Forecasting the number of road traffic fatalities in Greece

Constantinos Antoniou1, George Yannis1, Eleonora Papadimitriou2, Sylvain Lassarre2, Jacques F. Commandeur3, Frits Bijleveld3, Emmanuelle Dupont3

1 National Technical University of Athens, Athens, Greece
2 IFSTTAR, Paris, France
3 SWOV Institute for Safety Research, the Netherlands
4 Belgian Road Safety Institute, Brussels, Belgium

Introduction

A number of approaches for modelling road safety developments have been proposed. The objective of this paper is to apply the DaCoTA methodology for the development of structural time series models for Greece, in order to forecast road traffic fatalities for the period 2011-2020.

Analysis methods

Structural time-series models: Local Linear Trend (LLT) and Latent Risk Time-Series (LRT)

A basic concept in road safety is that the number of fatalities is a function of the road risk and the level of exposure of road users to this risk. In order to model the evolution of fatalities it is required to model the evolution of two parameters: a road safety indicator and an exposure indicator.

\[ \text{Traffic volume} = \text{Exposure} \times \text{Risk} \]

The latent variables (log [exposure] and log [risk]) need to be further specified by “state” equations, describing the development of the latent variable.

1. LLT model

\[ \text{Measurement equation} \]

\[ \log \text{Number of Fatalities} = \log \text{LatentFat} + \varepsilon \]

\[ \log \text{Traffic Volume} = \log \text{Exposure} + \varepsilon \]

2. State equations

LatentFat = LogLatentFat + \varepsilon

TrafficVolume = LogTrafficVolume + \varepsilon

The Equation now includes the Risk (not the fatalities)

DaCoTA Model selection logic

1. Investigate exposure:
   - Do the available exposure data make sense?
   - Can any sudden changes in the trend or slope be explained from some real events?

2. Establish whether the two series are statistically related: a SUTSE model is developed and based on the diagnostics, the modeler needs to decide whether the two-time-series are correlated.

3. Determine whether an LLT or an LRT model should be pursued:
   - If one or more of the null-hypotheses regarding the correlation of the disturbances is rejected, the time-series may be related and therefore an LRT can be estimated.
   - If, on the other hand, none of the hypotheses can be rejected, then there is no evidence that the two-time-series are correlated and therefore an LLT model would be more appropriate.

Results

SUTSE Model

- The correlation between the two levels (p=0.33) and two slopes (p=0.77) is not significant. The value of the correlation is 0.35 between the two levels and 0.24 between the two slopes.
- The measurement errors for exposure and fatalities are correlated at 0.46. The investigation of the SUTSE model indicates that a relation between vehicle fleet and fatalities in Greece is not present. Therefore an LLT model fit is for Greece.

LLT model

- The full model (LL1) was run first. None of the residual tests indicated a violation of the underlying assumptions. Furthermore, the level and slope components were significant.
- A new model (LL2) with additional interventions was estimated. While the fit of this model improved compared to the original model, the slope component became insignificant.
- Therefore, a third model (LL3) was also run, with the interventions, but keeping the slope of the fatalities fixed.

Intervention variables

- Change in fatality recording system (slope of fatalities 1996)
- Financial crisis (level of fatalities 1986)
- Introduction of car scrappage system (level of fatalities 1991)

Forecasts

The forecasts provide an indication of the fatality numbers that could be expected in Greece between 2011 and 2020 provided that the current trends keep on following throughout these years.

Short-term forecasts validation

- The estimated DaCoTA forecasts in all European countries appear to be realistic and within acceptable confidence intervals.
- The forecasts are based on “business-as-usual” scenarios.
- In Greece the economic recession effect is visible at the end of the fatalities series, which in turn affects the final forecasts. A scenario in which the forecasted value for 2020 is somewhat increased, may in this case provide a more realistic picture of future developments, as it takes into account the fact that the recession will end sooner (while in the baseline “business-as-usual” scenario, the effect of the recession is assumed to continue in the future).

Acknowledgments

This paper is based on work carried out within the scope of the DaCoTA (Data Collection Transfer and Analysis) project of the 7th framework RTD Program of the European Commission. The author would like to acknowledge the important contributions of all the partners involved in the DaCoTA with “Decision Support” in the development of the methodology and system used in this paper.