Explaining the road accident risk : Weather effects

Ruth Bergel Hayat, Mohammed Debbarh, Constantinos Antoniou, George Yannis

Explaining the road accident risk: weather effects

Ruth Bergel-Hayat[1*], Mohammed Debbarh[1],
Constantinos Antoniou[2], George Yannis[2]

[1] Paris-Est University, IFSTTAR French Institute of Science and Technology for Transport, Development and Networks
[2] NTUA National Technical University of Athens

* corresponding author: ruth.bergel@ifsttar.fr

Abstract
This research aims to highlight the link between weather conditions and road accident risk at an aggregate level and on a monthly basis, in order to improve road safety monitoring at a national level. It is based on some case studies carried out in Work Package 7 on "Data analysis and synthesis" of the EU-FP6 project "SafetyNet - Building the European Road Safety Observatory", which illustrate the use of weather variables for analysing changes in the number of road injury accidents. Time series analysis models with explanatory variables that measure the weather quantitatively were used and applied to aggregate datasets of injury accidents for France, the Netherlands and the Athens region, over periods of more than twenty years. The main results reveal significant correlations on a monthly basis between weather variables and the aggregate number of injury accidents, but the magnitude and even the sign of these correlations vary according to the type of road (motorways, rural roads or urban roads). Moreover, in the case of the interurban network in France, it appears that the rainfall effect is mainly direct on motorways - exposure being unchanged, and partly indirect on main roads – as a result of changes in exposure. Additional results obtained on a daily basis for the Athens region indicate that capturing the within-the-month variability of the weather variables and including it in a monthly model highlights the effects of extreme weather. Such findings are consistent with previous results obtained for France using a similar approach, with the exception of the negative correlation between precipitation and the number of injury accidents found for the Athens region, which is further investigated. The outlook for the approach and its added value are discussed in the conclusion

Key-words
Weather effects, precipitation, rainfall, temperature, frost, aggregate risk, road safety, injury accidents, France, the Netherlands, Athens region.

1. Introduction

Weather conditions are considered to be a factor that affects the number of road accidents and casualties significantly, with different effects according to the type of road (motorways, rural roads or urban roads). Moreover, as the weather also affects mobility, it is to be expected that the effects of weather on the number of injury accidents and casualties are partly due to the changes in mobility occurring at the same time.

This research is based on time series accident analyses conducted in Work Package 7 on "Data analysis and synthesis" of the EU-FP6 project "SafetyNet - Building the European Road Safety Observatory and which illustrate the use of weather variables for analysing changes in the number of road injury accidents (Stipdonk (Ed.), 2008).. Time series analysis models were fitted to monthly aggregated datasets of injury accidents for France, the Netherlands and the Athens region and included explanatory variables that measure the weather quantitatively. Long periods were considered: 1975-2000 for France, 1987-2005 for the Netherlands and 1985-2005 for Athens region.

The paper is organized as follows. In Section 2, the main questions raised by the inclusion of a measure of the climatic factor in a time series analysis of road safety trends at an aggregate level are recalled, while Section 3 provides a brief survey of the results found in the literature. The approach adopted and, in particular, the handling of the available weather data, are explained in Section 4. The results obtained with averaged data and extreme weather data are detailed in Sections 5 and 6, and are further discussed in Section 7. Finally, the outlook for the approach and its added value are discussed in Section 8 which concludes the paper.

2. The research problem

Time series analysis techniques are widely used for analysing changes in road safety trends as observed at national level, with the use of meteorological variables to capture short-term changes in road safety indicators. Fixing the time scale for the analysis is very important, as the effects of weather conditions on safety indicators may be expected to differ according to the time scale, for example depending on whether the day or the month is considered. As daily monitoring requires a large number of daily accident data counts, it is rarely used at the national level. Instead, monthly monitoring is preferred as it facilitates international benchmarking. In France, for instance, short-term monitoring is performed on monthly sets of accident data (ONISR, 2010). This is the approach applied in this paper.

According to the literature, weather may explain about 5% of monthly accident/fatality variability (see Fridstrom et al., 1995, Hermans et al., 2006). Therefore, including the influence of the weather conditions in the analysis of road accident trends at an aggregate level is especially appropriate for short-term analysis -whether for the past or the near future - as it provides a short-term trend which will make it easier to identify the effects of safety policy.

Capturing and measuring the climatic factor is not straightforward. It raises a number of issues, most importantly, how can we determine which meteorological phenomena significantly influence the road risk level, which variables are to be chosen for measuring them and also on which time scale should the weather be measured to be significant on a monthly time scale.

The number of injury accidents that are recorded during inclement weather has led rain to be considered as the major meteorological explanatory factor for road accident risk (Brodsky and Hakkert, 1988). In France for instance, for the period 1990-2000, 14 % of all injury accidents took place during rainy weather and 1%, at the most during fog, frost or snow/hail (Aron et al., 2007). As the aggregate number of injury accidents depends on the aggregate exposure level, the next step is to identify the weather situations that increase or reduce mobility. Fair weather favours mobility, while cold spells reduce it, and it seems natural to consider temperature as an appropriate variable for representing the regular changes in mobility during the year (the general cycle, from January to December, that is repeated during every 12-month period), as well as the atypical changes in mobility that occur in the event of unusually hot or cold weather. Examples of models accounting for weather effects on the mileage in France have been provided by Bergel (1992), Jaeger (1998), and more generally at international level by Gaudry and Lassare (Eds) (2000). In this research, we have therefore decided to focus on three climatic situations: rainfall (as a risk factor), fair weather (as a mobility factor) and frost (which plays a role in both risk and mobility). The two other issues; the construction of the weather variables and the time scale to be used to capture significant weather information will be discussed in Section 4.

3. Current state of knowledge

The influence of temperature and rainfall on the aggregate numbers of injury accidents and casualties were first studied by Scott (1986) who modelled the changes in their monthly number in the UK from 1970 to 1978. Since then, numerous attempts have been made to build an explanatory model to account for the influence of climate using these two variables independently or simultaneously, which we recall and summarize in this section. But the results of these attempts vary according to the time scale of the analysis (the day or the month) and the construction of the weather variables (the mean values or the extreme values for the considered time scale). In addition, in cases where the weather variables are significant, the magnitude and sometimes even the sign of their impact may differ.

As regards the methodology adopted to construct the weather variables, meteorological data including both average and atypical values during the month was first used by (Fridstrom & Ingebrigtsen 1991) and (Fridstrom et al, 1995) who analysed the changes in the monthly number of accidents of four countries with similar characteristics (Denmark, Finland, Norway and Sweden,). Monthly aggregate models exploiting weather data including both average and atypical values during the month were also fitted to accident data for France, with disaggregation according to the type of road (Bergel and Depire, 2004a & b).

The relationship between weather and road accidents has been investigated intensively in recent years, and research results that focus on very short-term links (daily level) and on short-term links (at a monthly level) have been published in (Karlaftis and Yannis, 2010).

The influence of weather conditions on road crashes in Belgium has been demonstrated on a monthly time scale with the hypothesis of a stochastic trend (Hermans et al., 2006) and on a daily time scale after the serial temporal correlation was accounted for (Brijs et al., 2008). Although the use of the daily and monthly variability of the weather information was not taken into account in a single model, both these studies appear to show the importance of which time series model is chosen for the analysis. This is now accepted in the road safety research field at international level, as the use of appropriate techniques for analysing road safety trends has been recommended (Dupont and Martensen (Eds.), 2007).

As regards the link between each of these weather variables - rainfall and temperature - and the accident frequencies, the results found in the literature sometimes differ.

Increased rainfall in often linked to increased accident frequencies (Fridstrom and Ingebrigtsen 1991; Fridstrom et al. 2005; Chang and Chen 2005; Caliendo et al. 2007, Shankar et al. 1995; Keay and Simmonds 2006; Hermans et al. 2006). However, increases in rainfall have also been found to reduce the number of accidents (Karlaftis and Yannis, 2010), a result that may be attributed to driver risk compensation behaviour or to a simultaneous decrease of exposure. The time-varying effects of rainfall (and precipitation) have also been investigated and Eisenberg (2004) has shown that the impact of precipitation on a given day is reduced when precipitation was observed in the previous days, which is possibly due to driver adaptation. In the same vein, Brijs et al. (2008) have confirmed Eisenberg's finding that the longer the "dry spell" (i.e. the number days since the previous rainfall), the higher the number of accidents when rainfall occurs.

Higher temperatures increase accident frequencies (Scott, 1986) and this finding established by Scott on data from the UK has been confirmed at international level. The number of hours of sunlight appears to increase road accidents (Fridstom et al, 1995, Hermans et al. 2006), while deviations from mean daily or monthly temperatures have also been found to increase road accidents (Brijs et al., 2008; Stipdonk (Ed.), 2008). Extreme temperatures (low in winter and high in summer) are positively correlated with road accidents (Malyshkina et al., 2008), but when the monthly number of days with sub-zero temperatures increases, the number of road accidents is reduced, possibly due to lower exposure (Hermans et al., 2006; Stipdonk (Ed.), 2008).

However, this brief review of research findings relates to both very short-term and short-term links between precipitation and temperature and aggregate accident frequencies. In addition, some of the links were estimated while controlling for exposure and some were not. It is to be expected that some of the short-term results mentioned above, that appear to be significant at the daily level, will not remain significant at the monthly level or, if they still do, will change in magnitude and maybe even in sign.

A methodological framework for accounting for the climatic factor in the systemic approach for risk modelling, for monthly aggregated data, has been provided by the DRAG approach (Gaudry and Lassarre (Eds.), 2000). Still, as the mileage is systematically included in the explanatory model for monthly accident frequencies, this approach is intended to provide comparative results at national level while controlling for exposure. The link between precipitation and rainfall and the aggregate numbers of injury accidents under constant exposure has thus been identified for several countries.

Other studies, which associated mobility data (traffic volume or vehicle-kilometres of travel) with weather variables, have generally revealed an increase in accident frequencies as a result of increases in temperature and rainfall, under constant exposure (Scott, 1986; Eisenberg, 2004; Chang and Chen, 2005; Brijs et al., 2008; Malyshkina et al., 2008). The indirect effect of these weather variables on the accident frequencies is estimated with models analysing mileage changes. References to the estimation of weather effects on both road travel demand and road risk at aggregate level in the case of France are given by Bergel (2008).

4. Methodology

4.1. Approach

A two-stage approach was adopted in order to take account of the climatic factor. First, an analysis was performed for France, the Netherlands and the Athens region, with the aim of comparing the effects of monthly weather averages. Second, the results for the Athens region, which differed significantly from the results for both France and the Netherlands as regards the effects of rainfall, were further investigated and extreme weather variables were also used, on both a daily and a monthly time scale.

4.2 Accident data

The numbers of injury accidents were collected from police sources, for the following long periods: 1975-2000 for France, 1987-2005 for the Netherlands and 1985-2005 for Athens region. Figures 1 and 2 show the monthly numbers of injury accidents for France, the Netherlands and the Athens region for the period 1987-2000. Although the datasets differ in their values and in their trends, there is a visible seasonal pattern in each of them, which repeats every year more or less regularly. The lowest values in their monthly numbers are registered in the winter and the largest ones in the summer. However, that seasonal pattern (or seasonal component) is partly periodic and partly irregular. It is therefore natural to try to capture the part of the seasonal component due to atypical weather conditions - that differ significantly from the seasonal average (unusually heavy rain or dry weather for the month, extremely hot or cold weather for the month).

4.3. Measuring the weather

The approach we adopted for measuring the weather was quantitative and focused on three specific situations: rain, fair weather and frost; as has already been presented in the case of France (Bergel and Depire, 2004a and 2004b). This research extends this approach, applying it to the Netherlands and the Athens region. The climatic variables considered are precipitation, in order to measure rain as a risk factor, and temperature, in order to measure fair weather conditions as mobility factors. Frost (negative temperature) was also considered as it plays a role in both risk and mobility.

These meteorological factors were measured on a daily basis by three variables: the rainfall, measured in mm, the maximum temperature during the day, measured in °C, and the occurrence of frost by means of a dummy variable (1/0). The average and extreme values of these three weather variables were both obtained.

For France, the daily climatic variables were calculated by averaging the daily variables measured at the hundred meteorological observation points that are spread throughout the country. For the Netherlands, the daily variables were measured at a single centrally located observation point, the "de Bilt" station, and for the region of Athens the daily variables were measured at a single observation point in the old airport of Hellenikon, in the southern suburbs of Athens. These daily variables were then aggregated or averaged in order to construct monthly variables.
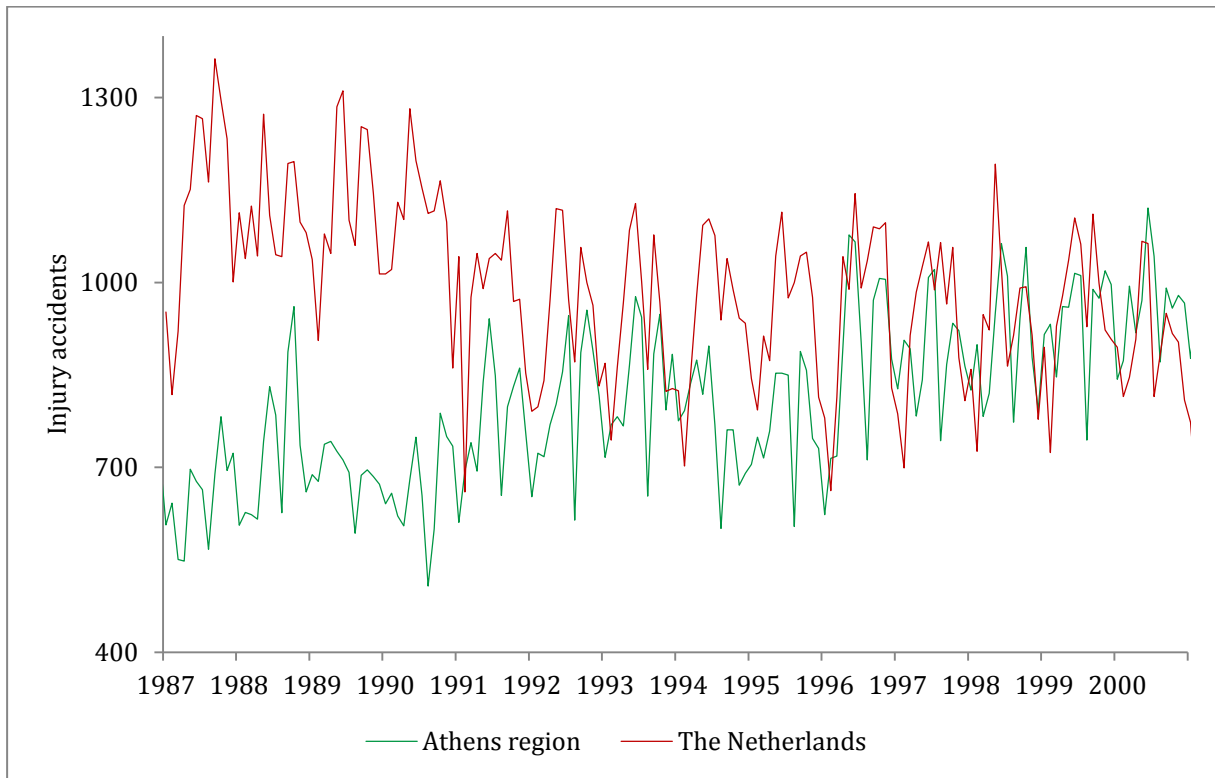
Figure 1: Monthly number of injury accidents in the Netherlands and the Athens region from 1987 to 2000
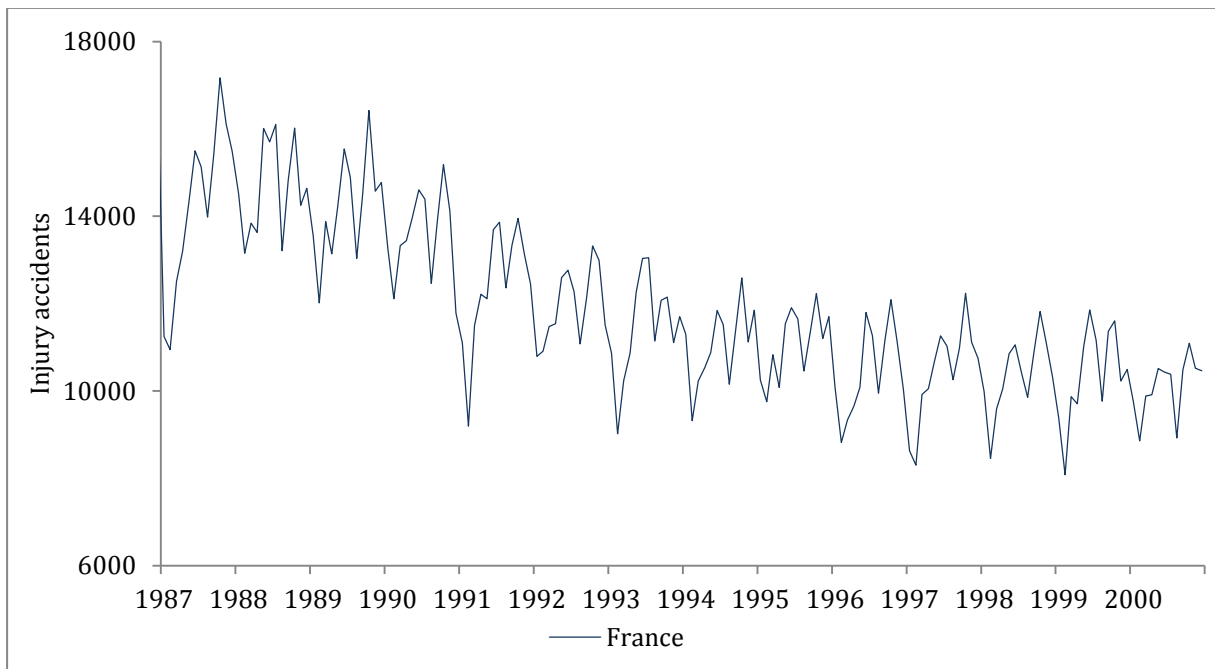

Figure 2: Monthly number of injury accidents in France from 1987 to 2000

Figures 3 to 5 show the monthly weather variables used for the Netherlands, France and the Athens region for the period 1987-2000. Obviously, there are differences between the seasonal patterns (i.e. the periodic variation within the year) for different years, in particular for the temperature and the occurrence of frost. For instance in France, the maximum temperature measured during the day (averaged over the hundred meteorological stations and over the month) is only 2-3 °C during cold winters for the months of January/February, whereas the seasonal average is 7-8 °C for these months. The number of days of frost (the occurrence of frost, averaged over the hundred stations and over the month) reaches 16-17 during the same period, whereas the seasonal average is 12-13 for these months. The amount of precipitation, subject to more variability than the other variables, was, for example, higher than usual during 1991, compared with the seasonal average for that period of the year. These divergences from an average seasonal pattern constitute the information of interest, as their influence on the number of injury accidents cannot be captured by the seasonal component in a time series analysis of the observed series.
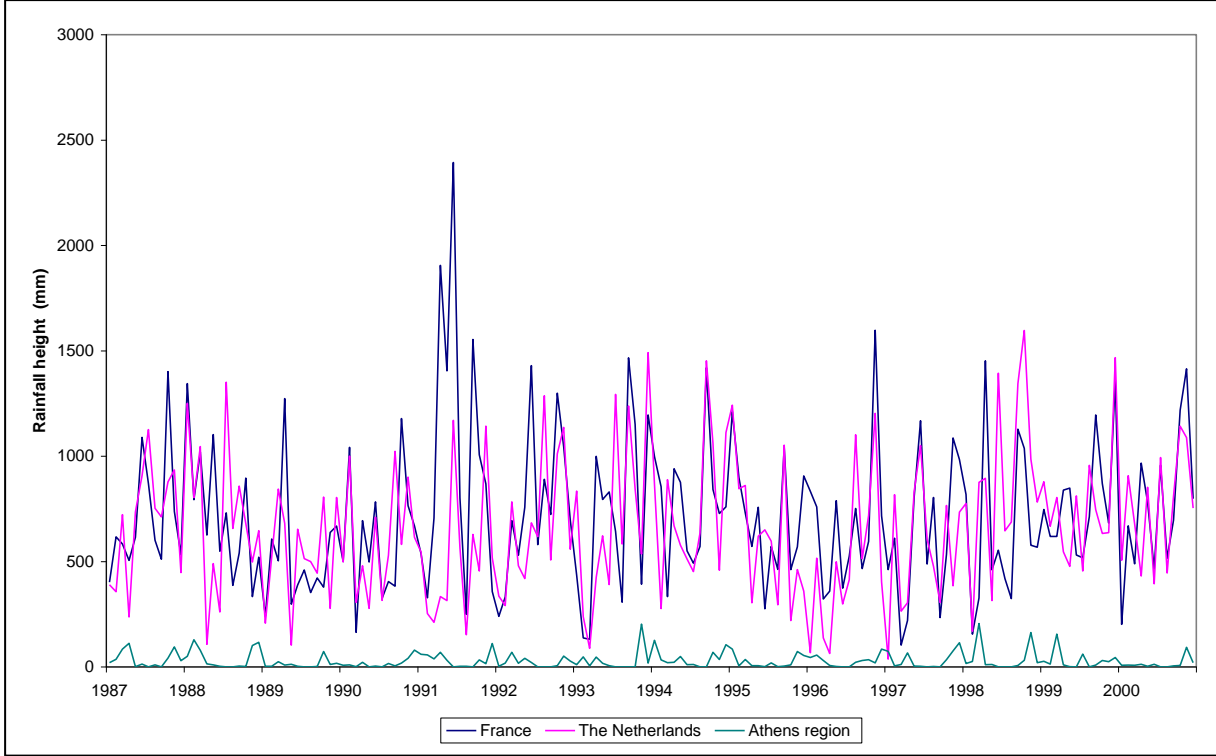

Figure 3: Monthly rainfall in France, the Netherlands and the Athens region (mm) from 1987 to 2000
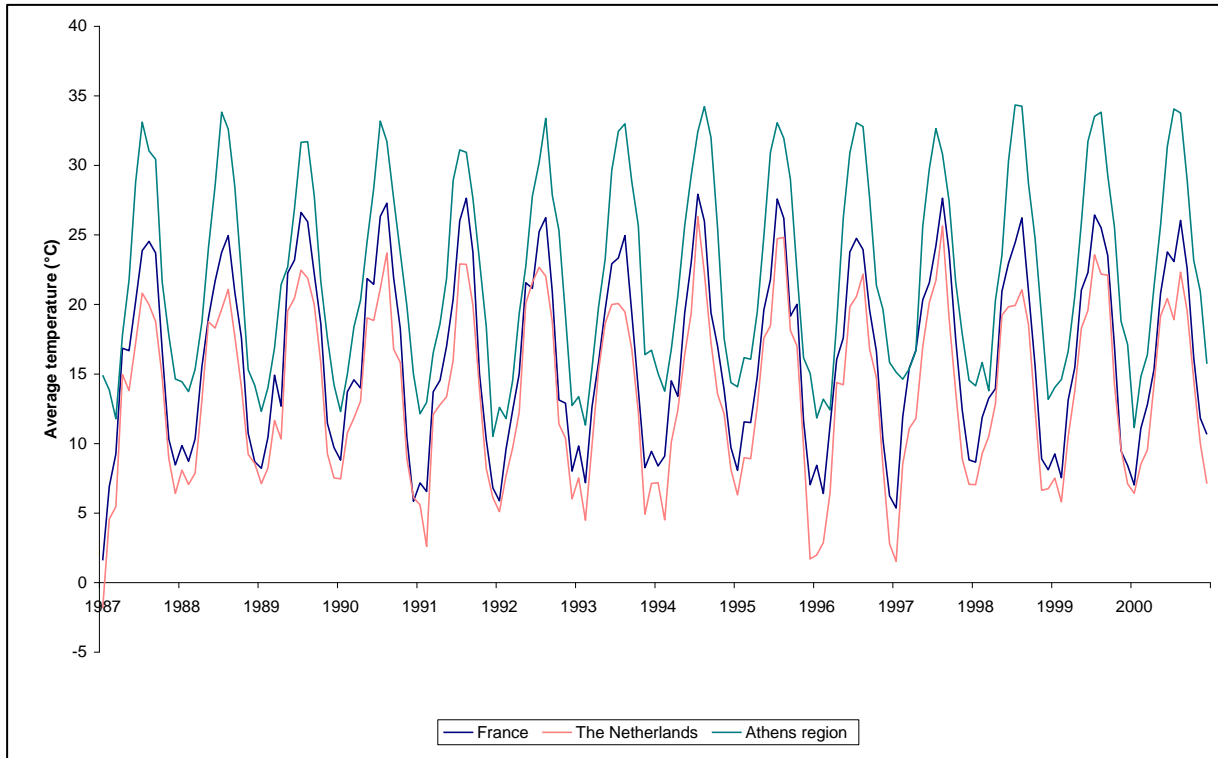
Figure 4: Average monthly temperature in France, the Netherlands and the Athens region (0.1°C) from 1987 to 2000.
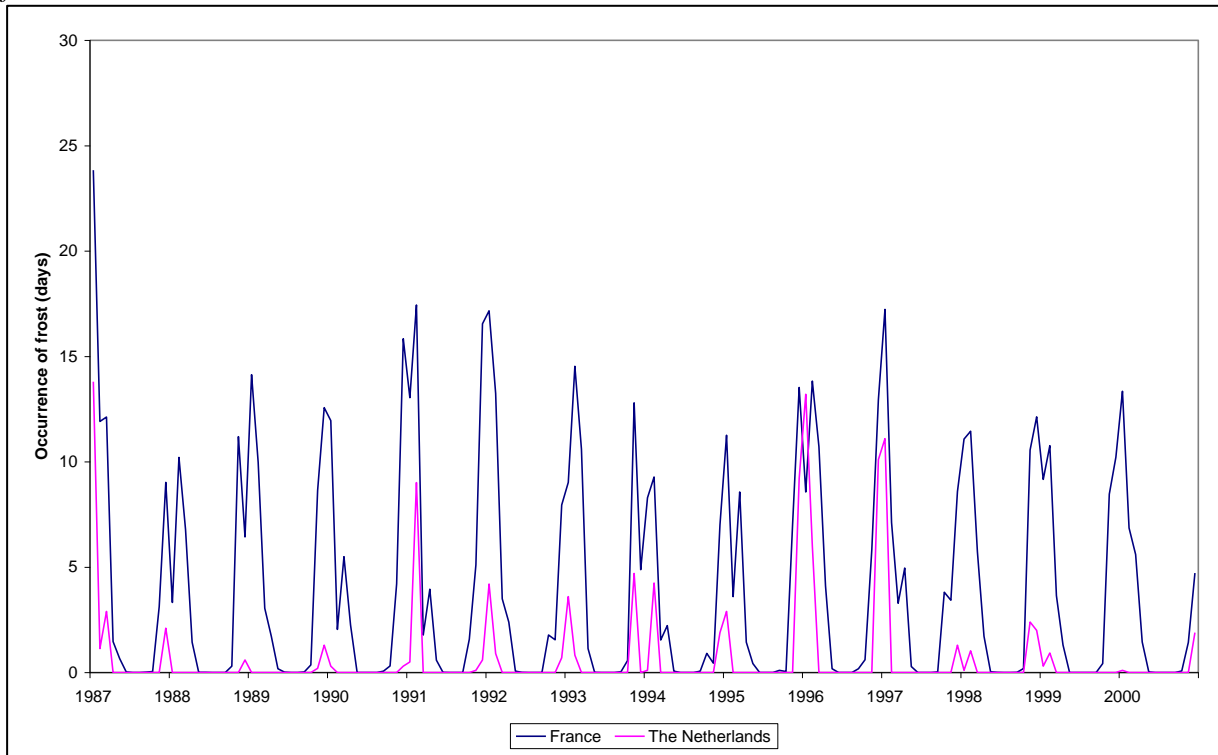
Figure 5: Monthly number of frost days in France and the Netherlands from 1987 to 2000

In the case of the Athens region, the weather effects were further investigated using daily data that included daily mean temperature (in °C) and daily precipitation (in mm). Figure 6 provides an overview of some key variables of the original daily data-set (1985-2005) used

for the Athens region. Periodicity over the years (at monthly level) is apparent, especially with respect to temperature and − to a lesser degree- precipitation that are typically reproduced each year. Unusually high precipitation was mainly recorded during the winter of 2002-03, following two relatively dry winter periods. The data reveal a significant improvement (reduction) in injury accidents during this period.

In order to examine the impact of extreme weather effects (as for example in Bergel and Depire, 2004b), atypicality variables were also constructed. First, the typical daily average temperature for each day in the year was computed. For example, the average temperature on March 4th was computed based on the average temperature on March 4th in each of the 21 years that were studied. The standard deviation of this temperature for each day in a typical year was also computed. These were then used to develop dummy variables capturing the days for which the daily mean temperature was lower (or higher) than the typical daily average for this day in the year plus one or two standard deviations. The same process was repeated for precipitation, but obviously in this case only daily precipitation one or two standard deviations above the mean for the typical day is meaningful. These dummy variables were then summed over the month. Finally, atypicality variables that count the number of days in a month with unusually low (or high) temperatures and with unusually large amount of precipitation were constructed.
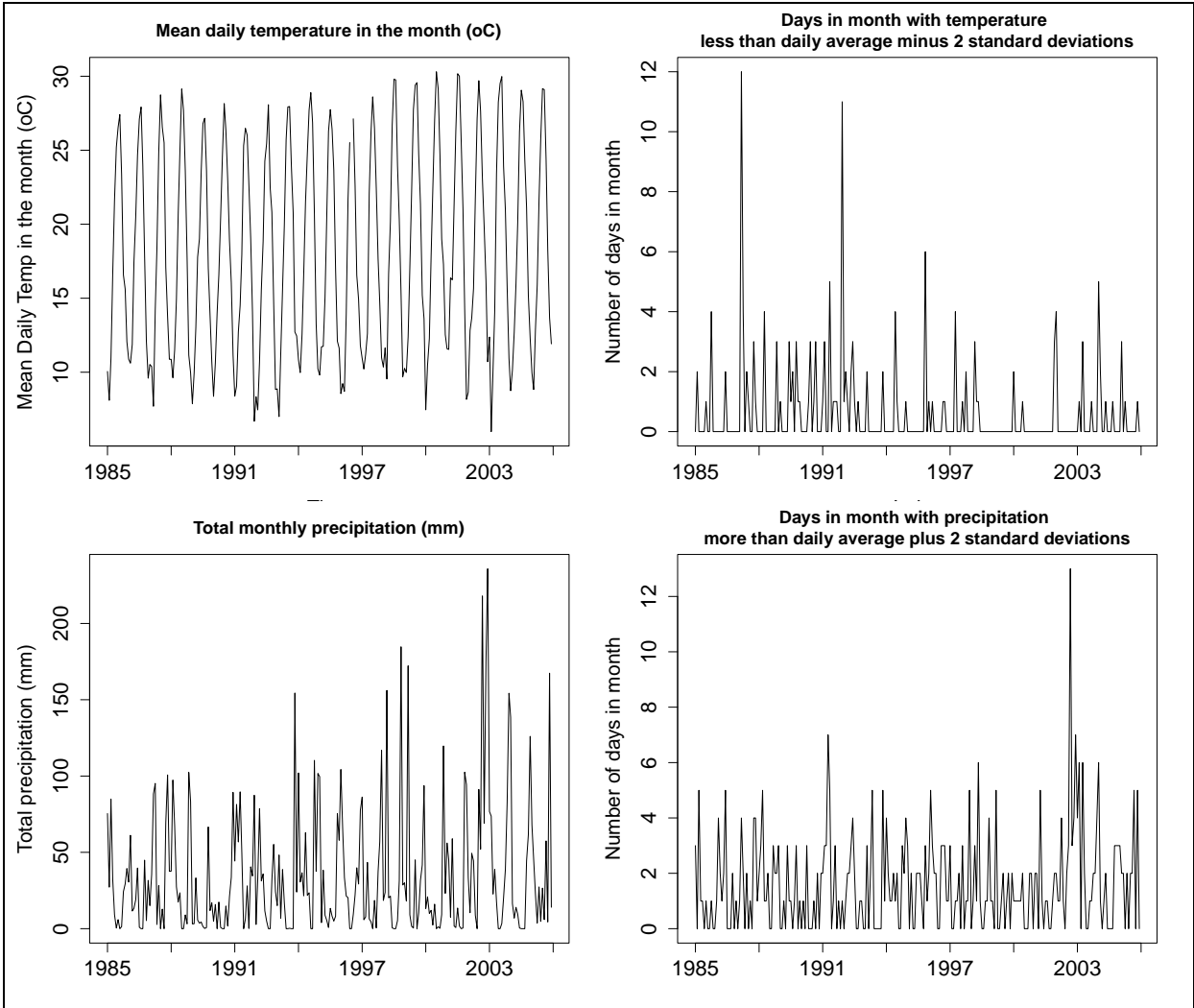

Figure 6: Descriptive graphs of the monthly dataset used for the analysis for the Athens region

5. Applications that provide averaged weather effects

The following structural model was used for each[1] dataset of monthly numbers of injury accidents (Durbin and Koopman, 2001, Commandeur and Koopman, 2007). It can take two forms, depending on whether or not a risk exposure factor in addition to the climate is included:

$$LogA_t = \sum_{k=1}^{K} a_k Z_{kt} + \mu_t + \gamma_t + \varepsilon_t \qquad LogA_t = aLogE_t + \sum_{k=1}^{K} a_k Z_{kt} + \mu_t + \gamma_t + \varepsilon_t$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \qquad\qquad\qquad \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$

$$\beta_t = \beta_{t-1} + \xi_t \qquad (5.1) \qquad \beta_t = \beta_{t-1} + \xi_t \qquad (5.2)$$

$$\gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \qquad\qquad\qquad \gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$$

where:

A_t is the monthly number of injury accidents,

E_t is a risk exposure indicator,

$Z_{k,k=1toK}$ are the K meteorological variables,

a and $a_{k,k=1toK}$ are parameters,

β is the slope of the trend μ,

γ is the seasonal component fitted with a dummy model,

ε, η, ζ and ω_{jt} are the white noises with the variances σ_ε^2, σ_η^2, σ_ζ^2 and σ_ω^2, which are not correlated with each other.

Specification (5.1) describes a locally linear trend and seasonal model for the monthly number of injury accidents, once it has been stabilized with a log transformation and corrected for the influence of the weather variables described in Section 4.3 (which measure rainfall, temperature and frost). The level and slope of the trend, and the seasonal component are considered as stochastic and defined each by a constraint that includes an error term. Four error terms are thus considered, instead of one only in the case of a deterministic linear trend and seasonal model, and their variances estimated in addition to the parameters of the weather variables, which are used to quantify their effects.

Specification (5.2) describes the same model with an additional variable that measures exposure, that has also been stabilizes with a log transformation. Exposure is thus controlled for and the parameters of the weather variables are used to quantify their effects under constant exposure.

All the explanatory variables were retained in the model, whether or not significant or not, in order to compare their significance within the same model, or between similar models. For each risk indicator (corrected for the effects of the explanatory variables), the level, slope and seasonal component were obtained for the whole period.. The variances of the four error terms were estimated, and the null hypothesis of the variance of the error was tested for the level, slope and seasonal component. The usual tests were performed to check the validity of

---

[1] When two series of accident data were considered simultaneously, a bivariate model was fitted to the vector of these two components (which was the case when the model was applied for both motorways and main roads in France, or for both motorways and rural roads in the Netherlands). However, although the bivariate approach improves each part of the model fitted to a single component, the estimated weather coefficients do not differ significantly from those obtained with two univariate models fitted to each component.

the residuals, with satisfactory results. The numerical results for the coefficients of the exogenous variables, estimated with (5.1) and (5.2), are given in Tables 1 and 2.

The following comments are restricted to the parameters with a 95% significance level).

At a national level, the overall effects of the weather on the number of injury accidents are consistent for France and the Netherlands:
- Rainfall is positively correlated with the number of injury accidents (100 mm of additional rainfall during a month increases the number of injury accidents in that month by 0.2%-0.3%)
- Temperature[2] is positively correlated with the number of injury accidents (1°C of additional average temperature during a month increases the number of injury accidents in that month by 1%-2%).
- The occurrence of frost is negatively correlated with the number of injury accidents (1 additional day of frost during a month decreases the number of injury accidents in that month by 0.3%-0.6%).

At the level of a specific type of road, these effects may differ between the three observed regions, not only in their magnitude but even in their sign:
- On motorways, rainfall is positively correlated with the number of injury accidents both in France and the Netherlands (100 mm of additional rainfall during a month increases injury accidents by about 0.5% on main roads and motorways in France, and 0.8% on motorways in the Netherlands). On the contrary, rainfall is negatively correlated with the number of injury accidents in the region of Athens, where the network mainly consists of urban roads (100 mm additional rainfall during a month decreases injury accidents by about 5 %).
- Temperature is positively correlated with the number of injury accidents, but with different magnitudes on main roads and motorways in France (1°C of additional average temperature during a month increases injury accidents by 0.4% and about 2% on main roads and motorways respectively), as is also the case on rural roads and motorways in the Netherlands (1°C of additional average temperature during a month increases injury accidents by about 1% and 2%-3% on rural roads and motorways respectively). In the region of Athens, on the contrary, the impact of temperature is only significant in the winter (1°C of additional average temperature during a month increases injury accidents by 2%).
- The occurrence of frost does not appear to have any significant effect on the number of injury accidents on either of the two networks considered (in France and the Netherlands).

Finally, an attempt to take account of risk exposure by means of proxies (oil sales) for France, the number of driver-kilometres provided by surveys for the Netherlands, the number of vehicles recorded at a major motorway toll station near Athens), was unsuccessful. Either the variable could not be considered as representative of true mobility (the elasticity value of the risk indicator with respect to mobility being too small) or the variable was not significant in the model fitted to the risk indicator (as was the case for the region of Athens).

For the two types of roads considered for France (main roads and motorways) on which the distance travelled is accurately known, sensor-based measurements of the number of vehicle-

---

[2] Two variables : the summer temperature (April to September) and the winter temperature (October to March) were used in the model

kilometres were significant and thus included in the models, which allowed to control for exposure. The changes in the number of injury accidents due to changes in the weather variables under constant exposure (direct weather effects) can thus be estimated. Due to the correlation between the temperature and the distance travelled, it is not possible to interpret the coefficients related to the average temperature as direct temperature effects. This is not the case for the two other weather variables. The most significant result relates to rain: rainfall has a positive (increasing) direct effect and a negative (decreasing) indirect effect on the number of injury accidents both on main roads and motorways (100 mm of additional rainfall in the month increases injury accidents by 1% on both networks - as exposure is unchanged, and decreases them by 0.4 and 0.2% on main roads and motorways - due to a decrease in the distance travelled); the rainfall effect thus appears to be mainly direct on motorways – as exposure is unchanged, and partly indirect on main roads – as a result of changes in exposure. Finally, the occurrence of frost has a significant positive effect on motorways (1 additional day of frost in the month increases injury accidents on motorways by 0.9%).

	RF(mm)	ST(°C)	WT(°C)	DF(days)
France	0.002% ***	0.7% ***	1.2% ***	-0.3% ***
The Netherlands	0.002% **	2.4% ***	1.7% ***	-0.6% ***
Athens region	-0.050% ***	0.3% *	2.1% ***	
Motorways in France	0.008% ***	2.1% ***	0.8% *	0.5% **
Main roads in France	0.006% ***	0.2% *	0.4% *	-0.3% **
Motorways in the Netherlands	0.010% ***	1.6% ***	0.3% *	0.1% *
Rural roads in the Netherlands	0.000% **	2.8% ***	1.2% **	-0.1% *

*Significance of parameter: ***(t-ratio > 2), **(1< t-ratio <2), *(t-ratio < 1)*
<u>Table 1 :</u> *Weather coefficients in model 5.1, of the log number of injury accidents for France for the period 1975-2000, for the Netherlands for the period 1987-2005 and for the Athens region for the period 1985-2005*

	Veh-km	RF(mm)	ST(°C)	WT(°C)	DF(days)
Motorways in France	1.04 ***	0.010% ***	1.9% ***	0.5% *	0.9% ***
Main roads in France	0.69 ***	0.010% ***	0.2% *	0.0% *	-0.2% *

*Significance of parameter: ***(t-ratio > 2), **(1< t-ratio <2), *(t-ratio < 1)*
<u>Table 2:</u> *Weather coefficients in model 5.2 of the log number of injury accidents for France for the period 1975-2000*

Key: RF Rainfall, ST : Summer temperature (April to September), WT: Winter temperature (October to March),DF: Days of frost,

6. Applications focusing on extreme weather effects

The magnitude and non-linearity of the effect of the two weather variables - mean temperature and total precipitation - on road safety indicators was demonstrated on a daily time scale through the use of a dataset of injury accidents in the Athens region for the period 1985-2005. The data include daily mean temperature (in °C) and daily precipitation (in mm), and atypicality dummies for mean temperature and precipitation, as described in Section 4.3.

Generalized linear models (GLM) with a negative binomial distribution and a log link function have been specified and estimated (Yannis et al., 2007, Hermans et al., 2006). The daily data variables were included in the model as factors, retaining the lowest level as the reference level. For example, in the temperature grouping, the reference level was the number of days with an average temperature of less than $5^{o}C$, and each of the other levels was entered into the formulation as a factor. Similarly, for the precipitation variables, the number of days without any precipitation was used as the reference case. The results of the final daily model are given in the upper part of Table 3. The positive correlation between mean temperature and the number of injury accidents, and the negative correlation between precipitation and the number of injury accidents were both confirmed, and in both cases the magnitude of the effects increased in the case of higher values. It should be noted that due to the log link in the model, the effects are given by the exponent of the coefficients.

In order to identify the impact of extreme conditions on a monthly basis, the same dataset of injury accidents in the Athens region was aggregated at a monthly level, and also modelled using a GLM with a binomial negative function and a log link function (the results for which are given in the lower part of Table 3). The variables of interest include the number of days with unusually low temperatures and very high precipitation in the month (number of atypical days with respect to temperature and precipitation), as already defined in Section 4.3, and the mean monthly temperature (average of the daily means in the month). Dummies for the months of August and December were also included.

With respect to the following interpretation of the weather coefficients, it should be noted that due to the log link in the model, the effects are given by the exponent of the coefficients. One additional day in the month with an extremely low temperature (i.e. more than two standard deviations below the expected typical average for that day) results in 2% fewer injury accidents in the month (note again the log link and hence the need to use the exponent of the coefficient). Similarly, each additional day with extremely high precipitation results in more than 1% fewer injury accidents in that month. For each degree Celsius that the average monthly temperature increases, injury accidents increase by approximately 1%. Furthermore, all other things being equal, there are 27% fewer accidents in the month of August and 6% more accidents in the month of December: these impacts are probably not due to weather (as this has been captured by the other coefficients), but most likely to exposure.

To sum up, the results from the model suggest that low temperatures are negatively correlated with the number of injury accidents (which is very similar to the negative correlation between the occurrence of frost and the number of injury accidents in France) and that unusually high precipitation (for Greece) is negatively correlated with the number of injury accidents. Binary (0/1) dummy variables for different months were also considered. Accidents increase in June, October, November and December- but these increases do not appear to be significant at the monthly level at the usual confidence level - and decrease significantly in August: these results are no doubt linked to simultaneous changes in exposure.

Daily model	Estimate	Std.Error	z value
(Intercept)	2.975	0.024	122.274
MeanTemp<5 ºC	---*	---*	---*
5<MeanTemp<10 ºC	0.194	0.026	7.550
10<MeanTemp<20 ºC	0.280	0.025	11.389
20<MeanTemp<30 ºC	0.334	0.025	13.353
MeanTemp>30 ºC	0.385	0.030	12.959
TotalPrecipitation == 0	---*	---*	---*
0<TotalPrecipitation<5mm	-0.024	0.010	-2.320
5<TotalPrecipitation<10mm	-0.073	0.023	-3.180
10<TotalPrecipitation<20	-0.097	0.023	-4.226
20<TotalPrecipitation<40	-0.190	0.031	-6.038
40<TotalPrecipitation<60	-0.458	0.074	-6.147
TotalPrecipitation>60mm	-0.557	0.103	-5.413
June	0.068	0.013	5.288
August	-0.257	0.013	-19.126
OctNovDec	0.047	0.007	6.511
Null deviance:	8754.5	(7659 d.o.f.)	
Residual deviance:	7768.4	(7646 d.o.f.)	
AIC:	51847		

Note: * indicates reference level

Monthly model	Estimate	Std. Error	z value
Intercept	6.535	0.033	196.352
Days in month with unusually* low temp	-0.017	0.006	-2.699
Mean monthly temperature	0.010	0.0016	6.377
Days in month with unusual* precipitation	-0.010	0.0054	-2.012
August dummy variable	-0.310	0.036	-8.529
December dummy variable	0.061	0.034	1.767

* unusual is defined as two standard deviations above/below mean daily value

Null deviance:	354.78	(250 d.o.f.)	
Residual deviance:	251.62	(245 d.o.f.)	
AIC:	3095.9		

Table 3. Model specification and estimation results for accidents (top: daily model with clustered data, bottom: monthly model with extreme weather variable data – GLM, negative binomial distribution using log link)

7. Discussion of the results

7.1. Averaged weather effects

We have highlighted significant correlations between average weather variables and aggregate numbers of injury accidents for several regions of Europe: France, the Netherlands and the Athens region, on a monthly time scale over a long period. Similarities were found between France and the Netherlands but the results for rainfall effects differed for the Athens region. However, the magnitude and even the sign of the correlations vary according to the type of network considered.

Rainfall is positively correlated with the number of injury accidents both in France and in the Netherlands. This finding holds for main roads and motorways in France and for rural roads and motorways in the Netherlands. The opposite was found for the Athens urban region. The inclusion of exposure to risk in the models, which was significant on main roads and motorways in France, led to the conclusion that the rainfall effect on the interurban network in France is mainly direct on motorways and partly indirect on main roads. Temperature was positively correlated to the number of injury accidents in all the cases we examined: France, the Netherlands, Athens (only in the winter), and for each type of network considered. As temperature was considered separately for two 6-month periods of the year ('summer' and 'winter'), the statistical significance of the correlation appears to vary according to the period considered: the significance is larger in the winter for the Netherlands and in the summer for France.

As regards the results relating to France, it is noted that the coefficients of the weather variables are extremely close to those estimated with a ARIMA model previously fitted to the same dataset (Bergel, Depire, 2004b); this is not surprising as equivalences between types of structural models and types of ARIMA models hold, but plays in favour of their validity.

7.2. Extreme and averaged weather effects

In addition to the results obtained with average weather variables, the effects of extremely low temperatures and extremely high precipitation could be highlighted for the Athens region. These results are consistent with previous results obtained for France with the same method and with an ARIMA model, with the exception of the negative correlation between precipitation and the number of injury accidents for the Athens region.

When fitting the ARIMA model to the numbers of injury accidents in France, extremely low temperatures in the winter were found to be even more significant than average temperatures, and also more significant than the occurrence of frost. When restricting the ARIMA model to main roads and motorways, extremely low temperatures in the winter and extremely high occurrence of frost were found to be mainly significant of the changes in the number of vehicle kilometers driven on these two networks. However, these are indicators of an indirect effect through changes in exposure, and therefore did not bring an additional explanation of the direct effect on the risk level (the number of injury accidents under constant exposure).

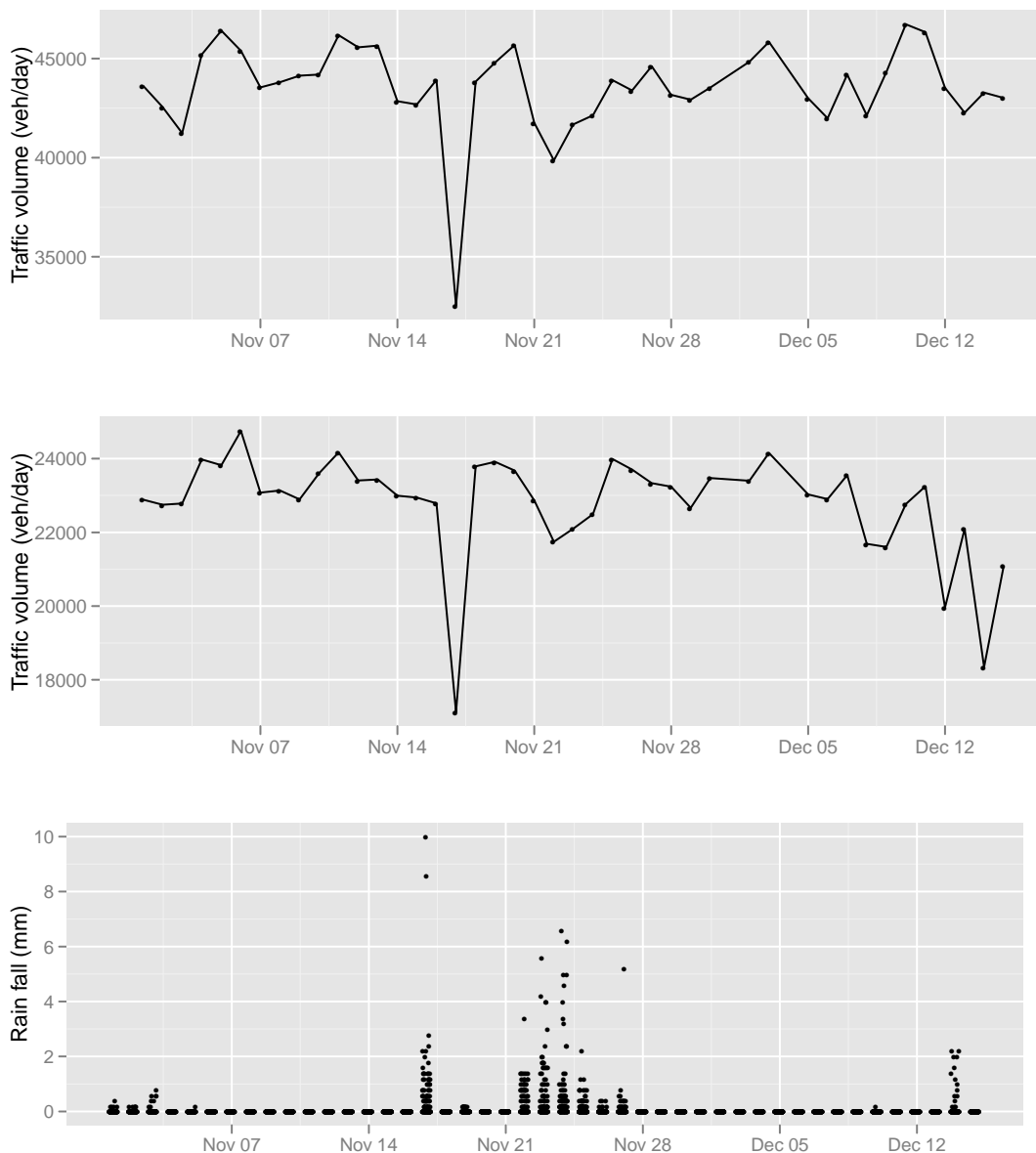## 7.3. The effect of rainfall in the Athens region: an exception

Figure 5: Traffic volume (from 3 main roads) and rainfall data for Athens in late 2005

The results of the model estimation for Athens show that precipitation reduces the number of traffic accidents. This finding could be explained by a reduction in exposure because of a decrease in traffic volume due to the rain. To validate this hypothesis, traffic and weather data were collected for the period of interest. While historical traffic data are available, the only available weather data (at http://hoa.ntua.gr) are for the period since the summer of 2005. Therefore, the first large rain events considered for this analysis were those in the winter of 2005-6, which are at the very end of the period analysed in the model. The traffic data for three main roads in Athens (Kifisias Ave., Alexandras Ave., and Athinon Ave., inbound) are presented in the top two graphs in Figure 5, along with rainfall data (mm per 10 minutes) from the meteorological station of Zografou (near the NTUA campus). A negative correlation between rainfall and traffic volume is apparent from visual inspection. The highest precipitation values on November 17 led to a sharp decrease in flow. A second period of rainy days between November 22 and November 27 also resulted in a decrease in flow, albeit not so

17

sharp. This may be partly due to the fact that the precipitation during these rainfall events was not as heavy as in the previous event. Furthermore, one can observe that traffic flow dropped during the first days of rain (November 22), but rose again afterwards to near its usual level. This can be explained by the fact that people initially cancelled their activities, resulting in lower flow. However, after a couple of days they were no longer able to keep postponing (some of) their activities and were forced to travel. A final rainfall event on December 14 produced a sharp decrease in the flow at one of the sensor locations, while its impact is not so clear at the other two locations: this could be attributed to variability in rainfall across the Athens basin. A smaller rainfall event on November 3rd is also clearly reflected in flow reductions at all three sensor locations.

A more rigorous statistical procedure was also used to validate this correlation, in particular computing Pearson's product moment correlation coefficient (between daily flow and rainfall). The results – summarized in Table 4 - confirm the presence of (negative) correlation. Similar results were also obtained based on Kendall's tau and Spearman's rho (Hollander and Wolfe, 1973).

Sensor location	Pearson's correlation coefficient	t-test	p-value
Kifisias Ave. (inbound traffic)	-0.63	-5.212	5.65E-06
Alexandras Ave. (inbound traffic)	-0.48	-3.509	0.00115
Athinon Ave. (inbound traffic)	-0.62	-5.015	1.07E-05

Table 4. Statistical analysis of correlation between traffic volume and rainfall

In sum, the negative correlation between rainfall and traffic volumes validates the hypothesis that rain reduces exposure in Athens region. Although it cannot be tested on all types of network or geographic area, this result may logically apply to other cases. A similar negative correlation, expressed as a negative elasticity value of mileage with respect to precipitation, between rainfall and traffic volumes was also found at an aggregate level with the DRAG-approach when total aggregate mileage was modelled in relation to weather data (Gaudry, Lassarre, 2000). However, it is rare to have monthly data for a long period of time for both rainfall and mileage at national or region level, and significant results are generally only obtained for the level where mileage data are available. For France, it has been shown that 100 mm of additional rain during a month results in a decrease in traffic of about 0.1% on the main road network which gathers main roads and on motorways, and about 0.2% on toll motorways alone (Bergel, Depire, 2004b).

Finally, the influence of rain on the risk level (the number of injury accidents under constant exposure) needs to be investigated, as it might also be the case that in the Athens urban network, where rainfall is less frequent and the road users distribution differ, the overall risk level decreases in case of rain.

8. Conclusion

In addition to the discussion of the results, the main points of discussion relate to the choice of the method, in particular the appropriate time series technique to use, the comparability of the results between European countries or regions, and the possibility of extending the model to other European countries or regions. The choice of the type of model used should not impact these results as, for instance for France, it has been demonstrated that different ARIMA models used with the same weather variables for fitting the aggregate number of injury accidents provided consistent results. Concerning the comparability of the results, the similarity of the results for France and the Netherlands could be extended to other European countries or regions with similar climate. On the contrary, it has been observed that rainfall leads to a decrease of road accidents in the Athens urban area where rainfall is less frequent, and this should be tested for other urban networks.

Furthermore, controlling for exposure by including the number of vehicle kilometres driven in the models improves the interpretability of the model. In the case of the interurban network in France, for instance, the rainfall effect appears to be mainly direct on motorways – as exposure is unchanged, and to be partly indirect on main roads – as a result of changes in exposure. However, it is only possible to control for exposure if aggregate risk exposure datasets are available, and recommendations for enriching national risk exposure databases have been made (Yannis et al, 2005). There is certainly a need for further research that explores the combined effect of weather and traffic (volume, speed, etc.) on the frequency and severity of road traffic accidents on the various types of road in Europe and worldwide.

The current limitation is that the effects of extreme weather (heavy rain, low and high temperatures) are not yet modelled in a systematic manner, and this should be tested on groups of countries with similar climates. The possibility of including all the average data and extreme value data recorded in a month in a single injury accident data model has been investigated, and research in this direction should be continued as it may significantly improve the model's explanatory fit and our understanding of how the weather affects road accident risk at a monthly scale.

Although the weather itself cannot be changed, understanding how it affects road safety is useful for both analysis and prevention. Prevention is obvious an important application, and takes different forms: information campaigns, road improvements, local warning systems. Analysis of the impact of adverse weather conditions on the accident process (be it via the infrastructure, the vehicle or driver behaviour) is performed at a disaggregate level by means of observation or simulation. But at an aggregate level, for example for a whole country, using the weather as an additional risk factor makes it possible to correct the risk indicators for its impact and thus better identify the impacts of other risk factors and safety policies. For national transport authorities in particular, monitoring the short-term trends of road safety indicators, once they have been corrected for transitory effects due to atypical weather conditions, is a new direction for identifying and anticipating changes in trends, and this approach is currently being implemented with data from France using the Giboulee model (ONISR, 2012).

Acknowledgements

References

Aron, M, Bergel-Hayat, R, Saint Pierre, Guillaume, Violette, E. (2007). *Added Risk by Rainy Weather on the Roads of Normandie-Centre Region In France*, Proceedings of 11th WCTR (CD-Rom), World Conference on Transport Research Society.

Bergel R. (1992). *Utilisation de modèles ARIMA avec variables externes dans le secteur des transports.* Proceedings of the 6th WCTR, World Conference on Transport research Society (in French).

Bergel-Hayat R. (2008). *Time series analysis with explanatory variables: applications to transport demand and to road risk*. Doctorat Thesis, University Paris-Est (in French).

Bergel R., Depire A. (2004a) Recherche d'une forme fonctionnelle d'un modèle agrégé de risque routier. Actes Inrets n° 90 du groupe de travail 2001 du Séminaire Modélisation du tafic, Arcueil (in French).

Bergel-Hayat R., Depire A. (2004b). *Climate, road traffic and road risk – an aggregate approach*. Proceedings of 10th WCTR (CD-Rom), World Conference on Transport Research Society.

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(3)*, 1180-1190.

Brockwell P.J., Davis R.A. (2002) *Introduction to time series and forecasting*, second edition, Springer Verlag.

Brodsky H., Hakkert S. (1988). Risk of road accident in rainy weather. *Accident Analysis and Prevention*, Vol.20, N°3, pp 161-176.

Caliendo C., Guida M., Parisi A., (2007). A crash-prediction model for multilane roads. Accident Analysis and Prevention 39, pp. 657-670.

Chang L.Y., Chen W.C. (2005). Data mining of tree-based models to analyze freeway accident frequency. Journal of Safety Research 36, pp. 365-375.

Commandeur J.J.F., Koopman S.J. (2007). *An introduction to state space time series analysis,* Oxford University Press.

Dupont E., Martensen H.,(Eds) (2007) *Multilevel modelling and time series analysis in traffic safety research - Methodology and Manual*. Deliverables D7.4 and D7.5 of the EU FP6 project SafetyNet, http://www.erso.eu

Durbin J, Koopman S.J. (2001) *Time Series Analysis by State Space Methods,* Oxford UniversityPress

Eisenberg, D. (2004). The Mixed Effects of Precipitation on Traffic Crashes, *Accident Analysis and Prevention,* Vol. 36, pp. 637-647.

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), pp. 1-20.

Fridstrom, L., Ingebrigtsen, S. (1991). An aggregate accident model based on pooled, regional time-series data. Accident Analysis and Prevention 23(5), pp. 363-378.

Gaudry M., Lassarre S. (Eds.) (2000). *Structural Road Accident Models - The International DRAG Family*, Pergamon.

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*, Vol.9 n°1, pp.63-76.

Jaeger L. (1998). *L'évaluation du risque dans le système des transports routiers par le développement du modèle TAG*. Thèse de doctorat ès sciences économiques, Université Louis Pasteur, Strasbourg (in French).

Karlaftis M, Yannis G. (2010) Weather Effects on Daily Traffic Accidents and Fatalities: A Time Series Count Data Approach, Proceedings of the 89th Annual meeting of the Transportation Research Board, Washington.

Keay K., Simmonds I. (2006). Road accidents and rainfall in a large Australian city. Accident Analysis and Prevention 38, 445-454.

Malyshkina N.V., Mannering F.L., Tarko A.P. (2008). Markov switching negative binomial models: An application to vehicle accident frequencies. Accident Analysis & Prevention 41(2), 217-226.

Myles Hollander & Douglas A. Wolfe (1973), Nonparametric Statistical Methods. New York: John Wiley & Sons. Pages 185-194

ONISR (2012) *La sécurité routière en France: bilan de l'année 2011*. Paris, France: La Documentation Française.

Scott P. P. (1986). Modelling Time-Series of British Road Accident Data. *Accident Analysis & Prevention*, Vol. 18 n°2, pp.109-117.

Shankar V., Mannering F., Barfield W. (1995). Effect Of Roadway Geometrics And Environmental Factors On Rural Freeway Accident Frequencies. Accident Analysis & Prevention 27 (3), pp. 371-389.

Stipdonk H. (Ed.) (2008). *Time series applications on road safety developments in Europe.* Deliverable D7.10 of the EU FP6 project SafetyNet   http://www.erso.eu

Yannis G et al. (2005). *State of the Art Report on Risk and Exposure Data*. Safetynet Deliverable 2.1, http://www.erso.eu

Yannis, G, Antoniou, C, Papadimitriou, N. (2007). Road Casualties and Enforcement: Distributional Assumptions Of Serially Correlated Count Data. *Traffic Injury Prevention,* Vol. 8:3, pp. 300-308.

_____