Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 4): pp. 3876-3891
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.04.423

Analyzing Road Traffic Crash Severity In Uttarakhand Region

Preeti Malik[1*], Ilarika Rautela1, Varsha Mittal[1], Dr. Hemant Chauhan[2]

[1]Graphic Era Deemed to be University, Dehradun, INDIA.

[2]Associate Professor, School of Management, Graphic Era Hill University, Dehradun.

*Email: preetishivach2009@gmail.com

Abstract:

In the current era of modernization as we move ahead towards a future which is more and more computerized and automated there are many events happenings around us which needs to be solved and handled carefully. Road accidents is one of them. It is an essential component of human life. Due to the inadequacy of conventional approaches to handle the complexity of road traffic characteristics and relationships, there has been a surge in interest in neural network applications in the field of transportation research in recent years. In this paper, we are going to study different pre introduced algorithms and models based on traffic analysis and also the causes that effects these accidents to happen and how we can prevent them. We are also going to use case study approach to analyse the past five years data based on road traffic.

Keywords: Traffic analysis, field of transportation, neural network, deep learning, motion detection.

1. Introduction

Millions of individuals lose their lives as a consequence of traffic accidents on the roads each year. Traffic accidents result in serious injuries, as well as financial damages for victims and their families. Most nations lost 3% of their gross domestic product as a result of this catastrophe (road traffic collision). 90% of road traffic deaths occur in low- and middle-income countries. There are multiple causes of these disastrous events like rain, high density area, poor road conditions etc. Some vehicles like trucks weighs over 10,000 pounds is a vital component for carrying goods logistics and other raw materials but they are very fatal in terms of crash in traffic frequency and causing serious property damage as well as human lives. Traffic analysis is a way of examining and intercepting messages to deduced information from patterns in communication. In traffic analysis, the greater the number of messages observed, intercepted and stored, the more it can be inferred from the traffic. There are already researches going on or had done in the past on not only traffic analysis but other causes to prevent this from happening or even lower the risk that is caused by road traffic accidents. Neural networks have also been used in traffic prediction in recent years and have been successful in this task. We will study the pre introduced algorithms and models and draw conclusions from it.

Traffic analysis is a way of collecting, storing, and analysing traffic. The data is collected so that we have an updated information about what is happening. This allows us to take action according to the problem. We can also store this data for future use. There are many factors responsible for traffic accidents such as environmental factors, location of accidents and other uncertain and complex factors. Traffic safety has become one of the major focuses of social issues. Thus, traffic analysis raises the demands for technological solutions and algorithms. It leads to the report of traffic rules violation, accidents and also the casual behavior of people on the road.

Field of Transportation includes air, land, water, cable, pipeline and space. This can be further divided into vehicles and operations. Our main focus on this paper would be on land (rail and road). In recent years, the field of transportation studies has seen an increased interest in neural network applications. It is because the traditional methods are unable to analyse the complexities of road traffic traits. Comparable and appropriate results have been achieved using neural networks in areas such as classification and detection of vehicles, traffic control analysis, driver's behaviour prediction, congestion detection etc.

Motion detection is the process of detecting a change in the position of an object relative to its surroundings or vice versa. It is a software algorithm designed to detect the moving objects in a specific area or to create an alert. Vehicle motion is the foundation of connected vehicles and safe driving. Traffic control systems either contain sensors or is done manually. The traffic detection by sensors is more accurate but need to be protected.

1.1. Preventions

Road accidents can be avoided. There are many efficient road safety measures available, thus the issue can be solved using a scientific system approach.

Vehicles	• Vehicles with good lighting, tyres, proper breaks etc. • Older vehicles and highly polluting vehicles should be eliminated [1]. • Vehicles should have seat belts and other essential safety amenity like airbags. • Vehicles should be serviced time to time so that there is no brakes or steering failure and the headlights should also have proper bulbs. • The vehicles should not be overloaded.

Road Conditions	• Roads should be well maintained with proper markings of road safety signs. • Proper footpaths for walkers. • Lanes separated for fast-moving and slow-moving vehicles. • Appropriate lightening of the roads for clear sight.
Road Obstacles	• Remove the obstacles (like poles or trees) • Proper warning signs or markings should be done [2] • Potholes should be properly covered. • Damaged roads should be reconstructed. If not so, then proper warnings or sign boards should be provided.
Adequate Training of Drivers	Drivers can notably contribute to reducing of the accidents [3]. • Drivers should be trained properly and hold a valid driving license. • Issuing of the license should be based on the whether the driver has learnt from a designated driving school. • The drivers should also go through a psychological test before obtaining a license. • The focus should be not only on the skills of driving a vehicle, but also on those related with the correct behavior on busy roads • Periodic medical check-ups should be conducted for the drivers especially for one's vision and hearing capabilities.
Driver's State of mind	• Driver should not have consumed alcohol or any other drug. • There should be no distractions (like talking on the mobile phone) while driving. • Driver should not be tired and should also have a good physical and mental health [1]. • Driver should have an experience of driving on particular road and should be familiar with the conditions
Management of accident victims	• Highways and busy roads should have first aid facilities. • Common people should be aware about how to treat accident victims and should not get scared so that the deaths can be reduced [4].

Strict Penalties	• Strict penalties should be there for lack of attention of the driver. • Tighten the rules for not following the speed limits. • Increase the awareness of drivers so that they understand the consequences of excessive speed.
Common People	• Pedestrians should not cross from wrong places and should not walk carelessly. • They should not board the vehicle from wrong side. • They should not run after a vehicle and should also not get in or out from a moving vehicle. • Passengers inside the vehicle should not project their body outside vehicle

Table 1: Suggested Preventions

2. Literature Review

Various studies have addressed the different aspects of road traffic accidents. Many studies and past researches have done in this field impacted the rate of these incidents and many models have also been developed to prevent these terrifying incidents from happening. To name one of the models that has been developed and used now a days is statistical model. Statistical modelling and descriptive analysis were used to determine the effects of work zones on traffic accidents. Enhanced datasets that may be utilised to identify factors influencing accident severity were used in the development of statistically robust models [5]. In order to separate work zone-specific factors, severity modelling parameters for accidents that occurred outside of a work zone were compared to those for those incidents. The logistic regression approach was used to analyse the best model. Based on odds-ratio values, modelling results were independently interpreted for work zone and non-work zone situations [6].

The primary distinctions between various logistic regression models and crash data aggregation techniques that are often used in road safety literature have been the subject of several research [7]. Three authorised logistic models—sequential binary logit models, ordered logit models, and multinomial logit models—as well as three data aggregation techniques—occupant-based, collision-based, and vehicle-based—are compared in this study. We utilised collision data from 31 highway routes in the province of Ontario, Canada, from 2001 to 2006. The results from occupant-based data are more trustworthy than those from vehicle- and collision-based data, with the multinomial logit model having the best match to the data of the three models.

According to research, models like partial proportional odds (PPO) were created for various driver groups to determine the risk variables associated with the severity of crashes. The outcomes of the model evaluation demonstrated that the risk variables and their effects on various driver categories were not the same. While driving behaviours like carelessness, driving while intoxicated or stoned, and fatigue were significantly associated with the high-risk group only, the low-risk groups had fewer risk factors and minor marginal effects. Weather conditions, road conditions, rural areas, and various crash manners were significantly associated with the crash severities in all driver groups. Different models have been created to improve traffic safety. These models just contain temporal parameters and have a straightforward mathematical structure. The first model predicts that death rates have a negative exponential function, while the second model predicts that traffic volumes have a logistic kind of saturation model. In the Netherlands, the United States, West Germany, and Great Britain, the models are applied to death rates and traffic volumes. The models' consistency with the data demonstrates that the linearity of the logarithmic mortality and volume rates is a fair assumption. The data linking the safety results to changes in traffic volume also points to a relationship between the two models. On the basis of temporal parameters alone, future estimates for mortality rates and traffic volume are made from the models.

3. Case Study

The Government of India: Ministry of Road Transport maintains a database of all of the road accidents that have been reported on Indian highways [9]. The database includes detailed information on all the accident including accident location, severity level, vehicle information, weather type and other characteristics that became the cause of the accidents. The data of Uttarakhand was collected and the analysis was performed on the characteristics such as road features, weather conditions and impacting vehicles. The data has been taken from [9].

Year	Number of Accidents	Number of persons killed	Person injured
2014	1081	702	1226
2015	1520	720	820
2016	1203	691	1352
2017	1600	642	658
2018	1470	601	869

Table 2: Number of accidents in Uttarakhand due to weather conditions

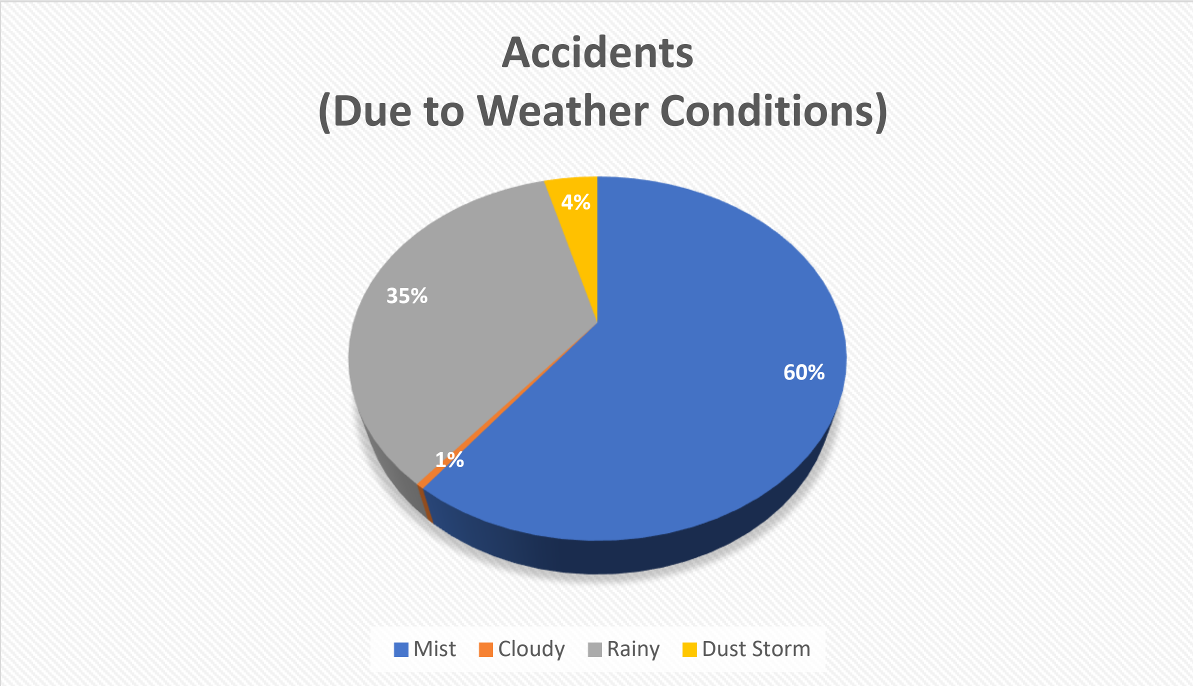

Figure 1: Pie Chart on Accidents occurred in 2016 due to different weather conditions

Year	Number of Accidents	Number of people Killed	Number of people injured
2014	1786	980	1575
2015	1650	1055	954
2016	1556	947	709
2017	1603	942	715
2018	1420	920	550

Table 3: Number of accidents in Uttarakhand due to Road Conditions

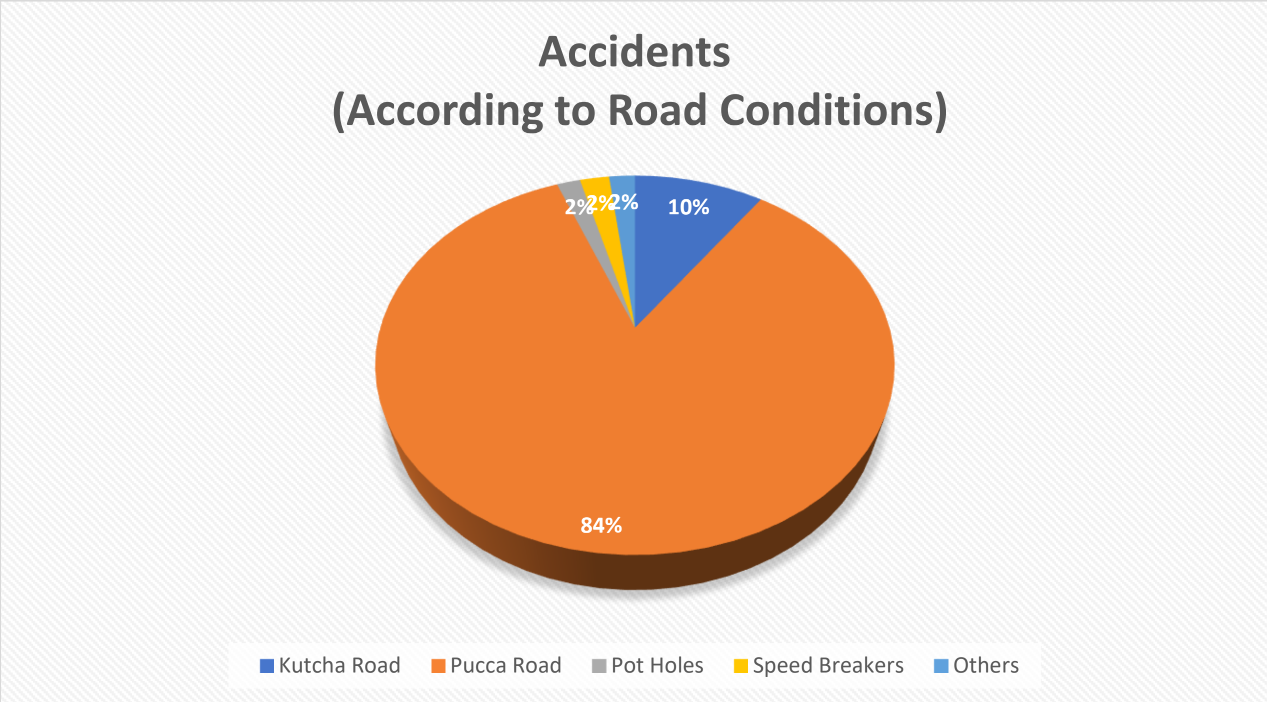

Figure 2: Pie Chart on Accidents occurred in 2016 According to different road conditions

Year	Number of Accidents	Number of people Killed	Number of people injured
2014	7050	4390	2660
2015	8700	3658	5042
2016	9000	4587	4413
2017	8800	2564	6236
2018	1073	256	817

Table 4: Total number of accidents in Uttarakhand from all possible causes

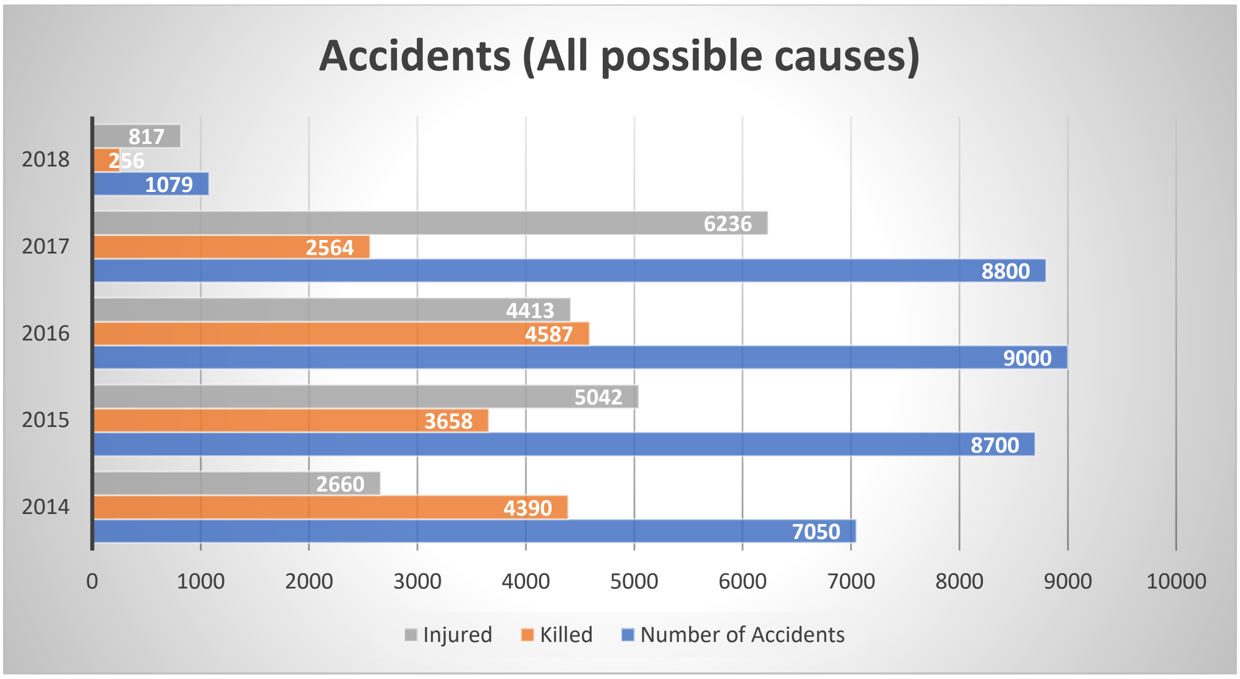

Figure 3: Pie Chart on Accidents with all possible causes

4. Comparative study

There are multiple models [5, 7, 8] that have already been developed and being used in the traffic analysis in this part we are going to study and compare some of the pre introduced models. One of the models that has been developed is known as Urban traffic modelling and analysis it is a component of sophisticated traffic intelligent management technologies, which has grown in importance in the field of traffic management and control. Its major goal is to forecast traffic congestion in a certain urban transportation network and make recommendations for traffic network upgrades. Three types of data are used in research. A model of the transport network architecture and methods related to both spatial and temporal dimensions, as well as historical and recent information on a traffic network's density and flow. The ultimate goal is to improve traffic infrastructure, such as traffic signals, through greater optimization.

Some of the algorithm that has been already in use are forecast algorithm where we only utilize data from similar instances in the past. As a result, they can forecast a density at a given place. This simple past-data technique is used by the following methods and algorithms; however, they are distinguished by distinct criteria. Statical methods are based on auto-regression and moving average approaches, to begin with. ARMA, ARIMA (Integrated ARMA), and SARIMA are examples of algorithms that use them (seasonal ARIMA). There are several more statistical approaches based on non-parametric regression and variations. Machine learning is being used to anticipate traffic models using a variety of methods such as Vector regression (SVR), time-delay neural network (TDNN), and Bayesian network.

5. Experimental Work

5.1. Based on different Road Conditions:

The Summary Output tells how well the calculated linear regression equation fits the source data.

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.997247
R Square	0.994502
Adjusted R Square	0.993717
Standard Error	24.52396
Observations	9

The **Multiple R correlation coefficient** gauges how strongly two variables are related linearly. The correlation coefficient's absolute value, which ranges from -1 (strongly negative correlation) to 1 (strongly positive correlation), represents the strength of the link. Given that the Multiple R value is 0.997247, the association is strong.

The **Coefficient of Determination**, or R Square, is a measure of how well a model fits the data. How many points fall on the regression line is indicated. R^2 in this instance is 0.99 (rounded to two digits), which is respectable. This indicates that the regression analysis model fits 99% of our data. In other words, the independent factors may account for 99.9% of the dependent variables.

Adjusted R Square is the R square that has been adjusted for the number of independent variables in the model is known as the adjusted R square. In the context of multiple regression analysis, we wish to substitute this number for R square.

Another goodness-of-fit metric that demonstrates the accuracy of the regression analysis is **standard error**, which is an absolute measure of the average separation between the data points and the regression line.

The quantity of data in your model is called **observations**.

The ANOVA (Analysis of Variance) splits the sum of squares into individual components that give information about the levels of variability within the regression model.

ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	761522	761522	1266.197	3.59E-09	

Residual	7	4209.971	601.4244			
Total	8	765732				

df is the number of the degrees of freedom associated with the sources of variance. SS is the sum of squares. The smaller the Residual SS compared with the Total SS, the better your model fits the data. MS is the mean square. F is the F statistic, or F-test for the null hypothesis. It is used to test the overall significance of the model. Significance F is the P-value of F. The Significance F value gives an idea of how reliable (statistically significant) your results are.

The coefficient is the most useful component. It enables you to build a linear regression equation: Y = bx + a

	Coeffic ients	Standar d Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	8.075 969	9.68169 5	0.834 148	0.431 739	- 14.81 76	30.96 954	- 14.817 6	30.969 54
Number of accidents	0.614 392	0.01726 6	35.58 367	3.59 E-09	0.573 564	0.655 22	0.5735 64	0.6552 2

For our data set, where y is the person killed and x is number of accidents, our linear regression formula goes as follows:

Y = Number of Accident coefficient * x + Intercept

Equipped with a and b values rounded to three decimal places, it turns into:

Y=0.614 *x- 8.075

For example, with the average accidents equal to 1003, the person killed would be approximately 607.767:

0.614*1003- 8.075=607.767

If we compare the estimated and actual number of person killed corresponding to the number of accidents of 976 people, we will see that these numbers are slightly different that is the estimated value is 607.767 whereas the actual value is 562. There is a difference because independent variables are never perfect predictors of the dependent variables. And the residuals can help you understand how far away the actual values are from the predicted values:

RESIDUAL OUTPUT		

Observation	Predicted Person Killed	Residuals
1	607.7227	-45.7227
2	134.6407	31.35926
3	21.5926	1.407405
4	17.29185	-10.2919
5	24.05016	-9.05016
6	15.44867	-7.44867
7	24.66456	9.335444
8	49.85463	2.145367
9	838.7341	28.26589

For the first data point, the residual is approximately -45.722. So, we add this number to the predicted value, and get the actual value: 607.767– 45.722= 562.045 (approx. 562).

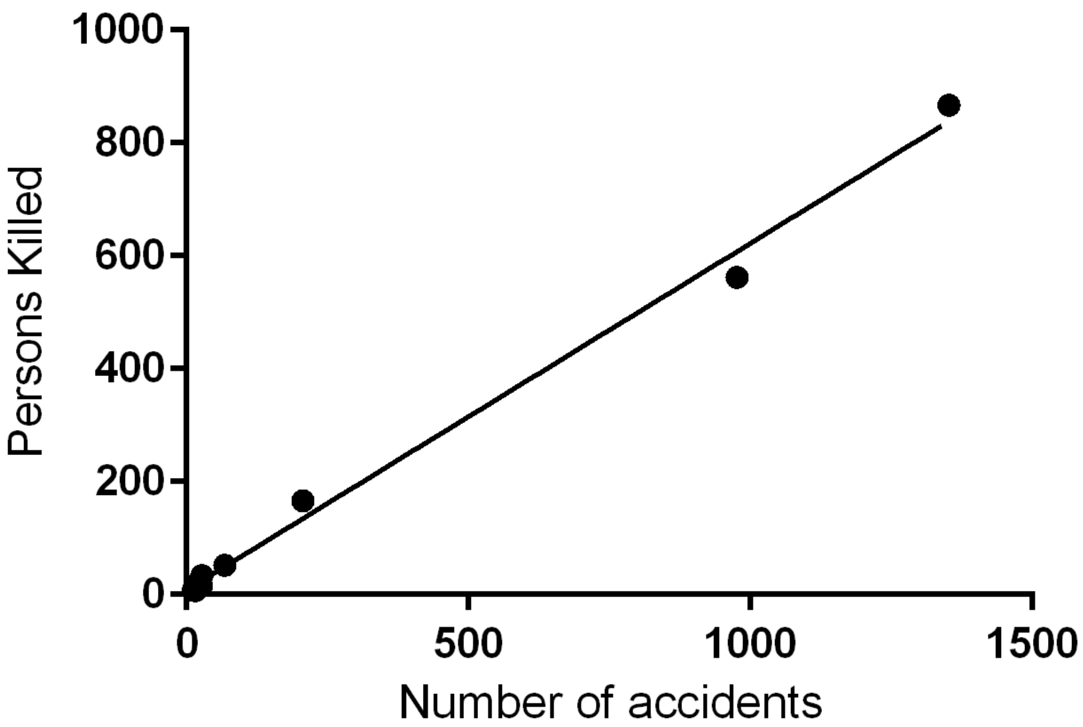

Figure 4: Road Conditions

5.2. Based on the Weather Conditions:

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.999375

R Square	0.998749
Adjusted R Square	0.998333
Standard Error	18.79389
Observations	5

ANOVA

	df	SS	MS	F	Significance F
Regression	1	846277.6	846277.6	2395.958	1.88E-05
Residual	3	1059.631	353.2105		
Total	4	847337.2			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-2.16217	10.08209	-0.21446	0.843941	-34.2479	29.92354	-34.2479	29.92354
Person killed	1.57187	0.032113	48.94853	1.88E-05	1.469673	1.674066	1.469673	1.674066

RESIDUAL OUTPUT

Observation	Predicted Number of accidents	Residuals
1	1088.715	-1.71527
2	136.1623	25.83766

3	57.56887	-6.56887
4	0.981566	1.018434
5	68.57195	-18.572

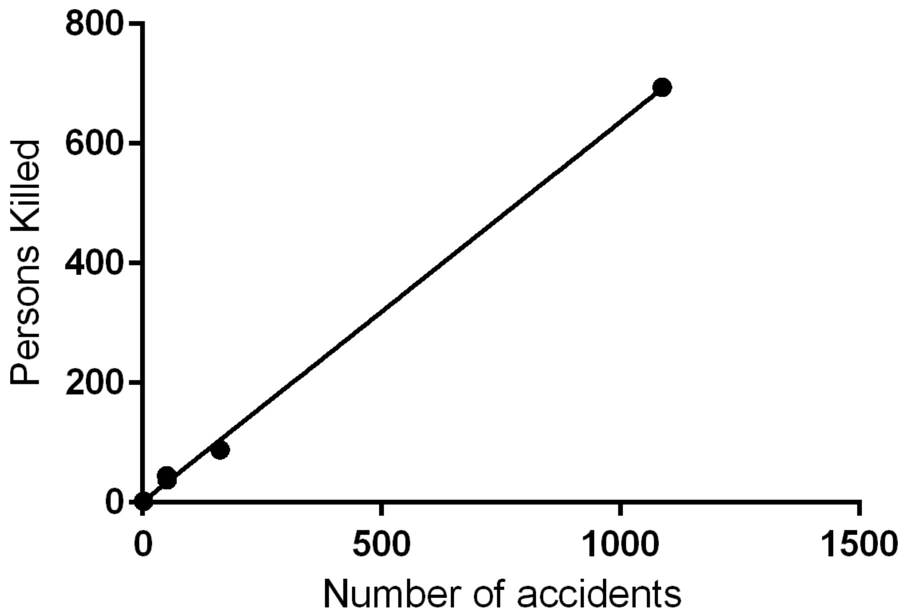

Figure 5: Weather Conditions

5.3. Based on the Impacting Vehicles

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9702
R Square	0.941288
Adjusted R Square	0.932901
Standard Error	50.26149
Observations	9

ANOVA

	df	SS	MS	F	Significance F
Regression	1	283508	283508	112.2263	1.46E-05
Residual	7	17683.52	2526.217		
Total	8	301191.6			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-4.41132	22.22062	-0.19852	0.848278	-56.9548	48.1321	-56.9548	48.1321
Person killed	1.605193	0.151523	10.59369	1.46E-05	1.246897	1.963489	1.246897	1.963489

RESIDUAL OUTPUT

Observation	Predicted Number of Accidents	Residuals
1	326.2583	53.74166
2	42.13926	-2.13926
3	544.5645	19.43548
4	27.69253	1.307474
5	107.9522	20.04785

6	24.48214	15.51786
7	6.825023	8.174977
8	10.03541	0.964591
9	262.0506	-117.051

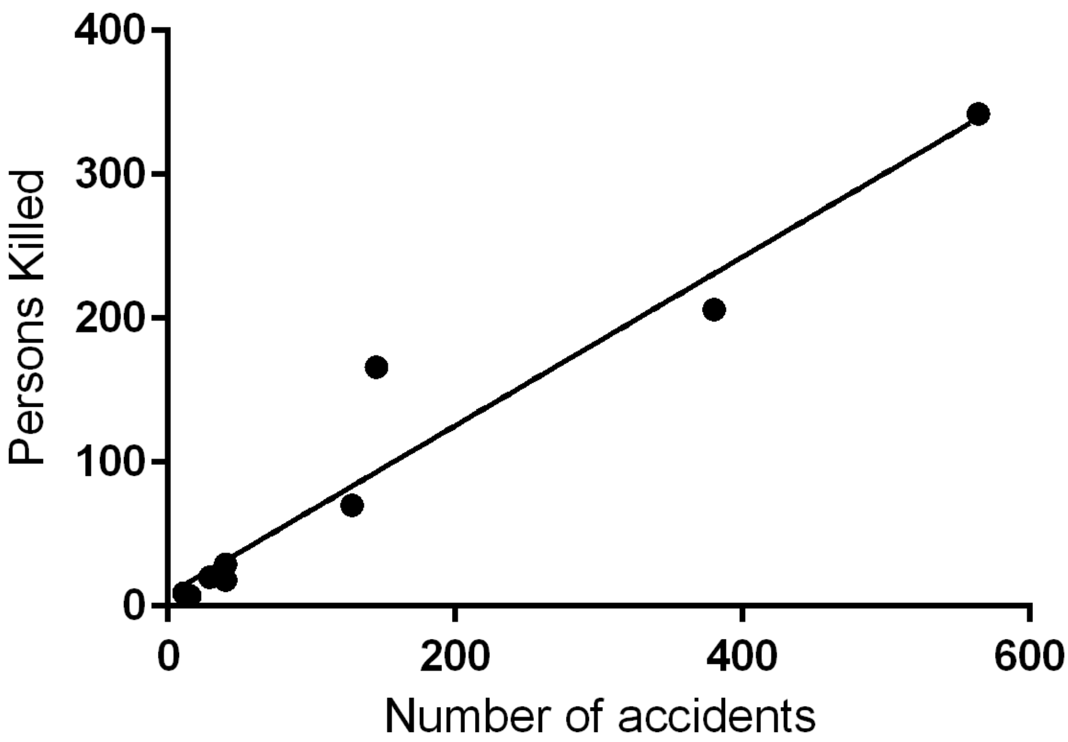

Figure 6: Impacting Vehicles

6. Conclusion

The socioeconomic considerations that are typically the focus of road planning and environmental impact assessments cause a mismatch between planning scales and the spatial scales at which biological systems operate. The fact that there are little legal incentives or disincentives to take environmental implications outside of political domains into account means that decision-making is mostly local, which is one of the causes of this mismatch. Roads frequently have ecological effects that go beyond the purview of local planning borders and are frequently far more significant than the road itself. In this paper we have discussed about the factors that affects road accidents and the causes that is rapidly increasing. We have seen that how these factors are causing multiple severe injuries and fatalities.

7. References

[1] Cuicui Xiao, Miao Chang, Peikun Guo, Qing Chen, Xin Tian, "Comparison of the cost-effectiveness of eliminating high-polluting old vehicles and imposing driving restrictions to reduce vehicle emissions in Beijing", In Transportation Research Part D: Transport and Environment, Vol. 67, pp. 291-302, 2019.

[2] Malik, Faheem, "Road Accidents and Prevention" International Journal Of Engineering Development And Research, Vol. 5. 40-46, 2017.

[3] S. Gopalakrishnan, "A Public Health Perspective of Road Traffic Accidents", J Family Med Prim Care, 1(2), pp 144–150, 2012.

[4] Md. Kamruzzaman, Md. Mazharul Haque, Bayes Ahmed, Tahmina Yasmin, "Analysis of traffic injury severity in a mega city of a developing country", **C**onference: 4th Road Safety International Conference At: Sydney, Australia, 2013.

[5] Ozgur Ozturk, Kaan Ozbay, Hong Yang, "Investigating the Impact of Work Zones on Crash Severity by Comparative Analysis", Transportation Research Board 94th Annual Meeting, 2015.

[6] Yingfeng Li, Yong Bai, "Development of crash-severity-index models for the measurement of work zone risk levels", Accident Analysis & Prevention, Vol. 40, Issue 5, September 2008, Pages 1724-1731.

[7] Taimur Usman, Liping Fu, Luis F. Miranda-Moreno, "Injury severity analysis: comparison of multilevel logistic regression models and effects of collision data aggregation" Article in Journal of Modern Transportation, 2016.

[8] Tibebe Beshah, Shawndra Hill, "Mining Road Traffic Accident Data to Improve Safety: Role of Road- Related Factors on Accident Severity in Ethiopia", AAAI Spring Symposium Artificial Intelligence for Development, pp. 14-19, 2010.

[9] Data Source: www.morth.nic.in (Government of India Ministry of Road Transport and Highways Transport Research Wing, New Delhi)