ZAGADNIENIA
EKSPLOATACJI
MASZYN
Zeszyt 1 (149) 2007

R. KRYSTEK*, J. ŻUKOWSKA*

Time series – the tool for traffic safety analysis

Key-words

Road safety, time-series, modeling, risk, road accidents.

Słowa kluczowe

Bezpieczeństwo na drogach, szeregi czasowe, modelowanie, ryzyko, wypadki drogowe.

Summary

The main objective of the paper is to present modelling technique and applied models for analysing and forecasting how road fatalities change in Poland. To that end the theory of structural time-series models is used under the assumption that both road traffic and road safety are dynamic processes where the past has a significant effect on the system's present and future.

1. Introduction

Poland's traffic fatality risk rates are on European average level, reaching about 14 killed per 100,000 population and 4 killed per 10,000 vehicles. Nevertheless there is still need for research and intensified systemic and preventive work in particular after Poland's accession to the European Union where the safest countries have reached a level of 6 killed per 100,000 people and less then 2 killed per 10,000 vehicles (Fig. 1).

* Politechnika Gdańska, Wydział Inżynierii Lądowej i Środowiska, Katedra Inżynierii Drogowej, ul. Narutowicza 11/12, 80-952 Gdańsk, tel.:(058) 347-11-47 e-mail: rkrystek@pg.gda.pl

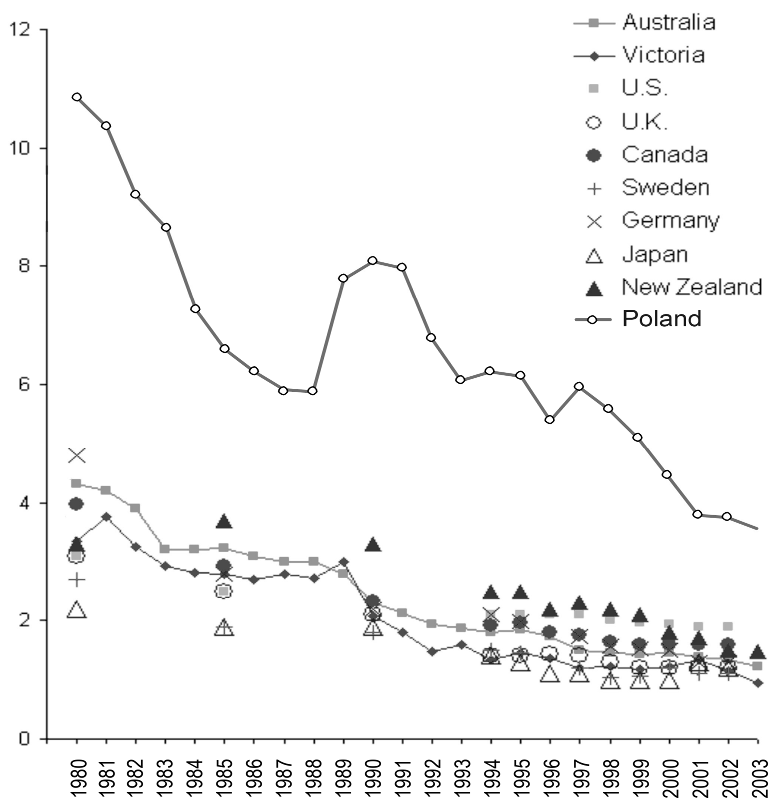

Fig. 1. Road fatalities per 10,000 vehicles in the safest countries vs. Poland
Rys. 1. Śmiertelność z powodu wypadków drogowych na 10.000 pojazdów w krajach najbezpieczniejszych w porównaniu z Polską

This is no doubt a big achievement whose source lies in methodical research and projects such as the program [1]. Work like this is aimed at the attainment of a common goal defined in the "The European Road Safety Action Program" – a reduction in the number of road deaths from the estimated 40,000 to less than 20,000 in 2010 [2]. Increasing road user and policy maker risk awareness is a fundamental task of road safety policies. The experience of successful countries shows that success can only be achieved through integrated action, e.g. by following the 3E rule (Education, Enforcement, Engineering). All policies should have methodical research as the foundation allowing interpretation of the phenomena and how they are produced, analysis of trends as they change, conclusions about the effectiveness of improvement measures, comparing the situation and effects in different countries and estimating future safety trend developments. The knowledge from research helps improving the methods and safety measures. There is no doubt that in order to develop realistic quantitative safety targets, and then to design effective strategies and plans, one has to be able to measure safety developments and to understand the underlying processes and their causes. This, in turn, requires extensive and reliable data, possibly over a long period of time, and modelling techniques that are suitable for describing, interpreting and – ideally – forecasting safety developments.

Before realistic quantitative targets for EU countries could be defined, the actual road safety situation had to be evaluated and its processes understood. The same was done in Poland [3]. The new national program for Poland made it clear how important the monitoring of assumed goals is and how important detailed probability analyses based on the current accident statistics are. This realization encouraged authors to begin exploring the road safety characteristics analyses. The analyses presented in this paper are part of the result of the work.

2. Modelling methodology and goals

The main objective of the authors' research presented in this paper was to build models for analysing and forecasting how the road fatalities may change in Poland. To that end the theory of structural time series models was used under the assumption that both road traffic and road safety are dynamic processes where the past has a significant effect on the system's present and future.

The models are based on monthly road deaths in Poland between 1992 and 2003. This is because the period after 1989, and especially after 1992, is considered reliable due to the social and economic stability which followed Poland's political transition. The assumption is critical for the conclusions drawn and how correct they are about the causes of change in the series and the estimation of the series' future values.

The models are designed to determine the effect the unemployment rate has on the number of fatalities in Poland over the analysed period. By assessing the effect authors will be able estimate how the effect of changes in the economy may have on Poland's road safety levels.

3. Structural time series analysis

Because of the nature of road traffic safety, it lends itself very well to modelling using time series. Observations of a series available in moment t to forecast its future value t+l is the basis for planning in economics, trade and production control [4]. Effective forecasting a trend using historical data requires good quality data and models based on realistic assumptions. In case of traffic, we assume that any changes in its future state and safety will occur in similar socio-economic conditions, i.e. no unexpected events are taken into account [5].

The time series can be defined as a certain (stochastic) process where the subsequent observations change in time randomly. The observation may be e.g. the number of killed in road traffic, injured or the total number of accidents over any discrete time, e.g. over a month, quarter or year. This creates the time series which we then use to build the model.

A structural model of a discrete time series consists of two basic components: the trend and irregular interference. It can then be extended according to the needs of the situation. The trend represents long-term changes in the series by identifying the direction of the series is taking. Consequently, it is that part of the series that changes relatively slowly (evolves) in time [6]. Trend is the general direction a series will take, a tendency. There are two parts of the trend: level – the actual value of the trend and slope which may or may not to be present. From the perspective of the forecasting procedure, trend is that part of a series which when extrapolated gives a better indication of how the series will behave in the future.

Many time series shows recurring changes up and down, i.e. seasonal changes that are modelled using the seasonality component. Seasonality is that part of the series which when extrapolated is repeated so many times, and for the entire series averages to zero. Random disturbance is the so called "white noise", a sequence of uncorrelated random variables of a constant mean and variation. The structure of series presents the equation (1) and Fig. 2. Below is an illustration:

$$\text{Observed series} = \text{Trend} + \text{Seasonal} + \text{Irregular} \tag{1}$$

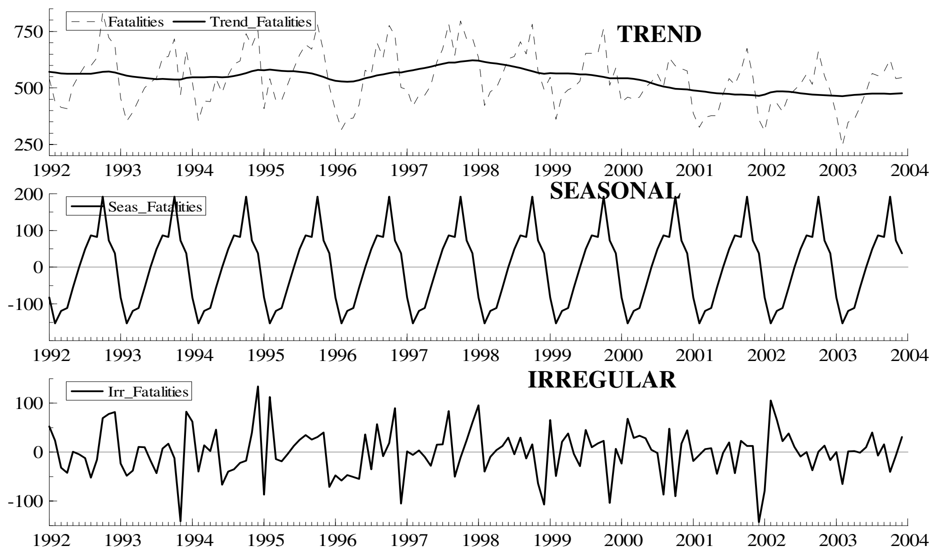

Fig. 2. An example of a structure of the time series
Rys. 2. Przykład struktury szeregu czasowego

The three series components: trend, seasonal and irregular are in actual fact unobservable. All that we are able to observe is the full series of e.g. number of monthly fatalities. The advantage of structural series is that the components can be isolated and analysed separately. That way we know about how the components behave in time which makes the assessment of the impact of the factors on the series easier. A change of the highway code could e.g. substantially distort the random component. The introduction of a speed limit for example or road safety measures will result in its intensification. A regular seasonal diagram implies that certain months have a relative level of safety, for example winter months have a lower number of road deaths than summer months. An important advantage of structural models is that we can add explanatory variables to them making the model a mixed model comprising the features of time series and regression. As a result, the model can be used for both forecasting and explaining change in the series in the past.

4. Example of application of the time series theory

A number of international studies argue that there is a correlation between the number of traffic fatalities and the degree of public activity. The studies use the unemployment rate to support that argument [7]. As unemployment grows, miles traveled fall, a factor known to affect road safety. This relationship seems to be true for Poland, as well.

The model was intended to detect the relation between the monthly number of fatalities in Poland in the period 1992-2003 (Fig. 3) and the unemployment rate within that period (Fig. 4). More generally the influence of the economy changes on road safety in the period in Poland was investigated and the level of national economy condition was measured by the rate of unemployment. After the preliminary analysis authors have proposed following structure of the model:

LFatalities = Level + Fixed seasonal + Explanatory variables + Interventions + Irregular (2)

The number of fatalities (Fatalities) has been transformed into logs (LFatalities) to simplify analysis. The model written in mathematical formulas takes a form:

$$y_t = \mu_t + \gamma_t + \beta_t x_t + \lambda_{t1} w_{t1} + \lambda_{t2} w_{t2} + \varepsilon_t \text{, where } \varepsilon_t = N(0, \sigma_\varepsilon^2) \quad (3)$$

$$\mu_{t+1} = \mu_t + \eta_t \qquad \eta_t = N(0, \sigma_\eta^2) \quad (4)$$

$$\gamma_{t+1} = \gamma_t + \omega_t \qquad \omega_t = N(0, \sigma_\omega^2) \quad (5)$$

$$\beta_{t+1} = \beta_t + \tau_t \qquad \tau_t = N(0, \sigma_\tau^2) \quad (6)$$

$$\lambda_{t+1} = \lambda_t + \varsigma_t \qquad \varsigma_t = N(0, \sigma_\varsigma^2) \quad (7)$$

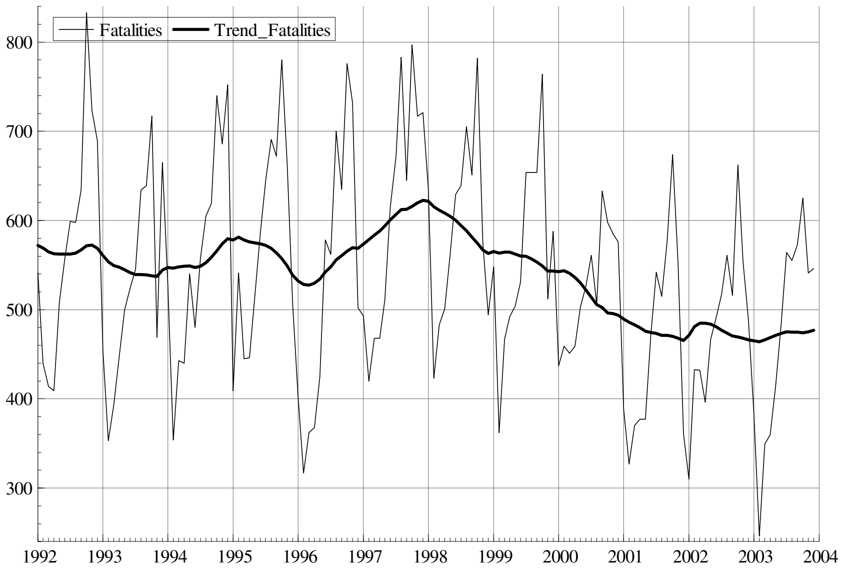

Fig. 3. Monthly number of fatalities in Poland in the period 1992–2003 and its trend
Rys. 3. Miesięczna liczba wypadków śmiertelnych w Polsce w latach 1992–2003 i jej trend

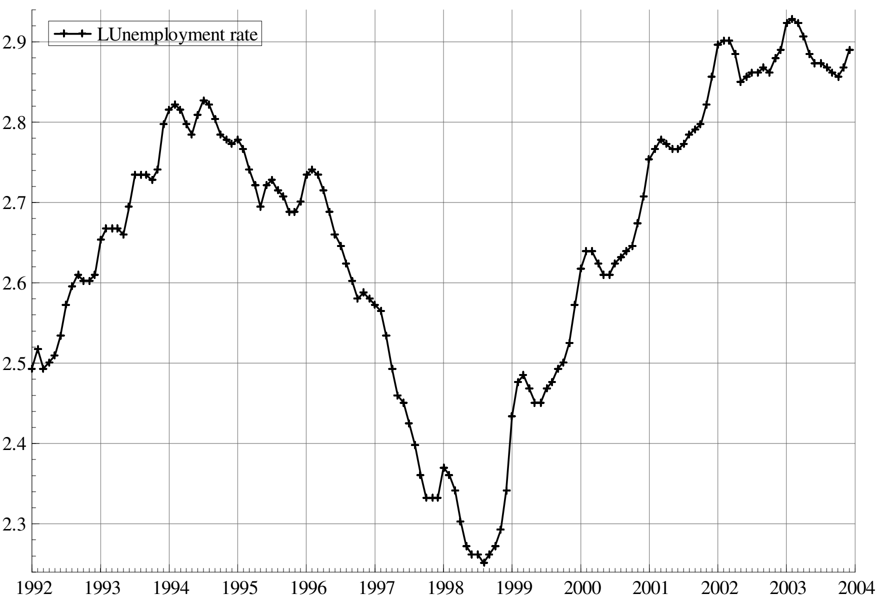

Fig. 4. The unemployment rate in Poland in the period 1991–2003
Rys. 4. Stopa bezrobocia w Polsce w latach 1991–2003

for $t = 1,...,n$. The state errors τ_t, ω_t and ς_t are fixed on zero. μ_t is the unobserved state in time t, ε_t is the residual, η_t is so called state error. γ_t is a seasonal component, β_t is a regression coefficient and x_t is a regressor, λ_t a coefficient which measures the changes in the series level and w_t is an intervention variable [8].

The model is an extended version of the basic structural model presented in equation (1). There are few more variables added: one explanatory variable (an unemployment rate) and two intervention variables (there was a need to include in the model some information about the possibility of influence of the sudden economical break observed in Poland in 2001 on the road safety in our country.

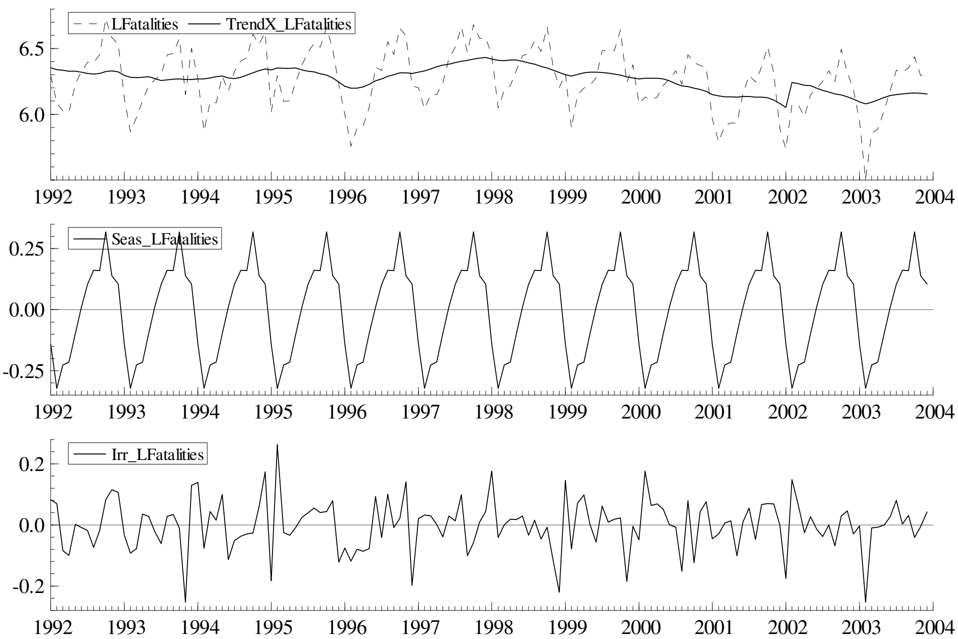

Fig. 5. Structural time series model of the number of road fatalities in Poland in the period 1992–2003
Rys. 5. Model strukturalnych szeregów czasowych liczby śmiertelnych wypadków drogowych w Polsce w latach 1992–2003

To investigate the influence of it on the number of fatalities the intervention variables were added to the model). The moments of the interventions were detected by auxiliary residuals analysis. The greatest values of them were observed in December 2001 (outlier) and in February 2002 (structural break). After the estimation procedure subsequent coefficients of explanatory variables and interventions were achieved:

Table 1. Estimated coefficients of the modeled explanatory and intervention variables
Tabela 1. Estymowane współczynniki modelowanych zmiennych wyjaśniania i interwencji

Variable	Coefficients	Rmse	t-value
LUnemployment	-0.3657	0.1868	-1.9574 [0.0523]
Irr 2001.12	-0.2973	0.1106	-2.6868 [0.0081]
Lvl 2002.2	0.1925	0.0762	2.5269 [0.0126]

The negative value of β *(LUnemployment)* indicates a negative relationship between the number of fatalities and the unemployment rate in Poland. Thus the hypothesis on existence as such was proved. The lower unemployment rate is associated with the higher fatalities notes. The value of $\beta = -0.36$ indicates that 1% increase in the unemployment rate resulted in a 0.3% decrease of the number of fatalities.

Concerning interventions since $1 - e^{-0.2973} = 0.2572$, the number of fatalities decreased in December 2001 by 25%. Because this intervention was modeled by the impulse variable it means that the decrease was only for one moment. The second intervention was modeled as structural break in the level of the series. The value of intervention coefficient is 0.1925, which means that the level of fatalities trend has increased after the intervention by 21%. The explanation of this increase may be the return of the economic prosperity after the short downturn in 2001.

The structural time series model may also be used as a basis for short term forecasting. Our model has generated a prognosis for the years 2004 and 2005. It is pictured in the Fig. 6. In the result of the modeling the prognosed value of the number of fatalities in 2004 is 5606 persons and in 2005–5365 persons.

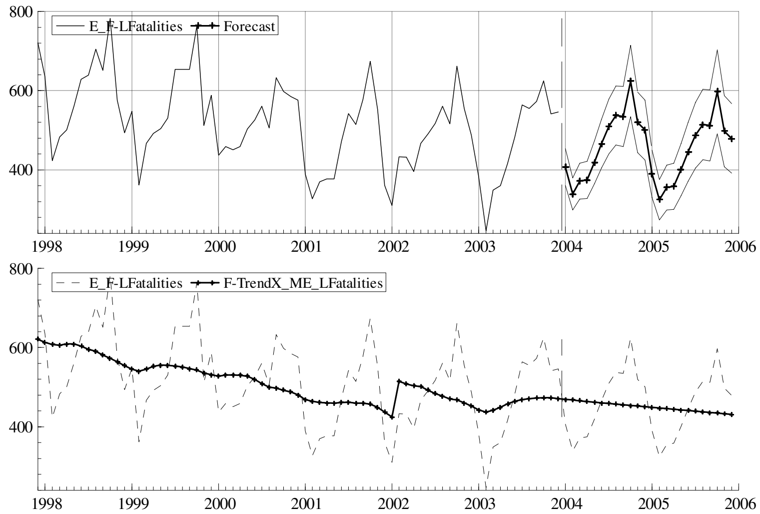

Fig. 6. Forecast of the number of fatalities in Poland in the years 2004 and 2005 made on the basis of the structural model
Rys. 6. Prognoza liczby śmiertelnych wypadków drogowych w Polsce w roku 2004 i 2005 wykonana na podstawie modelu strukturalnego

5. Conclusions

In 1999 we were preparing a forecast of the number of fatalities in Poland up to 2010 for the National Road Safety Program GAMBIT 2000. It was based on the assumption of a monotone increase of public activity in our country in forthcoming years [9]. What we found, however, was a sudden decrease of the activity caused by the short term economic crisis that reached Poland in 2001. In spring 2001 the government informed the public about Poland's budget deficit. The estimates of the deficit kept changing and ranged between 40 and 80 billion PLN. In a consequence the falling public activity reduced the demand for transport leading to a reduction in miles traveled. This in turn resulted in fewer accident fatalities.

The modeling technique presented in this paper has successfully proved the influence. Moreover, a prognosis developed on the basis of this technique has showed almost exact results with the real data from 2004. So is it safe to say that the subsequent years with their projected economic development and a drop in unemployment may see an increase in traffic risk? It probably is, unless additional activities forward road safety improvement in Poland will undertaken.

Praca wpłynęła do Redakcji 24.03.2006 r.

References

[1] COST 329. Models for traffic and safety development and interventions. European Cooperation in the field of Scientific and Technical Research. European Commission Directorate General for Transport. Brussels 1996.

[2] ETCS. The III European Road Safety Action Programme. Brussels 2003.

[3] Gerondeau, C.: Road Safety in Poland. Study of road safety in Central and Eastern Europe Countries. Report of Expert's Group. World Bank European Community, Paris 1992.

[4] Box, G.E.P., Jenkins, G.M.: Analiza szeregów czasowych. Prognozowanie i sterowanie. Państwowe Wydawnictwo Naukowe. Warszawa 1983.

[5] Oppe, S.: Traffic safety developments in Poland. A research note. SWOV, Leidschendam 2000.

[6] Harvey, A.C.: Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge, Cambridge University Press, 1989.

[7] Koornstra, M.J.: Prognostic analysis of road safety in Poland. Raport nr R-96-41. SWOV Institute for Road Safety Research. Leidschendam. The Netherlands 1996.

[8] STAMP. Structural Time Series Analyzer, Modeler and Predictor. Timberlake Consultants Ltd., London, UK 2000.

[9] Krystek, R., Żukowska, J.: Zmiana postaw społeczeństwa i władz wobec bezpieczeństwa ruchu drogowego – 10 lat polskich doświadczeń. IV Międzynarodowa Konferencja „Globalne Bezpieczeństwo". Bled. Słowenia 2000.

Szeregi czasowe – narzędzie analizy bezpieczeństwa ruchu drogowego

S t r e s z c z e n i e

Głównym celem artykułu jest przedstawienie sposobu modelowania i modeli stosowanych w analizach i prognozowaniu odnośnie do zmian śmiertelności w wypadkach drogowych w Polsce. W tym celu zastosowano teorię modeli strukturalnych szeregów czasowych przy założeniu, że zarówno ruch drogowy, jak i bezpieczeństwo na drogach są procesami dynamicznymi, w których przeszłość ma znaczący wpływ na teraźniejszość i przyszłość systemu.