors which significantly increase the likelihood of an FSI crash. These factors can be used by future policy makers who have the goal of reducing FSI crashes. Some suggested changes to help achieve this have also been identified, such as improved lighting and reflector strips for the wrong way driving. The final logistic model only examined two vehicle crashes. Because of this the results can only be applied to crashes which involve two vehicles. This is not a problem though, because the descriptive analysis indicated that the largest proportion of crashes for both the FSI and the minor injury crashes involved two vehicles. This may suggest that this problem, as well as the others identified need to be targeted in Victoria.

Acknowledgements

The author would like to thank Denny Meyer and Piyush Madhamshettiwar for their help throughout the study.

References

Asgarzadeh, M., Fischer, D., Verma, S. K., Courtney, T. K., & Christiani, D. C. (2018). The impact of weather, road surface, time-of-day, and light conditions on severity of bicycle-motor vehicle crash injuries. *American Journal of Industrial Medicine*, *61*(7), 556-565. DOI: 10.1002/ajim.22849

Bergel-Hayat, R., Debbarh, M., Antoniou, C., & Yannis, G. (2013). Explaining the road accident risk: weather effects. *Accident Analysis & Prevention*, *60*, 456-465.

Boufous, S., de Rome, L., Senserrick, T., & Ivers, R. (2012). Risk factors for severe injury in cyclists involved in traffic crashes in Victoria, Australia. *Accident Analysis & Prevention*, *49*, 404-409. doi.org/10.1016/j.aap.2012.03.011

Brodie, L., Lyndal, B., & Elias, I. J. (2009). Heavy vehicle driver fatalities: Learning's from fatal road crash investigations in Victoria. *Accident Analysis & Prevention*, *41*(3), 557-564. doi:10.1016/j.aap.2009.02.005

Chen, H. Y., Senserrick, T., Martiniuk, A. L. C., Ivers, R. Q., Boufous, S., Chang, H. Y., & Norton, R. (2010). Fatal crash trends for Australian young drivers 1997-2007: Geographic and socioeconomic differentials. *Journal of Safety Research*, *41*(2), 123-128. doi:10.1016/j.jsr.2009.12.006

CrashStats. (2013). *Road Crash Statistics Victoria, 2013 Edition*. VicRoads.

Dissanayake, S., & Roy, U. (2014). Crash severity analysis of single vehicle run-off-road crashes. *Journal of Transportation Technologies*, *4*, 1-10. DOI:10.4236/jtts.2014.41001

Edwards, J. B. (1998). The relationship between road accident severity and recorded weather. *Journal of Safety Research*, *29*(4), 249-262.

Jung, S., Jang, K., Yoon, Y., & Kang, S. (2014). Contributing factors to vehicle to vehicle crash frequency and severity under rainfall. *Journal of Safety Research*, *50*, 1-10. http://dx.doi.org/10.1016/j.jsr.2014.01.001

Jung, S., Qin, X., & Noyce, D. A. (2010). Rainfall effect on single-vehicle crash severities using polychotomous response models. *Accident Analysis & Prevention*, *42*(1), 213-224. doi:10.1016/j.aap.2009.07.020

Langford, J. & Newstead S. (2008). Road accidents in Australia. *IATSS Research*, *32*(2), 122-125.

Liu, L., & Dissanayake, S. (2009). Factors affecting crash severity on gravel roads. *Journal of Transportation Safety & Security*, *1*(4), 254-267. doi.org/10.1080/19439960903381669

Martensen, H., & Dupont, E. (2013). Comparing single vehicle and multivehicle fatal road crashes: A joint analysis of road conditions, time variables and driver characteristics. *Accident Analysis & Prevention*, *60*, 466-471. doi.org/10.1016/j.aap.2013.03.005

Naik, B., Tung, L. W., Zhao, S., & Khattak, A. J. (2016). Weather impacts on single-vehicle truck crash injury severity. *Journal of Safety Research*, *58*, 57-65. doi.org/10.1016/j.jsr.2016.06.005

O'Hern, S., & Oxley, J. (2018). Fatal cyclist crashes in Australia. *Traffic Injury Prevention*, *19*(2), 27-31. doi.org/10.1080/15389588.2018.1497166

Palamara, P., Broughton, M., & Fraser, M. (2009). An investigation of urban area run off road crashes in Western Australia 2005-2009. Report No.10-005. Curtin-Monash Accident Research Centre (C-MARC), Australia.

Pour-Rouholamin, M., & Zhou, H. (2016). Analysis of driver injury severity in wrong-way driving crashes on controlled-access highways. *Accident Analysis & Prevention*, *94*, 80-88.

Tay, R., & Rifaat, S. M. (2007). Factors contributing to the severity of intersection crashes. *Journal of Advanced Transportation*, *41*(3), 245-265. doi.org/10.1002/atr.5670410303

Thompson, J. P., Baldock, M. R., & Dutschke, J. K. (2018). Trends in the crash involvement of older drivers in Australia. *Accident Analysis & Prevention*, *117*, 262-269. doi.org/10.1016/j.aap.2018.04.027

Transport accident commission. *A DECADE OF REDUCING ROAD TRAUMA Road Safety Statistics from 2000 to 2009.* https://www.tac.vic.gov.au/files-to-move/media/upload/a-decade-of-reducing-road-trauma.pdf

Appendix A

Table 1: *Variables used in each analysis*

Variable Name	Random forest and gradient boosting	LASSO Regression	Stepwise Logistic Regression
FSI_OR_MINORINJURY	Y	Y	Y
ACCIDENT_TYPE	Y	Y	Y
DAY_OF_WEEK	Y	Y	Y
LIGHT_CONDITION	Y	Y	Y
ROAD_GEOMETRY	Y	Y	Y
SPEED_ZONE	Y	Y	Y
Deg.Urban.Name	Y	Y	Y
Postcode.No	Y	Y	Y
SURFACE_DRY	Y	Y	Y
SURFACE_WET	Y	Y	Y
SURFACE_MUDDY	Y	Y	Y
SURFACE_SNOWY	Y	Y	Y
SURFACE_ICY	Y	Y	Y
ATMOSPHERIC_CLEAR	Y	Y	Y
ATMOSPHERIC_RAINING	Y	Y	Y
ATMOSPHERIC_SNOWING	Y	Y	Y
ATMOSPHERIC_FOG	Y	Y	Y
ATMOSPHERIC_SMOKE	Y	Y	Y
ATMOSPHERIC_DUST	Y	Y	Y
ATMOSPHERIC_STRONG.WIND	Y	Y	Y
VEHICLE_YEAR_MANUF.A	Y	Y	N
ROAD_SURFACE_TYPE.A	Y	Y	N
TRAFFIC_CONTROL.A	Y	Y	Y
DRIVER_INTENT.A	Y	Y	Y
VEHICLE_MOVEMENT.A	Y	Y	Y
VEHICLE_COLOUR_1.A	Y	Y	Y
VEHICLE_YEAR_MANUF.B	Y	Y	Y
ROAD_SURFACE_TYPE.B	Y	Y	Y

TRAFFIC_CONTROL.B	Y	Y	Y
DRIVER_INTENT.B	Y	Y	Y
VEHICLE_MOVEMENT.B	Y	Y	Y
VEHICLE_COLOUR_1.B	Y	Y	Y
VEHICLE_YEAR_MANUF.E	Y	N	N
ROAD_SURFACE_TYPE.E	Y	N	N
TRAFFIC_CONTROL.E	Y	N	N
DRIVER_INTENT.E	Y	N	N
VEHICLE_MOVEMENT.E	Y	N	N
VEHICLE_COLOUR_1.E	Y	N	N
VEHICLE_YEAR_MANUF.C	Y	N	N
ROAD_SURFACE_TYPE.C	Y	N	N
TRAFFIC_CONTROL.C	Y	N	N
DRIVER_INTENT.C	Y	N	N
VEHICLE_MOVEMENT.C	Y	N	N
VEHICLE_COLOUR_1.C	Y	N	N
VEHICLE_YEAR_MANUF.D	Y	N	N
ROAD_SURFACE_TYPE.D	Y	N	N
TRAFFIC_CONTROL.D	Y	N	N
DRIVER_INTENT.D	Y	N	N
VEHICLE_MOVEMENT.D	Y	N	N
VEHICLE_COLOUR_1.D	Y	N	N
VEHICLE_YEAR_MANUF.G	Y	N	N
ROAD_SURFACE_TYPE.G	Y	N	N
TRAFFIC_CONTROL.G	Y	N	N
DRIVER_INTENT.G	Y	N	N
VEHICLE_MOVEMENT.G	Y	N	N
VEHICLE_COLOUR_1.G	Y	N	N
VEHICLE_YEAR_MANUF.H	Y	N	N
ROAD_SURFACE_TYPE.H	Y	N	N
TRAFFIC_CONTROL.H	Y	N	N
DRIVER_INTENT.H	Y	N	N
VEHICLE_MOVEMENT.H	Y	N	N

VEHICLE_COLOUR_1.H	Y	N	N
VEHICLE_YEAR_MANUF.F	Y	N	N
ROAD_SURFACE_TYPE.F	Y	N	N
TRAFFIC_CONTROL.F	Y	N	N
DRIVER_INTENT.F	Y	N	N
VEHICLE_MOVEMENT.F	Y	N	N
VEHICLE_COLOUR_1.F	Y	N	N
VEHICLE_YEAR_MANUF.I	Y	N	N
ROAD_SURFACE_TYPE.I	Y	N	N
TRAFFIC_CONTROL.I	Y	N	N
DRIVER_INTENT.I	Y	N	N
VEHICLE_MOVEMENT.I	Y	N	N
VEHICLE_COLOUR_1.I	Y	N	N
VEHICLE_YEAR_MANUF.J	Y	N	N
ROAD_SURFACE_TYPE.J	Y	N	N
TRAFFIC_CONTROL.J	Y	N	N
DRIVER_INTENT.J	Y	N	N
VEHICLE_MOVEMENT.J	Y	N	N
VEHICLE_COLOUR_1.J	Y	N	N
VEHICLE_YEAR_MANUF.K	Y	N	N
ROAD_SURFACE_TYPE.K	Y	N	N
TRAFFIC_CONTROL.K	Y	N	N
DRIVER_INTENT.K	Y	N	N
VEHICLE_MOVEMENT.K	Y	N	N
VEHICLE_COLOUR_1.K	Y	N	N
VEHICLE_YEAR_MANUF.L	Y	N	N
ROAD_SURFACE_TYPE.L	Y	N	N
VEHICLE_TYPE.L	Y	N	N
TRAFFIC_CONTROL.L	Y	N	N
DRIVER_INTENT.L	Y	N	N
VEHICLE_MOVEMENT.L	Y	N	N
VEHICLE_COLOUR_1.L	Y	N	N
VEHICLE_YEAR_MANUF.N	Y	N	N

ROAD_SURFACE_TYPE.N	Y	N	N
TRAFFIC_CONTROL.N	Y	N	N
DRIVER_INTENT.N	Y	N	N
VEHICLE_MOVEMENT.N	Y	N	N
VEHICLE_COLOUR_1.N	Y	N	N
VEHICLE_YEAR_MANUF.M	Y	N	N
ROAD_SURFACE_TYPE.M	Y	N	N
TRAFFIC_CONTROL.M	Y	N	N
DRIVER_INTENT.M	Y	N	N
VEHICLE_MOVEMENT.M	Y	N	N
VEHICLE_COLOUR_1.M	Y	N	N
SEX.A.	Y	Y	Y
AGE.A.	Y	Y	Y
ROAD_USER_TYPE.A.	Y	Y	Y
SEX.B.	Y	Y	Y
AGE.B.	Y	Y	Y
ROAD_USER_TYPE.B.	Y	Y	Y
SEX.C.	Y	N	N
AGE.C.	Y	N	N
ROAD_USER_TYPE.C.	Y	N	N
SEX.E.	Y	N	N
AGE.E.	Y	N	N
ROAD_USER_TYPE.E.	Y	N	N
SEX.D.	Y	N	N
AGE.D.	Y	N	N
ROAD_USER_TYPE.D.	Y	N	N
SEX.F.	Y	N	N
AGE.F.	Y	N	N
ROAD_USER_TYPE.F.	Y	N	N
SEX.I.	Y	N	N
AGE.I.	Y	N	N
ROAD_USER_TYPE.I.	Y	N	N
SEX.H.	Y	N	N

AGE.H.	Y	N	N
ROAD_USER_TYPE.H.	Y	N	N
SEX.G.	Y	N	N
AGE.G.	Y	N	N
ROAD_USER_TYPE.G.	Y	N	N
SEX.J.	Y	N	N
AGE.J.	Y	N	N
ROAD_USER_TYPE.J.	Y	N	N
DCA_CAT	Y	Y	Y
VEHICLE_TYPE.A_RECODED	Y	Y	Y
VEHICLE_TYPE.B_RECODED	Y	Y	Y
VEHICLE_TYPE.C_RECODED	Y	N	N
VEHICLE_TYPE.D_RECODED	Y	N	N
VEHICLE_TYPE.E_RECODED	Y	N	N
VEHICLE_TYPE.F_RECODED	Y	N	N
VEHICLE_TYPE.G_RECODED	Y	N	N
VEHICLE_TYPE.H_RECODED	Y	N	N
VEHICLE_TYPE.I_RECODED	Y	N	N
VEHICLE_TYPE.J_RECODED	Y	N	N
VEHICLE_TYPE.K_RECODED	Y	N	N
VEHICLE_TYPE.L_RECODED	Y	N	N
VEHICLE_TYPE.M_RECODED	Y	N	N

Note: Y = variable included in analysis, N = variables excluded from analysis

Table 2: *Variables with significant odds ratios from stepwise logistic regression*

Variable Name	Odds Ratio	95%CI	P-Value
LIGHT_CONDITION3	1.33	1.25 and 1.42	<.001
LIGHT_CONDITION5	1.50	1.24 and 1.83	<.001
LIGHT_CONDITION9	0.63	0.50 and 0.78	<.001
ROAD_GEOMETRY2	1.08	1.02 and 1.15	0.014
ROAD_GEOMETRY5	1.14	1.05 and 1.22	<.001
SPEED_ZONE70	1.23	1.14 and 1.32	<.001
SPEED_ZONE80	1.50	1.40 and 1.60	<.001
SPEED_ZONE90	2.11	1.37 and 3.29	<.001
SPEED_ZONE100	1.68	1.51 and 1.88	<.001
Deg.Urban.NameMELB_URBAN	1.13	1.04 and 1.24	0.006
Deg.Urban.NameSMALL_CITIES	0.86	0.75 and 0.97	0.017
Postcode.No	1.00	1.00 and 1.00	0.036
SURFACE_DRY	1.27	1.09 and 1.48	0.002
ATMOSPHERIC_CLEAR	1.68	1.49 and 1.90	<.001
ATMOSPHERIC_RAINING	1.82	1.54 and 2.15	<.001
ATMOSPHERIC_FOG	1.49	1.02 and 2.17	0.038
ATMOSPHERIC_STRONG.WIND	2.19	1.61 and 2.98	<.001
VEHICLE_MOVEMENT.A2	0.78	0.72 and 0.84	<.001
VEHICLE_MOVEMENT.A3	0.72	0.63 and 0.82	<.001
VEHICLE_MOVEMENT.A4	0.60	0.47 and 0.76	<.001
VEHICLE_MOVEMENT.A9	0.64	0.43 and 0.93	0.021
VEHICLE_MOVEMENT.A12	2.24	1.03 and 5.09	0.046
VEHICLE_MOVEMENT.A13	0.71	0.54 and 0.94	0.018
VEHICLE_MOVEMENT.A15	0.73	0.66 and 0.81	<.001
VEHICLE_MOVEMENT.A16	0.21	0.05 and 0.66	0.016
VEHICLE_MOVEMENT.A17	0.74	0.67 and 0.82	<.001
VEHICLE_MOVEMENT.A18	1.71	1.51 and 1.94	<.001
VEHICLE_MOVEMENT.A19	2.14	1.41 and 3.35	<.001
VEHICLE_MOVEMENT.A99	1.53	1.18 and 1.98	0.001
VEHICLE_COLOUR_1.AGRN	1.25	1.12 and 1.41	<.001
VEHICLE_COLOUR_1.APUR	0.70	0.48 and 0.99	0.048

ROAD_SURFACE_TYPE.B3	1.27	1.06 and 1.51	0.008
TRAFFIC_CONTROL.B9	0.55	0.48 and 0.61	<.001
TRAFFIC_CONTROL.B11	0.76	0.69 and 0.84	<.001
TRAFFIC_CONTROL.B99	0.75	0.64 and 0.86	<.001
VEHICLE_MOVEMENT.B2	0.78	0.72 and 0.85	<.001
VEHICLE_MOVEMENT.B3	0.72	0.61 and 0.84	<.001
VEHICLE_MOVEMENT.B4	0.65	0.48 and 0.88	0.005
VEHICLE_MOVEMENT.B5	0.73	0.55 and 0.98	0.035
VEHICLE_MOVEMENT.B9	0.50	0.30 and 0.81	0.007
VEHICLE_MOVEMENT.B10	0.53	0.32 and 0.86	0.011
VEHICLE_MOVEMENT.B15	1.19	1.08 and 1.31	<.001
VEHICLE_MOVEMENT.B18	1.31	1.09 and 1.57	0.004
SEX.A.M	1.05	1.01 and 1.10	0.026
AGE.A.	1.01	1.01 and 1.01	<.001
SEX.B.M	1.06	1.01 and 1.11	0.017
AGE.B.	1.01	1.01 and 1.01	<.001
DCA_CATB	0.39	0.25 and 0.59	<.001
DCA_CATC	0.47	0.30 and 0.71	<.001
DCA_CATD	0.20	0.13 and 0.31	<.001
DCA_CATE	0.32	0.20 and 0.49	<.001
DCA_CATF	0.40	0.24 and 0.67	<.001
DCA_CATG	0.27	0.17 and 0.43	<.001
DCA_CATH	0.51	0.31 and 0.81	0.005
DCA_CATI	0.44	0.20 and 0.99	0.042
DCA_CATJ	0.31	0.15 and 0.63	0.001
VEHICLE_TYPE.A_RECODED2	1.46	1.26 and 1.69	<.001
VEHICLE_TYPE.B_RECODED2	1.26	1.10 and 1.44	<.001
VEHICLE_TYPE.B_RECODED4	1.59	1.44 and 1.75	<.001
VEHICLE_TYPE.B_RECODED5	1.51	1.06 and 2.15	0.024
VEHICLE_TYPE.B_RECODED9	2.08	1.07 and 3.97	0.027