Road Safety Forecasts in Five European Countries Using Structural Time-Series Models

Constantinos Antoniou, Eleonora Papadimitriou & George Yannis - National Technical University of Athens, Greece

Scan this QR code to access this poster online

Abstract

Modeling road safety development is a complex task, which needs to consider both the quantifiable impact of specific parameters, as well as the underlying trends that cannot always be measured or observed. The objective of this research is to apply structural time series models for obtaining reliable medium- to long-term forecasts of road traffic fatality risk, using data from five countries with different characteristics from all over Europe (Cyprus, Greece, Hungary, Norway and Switzerland). Two structural time series models are considered: (i) the local linear trend model and the (ii) latent risk time-series model. Furthermore, a structured decision tree for the selection of the applicable model for each situation (developed within the DACOTA research project) is outlined. First, the fatality and exposure data that are used for the development of the models are presented and explored. Then, the modeling process is presented, including the model selection process, the introduction of intervention variables and the development of mobility scenarios. The forecasts using the developed models appear to be realistic and within acceptable confidence intervals. The proposed methodology is proved to be very efficient for handling different cases of data availability and quality, providing an appropriate alternative from the family of structural time series models in each country. A concluding section providing perspectives and directions for future research is finally presented.

Background & Objectives

- A number of approaches for **modelling road safety developments** have been proposed
- During the last decade, the modeling approach of **structural time-series models**, such as those proposed by Harvey & Shephard (1993) is applied by several researchers
- In this approach, which belongs to the family of unobserved component models, latent variables are decomposed into components which are incorporated into the structural models.

Objectives

to apply structural time series models for obtaining reliable medium- to long-term forecasts of fatality risk

- **to develop models** for modeling the relationship between mobility and risk and examine the effect of mobility on risk.
- **to develop a structured methodology** for the selection of the optimal forecasting models, based on a number of criteria, diagnostics and measures of goodness of fit.
- **to demonstrate that the developed approach is robust** and applicable to different conditions and environments, by applying it to data from five European countries with very different characteristics.

Methodology

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

$$Log\ Traffic\ volume = \log exposure + random\ error\ in\ traffic\ volume$$
$$Log\ Number\ of\ fatalities = \log exposure + \log risk + random\ error\ of\ fatalities$$

- The latent variables [log (exposure) and log (risk)] need to be further specified by **"state" equations**, which, once inserted in the general model, describe 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**
  - Measurement equations $$\log TrafficVolume_t = \log Exposure_t + \varepsilon_t{}^e$$ $$\log Number\ of\ Fatalities_t = \log Exposure_t + \log Risk_t + \varepsilon_t{}^f$$
  - State equations $$Trend(\log Risk_t) = Level(\log Risk_{t-1}) + Slope(\log Risk_{t-1}) + \xi_t{}^r$$ $$Slope(\log Risk_t) = Slope(\log Risk_{t-1}) + \zeta_t{}^r$$ $$Level(\log Exposure_t) = Level(\log Exposure_{t-1}) + Slope(\log Exposure_{t-1}) + \xi_t{}^e$$ $$Slope(\log Exposure_t) = Slope(\log Exposure_{t-1}) + \zeta_t{}^e$$

The Equation now includes the Risk (and not the fatalities)

- **SUTSE (Seemingly Unrelated Time Series) model**
  A third class of models, used as a preliminary step in establishing whether the two time-series may be correlated.

Methodology (cont.)

- **Model selection logic**

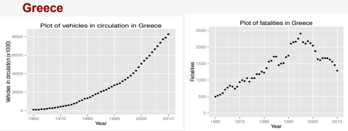
- The family of structural time-series models lends to a large number of assumptions that distinguish the resulting models into different categories.
- Within the framework of the **DaCaTA research project**, a decision process and model selection logic has been developed, in which the following steps are considered:

1. **Investigate exposure**: the first step in every modeling effort is to assess the quality and characteristics of the underlying data.
   - 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, then 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.

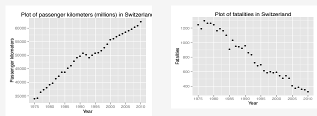
Model application

- **Data collection & analysis**

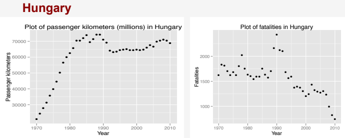
Greece



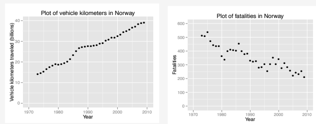
Switzerland



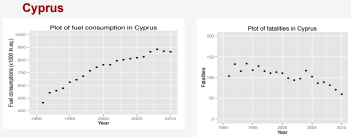
Hungary



Norway



Cyprus



- **Models by country**
- *Model selection process for Switzerland*

Model type	LRT full	LRT restricted	LRT restricted with interventions
Model Criteria			
ME10 Fatalities	-6037	-5374	-4918
MSE10 Fatalities	5.56827	4.79550	4.35124
log likelihood	18156	17675	17071
AIC	-36262	-35322	-34115
Variance of state components			
Level exposure	1.61E-04	-	-
Level risk	5.84E-04	7.66E-04 *	7.79E-04 *
Slope exposure	6.46E-06	4.15E-05 *	6.84E-06 *
Slope risk	9.41E-06	-	-
Correlations between state components			
level-level	0.64	-	-
slope-slope	1	-	-
Observation variance			
Observation variance exposure	2.95E-06	5.95E-05 *	7.32E-05 *
Observation variance risk	4.18E-06	2.99E-04	2.47E-04
Interventions			
(1993 exposure level)	-	-	-0.0501062 *
Model Quality			
Box-Ljung test 1 Exposure	0.228	121.897	136.467
Box-Ljung test 2 Exposure	0.801	241.477	503.337
Box-Ljung test 3 Exposure	0.8525	329.751	583.505
Box-Ljung test 1 Fatalities	216.579	286.154	263.737
Box-Ljung test 2 Fatalities	255.335	316.426	265.737
Box-Ljung test 3 Fatalities	311.375	376.553	33.562
Heteroscedasticity Test Exposure	0.386	0.454	0.807
Heteroscedasticity Test Fatalities	269.171	302.679	280.834
Normality Test standard Residuals Exposure	5.99954*	132.338	329.738
Normality Test standard Residuals Fatalities	0.0189	0.312	0.525
Normality Test output Aux Res Exposure	0.0439	0.458	353.243
Normality Test output Aux Res Fatalities	124.914	159.349	183.043
Normality Test State Aux Res Level exposure	338.426	307.695	0.0385
Normality Test State Aux Res Slope exposure	129.975	0.706	0.183
Normality Test State Aux Res Level risk	3.574	8.381*	7.704*
Normality Test State Aux Res Slope risk	0.068672	3.92E-05	3.37E-05

- The SUTSE model revealed a strong correlation between the fatality and the exposure series
- The full LRT model suggests that both the level and slope of both components are non significant (& common components)
- Best fitting restricted LRT: fixed level exposure & slope risk
- Intervention variables: a change in exposure level on 1993

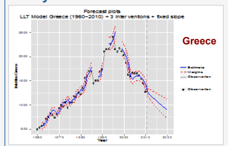
- *Final models for the other countries*

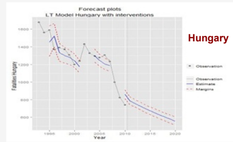
Country	Greece	Hungary	Norway	Cyprus
Model Type	LLT restricted with interventions	LLT deterministic with interventions	LRT restricted	LRT full
Model Criteria				
ME10 Fatalities	-251.5	196297	24	-2.59
MSE10 Fatalities	70572.97	58253.62	967.3	118.25
log likelihood	65.82	167835	156.941	52.96
AIC	-131.55	-324559	-313.612	-105.02
Variance of state components				
Level exposure	-	-	-	9.22E-05
Level risk	2.67E-03*	-	3.84E-03 *	6.53E-04
Slope exposure	-	-	3.16E-04 *	1.08E-04 *
Slope risk	-	-	-	8.10E-06
Correlations between state components				
level-level	-	-	-	-1
slope-slope	-	-	-	1
Observation variance				
Observation variance exposure	-	-	1.45E-06	3.60E-04
Observation variance risk	1.00E-09	1.88E-03 *	5.40E-04	1.11E-03
Intervention and explanatory variables tests				
(slope fat 1996)	-0.080 *	-	-	-
(level fat 1986)	-0.211 *	-	-	-
(level fat 1991)	0.147 *	-	-	-
(level fat 2002)	-	0.220 *	-	-
(level fat 2008)	-	-0.259 *	-	-
Model Quality				
Box-Ljung test 1 Exposure	-	-	0.15	4.70*
Box-Ljung test 2 Exposure	-	-	1.34	5.3
Box-Ljung test 3 Exposure	-	-	2.35	5.67
Box-Ljung test 1 Fatalities	0.29	150.267	0.42	1.62
Box-Ljung test 2 Fatalities	2.78	188.584	0.42	1.91
Box-Ljung test 3 Fatalities	4.03	322.822	1.91	2.27
Heteroscedasticity Test Exposure	-	-	0.34	0.47
Heteroscedasticity Test Fatalities	0.76	263.094	1.1	2.45
Normality Test standard Residuals Exposure	-	-	1.63	1.98
Normality Test standard Residuals Fatalities	2.06	182.026	1.35	5.89
Normality Test output Aux Res Exposure	-	-	0.84	0.92
Normality Test output Aux Res Fatalities	1.17	118.117	0.55	3.74
Normality Test State Aux Res Level exposure	-	-	0.76	14.54***
Normality Test State Aux Res Slope exposure	-	-	1.71	0.16
Normality Test State Aux Res Level risk	1.1	0.943	1.76	2.69
Normality Test State Aux Res Slope risk	0	145.961	0.06	0.08

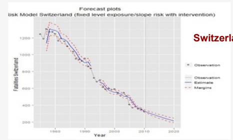
Note: * denotes significant at 95% level, *** denotes significant at 99.9% level

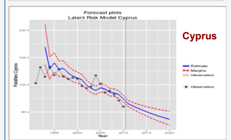
Model application (cont.)

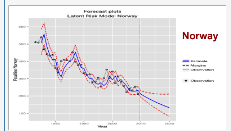
- **Synthesis & Forecasts**









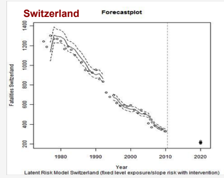


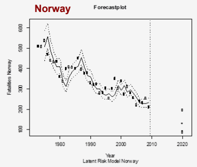
- In **Greece**, there were approximately 1300 fatalities on 2010, and the forecast for 2020 is 898 fatalities (95% confidence interval: 585-1379 fatalities).
- In **Hungary**, there were 740 fatalities on 2010, and the forecast for 2020 is 555 fatalities (95% confidence interval: 472-653 fatalities).
- In **Switzerland**, there were 329 fatalities on 2010, and the forecast for 2020 is 216 fatalities (95% confidence interval: 167-278 fatalities). The number of vehicle kilometres is expected to increase up to 70.8 billion in 2020, compared to 62.3 in 2010.
- In **Norway**, there were 212 fatalities on 2009, and the forecast for 2020 is 132 fatalities (95% confidence interval: 53-333 fatalities). The number of vehicle kilometres is expected to increase up to 42 billion in 2020, compared to approximately 40 in 2009.
- In **Cyprus** there were 60 fatalities on 2010, and the forecast for 2020 is 37 (95% confidence interval: 53-333 fatalities). The fuel consumption is expected to increase up to 894 million tn.eq. in 2020, compared to 860 million in 2010.

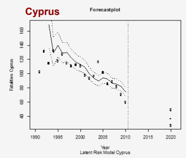
Provided that the trends will keep on following the developments that they have shown in the past, and no principal changes occur in the meantime ("business as usual" assumption)

- **Mobility scenarios**

- Fatality forecasts on the basis of three different scenarios for exposure: the exposure as predicted from the selected LRT model plus/minus one standard deviation.







- **Overview for the five countries**

	Cyprus	Greece	Hungary	Norway	Switzerland
Data available	1990-2010	1960-2010	1970-2010	1970-2009	1975-2010
Exposure	Fuel consumption	Vehicle fleet	Passenger kilometres	Vehicle kilometres	Vehicle kilometres
Recession effect	Yes	No	Yes	No	No
Information on inverventions	No	Yes	Yes	No	No
Data used	1990-2010	1960-2010	1993-2010	1970-2009	1975-2010
Model type	LRT	LLT	LT	LRT	LRT
Interventions	No	Yes	Yes	No	Yes
Forecast 2020	Yes	Yes	Yes	Yes	Yes
Mobility scenario	Yes	No	No	Yes	Yes

Conclusions

- The proposed methodology contributes **meaningful steps for model selection**, starting with SUTSE modeling and proceeding to LLT / LRT, full or restricted, on the basis of sound criteria in each case.
- Nevertheless, **a good knowledge of the road safety and socioeconomic situation** in the examined countries was still necessary, not only for understanding the description and forecasts of the developments, but also for making decisions in data handling, introduction of intervention variables etc.
- The proposed methodology was proved to be very efficient for handling **different cases of data availability and quality**, providing an appropriate alternative from the family of structural time series models in each case.
- The estimated forecasts in all 5 countries appear to be **realistic and within acceptable confidence intervals**.
- These results may be useful for **understanding past developments**, the dynamics and particularities of the relationship between exposure and fatality risk, and the effects of safety interventions or other socio-economic events on mobility and road safety.
- The estimated forecasts reflect the future situation if the existing policy efforts and the socio-economic context extent to the future, and this may be **motivating for devoting additional efforts in outperforming these forecasts**.

Key references

Commandeur J.F.C., Bijleveld, F., Bergel R., Antoniou C., Yannis G., Papadimitriou E. (2012). On statistical inference in time series analysis of the evolution of road safety. Article In Press, Accident Analysis & Prevention, Special Issue on SafetyNet.

Dupont E. & Martensen H. (Eds.) (2012). Forecasting road traffic fatalities in European countries. Deliverable 4.4 of the EC FP7 project DaCoTA.

Harvey, A. C. & Shephard N. (1993). Structural Time Series Models. In G. S. Maddala, C. R. Rao and H. D. Vinod, eds., Handbook of Statistics, Vol. 11, pp. 261-302, Elsevier Science Publishers, B. V.

Lassarre, S. (2001). Analysis of progress in road safety in ten European countries. Accident Analysis and Prevention 33, pp. 743 – 751.

Yannis G., Antoniou, C., Papadimitriou E., Katsochis D. (2011). When may road fatalities start to decrease? Journal of Safety Research 42, pp. 17-25.