Weather Effects on Daily Traffic Accidents and Fatalities:
A Time Series Count Data Approach

By George Yannis and Matthew G. Karlaftis
Department of Transportation Planning and Engineering
National Technical University of Athens
Athens, Greece
Tel: +30.210.7721326; +30.210.7721280
Email: geyannis@central.ntua.gr; mgk@central.ntua.gr

Abstract
The impact of weather conditions on traffic safety is a topic that has attracted considerable interest in the literature. In this research, an integer autoregressive model (INAR) is used to estimate the effects of weather conditions on four traffic safety categories: vehicle accidents, vehicle fatalities, pedestrian accidents and pedestrian fatalities, using 21 years of daily count data for Athens, Greece. The results suggest that the most consistently significant and influential variable is mean daily precipitation height along with its lagged value. It is found that, contrary to much previous research, increases in rainfall reduce the total number of accidents and fatalities as well as the pedestrian accidents and fatalities, a finding that may be attributed to the safety offset hypothesis resulting from more cautious and less speedy driver behaviour. Similarly, temperature increase was found to lead to increased accidents.

Keywords: time series count data, traffic accidents, casualties, pedestrians

1. Introduction

Weather conditions, such as air temperature and precipitation, are associated with considerable impacts on road safety, mainly through their influence on both the exposure and the behavior of road users. The interaction between weather effects and the effects of other road safety factors, including roadway, driver, vehicle and intervention variables on road accident frequency is certainly a complex phenomenon that attracts increasing attention by researchers. Stipdonk (2008) underlines that weather effects need to be controlled for in any multivariate analysis aiming to explain changes in road safety outcomes. Koetse and Rietveld (2009) further emphasize this need within the climate change context.

Several studies included or focused on the effects of weather conditions on road accidents occurrence and severity, attempting to capture these often complex effects. A thorough review of mostly earlier studies on weather effects can be found in Eisenberg (2004).

Weather condition data relevant to road accidents are typically recorded at the accident scene as the prevailing conditions during the accident. In several studies, weather conditions during the accident are associated with the accident outcomes usually through the calculation of casualty risk ratios using a control group (Ivey et al. 1981; Brodsky and Hakkert 1988; Majdzadeh et al. 2008), and the results indicate increased casualty risks in adverse weather conditions. In several cases, particular groups of drivers such as older drivers (Baker et al. 2003), motorcyclists (Pai and Saleh 2008) and truck drivers (Young and Liesman 2007) are examined. Golob and Recker (2004) found unique profiles in terms of the type of accidents that are most likely to occur in different weather and traffic conditions.

On the other hand, continuous time series of meteorological data are generally gathered by means of permanent and appropriately localized measurement stations. Several studies used such data for road safety analysis, within both spatial and time series frameworks. In the first case, the spatial distribution of road accident counts is associated with meteorological phenomena. Edwards (1996) and Khan et al. (2008) showed that the occurrence of road accidents in hazardous weather conditions (rainfall, fog, snowfall and wind) broadly follows the regional weather patterns for those conditions. Geurts et al. (2005) reported a significant spatial association of road accidents at hazardous locations with rainfall. Regarding accident severity, a positive spatial effect of fog in rural areas and a negative overall spatial effect of rainfall were identified (Edwards, 1998). Aguero-Valverde and Jovanis (2006) introduced hierarchical models with random spatial and time effects and found that rainfall may increase road accident frequencies.

However, a large part of existing research has involved time series data that may capture both global and seasonal effects. These studies are summarized in Table 1; they range from yearly to daily analyses and from national to local level, while they use approaches ranging from generalized linear modeling techniques (i.e. Poisson-family models) to advanced, dedicated time series analysis techniques. Moreover, several additional variables are often controlled for, such as exposure, roadway design, demographics and interventions.

Higher temperatures appear to have a decreasing effect on accident frequencies and severity both at daily, weekly and monthly bases (Scott, 1986; Brijs et al. 2008). The hours of sunlight appear to increase road accidents (Hermans et al. 2006; Brijs et al. 2007), while deviations from mean daily or monthly temperatures were found to increase road accidents (Brijs et al. 2008; Stipdonk 2008). Malyshkina et al. (2008) found that extreme temperatures (both low during winter and high during summer) are positively correlated with road accidents; on the other hand, when the monthly number of days with temperature below zero increases, road accidents are reduced possibly due to reduced exposure (Hermans et al. 2006; Stipdonk 2008).

Findings regarding rainfall are extensive and quite consistent. Increased daily, monthly or even yearly rainfall appears to increase accident frequencies (Fridstrom & Ingebrigtsen 1991; Fridstrom et al. 2005; Chang and Chen 2005; Caliendo et al. 2007). A similar effect is obtained when examining the monthly number of days with rainfall (Shankar et al. 1995; Keay & Simmonds 2006; Hermans et al. 2006). Brijs et al. (2007) proposed a rainfall intensity indicator, defined as the centimeters of rainfall divided by its duration, which was found to increase the daily number of accidents. Further, lagged effects of rainfall (and precipitation in general) are often investigated; Eisenberg (2004) showed that the impact of precipitation on a given day is reduced when precipitation was observed in the previous days. Similar to this, Brijs et al. (2008) found that, the longer a "dry spell" (i.e. days from the previous rainfall), the higher the number of accidents in rainfall.

In several studies, it was possible to interpret the positive effect of rainfall on road accidents. Keay and Simmonds (2006) showed that increased rainfall in centimeters results in decreased daily traffic volume, both at daytime and nighttime, winter and spring. Bergel-Hayat and Depire (2004) decomposed the global effect of monthly rainfall in two components: a direct effect on the number of injury accidents and fatalities, and an indirect effect on traffic volume. In Stipdonk (2008), the indirect effect was confirmed, leading to a recommendation for estimating weather effects on road accidents under constant traffic conditions. Further, they also suggest that reduced traffic may lead to increased travel speeds that result in increased accident risk.

In Table 1, one can notice that the variables used to express each meteorological factor are quite diverse and in a few cases different results are obtained. For example, temperature may express either heat or frost conditions, whereas precipitation mainly refers to rainfall. Depending on the specification of the variables in each case, a correlation between temperature and precipitation variables may be more or less pronounced. In general, results from previous studies are rather consistent with regards to rainfall effects, but somewhat less consistent with regards to temperature effects. It is also important to note that, although most existing studies control for exposure, either through traffic measurements or through a proxy measures (e.g. petrol sales, vehicle fleet, and so on), only in few studies are the weather effects interpreted through their effects on exposure.

The objective of this research is to further investigate the impact of weather conditions on traffic safety using detailed data for daily accident, traffic and weather information. An integer autoregressive model (INAR) is used for the estimation of the effects of weather conditions on four traffic safety categories: vehicle accidents, vehicle fatalities, pedestrian accidents and pedestrian fatalities, using 21 years of daily count data for Athens, Greece.

2. Data and Methodology

2.1 The Data

In order to meet the research objectives, a large data set of daily traffic accident and weather conditions was used. Daily numbers of fatalities and injury accidents were used, together with the number of pedestrians killed or injured in these accidents. These data concern the region of Athens, the capital of Greece, with a population of 3,13 million inhabitants and an area of 411 km^2. The daily data set refers to a complete 21 year period from 1.1.1985 to 31.12.2005, a total of 7670 cases. These information were extracted from the database with disaggregate data maintained at the National Technical University of Athens based on data collected by the police and coded by the National Statistical Service of Greece.

In addition, the respective weather condition data set was used, referring to mean daily temperature (in Celsius degrees) and mean daily precipitation height (in cm) for exactly the same period (1.1.1985 to 31.12.2005). These data were extracted from the data file with disaggregate data of the National Observatory of Athens and concerned data from a meteorological station representative of the weather conditions of Athens.

2.2 The Methodology

Much of the early work on the empirical analysis of accident data was done with the use of multiple linear regression models; as is well known, these models suffer from several methodological limitations and practical inconsistencies which have been pointed out repeatedly in the literature (see for example Washington et al. 2003). To overcome these limitations, several authors used Poisson regression models which are a reasonable alternative for events that occur randomly and independently over time. The Poisson model has a number of advantages over the normal regression model when dealing with *count* data (as, for example, accident count data). First, linear regression assumes a normal distribution of the dependent variable, an assumption which does not hold with count (accident) data. The Poisson model, on the other hand, recognizes the discrete nature of accident counts. Second, linear regression may produce negative estimates for the dependent variable, which is incorrect for accident counts.

In many cases during safety investigations, data are available on a time series dimension, i.e. the variables examined are available over a (long) period of time. A time series of count data is an integer value non-negative sequence of count observations over time. Several models for the analysis of time series of count data are available, but the INARMA class of models - evolved similarly to the continuous

ARMA models – have found wide applications in many research areas.[1,2] The most commonly encountered form of the INARMA model is the INAR(1) process that can be defined as

$$y_t = \alpha \circ y_{t-1} + \varepsilon_t, \; t = 2,...,T. \qquad (1)$$

where $\{\varepsilon_t\}$ is assumed i.i.d. Poisson with $E(\varepsilon_t) = \lambda > 0$ and independent of y_{t-1}; further, as McKenzie (1985) has shown, when $\alpha \in (0,1)$ and y_t is discrete self-decomposable, the AR(1) process is stationary. This model follows the 'usual' AR(1) model in that it explicitly models serial correlation as lags of the endogenous variables, but where the scalar multiplication is replaced by a binomial thinning operator (α). The operator, introduced by Van Harn and Steutel (1977), can be defined as $\alpha \circ y = \sum_{i=1}^{y} u_i$, where u_i is a sequence of binary random variables where each component *i,* either 'survives' with probability α (i.e. $u_i = 1$) or does 'not survive' with probability $(1-\alpha)$. This model form makes five basic assumptions: i. $E(u_i u_j) = E(u_i)E(u_j)$; ii. $E(u_i \varepsilon_t) = E(u_i)E(\varepsilon_t)$; iii. $E(u_i y_{t-1}) = E(u_i)E(y_{t-1})$; iv. $Cov(\varepsilon_t \varepsilon_s) = 0, \forall\, t \neq s$; and, v. $E(\varepsilon_t y_{t-1}) = E(\varepsilon_t)E(y_{t-1})$.[3]

A number of extensions for this basic model have been developed; these include general INARMA models (McKenzie, 1986; Al-Osh and Al-Zaid, 1991) and the INAR(p) model (Al-Zaid and Al-Osh, 1990; Jin-Guan and Yuan, 1991), while a finite mixture version of the Poisson regression was developed by Böckenholt (1998). However, among the most important developments from an empirical perspective was Brannas' (1995) extension to include explanatory variables in the basic model. In the case of safety analyses, λ may represent the mean monthly vehicle accident rate that depends on various characteristics such as weather, traffic and so on, that may in-turn also vary with time. Brannas (1995) suggests that explanatory variables could be introduced to the model as $\alpha_t \in (0,1)$ and $\lambda_t > 0$ which, using the logistic and exponential distributions, can be given as $\alpha_t = 1/\left[1 + \left(e^{\mathbf{x}_t\boldsymbol{\beta}}\right)\right]$ and $\lambda_t = e^{\mathbf{z}_t\boldsymbol{\gamma}}$.[4] Using these two modifications, Eq. (1) can be rewritten as[5]

$$y_t = \alpha_t \circ y_{t-1} + \varepsilon_t, \; t = 1,...,T. \qquad (2)$$

---

[1] Examples of time series model applications in the area of transportation and safety include Brijs et al. (2008), Levine et al. (1995), Miaou and Lord (2003), Shankar et al. (1998), Ulfarsson and Shankar (2003), Quddus (2007).

[2] The Poisson AR(1) model was first developed by Al-Osh and Al-Zaid (1987) and McKenzie (1985) and later generalized by Brannas (1994, 1995) and Joe (1996); see Brannas and Hellstrom (2002) for a survey.

[3] First and second order moments and properties for this model can be found in Brannas (1994 and 1995).

[4] Vectors $\mathbf{x}_t$ and $\mathbf{z}_t$ are considered fixed and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are vectors of parameters to be estimated.

[5] Moment relations for this model are given in Brannas (1995).

In this paper we use the model of Eq. (2) to estimate the effects of, primarily, weather on a variety of safety measures (such as vehicle and pedestrian accidents and casualties) based on a data base of 21 years of accidents measured on a daily basis.[6]

3. Empirical Estimation

We use the model described in the previous section (Eq. (2)) to estimate integer time series models for 21 years of data on four incident categories: vehicle accidents, vehicle fatalities, pedestrian accidents and pedestrian fatalities. The independent variables in all models are related to weather (mean daily temperature and mean daily precipitation height), traffic (through proxies for day of the week) and time parameters (monthly and annual dummy variables).[7] In preliminary investigations we examined and tested for some of the fundamental characteristics of the series; first, we examined the Poisson assumption for the data. In all four cases, overdispersion (over Poisson variance) in the data is small and the INAR(1) model can be used to describe the data (Böckenholt 1998 and Brijs et al. 2008).[8] Second, we examined the underlying correlation structure in the data; although it may be reasonable to a-priori assume the existence of significant correlation between successive incident counts, we employed the parametric tests in Jung and Tremayne (2001) that showed the existence of strong low order dependence. Third, we examined the appropriateness of the INAR(1) specification for the model. We used the Dickey-Fuller test of a random walk with Poisson distributed errors and rejected the null hypothesis of non-stationarity in favor of the stationary alternative (Hellstrom 2001); further, the tests of Jung and Tremayne (2001) employed in the previous step indicated the appropriateness of an AR process rather than an MA (Moving Average) or a mix ARMA process.[9]

Estimation results for all incident types appear in Table 2 (t-stats correspond to coefficient estimates divided by their asymptotic standard errors). In all four models (one for each dependent variable), the most consistently significant and influential variable is mean daily precipitation height along with its lagged value. This is an important finding particularly as it pertains to its consistently negative sign; while much previous research has determined the significance of rainfall to accident prediction (Fridstorm et al. 1995, Eisenberg 2004, Brijs et al. 2008), most research has indicated that increases in precipitation lead to increases in vehicle accidents. In this work we find the opposite result; that is, we find that increases in rainfall reduce all

---

[6] Here we present only the essential parts of the model specification; readers interested in additional information including testing and estimation issues can refer to Brijs et al. (2008) for an excellent discussion on the topic.

[7] It is well established that, in accident models, a measure of exposure – traffic volume most frequently – must be included in estimations, but we were not able to collect reliable or representative traffic data for our study area. However, as the literature has repeatedly indicated, day-of-the-week dummies are well suited for capturing variability in exposure and yield consistent estimation results and we thus use this approach to overcome the lack of direct exposure data (for more details on using dummy variables for capturing exposure variability see Brijs (2008), Martin (2002) and Levine et al. (1995)).

[8] We note here that the INAR model used in this paper makes the assumption of Poisson marginal distributions and does not account for dispersion; extensions that consider overdispersion can be found in Karlis and Xekalaki (2001) and Gourieroux and Jasiak (2004).

[9] We also note that as, Brijs et al. (2008) and Jin-Guan and Yuan (1991) showed, although it is straightforward to incorporate higher order lags into the INAR model, coefficient interpretation becomes complicated.

types of incidents. Although this appears counter to what would seem logical, it may be attributed to the offset hypothesis by which people have an acceptance of a level of safety; if they feel less safe, because of rainfall in this case, then they may drive with lower speed and more carefully to compensate; further, this may also be a characteristic of Southern European drivers who are not accustomed to driving in wet conditions and thus become overly cautious when raining. Finally, we note that our findings largely support Eisenberg's (2004) hypothesis of the existence of a significant lagged effect between precipitation and accidents.

The second weather related variable we examined concerns temperature's effects on incidents. Research has shown temperature to be an important factor in determining car accidents (Branas and Knudson 2001, Brijs et al. 2008, Fridstrom et al. 1995), particularly for temperatures below freezing and when combined with snow. Our results show mean daily temperature and its one period lag to be important determinants of vehicle accidents, where an increase in temperature leads to increased accidents; we do note, however, that our data did not contain any mean daily temperature readings below freezing while snowfall is very sparse. Again, the offset hypothesis may be an important factor in this finding. The 3-day moving temperature average was significant in determining pedestrian injuries (increase in average temperature increases pedestrian injuries); we attribute this result to the corresponding increase in pedestrian traffic associated with higher temperatures.

As previously noted, lack of dependable traffic data required the use of day-of-the-week dummy variables to capture traffic's variability. The results indicate that, for most categories, incidents are higher during the weekends than during the week, with the exception of pedestrian fatalities which appear to be higher on Mondays. The same general finding applies for Fridays, with the exception of pedestrian injuries. Although this finding is different from previous research (see, for example, Brijs 2008), it can be directly attributed to much higher – and higher risk - traffic during Friday and weekend nights.

In estimating the models, dummy variables were used to capture some of the temporal characteristics in the series. The monthly dummy variables (most of which were statistically significant at the 90% level) indicated that, for example, vehicle accidents were much lower in August (when traffic is also much lower due to an extensive vacation period) and were higher in January and May (these results confirm monthly accident statistics appearing in Table 3). The annual dummy variables were largely statistically significant without, however, indicating a significant trend (either upward or downward), and were thus left in the model as individual annual dummy variables (in Table 3 we only mention the number of dummy variables – out of a total of 21 – that were not significant). Finally, for each model we estimated the MAPE (Mean Absolute Percent Error measured in %) and MAE (Mean Absolute Error, measured in number of incidents); the MAPE values range from 17.6% in the case of vehicle accidents to 37.3% for pedestrian fatalities. These MAPE values are rather high – particularly for both pedestrian models - and certainly suggest the existence of additional exogenous factors that affect the phenomena studied and that have not been included in the models. Finally, Figures 3 – 5 depict actual versus predicted values for some of the models (ρ is the correlation coefficient between actual and predicted values, while the red dotted line suggests 'perfect' predictions); the vehicle fatalities model yields predictions without an apparent 'systemic' problem (Figure 3), the vehicle accidents model yields widely varying predictions (Figure 4), while the pedestrian fatalities model systematically errs in predicting low number of fatalities (Figure 5).

4. Conclusions

In this paper we revisit the problem of the effects of weather on four incident categories: vehicle accidents, vehicle fatalities, pedestrian accidents and pedestrian fatalities. Independent variables used are related to weather (mean daily temperature and mean daily precipitation height), traffic and time. Estimation results for all incident types suggest that the most consistently significant and influential variable is mean daily precipitation height along with its lagged value. While much previous research has determined the significance of rainfall to accident prediction, most research has indicated that increases in precipitation lead to increases in vehicle accidents. We find the opposite result that suggests that increases in rainfall reduce all types of incidents. Although this appears counter to what would seem logical, it may be attributed to the safety offset hypothesis (less speeding, etc.) and may also be a characteristic of Southern European drivers who are not accustomed to driving in wet conditions.

We also examined the effects of temperature on incidents. Our results show that mean daily temperature and its one period lag to be important determinants of vehicle accidents, where an increase in temperature leads to increased accidents. Again, the offset hypothesis may be an important factor in this finding; finally, the 3-day moving temperature average was significant in determining pedestrian injuries (increase in average temperature increases pedestrian injuries) and we attribute this result to the significant increase in pedestrian traffic associated with higher temperatures in Athens.

It is worth noting that in this paper we accounted for temporal correlation in the dependent variable through a first order integer autoregressive process. We do however recognize that there are two important methodological issues still unaccounted for that should be addressed in future research; first, model extensions that allow for possible overdispersion in the dependent variable should be tested. Second, higher order autoregressive as well as integer ARMA models can be developed and their fit should be compared to that of existing models.

As a final note, the effects of weather conditions on traffic accidents pass through the weather effects on driver behaviour and speeding and on traffic volume and its effects on traffic accidents. Positive and negative effects co-exist and before identifying their combined effects, further research is needed. Consequently, although we used daily data in this paper, weather variables could also be accounted for in smaller time intervals and with related in-depth studies that capture the effects of, particularly, the intensity of the weather phenomena to driver behaviour and traffic volume and their impact on traffic accidents.

5. References

1. Aguero-Valverde, J., Jovanis, P.P., 2006. Spatial analysis of fatal and injury crashes in Pennsylvania. Accident Analysis and Prevention 38, 618-625.
2. Al-Osh, M.A., Al-Zaid, A.A., 1987. First order integer valued autoregressive process. Journal of Time Series Analysis 8, 261-275.
3. Al-Zaid, A.A., Al-Osh, M.A., 1993. Some autoregressive moving average processesnwith generalized Poisson marginal distributions. Annals of the Institute Statistics Mathematics 45, 223-232.
4. Baker, T.K., Falb, T., Voas, R., Lacey, J., 2003. Older women drivers: Fatal crashes in good conditions. Journal of Safety Research 34, 399-405.
5. Bergel-Hayat, R., Depire, A., 2004. Climate, road traffic and road risk - an aggregate approach. Proceedings of the 10th World Conference on Transport Research (CD-ROM), World Conference on Transport Research Society.
6. Böckenholt, U., 1998. Mixed INAR(1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data. Journal of Econometrics 89, 17-338
7. Branas, K., Knudson, M., 2001. Helmet laws and motorcycle rider death rates. Accident Analysis and Prevention 33(5), 641-648.
8. Brännäs, K., Hellström, J., Nordström, J., 2002. A new approach to modelling and forecasting monthly guest nights in hotels. International Journal of Forecasting 18, 19-30
9. Brannas, K., 1994. Estimation and testing in Integer-valued AR(1) models. Umea Economic Studies 335 (revised), Sweden.
10. Brannas, K., 1995. Explanatory variables in the AR(1) count data model. Umea Economic Studies 381, Sweden.
11. Brijs, T., Karlis, D., Wets, G., 2008. Studying the effect of weather conditions on daily crash counts using a discrete time-series model. Accident Analysis and Prevention 40, 1180-1190.
12. Brijs, T., Vlassenroot, S., Broekx, S., De Mol, J., Panis, L.I., Wets, G., 2007. Driving with intelligent speed adaptation: Final results of the Belgian ISA-trial. Transportation Research Part A: Policy and Practice 41, 267-279
13. Brijs, T., Karlis, D., Weets, G., 2008. Studying the effects of weather conditions on daily crash counts using a discrete time-series model. Accident Analysis and Prevention 40, 1180-1190.
14. Caliendo, C., Guida, M., Parisi, A., 2007. A crash-prediction model for multilane roads. Accident Analysis and Prevention 39, 657-670.
15. Chang, L.Y., Chen, W.C., 2005. Data mining of tree-based models to analyze freeway accident frequency. Journal of Safety Research 36, 365-375.
16. Edwards, J.B., 1996. Weather-related road accidents in England and Wales: a spatial analysis. Journal of Transport Geography 4 (3), 201-212.
17. Edwards, J.B., 1998. The Relationship Between Road Accident Severity and Recorded Weather. Journal of Safety Research 29 (4), 249-262.
18. Eisenberg, D., 2004. The mixed effects of precipitation on traffic accidents. Accident Analysis and Prevention 36(4), 637-647.
19. Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., Thomsen, L.K., 1995. Measuring The Contribution Of Randomness, Exposure, Weather, And Daylight To The Variation In Road Accident Counts. Accident Analysis & Prevention 27 (1), 1-20.

20. Fridstrom, L., Ingebrigtsen, S., 1991. An aggregate accident model based on pooled, regional time-series data. Accident Analysis and Prevention 23(5), 363-378.
21. Fridstrom, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., Krogsgard, L., 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. Accident Analysis and Prevention 27(1), 1-20.
22. Geurts, K., Thomas, I., Wets, G., 2005. Understanding spatial concentrations of road accidents using frequent item sets. Accident Analysis and Prevention 37, 787-799.
23. Golob, T.F., Recker, W.W., 2004. A method for relating type of crash to traffic flow characteristics on urban freeways. Transportation Research Part A 38, 53-80.
24. Gourieroux, C., Jasiak, J., 2004. Heterogeneous INAR(1) model with application to car insurance. Insurance: Mathematics and Economics 34, 177-192.
25. Hellstrom, J., 2001. Unit Root testing in integer-valued AR(1) models. Economic Letters 70, 9-14.
26. Hermans, E., Wets, G., Van Den Bossche, F., 2006. Frequency and Severity of Belgian Road Traffic Accidents Studied by State Space Methods. Journal of Transportation and Statistics 9 (1), 63-76.
27. Ivey, D.L., Griffin, L., Newton, T.M., Lytton, R., Hankin, K.C., 1981. Predicting Wet Weather Accidents. Accident Analysis And Prevention 13, 83-99.
28. Jin-Guan, D., Yuan, L., 1991. The integer values autoregressive (INAR(p)) model. Journal of Time Series Analysis 12, 129-142.
29. Joe, H., 1996. Time series models with univariate margins in the convolution-closed infinitely divisible class. Journal of Applied Probability 33, 664-677.
30. Jung, R.C., Tremayne, A.R., 2003. Testing for serial dependence in time series models of counts. Journal of Time Series Analysis 24(1), 65-84.
31. Karlis, D., Xekalaki, E., 1999. Maximum likelihood estimation for the poisson-binomial distribution via the EM algorithm with application to discrete time series models. Technical Report No. 62, Department of Statistics, Athens University of Economics, Greece.
32. Keay, K., Simmonds, I., 2005. The association of rainfall and other weather variables with road traffic volume in Melbourne, Australia. Accident Analysis and Prevention 37, 109-124.
33. Keay, K., Simmonds, I., 2006. Road accidents and rainfall in a large Australian city. Accident Analysis and Prevention 38, 445-454.
34. Khan, G., Qin X Noyce, D.A., 2008. Spatial Analysis of Weather Crash Patterns. Journal of Transportation Engineering, May 2008, 191-202.
35. Koetse, M.J., Rietveld, P., 2009. The impact of climate change and weather on transport: An overview of empirical findings. Article in press, Transportation Research Part D.
36. Levine, N., Kim, K.E., Nitz, L.H., 1995. Daily fluctuations in Honolulu motor vehicle accidents. Accident Analysis and Prevention 27(6), 785-796.
37. Majdzadeh, R., Khalagi, K., Naraghi, K., Motevalian, A., Eshraghian, M.R., 2008. Determinants of traffic injuries in drivers and motorcyclists involved in an accident. Accident Analysis and Prevention 40, 17-23.

38. Malyshkina, N.V., Mannering, F.L., Tarko, A.P., 2008. Markov switching negative binomial models: An application to vehicle accident frequencies. Article in press, Accident Analysis & Prevention.
39. Martin, J.L., 2002. Relationship between crash rate and hourly traffic flow on interurban motorways. Accident Analysis and Prevention 34(5), 619-629.
40. McKenzie, E., 1985. Some simple models for discrete variable time series. Water Resources Bulletin 21, 645-650.
41. McKenzie, E., 1986. Autoregressive moving-average processes with negative binomial and geometric marginal distributions. Advances in Applied Probability 18, 679-695.
42. Miaou, S.P., Lord, D., 2003. Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. Transportation Research Record 1840, 31–40.
43. Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: An exploratory empirical analysis. Accident Analysis and Prevention 40, 260-266.
44. Pai, C-W., Saleh, W., 2008. Modelling motorcyclist injury severity by various crash types at T-junctions in the UK. Safety Science 46, 1234-1247.
45. Quddus, M.A., 2007. Time series count data models: An empirical application to traffic accidents. In Proceedings of the 86th Transportation Research Board Annual Meeting, Washington, D.C.
46. Scott, P., 1986. Modelling Time-Series Of British Road Accident Data. Accident Analysis & Prevention 18 (2), 109-117.
47. Shankar, V., Mannering, F., Barfield, W., 1995. Effect Of Roadway Geometrics And Environmental Factors On Rural Freeway Accident Frequencies. Accident Analysis & Prevention 27 (3), 371-389.
48. Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1998. Evaluating median cross-over likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. Transportation Research Record 1635, 44-48.
49. Stipdonk, H.L., (Ed.) 2008. Time series applications on road safety developments in Europe. Deliverable D7.10 of the EU FP6 project SafetyNet. Available on-line at: http://www.erso.eu/safetynet/fixed/WP7/SN_D7.10_final.pdf
50. Ulfarsson, G.F., Shankar, V.N., 2003. An accident count model based on multi-year cross-sectional roadway data with serial correlation. In Proceedings of the 82nd Transportation Research Board Annual Meeting, Washington, D.C.
51. Van Harn, K., Steutel, F.W., 1977. Generalized renewal sequences and infinitely divisible lattice distributions. Stochastic Processes and their Applications 5, 47-55
52. Washington, S., Karlaftis, M., Mannering, F., 2003. Statistical and econometric methods for transportation data analysis, Chapman & Hall/CRC, Boca Raton, Florida.
53. Young, R.K., Liesman, J., 2007. Estimating the relationship between measured wind speed and overturning truck crashes using a binary logit model. Accident Analysis and Prevention 39, 574-580.

Table 1. Summary of existing research on time series analysis of weather effects on road accidents

Author	Year	Level			Dependent			Method				Time				Weather variables effect on accidents								Other variables					Country	Period
		National	Regional	Local	Accident Frequency	Accident Severity	Traffic volume	State-space	ARMA	GLM (Poisson - Negative Binomial)	Other	Daily	Weekly	Monthly	Yearly	Temperature (degrees)	Deviation from mean temperature	Sunlight (hours)	Rainfall (mm)	Number of days with rainfall	Rainfall intensity (mm/duration)	Dry spell (days from previous rainfall)	Number of days below 0	Road design	Demographic	Economic	Traffic	Interventions		
Scott	1986	●				●			●					●		-			+					●			●	●	UK	1970-1978
Fridstrom & Ingebrigtsen	1991	●				●				●				●					+						●	●	●	●	Norway	1974-1986
Fridstrom et al.	1995	●			●	●								●					+						●	●	●	●	DK, SE, FI, NO	1975-1987
Shankar et al.	1995		●		●					●				●						+									Washington	1988-1993
Eisenberg	2004	●			●					●*		●		●					+			+					●		USA	1990-1999
Bergel-Hayat & Depire	2004	●			●		●		●					●					+					●			●		France	1975-1999
Chang & Chen	2005	●			●					●					●				+					●			●		Taiwan	2001-2002
Keay & Simmonds	2005			●			●			●		●															●		Melbourne, Australia	1989-1996
Hermans et al.	2006	●			●	●		●						●				+		+			-			●		●	Belgium	1974-1999
Keay & Simmonds	2006				●					●		●								+							●		Melbourne, Australia	1987-2002
Caliendo et al.	2007			●	●					●		●							+										Italy	1999-2003
Brijs et al.	2008		●		●				●			●				-	+	+	+		+	+					●		3 cities, Netherlands	2001
Malyshkina et al.	2008		●		●					●**			●			-			+					●			●		Indiana, USA	1995-1999
Stipdonk	2008	●	●		●			●						●			+		+				-	●			●		France	1975-2000

* with lagged variables ** Markovian

Table 2
Model Results

	Vehicle Fatalities		Vehicle Crashes		Pedestrian Fatalities		Pedestrian Injuries	
	coefficient	t-stat	coefficient	t-stat	coefficient	t-stat	coefficient	t-stat
constant	1.022	10.8	23.37	38.2	0.347	8.83	5.09	22.5
autoregressive term	0.022	1.92	0.201	17.9			0.047	4.1
mean temperature			0.097	2.1				
temperature lag 1			0.104	2.26				
3-day moving average	0.001	1.21					0.025	4.51
mean precipitation	-0.005	-1.89	-0.135	-9.51	-0.003	-1.67	-0.021	-3.58
precipitation lag 1	-0.009	-3.53	-0.107	-7.61	-0.002	-1.49	-0.019	-3.11
dummy variables								
weekend	0.138	2.83	0.941	3.89			1.702	15.39
Monday					0.052	2.51		
Tuesday								
Wednesday								
Thursday								
Friday	0.083	1.69	3.13	12.89	0.032	1.27		
January	0.331	1.51	2.114	4.93	-0.052	-1.57	-0.382	-2.41
Febraury	-0.111	-1.63			-0.101	-1.97	-0.3	-1.85
March	-0.153	-2.24	-1.08	-2.51	-0.116	-3.51	-0.448	-2.82
April	-0.132	-1.85	1.31	2.88	-0.098	-2.93	-0.592	-3.52
May	0.139	1.85	1.37	3.01	0.13	3.97	0.427	1.87
June	-0.223	-2.44			0.11	3.42	0.457	2.38
July	0.237	2.13	1.75	2.66	0.15	4.23	-1.657	-6.75
August	0.129	1.99	-8.03	-12.1	-0.145	-4.37	-3.048	-12.4
September	0.133	1.413	-1.41	-2.44	-0.09	-1.88		
October	0.177	2.11	1.24	2.51				
November	0.189	2.31						
Annual 1985 - 2005	*6 not significant*		*2 not significant*		*4 not significant*		*5 not significant*	
Number of data points	7670		7670		7670		7670	
Mean Absolute Percent Error	27.9		17.6		37.3		35.1	
Mean Absolute Error	0.78		4.37		0.488		2.08	

Table 3
Monthly Incident distribution (2007)

Month	% of Accidents	% of Casualties
January	9.2	10.8
February	7.2	7.5
March	8.9	7.8
April	7.9	6.7
May	9.2	9.1
June	8.2	5.6
July	9.3	10.2
August	6.6	8.3
September	8.2	8.3
October	9.0	9.1
November	8.3	9.4
December	8.0	7.0

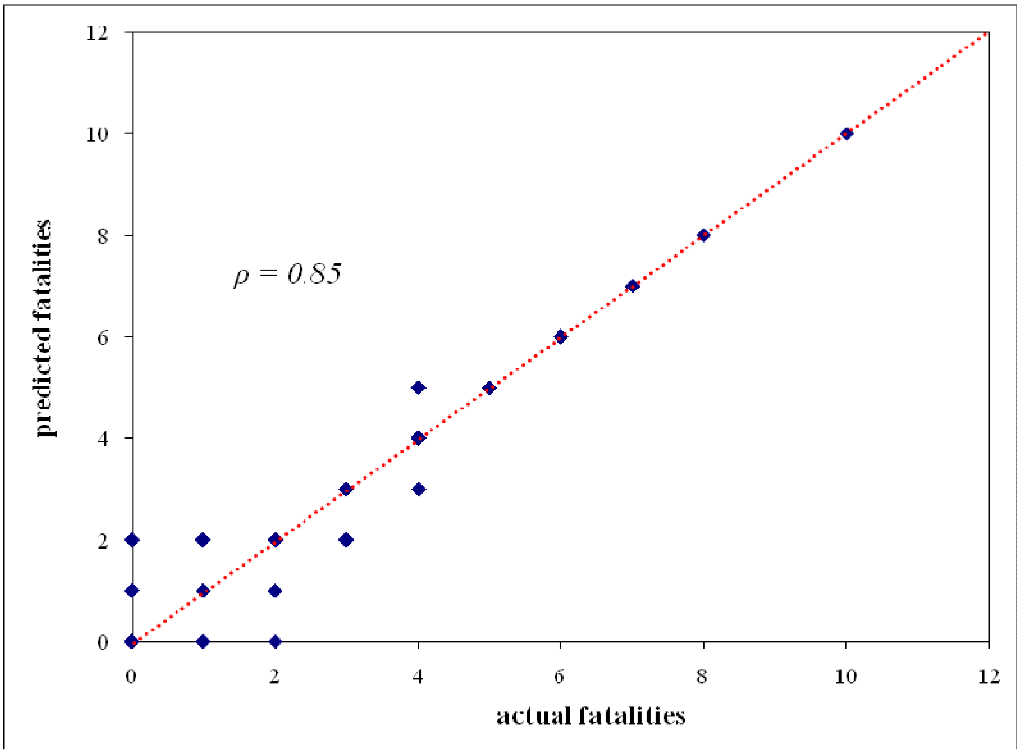

Figure 3. Actual versus Predicted diagram for vehicle fatalities

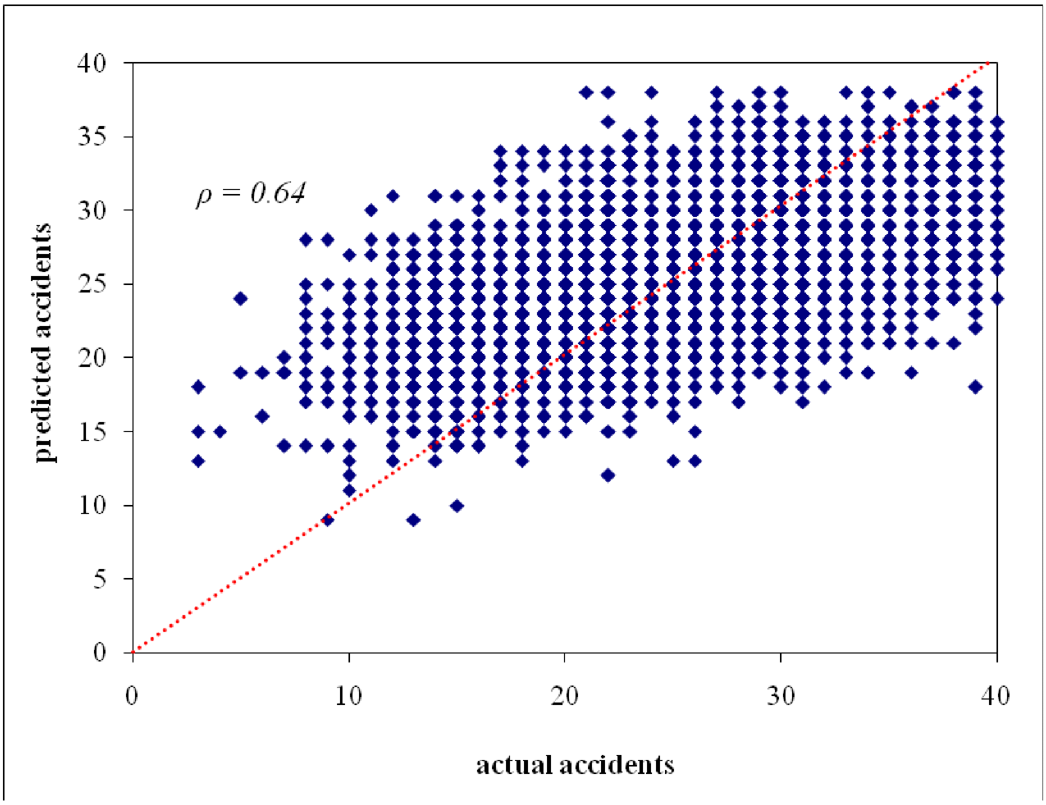

Figure 4. Actual versus Predicted diagram for vehicle accidents

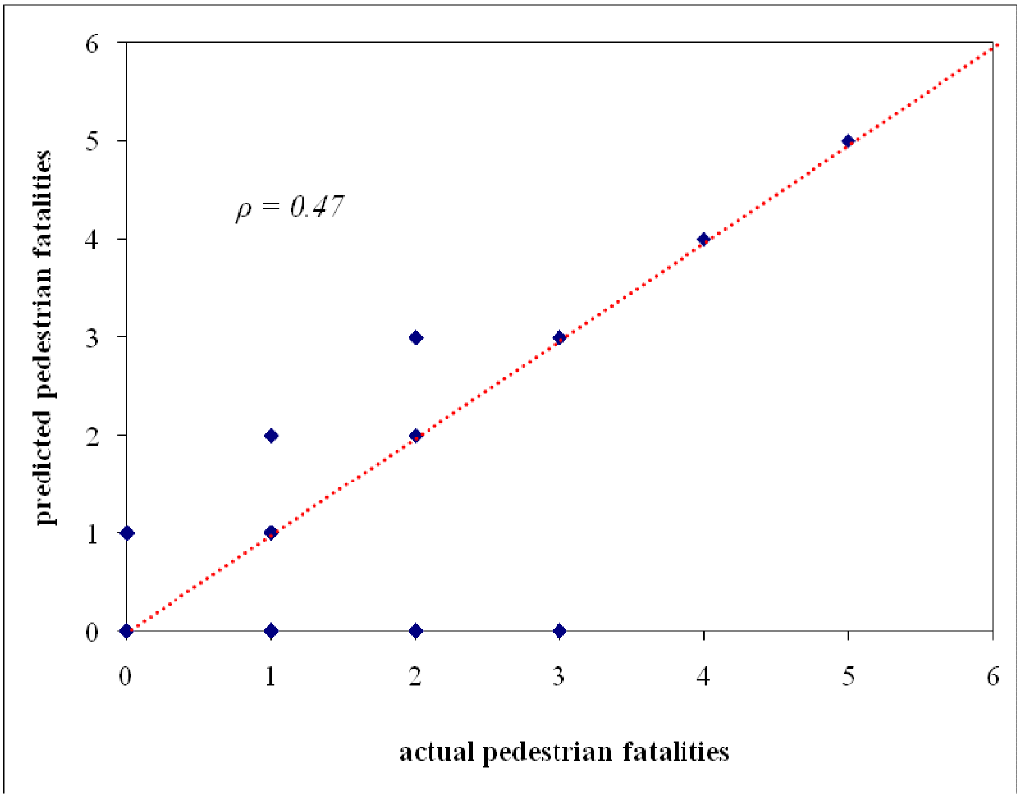

Figure 5. Actual versus Predicted diagram for pedestrian fatalities