sustainability

MDPI

Article

Analysis of Vehicle Breakdown Frequency: A Case Study of New South Wales, Australia

Sai Chand [1,*], Emily Moylan [2], S. Travis Waller [1] and Vinayak Dixit [1]

[1] Research Centre for Integrated Transport Innovation (rCITI), School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia; s.waller@unsw.edu.au (S.T.W.); v.dixit@unsw.edu.au (V.D.)

[2] School of Civil Engineering, The University of Sydney, Sydney, NSW 2006, Australia; emily.moylan@sydney.edu.au

* Correspondence: saichand.chakka@unsw.edu.au

Received: 8 September 2020; Accepted: 5 October 2020; Published: 7 October 2020

check for updates

Abstract: Traffic incidents such as crashes, vehicle breakdowns, and hazards impact traffic speeds and induce congestion. Recognizing the factors that influence the frequency of these traffic incidents is helpful in proposing countermeasures. There have been several studies on evaluating crash frequencies. However, research on other incident types is sparse. The main objective of this research is to identify critical variables that affect the number of reported vehicle breakdowns. A traffic incident dataset covering 4.5 years (January 2012 to June 2016) in the Australian state of New South Wales (NSW) was arranged in a panel data format, consisting of monthly reported vehicle breakdowns in 28 SA4s (Statistical Area Level 4) in NSW. The impact of different independent variables on the number of breakdowns reported in each month–SA4 observation is captured using a random-effect negative binomial regression model. The results indicate that increases in population density, the number of registered vehicles, the number of public holidays, average temperature, the percentage of heavy vehicles, and percentage of white-collared jobs in an area increase the number of breakdowns. On the other hand, an increase in the percentage of unrestricted driving licenses and families with children, number of school holidays, and average rainfall decrease the breakdown frequency. The insights offered in this study contribute to a complete picture of the relevant factors that can be used by transport authorities, vehicle manufacturers, sellers, roadside assistance companies, and mechanics to better manage the impact of vehicle breakdowns.

Keywords: breakdowns; incidents; negative binomial; random effects

1. Introduction

Investigating ways to reduce the impacts of road congestion is an increasingly important challenge as vehicle ownership and population increase across the world. Traffic congestion is divided into two categories, namely recurrent and non-recurrent [1]. Recurrent congestion is caused by demand chronically exceeding road capacity, and non-recurrent congestion is caused by random events, such as traffic incidents, adverse weather, and hazards [2]. Non-recurrent traffic congestion is non-trivial, and it was found to account for up to 60% of total congestion [3]. As the critical source of non-recurrent congestion, a traffic incident is defined as a non-recurring event that causes a reduction of roadway capacity or an abnormal increase in demand [4]. Crashes, breakdowns, police stops, and hazards are some examples of the non-planned incidents that impact the typical traffic conditions. The sudden and unpredictable nature of these incidents results in unreliable and fluctuating travel times [5]. Commuters tend to show varying behaviors, such as risk aversion, risk neutrality, and risk seeking [6].

Furthermore, non-recurrent congestion is also found to influence commuters' departure time, route, and mode choice [7].

The unplanned incidents drastically reduce the performance of a network through increased congestion and unreliability [8,9]. Furthermore, the primary incidents are likely to provoke secondary incidents, particularly in areas with more vehicular traffic [10]. This could be due to driver distraction, more increased congestion than usual, stop-and-go movements, etc. [11,12]. The secondary crashes are likely to increase by 2.8% for every minute the primary incident is not cleared and continues to be a hazard [13]. Identifying the key factors that influence unplanned incidents would allow us to suggest appropriate management strategies to mitigate the damage that road traffic inflicts on the environment [14–16].

Numerous studies focus on the prediction of crash frequencies, severities, rates, and durations using advanced statistical models [17]. For example, Negative Binomial and Poisson models are commonly used for crash frequencies [18], Tobit models are used for crash rates [19], Multinomial Logit, Ordered Logit, and Ordered Probit models are used for modeling crash severity [20], and Hazard-based models are used for crash duration [21]. Various data mining and empirical approaches, such as clustering [22,23], support vector machine [24], fuzzy logic [25], artificial neural networks [26], time-series analysis [27,28], and genetic algorithms [29], have been widely used to identify trends and patterns in large temporal and spatial crash datasets [30]. Furthermore, advanced statistical models, such as Random Parameter models, Latent Class models, Bayesian models, and Markov Switching models, are widely used to account for unobserved temporal and spatial heterogeneity [17].

However, research on the prediction of other incident types has been sparse. A contributing factor to this underattention is the lack of comprehensive data regarding other types of incidents, such as breakdowns. These incidents are not well reported to the transport authorities because their safety and property damage repercussions are small compared to those of crashes. However, traffic congestion induced by on-road breakdowns is non-trivial, as noted by Wang et al. (2005) [31].

Vehicle breakdown is a type of unplanned incident where a vehicle fails during operation on a roadway and is forced to stop. There can be many reasons for a vehicle breakdown, such as a flat battery, faulty electrical wiring, fuel pressure problems, tire puncture, driver error, etc. Most breakdowns can be resolved on the spot by self-repairing or calling a mechanic or roadside assistance company. Few breakdowns are so complex as to require towing. Vehicle breakdowns can result in traffic congestion, particularly when the road is partially or fully closed due to obstruction by the broken-down vehicle or the towing equipment. Furthermore, vehicle breakdowns are unusual events, and many drivers are uncertain about how to respond to the failure. This can lead to unsafe behavior and, in some cases, secondary incidents.

Despite the recent improvements in the quality of automobile designs that have improved safety and security, vehicle breakdowns still happen on roads. In fact, they constitute a major proportion of the road incidents, particularly on freeways [31]. There were around 20,000 yearly reported on-road vehicle breakdowns compared to the 27,000 yearly crashes in New South Wales (NSW) between the years 2012 and 2015. Vehicle breakdowns accounted for about 30% of all types of incidents in the state of New South Wales (NSW), Australia, 64% of all incidents on a motorway in the United Kingdom [31], and 32% of total incidents on urban freeways in South East Queensland, Australia [32]. As noted, there are only a handful of studies on vehicle breakdowns.

However, the total number of on-road breakdowns could be even higher because of under-reporting, including breakdowns that occur on local streets or can be easily relocated to a side street or parking lot. The relevant transport authorities may not be notified of such breakdowns, but they are also less likely to contribute significantly to congestion. On the other hand, on controlled-access highways, tunnels, and bridges, vehicle failure often requires the driver to stop the vehicle in a driving lane or a dedicated breakdown bay. Stopping of vehicles on such roadways contributes significantly to congestion, and these breakdowns are reported immediately to the concerned authorities for rapid response and clearance.

Vehicle breakdowns are non-trivial in number and significantly impact traffic congestion. However, the literature has not yet explored the factors behind vehicle breakdowns at a macro-level. The current study addresses the gap in the literature thanks to a more comprehensive, long-baseline dataset of unplanned incidents that includes on-road breakdowns across the state of NSW, Australia. The main objectives of the study are to investigate the impacts of various variables, including socioeconomic attributes, weather, heavy vehicles, exposure, and temporal attributes, on vehicle breakdown frequency.

2. Study Area and Data Description

2.1. Study Area

New South Wales (NSW), with an estimated population of 7.5 million, is the most populous state in Australia. Sydney, Newcastle, and Wollongong are the three largest cities in the state, accounting for almost 70% of the population of NSW. In order to curb congestion and improve safety, state transport authorities employ various forms of Intelligent Transportation System (ITS) infrastructures, such as Variable Message Signs (VMS), Variable Speed Limit Signs (VSLS), and Vehicle Detection Systems, to manage demand and safety on major roads. Crash fatalities have been declining over the last decade; however, incidents of different types and magnitudes continue to increase [33].

Statistical Area Level 4 (SA4) is the largest spatial unit defined under the Australian Statistical Geography Standard (ASGS). NSW is comprised of 28 SA4s, and each SA4 has a population of at least 100,000 (Figure 1). The SA4 classification was used in this study to model vehicle breakdowns. According to the Household Travel Survey of Sydney [34], most trips are short ones and are mainly for non-work purposes, such as education, shopping, socializing, and entertainment. In addition, the survey also showed that 76% of trips were less than 10 km. Although the average distance of commuting travel (usually the longest trip for a traveler) was around 15 km, it only accounted for approximately 15% of the total number of trips. Therefore, most of the trips happen within the SA4. Therefore, the SA4 aggregation is acceptable for the study, which provides a nexus between the sociodemographic, weather, and infrastructure attributes of the statistical areas and the vehicle breakdowns.

2.2. Data Description

For this study, historical incident data for 4.5 years, i.e., from the 1st of January 2012 to the 30th of June 2016, were obtained for New South Wales, Australia. The dataset includes information on time, location, duration, incident type, incident detection mechanism, incident severity, and a description of the nature of the incident. The dataset contains over 320,000 records of unplanned incidents, including accidents, breakdowns, hazards, police stops, towing, and fires. There are 90,159 records of reported on-road vehicle breakdowns in the dataset. The average breakdown duration is 43 min, the standard deviation is 80 min, and the median value is 24 min. An interesting observation from the dataset is that the breakdowns that occur near the Central Business District (CBD) and urbanized areas tend to have lower durations than the ones that occur in other locations.

Using basic text filtering of the description field, it was observed that 50.8% of breakdowns involved light vehicles (cars, vans, light commercial vehicles, motorcycles, and taxis), 24.8% involved heavy vehicles (buses, trucks, and tankers), and the remaining observations did not specify the vehicle type. According to the Australian Bureau of Statistics (ABS), light vehicles account for 96.6% of the total registered vehicles in NSW, whereas heavy vehicles account for only 3.4% [35]. Furthermore, the total vehicle-kilometers traveled (VKT) by heavy vehicles made up 8.3% of the total kilometers traveled in NSW during a one-year period starting from July 2015. These statistics indicate that despite the low proportionality of heavy vehicles, they are prone to more on-road breakdowns than regular passenger vehicles.

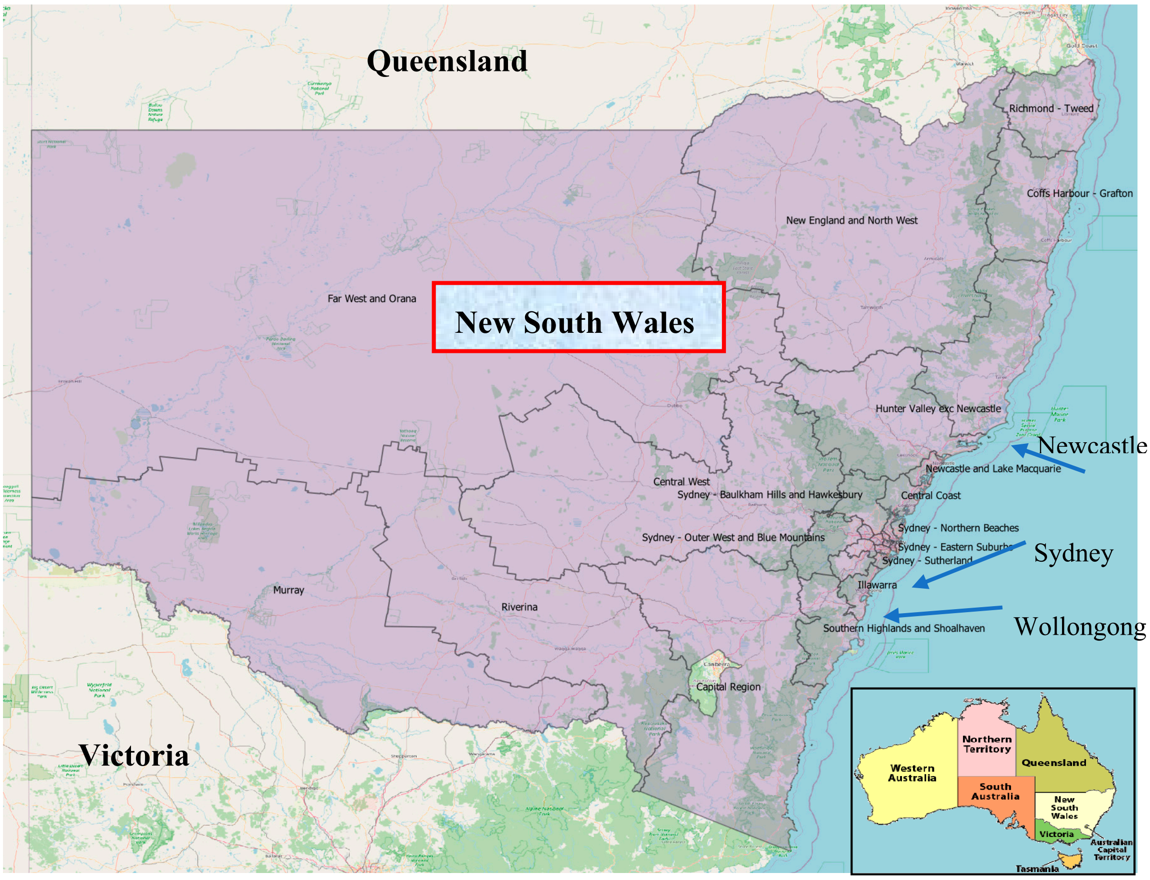

Figure 1. Map of New South Wales (NSW) with Statistical Area Level 4 (SA4) classification (map of Australia in the inset).

Less than 0.3% of the records mentioned fuel or petrol in the text description, so running out of fuel is not a major cause of breakdowns. Furthermore, 24.8% of breakdowns occurred on freeways, 48.6% occurred on non-freeways, and the rest did not identify a primary street for the breakdown location. However, freeways contribute to just 0.54% of the road length in NSW [36,37]. One potential reason for the higher proportion of breakdowns on freeways compared to the road length is the efficient detection of traffic incidents on motorways [38].

The individual breakdown records were arranged in a panel data format, consisting of monthly reported vehicle breakdowns in 28 SA4s in NSW. Data about the weather, holidays, socioeconomic attributes, and other relevant variables were collated from various sources for the model estimation. Only the relevant variables were included in the final model. However, the descriptive statistics of all the potential variables are shown in Table 1. The dependent variable, i.e., the frequency of vehicle breakdowns, ranges from 0 to 416. Furthermore, the variance of the number of breakdowns per month is greater than the observed mean, indicating a potential over-dispersion and suggesting the use of a negative binomial model.

Table 1. Descriptive statistics of all the considered variables.

Variable	Freq.	Mean	Std. Dev.	Min.	Max.
Dependent variable: Frequency of breakdowns	Monthly	60	81	0	416
Exposure variables					
Logarithm of population density (per sq. km.)	TI *	4.70	2.88	−1.04	8.5
Total registered vehicles (in 10,000)	Yearly	17.81	5.74	8.17	34.81
Socioeconomic variables					
Income (in 10,000 AUD)	TI	4.45	0.58	3.46	5.74
Percentage of income earners	TI	51.18	4.47	43.00	59.00
Percentage of managers and professionals	TI	35.26	9.25	23.30	57.50
Percentage of families with children aged less than 15 years	TI	47.54	4.81	35.05	57.41
Percentage of young adults (aged 19–29)	TI	13.39	3.35	8.00	25.00
Percentage of people born overseas	TI	28.06	14.31	11.80	52.40
Percentage of people who speak language other than English at home	TI	18.06	17.66	2.20	59.10
Vehicle characteristics					
Percentage of vehicles aged less than 5 years	Yearly	24.99	4.52	17.00	38.91
Percentage of vehicles aged between 5 and 10 years	Yearly	28.60	2.48	24.14	34.88
Percentage of vehicles aged greater than 5 years	Yearly	46.40	6.44	33.23	57.50
Percentage of heavy vehicles	Yearly	3.50	1.58	1.01	7.27
Percentage of vehicles operated by petrol	Yearly	80.81	7.40	64.29	91.25
Percentage of vehicles operated by diesel	Yearly	17.15	7.40	7.20	35.38
Percentage of vehicles operated by Liquefied Petroleum Gas (LPG)	Yearly	2.02	0.80	1.11	4.22
Weather variables					
Average temperature (°C)	Monthly	17.63	4.65	5.65	27.95
Average rainfall (cm)	Monthly	7.64	3.14	1.95	16.47
Driver licenses					
Percentage of drivers with learners' licenses	Quarterly	5.22	1.41	3.46	9.32
Percentage of drivers with P1 licenses	Quarterly	2.92	0.57	1.92	4.30
Percentage of drivers with P2 licenses	Quarterly	5.24	1.13	3.64	8.91
Percentage of drivers with unrestricted licenses	Quarterly	86.62	2.74	79.77	89.98
Other variables					
Number of school holidays	Monthly	4	5	0	18
Number of public holidays	Monthly	0	1	0	3

* TI—Time invariant.

3. Methodology

To date, the literature has not addressed the impact of various factors on vehicle breakdown frequency. However, there have been several studies in the recent past on macro-level safety models, where spatially aggregated accidents are modeled against area-wide variables. These studies have employed various aggregation levels, such as census tracts, traffic analysis zones, counties, cities, states, and countries [39,40]. Attributes of these spatial aggregations, such as population, density, income, land use characteristics, environmental variables, traffic characteristics, trip generation rates, road density, etc., are typically used to model crashes.

Panel datasets are widely used in the macro-level safety models to observe the effect of spatiotemporal variations of the explanatory variables on crash frequencies [40–42]. The crash counts in a region (city, state, country, etc.) will be correlated over time because the unobserved effects associated with a specific region will remain the same over time [43]. Similarly, there can be correlation over space because regions that are nearby may share unobserved effects. These correlations violate the assumptions of ordinary least squares regression and misestimate the errors on the model coefficients. To account for these correlations, random-effect (RE) and random-parameter (RP) models are considered [40,44–46]. In the case of the RE model, the common unobserved effects are assumed to be distributed across the spatial and temporal units according to some distribution, and shared unobserved effects are assumed to be uncorrelated with explanatory variables [43]. Therefore, the intercept term is represented by a distribution in RE models. In the case of RP models, each estimable parameter (including the intercept) of the model can vary across observations in the dataset. In this regard, the RP model can be considered as a more flexible extension of the RE model.

While RP models account for unobserved heterogeneity and offer a better fit than fixed-parameter models, they are time-consuming and complex to estimate due to the simulation-based likelihood estimation. Furthermore, the analyst is required to select the random parameters and their appropriate distribution. The RP approach may not necessarily improve predictability, and for studies with many explanatory variables, using an RP approach can be computationally intensive due to simulation-based Halton sequences, which are subject to errors in specification because the modeler needs to select the variables with distributed parameters and non-parsimonious because of the many parameters to be estimated [18,41,47,48]. Therefore, in the current study, a random-effect negative binomial model (RENB) is developed to model vehicle breakdowns.

Poisson and negative binomial (NB) models are the most generally espoused approaches for count data modeling. The NB model can handle over-dispersed count data and assumes that counts are independent for an entity for any time. The form of the RENB model is:

$$\lambda_{it} = e^{(\beta X_{it} + \varphi_i + \varepsilon_{it})}, \tag{1}$$

where λ_{it} represents the expected number of breakdowns in a SA4 i in month t; X_{it} is a vector of explanatory variables; β is a vector of estimable parameters; ε_{it} represents the error term for the ith SA4 at time t; φ_i is the unknown intercept term of segment i, which varies with the individual and time such that e^{φ_i} follows a gamma distribution with mean 1 and variance α. To illustrate the variation of the SA4 effect over time, the associated dispersion parameters are not supposed to be constant.

It should be noted that any generalized linear model (GLM), such as a Poisson or a negative binomial regression with a fixed dispersion parameter with random effects, is part of the broader class of generalized linear mixed models (GLMM). The model given in (1) is a specific version using the canonical log link function, but one can use other link functions, such as logit, probit, and complementary log-log [49,50].

It is noted that $1/\theta_i = \alpha_i$, the over-dispersion parameter in the NB model. Furthermore, it is assumed that $\frac{\theta_i}{1+\theta_i}$ or $\frac{1}{1+\alpha i}$ follows a beta distribution with parameters a and b. All the parameters (a, b, and β) are estimated by maximum likelihood techniques [51]. This approach of the RENB model estimation has been used in several studies of crash modeling [52–54]. The joint probability density function can be written for an RENB model as follows:

$$f(y_{it} | X_{it}) = \frac{\Gamma(a+b)\Gamma(a+\sum_T \lambda_{it})\Gamma(b+\sum_T \lambda_{it})}{\Gamma(a)\Gamma(b)\Gamma(a+b+\sum_T \lambda_{it}+\sum_T y_{it})} \prod_T \frac{\Gamma(\lambda_{it}+y_{it})}{\Gamma(\lambda_{it})\Gamma(1+y_{it})}, \tag{2}$$

where y_{it} represents the observed number of breakdowns in SA4 i in month t; T is the number of statistical segments.

As noted by Shankar et al. (1998), the main advantage of this RENB approach is that the over-dispersion parameter is not constrained to be constant across SA4s, as it is in the case of the cross-sectional negative binomial regression [53]. Moreover, a unique characteristic of this formulation is that within-group effects can be allowed to vary over time, even when the exogenous vector of attributes is constant, thereby better accounting for unobserved heterogeneity.

Before estimating the RENB model, a simple multiple linear regression model was estimated by including all the potential independent variables. The variance inflation factors (VIFs) were calculated and the variables with VIFs exceeding 10 were omitted. This was done to address the problem of multicollinearity [55].

Then, for the RENB model, a reverse stepwise approach was used to explore model specification. First, all the variables of interest (identified from the VIF procedure) were included in the model fitting. To address multicollinearity and lack of parsimony, the least significant variables (critical $p = 0.10$ in this study) were sequentially dropped. Furthermore, the goodness-of-fit measures, such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), were evaluated at each step.

4. Results

The results of the final model, including coefficients, standard errors, z-values, and elasticities (only for the continuous variables) of the significant variables, are shown in Table 2. In addition, the log-likelihood, AIC, and BIC values are presented. The positive sign of the coefficient indicates that as the variable increases, the number of breakdowns increases. The elasticity indicates the percent change (+ sign indicates increase and − sign indicates decrease) in the number of breakdowns for a 1% increase in the continuous variable when all other variables are held constant.

Table 2. Random-effect negative binomial (RENB) model results for total vehicle breakdowns.

Variable	Coefficient	Std. Err.	z	Elasticity (% Change)
Year 2012	Fixed	-	-	-
Year 2013	0.167	0.016	10.41	-
Year 2014	0.131	0.020	6.67	-
Year 2015	0.157	0.024	6.57	-
Year 2016	0.218	0.031	7.02	-
Logarithm of population density (per sq. km.)	0.191	0.045	4.26	1.33
Total registered vehicles (in 10,000)	0.068	0.012	5.65	1.40
Average temperature (°C)	0.009	0.001	6.35	0.17
Average rainfall (cm)	−0.017	0.003	−6.25	−0.14
Percentage of managers and professionals	0.051	0.010	5.18	2.17
Percentage of families with children aged less than 15 years	−0.062	0.016	−3.81	−2.87
Percentage of heavy vehicles	0.083	0.028	3.01	0.24
Percentage of drivers with unrestricted licenses	−0.074	0.027	−2.75	−6.21
Number of school holidays	−0.007	0.001	−6.09	−0.70 *
Number of public holidays	0.016	0.007	2.29	1.60 *
Intercept	8.611	2.737	3.15	-
Parameter, *a*	5.241	1.578		
Parameter, *b*	4.137	1.288		
Log-likelihood	−5150			
Akaike Information Criterion (AIC)	10,333			
Bayesian Information Criterion (BIC)	10,424			

* The percentage of change in total breakdowns with an increase of one holiday.

5. Discussion

The results of the RENB model shown in Table 2 offer some interesting insights into the contributing factors and their directions.

5.1. Temporal Variables

Four dummy variables for the year were included in the model to compare with the base year, i.e., 2012. The breakdowns in all the years show a significant increase (see the z-values) compared to the base year, with the most recent year being the highest. This relationship can be partly attributed to a decline in vehicle maintenance skills among drivers and growth in fragile technological components in the vehicles. According to the Royal Automobile Club (RAC) of Britain, owners are less likely to read their vehicle manuals than in the past, and a quarter of breakdown call-outs could be prevented if the owners consulted the manual [56].

Additionally, some technologies, such as DVD systems, keyless electronic ignitions, music players, and satellite navigation, put more strain on battery life, thereby resulting in battery-related vehicle breakdowns. Furthermore, according to the American Automobile Association (AAA), low-profile tires used in the latest vehicles are highly damage-prone and contribute to breakdowns through flat tires [57]. Another potential reason for the increase in the number of reported breakdowns could be accredited to the increased deployment of ITSs across NSW, particularly cameras on motorways, which detect the incidents that might not otherwise have been recorded [38]. This is evident from Figure 2a, showing a steep rise in the number of breakdowns on motorways as compared to a stable pattern of breakdowns on non-motorways (Figure 2b) over the same period. As camera technology becomes more pervasive, the increased reporting of breakdowns is expected for non-motorways as well, and may be suggested in the increase in non-motorway breakdowns from mid-2015.

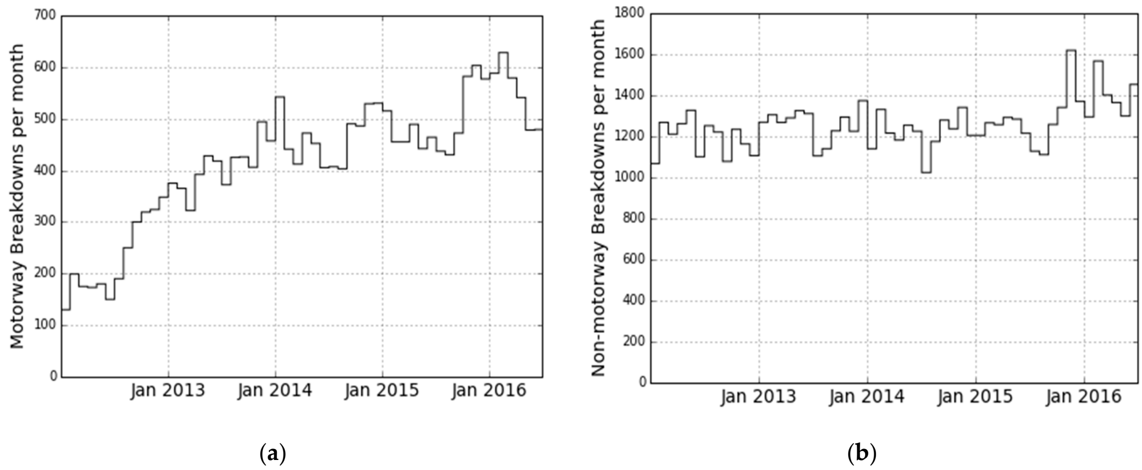

Figure 2. Comparison of increases in breakdowns on motorways (**a**) and non-motorways (**b**).

5.2. Exposure Variables

Population density is an indicator of congestion. Vehicles decelerate and accelerate more often in highly congested areas. These driving patterns put more strain on the vehicle, resulting in an increased expectation of brake failure or engine overheating. The model results support this hypothesis. The elasticity indicates that as the logarithm of population density increases by 1%, the total number of breakdowns in NSW increases by 1.33% when the remaining variables are equal to the original values.

The number of registered vehicles in SA4 acts as a proxy for vehicle use, which might also be measured as vehicle-kilometers traveled. Increased vehicle use is expected to increase chances of a breakdown due to the amount of time spent driving and wear and tear on the vehicle. In addition, the sheer fact that there are more cars in an SA4 could also result in the higher number of breakdowns. The model results indicate that a 1% increase in the number of registered vehicles could lead to a 1.40% increase in the number of on-road breakdowns.

5.3. Socioeconomic Variables

The percentage of families with children aged less than 15 shows a negative effect on the number of breakdowns. A 1% increase in such families could reduce the expected breakdowns by 2.87% when all the other variables are equal. This could potentially be due to the extra caution and care taken by the parents (drivers). Additionally, families are more likely to purchase vehicle makes and models known to be safe and reliable, which could impact the prevalence of vehicle breakdowns.

The increasing proportion of managers and professionals in a region shows a positive impact on breakdowns. These categories of people typically maintain a busy lifestyle with more work-related responsibilities and decreased capacity for vehicle maintenance. AAA states that 35% of Americans have delayed or skipped service or repairs that were recommended by a mechanic, and a significant number of breakdowns each year could be prevented with basic vehicle maintenance [57]. A related explanation is that managers and professionals have higher incomes, which might facilitate the purchase of new, technology-laden vehicles. However, according to the AAA and RAC, newer models are more prone to breakdowns, since they rely heavily on fragile electronic components [56,57]. Furthermore, the combination of technology-laden vehicles and higher mileage (because of the propensity of higher-income earners to travel long distances and make more trips) could lead to a greater number of breakdowns.

There are four categories of driving licenses. However, the number of people with unrestricted licenses is significantly more than any other license type (see Table 1). Only the unrestricted licenses were included in the model. As may be seen in Table 2, the increased percentage of unrestricted driving licenses results in a reduction of breakdowns. Unrestricted drivers are likely to be more experienced in both vehicle maintenance and operation and, due to being older on average, are more able to afford the costs associated with a reliable car and timely service and maintenance.

5.4. Heavy Vehicles

A 1% increase in the percentage of heavy vehicles registered in a region is observed to have a 0.24% rise of breakdown frequency. One contribution is the damage and deterioration caused by heavy vehicles to the roads, which, in turn, could result in more breakdowns. A second contributing reason is that a single heavy vehicle might be used by different drivers, who may not develop sufficient familiarity with the vehicle to monitor changes and address concerns before they evolve into failures. The maintenance strategy is typically post-active for commercial vehicles, such as trucks (primarily) and buses (to an extent), which means that a fault is fixed only after it has occurred [58]. One countermeasure for addressing the congestion caused by breakdowns is to support the adoption of predictive maintenance by heavy vehicle operators, i.e., forecasting that there is a need for maintenance before a vehicle breaks down. Some preventive maintenance recommendations include annual battery testing after it reaches three years of age, monthly tire pressure checks, filling the tank when the fuel level is below 1/4, inspecting battery terminals for corrosion, and multipoint inspections at a repair shop before a long road trip [57].

5.5. Weather Variables

NSW is typically warm, with summer temperatures reaching 40 °C. Increased breakdowns are observed in the dataset with increasing temperature. This can be explained by the overheating of engines and increased use of air conditioners. In contrast, more breakdowns are observed during cold weather in the UK, where heaters are used more often, thus placing extra strain on the vehicle's battery [56].

Rainfall has an adverse effect on the number of breakdowns. This could be due to the reduction in speeds and careful driving by motorists on wet pavements to avoid driving on potholes [59]. Furthermore, adverse weather deters travelers, decreasing vehicle-kilometers traveled (exposure) and, therefore, the expected frequency of breakdowns. However, rainfall also accelerates road damage, such as potholes, which might result in increased breakdowns. Additionally, the amount of rain over a month will operate on the frequency of breakdowns with different mechanisms than the presence of rain at the time of the breakdown, so it is important to consider the relevance of the explanation to the units of observation used in this model. The impact of weather on breakdowns is complex and is ripe for further study.

5.6. Holidays and Other Variables

Congestion is reduced during the school recess, with some commuters going on a vacation and some making fewer trips than usual due to caretaking responsibilities. This reduces the overall VKT exposure and, ultimately, the frequency of breakdowns. Moreover, people going for road trips during school holidays might assess the vehicle's condition to ensure comfort and safety. These factors impact breakdown frequency, as evidenced by the negative sign of the parameter. An additional school holiday every month could reduce the total breakdowns by 0.70% when the other variables do not change.

However, the number of public holidays was found to have a positive effect on breakdowns. The range of public holidays was narrower (minimum = 0 and maximum = 3) than the number of days of school holidays in any month in the dataset (minimum = 0 and maximum = 18). One explanation is that many residents take advantage of public holidays by staying at home and leaving their vehicle stationary. An RAC study found that more breakdowns occur during Monday morning peak hours, as commuters return to their cars after leaving them stationary for the weekend [56]. Public holidays may fill the same role as a weekend, causing vehicles to go unused for one or more days and resulting in an increase in breakdowns when the vehicles are called into action again.

5.7. Insignificant and Correlated Variables Not Considered in the Final Model

The variables indicating the percentage of vehicles by fuel type (petrol, diesel, and LPG) were found to be strongly correlated with other variables and, thus, had to be omitted in the final model. For example, LPG vehicles were positively correlated with population density. The majority of taxis in NSW use LPG, and there will be more taxis registered and operating in high-density areas. Furthermore, heavy vehicles use diesel-operated engines, and strong positive correlation occurs between these variables.

Furthermore, the percentage of vehicles by age was omitted because of correlation issues. For example, new vehicles are negatively correlated with heavy vehicles and old vehicles are negatively correlated with population density.

The information on the percentage of people born overseas or percentage of people who speak a language other than English at their homes was collected to test hypotheses related to the role of migrant status. Immigrants may be unfamiliar with the geography of NSW or may be less connected to social networks. These attributes might cause these drivers to exercise additional caution by checking and maintaining their vehicles regularly. However, they were insignificant and, also, as one would expect, highly correlated with each other. Furthermore, they were individually correlated with population density because high-density areas are associated with diverse populations.

The income and percentage of income earners were highly correlated with the percentage of managers and professionals. Furthermore, the percentage of young adults was negatively correlated with the percentage of unrestricted driving licenses and also the percentage of managers and professionals. The fraction of managers and professionals offers better explanatory power in the model as well as better interpretability, since this information captures a complex set of attributes surrounding income, career factors, and lifestyle.

6. Suggestions to Reduce Breakdowns

As cities grow in size, the average vehicle kilometers travelled increase. This leaves drivers spending an increasing fraction of their time either driving or working, leaving less time available for discretionary activities, such as vehicle monitoring and maintenance. Furthermore, changing automobile technology increases the reluctance among motorists to undertake vehicle maintenance themselves, as the technology has become more electronic and complex. One intervention to manage the impact of breakdowns on congestion is to educate drivers about vehicle maintenance procedures and also encourage them to read the vehicle manuals thoroughly. Another intervention is to initiate questions regarding routine maintenance and minor repair procedures during the vehicle registration process. This strategy takes advantage of the existing interaction between vehicle owners and the local transport authority in order to investigate and internalize the costs associated with preventable breakdowns. This proactive approach could reduce the breakdown frequency and durations, and could also reduce the unnecessary call-outs to roadside assistance companies.

One factor touched upon in these model results is the potential contribution of vehicle reliability. Some vehicle attributes, such as newer models with vulnerable electronic components, reduce the reliability of the vehicle. On the other hand, some vehicle makes and models have earned reputations for reliability. Providing information on breakdown frequency and magnitude is a valuable step towards educating purchasers about the potential risks associated with each vehicle—in the same way that vehicle manufacturers must report on the fuel economy of their products, they could be required to identify each model's vulnerability to breakdown. This kind of reporting is supported with the growing availability of relevant data, including big data analysis techniques and improved ITS deployment.

7. Conclusions and Future Work

Non-recurrent congestion that is caused by sudden and unpredictable occurrence of crashes, breakdowns, hazards, and adverse weather is non-trivial, and, in fact, constitutes a signficant proportion of total congestion delay. Numerous past studies have focused mainly on a single type of incident, i.e., crashes in terms of their duration, frequency, severity, and rate. Only a handful of studies have evaluated other types of incidents, which is likely due to the lack of availability of data at a large scale. In this context, the current study has been set up with an objective to understand the factors impacting vehicle breakdown frequency at a macro-level. This research was possible thanks to a large dataset of traffic incidents comprising 90,159 records over a period of 4.5 years (from January 2012 to June 2016) in the state of New South Wales, Australia.This study provides insight on various factors contributing towards regional-level breakdowns, and various interventions are suggested for the management of breakdowns and congestion. The results from this study can be employed by transport authorities, roadside assistance companies, automobile manufacturers, and mechanics to manage their involvement in on-road vehicle breakdowns. The scope of the study can be extended by including information on roadway characteristics and driving habits of the people at the SA4 level and also by better accounting for unobserved heterogeneity. The electric vehicle market is trivial in Australia, and so its impacts on breakdown frequency have not been studied in this paper.

Author Contributions: Conceptualization, S.C. and E.M.; methodology, S.C.; software, S.C.; validation, S.C.; formal analysis, S.C.; investigation, S.C. and E.M.; resources, E.M., S.T.W. and V.D.; data curation, S.C. and E.M.; writing—original draft preparation, S.C.; writing—review and editing, S.C., E.M., S.T.W. and V.D.; visualization, S.C.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors thank the three anonymous reviewers and the Editor for their thorough comments and feedback which improved the quality of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nair, D.J.; Gilles, F.; Chand, S.; Saxena, N.; Dixit, V. Characterizing multicity urban traffic conditions using crowdsourced data. *PLoS ONE* **2019**, *14*, e0212845.
2. Skabardonis, A.; Varaiya, P.; Petty, K.F. Measuring Recurrent and Nonrecurrent Traffic Congestion. *Transp. Res. Rec. J. Transp. Res. Board* **2003**, *1856*, 118–124. [CrossRef]

3. Ma, X.; Ding, C.; Luan, S.; Wang, Y.; Wang, Y. Prioritizing Influential Factors for Freeway Incident Clearance Time Prediction Using the Gradient Boosting Decision Trees Method. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 2303–2310. [CrossRef]
4. Neudorff, L.G.; Randall, J.; Reiss, R.A.; Gordon, R.L. *Freeway Management and Operations Handbook*; United States Federal Highway Administration, Office of Transportation: Washington, DC, USA, 2003.
5. Yin, Y.; Lam, W.H.; Ieda, H. New technology and the modeling of risk-taking behavior in congested road networks. *Transp. Res. Part C Emerg. Technol.* **2004**, *12*, 171–192. [CrossRef]
6. Yin, Y.; Ieda, H. Assessing Performance Reliability of Road Networks Under Nonrecurrent Congestion. *Transp. Res. Rec. J. Transp. Res. Board* **2001**, *1771*, 148–155. [CrossRef]
7. De Palma, A.; Rochat, D. Understanding individual travel decisions: Results from a commuters survey in Geneva. *Transportation* **1999**, *26*, 263–281. [CrossRef]
8. Chand, S.; Aouad, G.; Dixit, V.V. Long-Range Dependence of Traffic Flow and Speed of a Motorway: Dynamics and Correlation with Historical Incidents. *Transp. Res. Rec. J. Transp. Res. Board* **2017**, *2616*, 49–57. [CrossRef]
9. Tu, H.; van Lint, H.; van Zuylen, H. The effects of traffic accidents on travel time reliability. In Proceedings of the 2008 11th International IEEE Conference on Intelligent Transportation Systems, Beijing, China, 12–15 October 2008; pp. 79–84.
10. Park, H.; Haghani, A.; Samuel, S.; Knodler, M.A. Real-time prediction and avoidance of secondary crashes under unexpected traffic congestion. *Accid. Anal. Prev.* **2018**, *112*, 39–49. [CrossRef]
11. Park, H.; Haghani, A. Real-time prediction of secondary incident occurrences using vehicle probe data. *Transp. Res. Part C Emerg. Technol.* **2016**, *70*, 69–85. [CrossRef]
12. Ranney, T.A. *Driver Distraction: A Review of the Current State-of-Knowledge*; National Highway Traffic Safety Administration (NHTSA): Washington, DC, USA, 2008.
13. Khattak, A.; Wang, X.; Zhang, H. Incident management integration tool: Dynamically predicting incident durations, secondary incident occurrence and incident delays. *IET Intell. Transp. Syst.* **2012**, *6*, 204–214. [CrossRef]
14. Afrin, T.; Yodo, N. A Survey of Road Traffic Congestion Measures towards a Sustainable and Resilient Transportation System. *Sustainability* **2020**, *12*, 4660. [CrossRef]
15. Black, W.R. *Sustainable Transportation: Problems and Solutions*; Guilford Press: New York, NY, USA, 2010.
16. Haque, M.; Chin, H.C.; Debnath, A.K. Sustainable, safe, smart—three key elements of Singapore's evolving transport policies. *Transp. Policy* **2013**, *27*, 20–31. [CrossRef]
17. Mannering, F.L.; Bhat, C.R. Analytic methods in accident research: Methodological frontier and future directions. *Anal. Methods Accid. Res.* **2014**, *1*, 1–22. [CrossRef]
18. Pande, A.; Chand, S.; Saxena, N.; Dixit, V.; Loy, J.; Wolshon, B.; Kent, J.D. A preliminary investigation of the relationships between historical crash and naturalistic driving. *Accid. Anal. Prev.* **2017**, *101*, 107–116. [CrossRef] [PubMed]
19. Chand, S.; Dixit, V.V. Application of Fractal theory for crash rate prediction: Insights from random parameters and latent class tobit models. *Accid. Anal. Prev.* **2018**, *112*, 30–38. [CrossRef]
20. Milton, J.C.; Shankar, V.N.; Mannering, F.L. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accid. Anal. Prev.* **2008**, *40*, 260–266. [CrossRef]
21. Al Kaabi, A.; Dissanayake, D.; Bird, R. Response Time of Highway Traffic Accidents in Abu Dhabi. *Transp. Res. Rec. J. Transp. Res. Board* **2012**, *2278*, 95–103. [CrossRef]
22. Hu, L.; Bao, X.; Wu, H.; Wu, W. A Study on Correlation of Traffic Accident Tendency with Driver Characters Using In-Depth Traffic Accident Data. *J. Adv. Transp.* **2020**, *2020*, 9084245. Available online: https://www.hindawi.com/journals/jat/2020/9084245/ (accessed on 1 October 2020). [CrossRef]
23. Kumar, S.; Toshniwal, D. A data mining approach to characterize road accident locations. *J. Mod. Transp.* **2016**, *24*, 62–72. [CrossRef]
24. Li, X.; Lord, D.; Zhang, Y.; Xie, Y. Predicting motor vehicle crashes using Support Vector Machine models. *Accid. Anal. Prev.* **2008**, *40*, 1611–1618. [CrossRef]
25. Munyazikwiye, B.B.; Karimi, H.R.; Robbersmyr, K.G. Fuzzy logic approach to predict vehicle crash severity from acceleration data. In Proceedings of the 2015 International Conference on Fuzzy Theory and Its Applications (iFUZZY), Yilan, Taiwan, 18–20 November 2015; pp. 44–49.

26. Delen, D.; Sharda, R.; Bessonov, M. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accid. Anal. Prev.* **2006**, *38*, 434–444. [CrossRef] [PubMed]
27. Comi, A.; Persia, L.; Nuzzolo, A.; Polimeni, A. Exploring temporal and spatial structure of urban road accidents: Some empirical evidences from Rome. In *Data Analytics: Paving the Way to Sustainable Urban Mobility*; Advances in Intelligent Systems and Computing; Nathanail, E.G., Karakikes, I.D., Eds.; Springer: Cham, Switzerland, 2019; pp. 147–155.
28. Yannis, G.; Antoniou, C.; Papadimitriou, E. Autoregressive nonlinear time-series modeling of traffic fatalities in Europe. *Eur. Transp. Res. Rev.* **2011**, *3*, 113–127. [CrossRef]
29. Meng, Q.; Weng, J. A Genetic algorithm approach to assessing work zone casualty risk. *Saf. Sci.* **2011**, *49*, 1283–1288. [CrossRef]
30. Prato, C.G.; Bekhor, S.; Galtzur, A.; Mahalel, D.; Prashker, J. Exploring the potential of data mining techniques for the analysis of accident patterns. In Proceedings of the 12th WCTR Conference, Lisbon, Portugal, 11–15 July 2010.
31. Wang, W.; Chen, H.; Bell, M.C. Vehicle Breakdown Duration Modelling. *J. Transp. Stat.* **2005**, *8*, 75–84.
32. Hojati, A.T.; Ferreira, L.; Charles, P.; bin Kabit, M.R. Analysing freeway traffic-incident duration using an Australian data set. *Road Transp. Res. J. Aust. New Zealand Res. Pract.* **2012**, *21*, 19.
33. Transport for New South Wales. *NSW Road Fatalities Report*; Centre for Road Safety, Transport for New South Wales (TfNSW): Sydney, Australia, 2020.
34. *Transport for New South Wales Household Travel Survey Report: Sydney 2012/13*; Bureau of Transport Statistics, Transport for New South Wales (TfNSW): Sydney, Australia, 2014.
35. ABS. *Survey of Motor Vehicle Use, Australia*; Australian Bureau of Statistics: Australian Capital Territory, Australia, 2017.
36. New South Wales Government. NSW Spatial Data Catalogue. 2015. Available online: https://sdi.nsw.gov.au/nswsdi/catalog/search/resource/details.page?uuid=%7B10377F94-15B6-4F6D-8608-FEBCC59E373D%7D (accessed on 28 July 2020).
37. Transport for New South Wales. NSW Road Network. 2020. Available online: https://www.transport.nsw.gov.au/node/4888 (accessed on 27 July 2020).
38. Moylan, E.K.M.; Chand, S.; Waller, T. Framework for Estimating the Impact of Camera-Based Intelligent Transportation Systems (ITS) Technology on Incident Duration. *Transp. Res. Rec. J. Transp. Res. Board* **2018**, *2672*, 25–33. [CrossRef]
39. Lee, D.; Guldmann, J.-M.; Choi, C. Factors Contributing to the Relationship between Driving Mileage and Crash Frequency of Older Drivers. *Sustainability* **2019**, *11*, 6643. [CrossRef]
40. Truong, L.T.; Kieu, L.-M.; Vu, T.A. Spatiotemporal and random parameter panel data models of traffic crash fatalities in Vietnam. *Accid. Anal. Prev.* **2016**, *94*, 153–161. [CrossRef]
41. Chen, E.; Tarko, A.P. Modeling safety of highway work zones with random parameters and random effects models. *Anal. Methods Accid. Res.* **2014**, *1*, 86–95. [CrossRef]
42. Noland, R.B.; Quddus, M.A. Analysis of Pedestrian and Bicycle Casualties with Regional Panel Data. *Transp. Res. Rec. J. Transp. Res. Board* **2004**, *1897*, 28–33. [CrossRef]
43. Lord, D.; Mannering, F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transp. Res. Part A Policy Pract.* **2010**, *44*, 291–305. [CrossRef]
44. Coruh, E.; Bilgic, A.; Tortum, A. Accident analysis with aggregated data: The random parameters negative binomial panel count data model. *Anal. Methods Accid. Res.* **2015**, *7*, 37–49. [CrossRef]
45. Li, D.; Li, C.; Miwa, T.; Morikawa, T. An Exploration of Factors Affecting Drivers' Daily Fuel Consumption Efficiencies Considering Multi-Level Random Effects. *Sustainability* **2019**, *11*, 393. [CrossRef]
46. Xu, P.; Huang, H. Modeling crash spatial heterogeneity: Random parameter versus geographically weighting. *Accid. Anal. Prev.* **2015**, *75*, 16–25. [CrossRef]
47. Shugan, S.M. Editorial: Errors in the Variables, Unobserved Heterogeneity, and Other Ways of Hiding Statistical Error. *Mark. Sci.* **2006**, *25*, 203–216. [CrossRef]
48. Washington, S.P.; Karlaftis, M.G.; Mannering, F. *Statistical and Econometric Methods for Transportation Data Analysis*; CRC press: Boca Raton, FL, USA, 2010.
49. Bhowmik, T.; Yasmin, S.; Eluru, N. A multilevel generalized ordered probit fractional split model for analyzing vehicle speed. *Anal. Methods Accid. Res.* **2019**, *21*, 13–31. [CrossRef]

50. STATA. Multilevel Generalized Linear Models|Stata. 2020. Available online: https://www.stata.com/features/overview/multilevel-generalized-linear-models/ (accessed on 28 July 2020).
51. Hausman, J.; Hall, B.; Griliches, Z. Econometric Models for Count Data with an Application to the Patents-R&D Relationship. *Econometrica* **1984**, *52*, 909–938.
52. Chin, H.C.; Quddus, M.A. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. *Accid. Anal. Prev.* **2003**, *35*, 253–259. [CrossRef]
53. Shankar, V.N.; Albin, R.B.; Milton, J.C.; Mannering, F. Evaluating Median Crossover Likelihoods with Clustered Accident Counts: An Empirical Inquiry Using the Random Effects Negative Binomial Model. *Transp. Res. Rec. J. Transp. Res. Board* **1998**, *1635*, 44–48. [CrossRef]
54. Yan, Y.; Zhang, Y.; Yang, X.; Hu, J.; Tang, J.; Guo, Z. Crash prediction based on random effect negative binomial model considering data heterogeneity. *Phys. A Stat. Mech. Its Appl.* **2020**, *547*, 123858. [CrossRef]
55. Gujarati, D.N. *Basic Econometrics*; Tata McGraw-Hill Education: New York, NY, USA, 2009.
56. RAC. *Breakdown Britain*; Royal Automobile Club: Walsall, UK, 2006.
57. AAA. *Fact Sheet-Preventive Maintenance*; American Automobile Institution: Heathrow, FL, USA, 2015.
58. Prytz, R.; Nowaczyk, S.; Rögnvaldsson, T.; Byttner, S. Analysis of Truck Compressor Failures Based on Logged Vehicle Data. In Proceedings of the 9th International Conference on Data Mining, Las Vegas, NV, USA, 22–25 July 2013.
59. Dixit, V.V.; Gayah, V.V.; Radwan, E. Comparison of Driver Behavior by Time of Day and Wet Pavement Conditions. *J. Transp. Eng.* **2012**, *138*, 1023–1029. [CrossRef]