6ο Πανελλήνιο Συνέδριο Οδικής Ασφάλειας
www.nrso.ntua.gr/roadsafety2015

Αμφιθέατρο
Υπουργείου Οικονομίας, Υποδομών, Ναυτιλίας και Τουρισμού
12-13 Μαρτίου 2015

Forecasting the number of road traffic fatalities in Greece

Constantinos Antoniou[1], George Yannis[1], Eleonora Papadimitriou[1], Sylvain Lassarre[2], Jacques F. Commandeur[3], Frits Bijleveld[3], Emmanuelle Dupont[4]

[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.
- During the last decade, the modeling approach of **structural time-series models** is applied by several researchers, in which latent variables are decomposed into components.
- The DaCoTA research project of the European Commission aimed to obtain **forecasts for the number of traffic fatalities in each of the European countries in 2020** in a similar way by means of the structural time series approach, using comparable data as much as possible.
  - ✓ to develop **robust models** for modeling the relationship between mobility and risk and examine the effect of mobility on risk.
  - ✓ to develop (and apply) a structured methodology for the **selection of the optimal forecasting models**, based on a number of criteria, diagnostics and measures of goodness of fit

Objectives

- **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:

$$Traffic\ volume = Exposure$$
$$Number\ of\ fatalilties = Exposure \times Risk$$

- When the logarithm of the Equations is taken (and the error term is explicitly written out) the "measurement equations" of the model can be rewritten as:

$$Traffic\ volume = Exposure$$
$$Number\ of\ fatalities = Exposure \times Risk$$

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

- LLT model
- Measurement equation

$$\log Number\ of\ Fatalities_t = \log LatentFat._t + \varepsilon_t$$

- State equations

$$Level(\log LatentFat_t) = Level(\log LatentFat_{t-1}) + Slope(\log LatentFat_{t-1}) + \xi_t$$
$$Slope(\log(LatentFat_t) = Slope(\log LatentFat_{t-1}) + \zeta_t$$

LRT model

$$\log Number\ of\ Fatalities_t = \log Exposure_t + \log Risk_t + \varepsilon_t^{\ f}$$
$$\log TrafficVolume_t = \log Exposure_t + \varepsilon_t^{\ e}$$

$$Trend(\log Risk_t) = Level(\log Risk_{t-1}) + Slope(\log Risk_{t-1}) + \xi_t^{\ r}$$
$$Slope(\log Risk_t) = Slope(\log Risk_{t-1}) + \zeta_t^{\ r}$$
$$Level(\log Exposure_t) = Level(\log Exposure_{t-1}) + Slope(\log Exposure_{t-1}) + \xi_t^{\ e}$$
$$Slope(\log Exposure_t) = Slope(\log Exposure_{t-1}) + \zeta_t^{\ e}$$

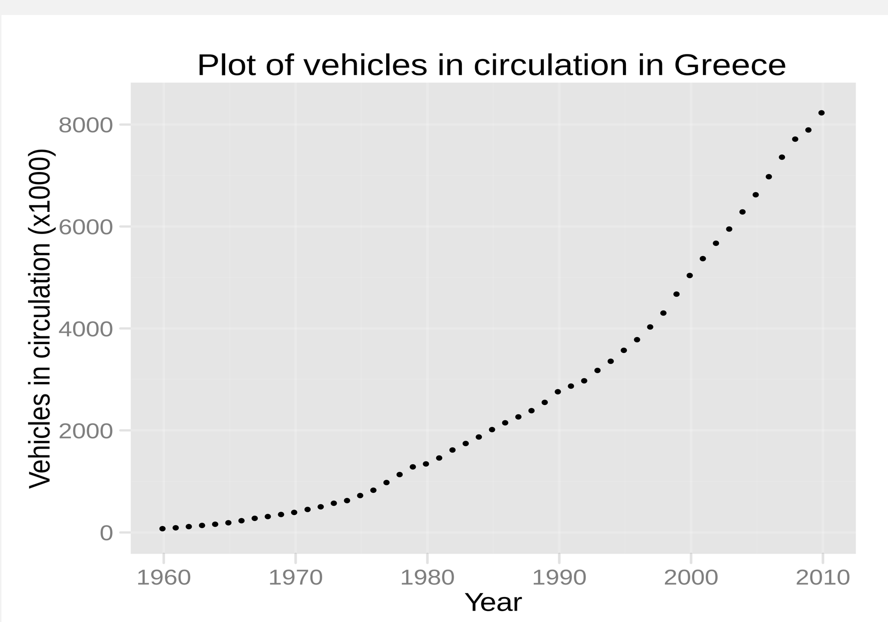
The Equation now includes the Risk (not the fatalities)

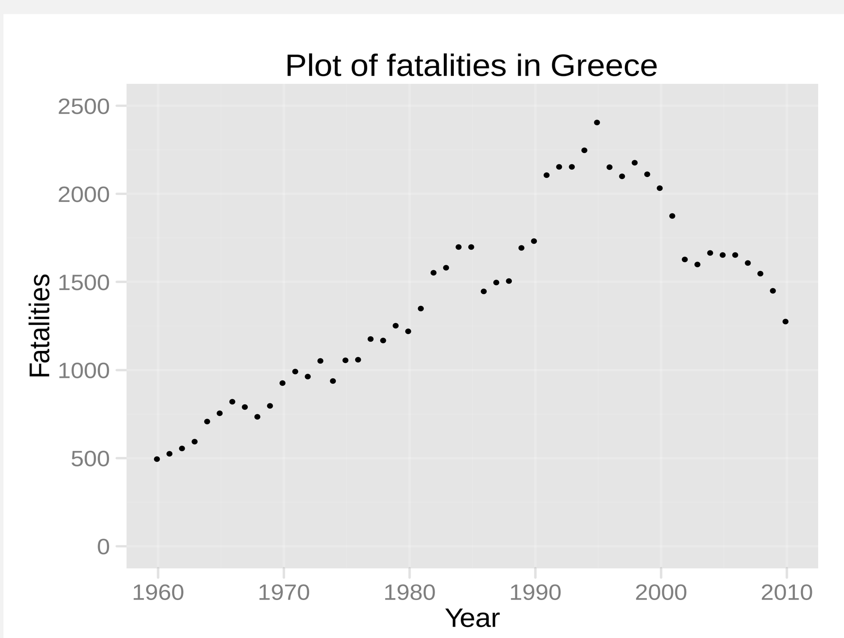
- SUTSE (Seemingly Unrelated Time Series) model
  A preliminary step in establishing whether the two time-series may be correlated.

DaCoTA Model selection logic

1. Investigate exposure:
   - Do the available exposure data make sense?
   - Can any sudden changes in the level 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.

Data





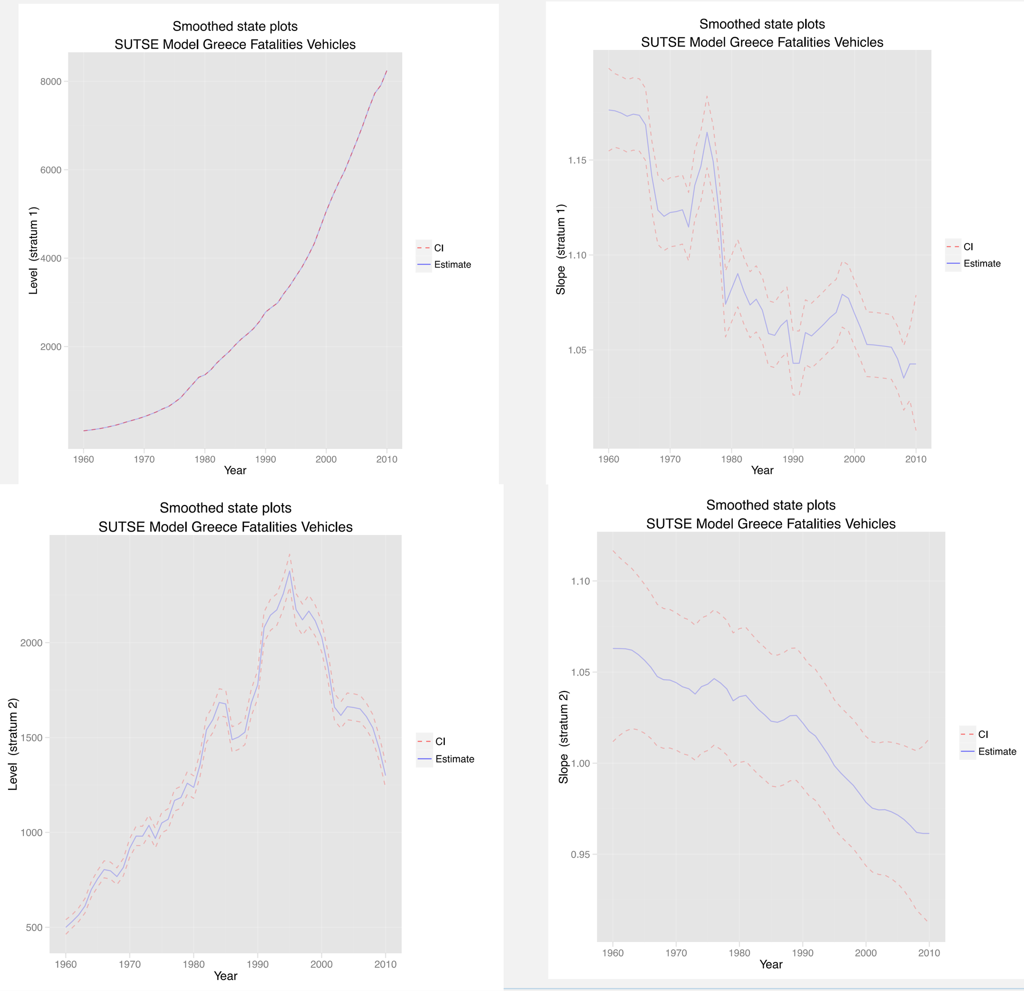
- The **vehicle fleet** is used as a "proxy" of the actual exposure
- source: Ministry of Infrastructure, Transport and Networks

- The **fatalities** (killed at 30 days from the accident)
- source: Hellenic Statistical Authority

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 6.4E-05.
- The investigation of the SUTSE model indicates that **a relation between vehicle fleet and fatalities in Greece is not present**. Therefore an LLT model is fit for Greece.



Model title	***SUTSEGreece1***
Model description	SUTSE full model
Model Criteria	
log likelihood	237.76
AIC	-475.17
Variance of the state components	
Level exposure	1.33E-04 nsc
Level fatalities	4.06E-03 *c
Slope exposure	2.17E-04 *c
Slope fatalities	1.09E-04 *c
Correlations between the state components	
level-level	0.35
slope-slope	0.24
Observation variance	
Observation variance exposure	1.014E-09 ns
Observation variance fatalities	1.689E-09 ns

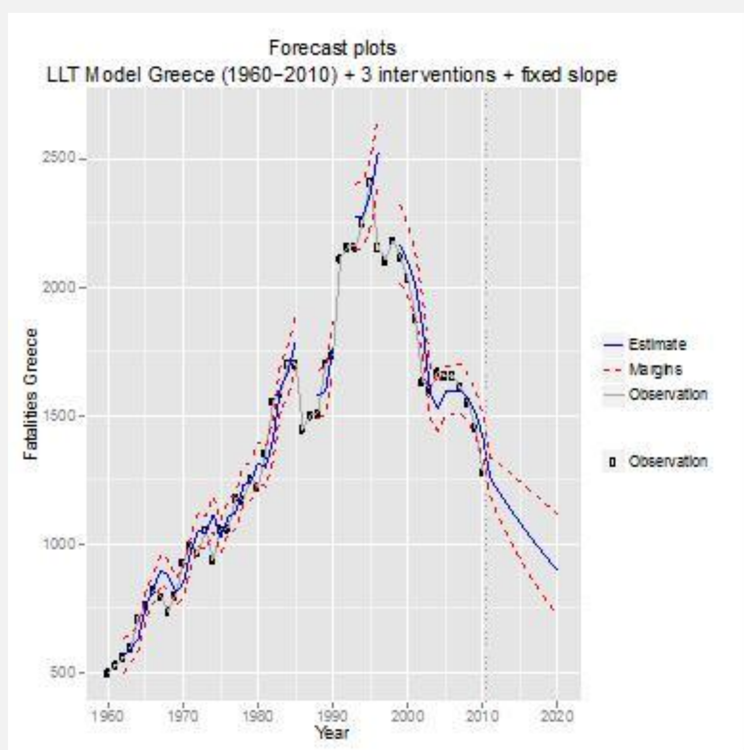
LLT model

- The full model (LLT1) 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 (LLT2) 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 (LLT3) 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)

Model title	***LLT 1***	***LLT 2***	***LLT3***
Model description	LLT for Greece – full model	LLT for Greece – with 3 interventions	LLT for Greece – with 3 interventions – fixed slope
Model Criteria			
ME4 Fatalities	-131	-61.4	-59.4
MSE4 Fatalities	28162.3	10047.9	9689.6
log likelihood	85.66	65.84	65.82
AIC	-171.21	-131.56	-131.55
Model Quality			
Box-Ljung test 1 Fatalities	2.73	2.96	0.29
Box-Ljung test 2 Fatalities	3.63	4.3	2.78
Box-Ljung test 3 Fatalities	5.82	4.33	4.03
Heteroscedasticity Test Fatalities	0.79	0.75	0.76
Normality Test standard Residuals Fatalities	0.8	1.95	2.06
Normality Test output Aux Res Fatalities	1.28	1.13	1.17
Normality Test State Aux Res Level risk	1.61	1.34	1.1
Normality Test State Aux Res Slope risk	0.05	0	0
Variance of state components			
Level risk	3.91E-03 *	2.61E-03 *	2.67E-03*
Slope risk	1.25E-04 *	6.92E-06 ns	-
Observation variance			
Observation variance risk	1.00E-09 ns	1.00E-09 ns	1.00E-09ns
Intervention and explanatory variables tests			
Change in fatality recording system (slope fat 1996)		-0.074 *	-0.080 *
Financial crisis (level fat 1986)		-0.209 *	-0.211 *
Introuction of car scrappage system (level fat 1991)		0.152 *	0.147 *

Forecasts

	Fatalities		
Year	**Predicted**	Confidence Interval	
2011	**1257**	1118	1414
2012	**1211**	1029	1426
2013	**1167**	953	1429
2014	**1124**	885	1427
2015	**1083**	824	1422
2016	**1043**	769	1415
2017	**1005**	717	1407
2018	**968**	670	1398
2019	**932**	626	1389
2020	**898**	585	1379



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

Year	Forecast fatalities	95% conf. interval (from – to)		Actual fatalities
2011	***1257***	1118	1414	**1141**
2012	***1211***	1029	1426	**1027**

Conclusions and discussion

- 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 authors would like to acknowledge the important contribution of all the partners involved in the DaCoTA WP4 "Decision Support" in the development of the methodologies and outputs used in this paper.