Stochastic Cusp Catastrophe Models with Traffic and Weather Data for Crash Severity Analysis on Urban Arterials

Transportation Research Board
96th Annual Meeting
January 8–12, 2017 ▪ Washington, D.C.

Athanasios Theofilatos, George Yannis, Eleni Vlahogianni, John Golias

Department of Transportation Planning and Engineering, School of Civil Engineering,
National Technical University of Athens

Introduction

- The effective treatment of crashes and the proactive transportation safety is a major concern to societies.
- Much research that utilized real-time collected traffic and weather data in freeways has been carried out recently.
- Crash injury severity is underrepresented.
- Alternative modeling techniques should also be considered.

Objectives

- The main objective is to **propose cusp catastrophe models for modeling crash injury severity.**
- Cusp catastrophe models are applied and compared with traditional statistical models.
- The potential existence of non-linearity in the system is examined.
- Real-time traffic and weather data from two major urban arterials in Athens, Greece are considered.
- The findings of the study are expected to extend previous research and add to current knowledge.

Data preparation

Within this research, a large collection and processing of data took place:

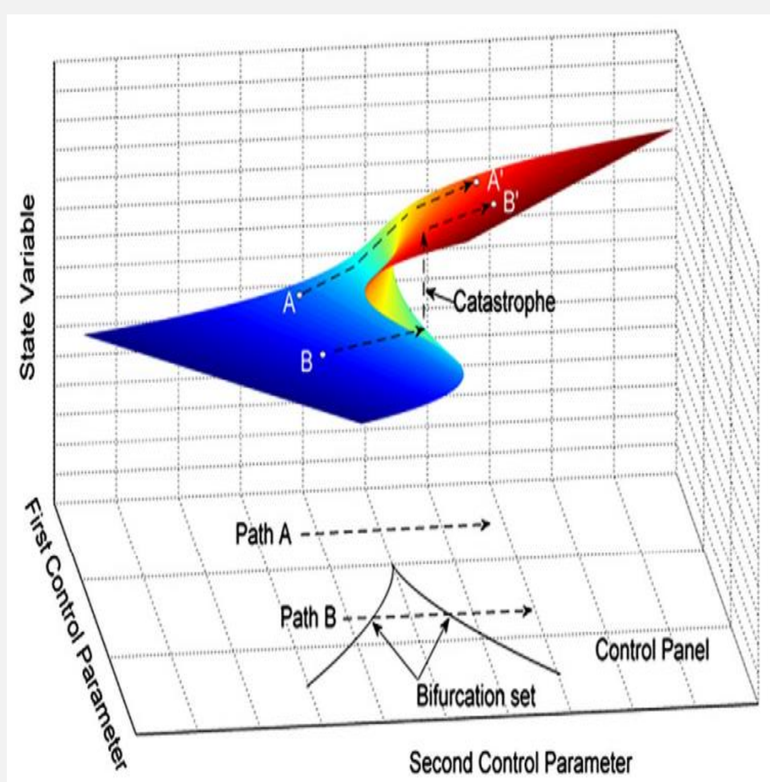
- ✓ The available dataset refers to the period 2006-2011 and come from two high demand urban arterials in the center of Athens (Greece).
- ✓ These two arterials have similar geometrical and traffic characteristics.
- ✓ Crash data were collected from the Greek accident database, SANTRA, which is provided by the National Technical University of Athens.
- ✓ Traffic data were extracted from the Traffic Management Centre (TMC) of Athens, which has been in operation since 2004 and covers several major roads in Athens. The TMC data included traffic flow, traffic occupancy and time speed every 1 minute.
- ✓ Traffic data from the adjacent upstream loop detector were considered.
- ✓ Data were further aggregated to 1-hour traffic information to obtain averages, standard deviations and so on, prior to a crash occurrence.
- ✓ Weather data were collected from the Hydrological Observatory of Athens, which is an online open-access database, covering more than 10 meteorological stations located in the greater Athens area.
- ✓ Data include rainfall, temperature, relative humidity, solar radiation, wind direction and wind speed.
- ✓ Each crash case was assigned to the closest meteorological station and then the relevant weather data had to be extracted.
- ✓ The 10-min raw weather data were aggregated over hour in order to obtain maxima, averages and standard deviations, in the time-slice of 1-hour, 2-hours, 6-hours and 12-hours prior to the time of the crash occurrence.
- ✓ Crash severity is defined in two ways:

$$\text{Severity\_1} = \frac{\textit{Number of severely injured and killed}}{\textit{Total number of persons involved}}$$

$$\text{Severity\_2} = \frac{\textit{Total number of persons invovled}}{\textit{Total number of vehicles involved}}$$

Method of Analysis

- The core analysis of this study is the **catastrophe theory.**
- **Cusp catastrophe** models are the most widely applied.
- Catastrophe theory examines the qualitative changes in the behavior of systems when the control factors that influence their behavioral state face smooth and gradual changes
- It assumes the existence of a dynamic system
- It explains the sudden transition between the system states.
- Two critical control parameters exist, namely α and β:



$$\alpha = \alpha_0 + a_1 X_1 + a_2 X_2 + \cdots + \alpha_n X_n$$

$$\beta = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_n X_n$$

where $X_1, X_2, \ldots, X_n$ is a set of measured independent variables

- The dependent variable can be a combination of other dependent variables $Y_1, Y_2, \ldots, Y_n$

$$y = w_0 + w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n.$$

- Due to the nature of the dependent variable (censored) the traditional linear regression model is not appropriate
- **The censored regression** is therefore applied.
- Comparison of results with the cusp models will reveal potential non-linearity in the system.

Findings (1/2)

- **Descriptive statistics**

Variable	Description	Unit	Mean	Std. deviation
Severity_1	Number of severely injured and killed divided by Total number of persons involved	unitless	0.086	0.259
Severity_2	Total number of persons involved divided by Total number of vehicles involved	unitless	0.885	0.500
Acc.Type	Acc.Type1 (Off road/Fixed object/Other)	unitless	155*	
	Acc.Type2 (Head-on)	unitless	36*	
	Acc.Type3 (Rear-end)	unitless	73*	
	Acc.Type4 (Side)	unitless	53*	
	Acc.Type5 (Sideswipe)	unitless	36*	
Q_avg_1h_up	1h average flow per lane upstream	veh/hour/lane	810.450	301.719
Q_stdev_1h_up	1h st.deviation of flow per lane upstream	veh/hour/lane	264.330	339.374
Q_median_1h_up	1h median of flow per lane upstream	veh/hour/lane	628.600	437.337
Q_cv_1h_up	1h coefficient of variation of flow per lane upstream	unitless	0.109	0.085
V_avg_1h_up	1h average speed upstream	km/h	47.340	18.959
V_stdev_1h_up	1h st.deviation of speed upstream	km/h	5.333	5.591
V_cv_1h_up	1h coefficient of variation of speed upstream	unitless	0.154	0.175
Occ_avg_1h_up	1h average occupancy upstream	percentage %	15.730	11.143
Occ_stdev_1h_up	1h st.deviation of occupancy upstream	percentage %	4.097	4.917
Occ_cv_1h_up	1h coefficient of variation of occupancy upstream	unitless	0.248	0.216
T_1h_max	1h maximum temperature	°C	19.240	7.710
T_1h_avg	1h average temperature	°C	18.700	7.714
T_1h_stdev	1h st.deviation of temperature	°C	0.397	0.335
Rain_1h_sum	1h sum of rainfall	mm	0.030	0.265
Rain_1h_st.dev	1h st.deviation of rainfall	mm	0.004	0.031
Rain_2h_sum	2h sum of rainfall	mm	0.068	0.618
Rain_2h_st.dev	2h sum of rainfall	mm	0.010	0.094
Rain_6h_sum	6h sum of rainfall	mm	0.152	0.921
Rain_6h_st.dev	6h st.deviation of rainfall	mm	0.013	0.083
Rain_12h_sum	12h sum of rainfall	mm	0.252	1.142
Rain_12h_st.dev	12h st.deviation of rainfall	mm	0.014	0.068
W.Sp_1h_max	1h maximum wind speed	m/sec	2.759	1.836
W.Sp_1h_avg	1h average wind speed	m/sec	2.204	1.688
W.Sp_1h_stdev	1h st.deviation of wind speed	m/sec	0.387	0.223
Sol_1h_max	1h maximum solar radiation	W/m^2	377.410	362.890
Sol_1h_avg	1h average solar radiation	W/m^2	307.550	321.884

* Distribution of crash types

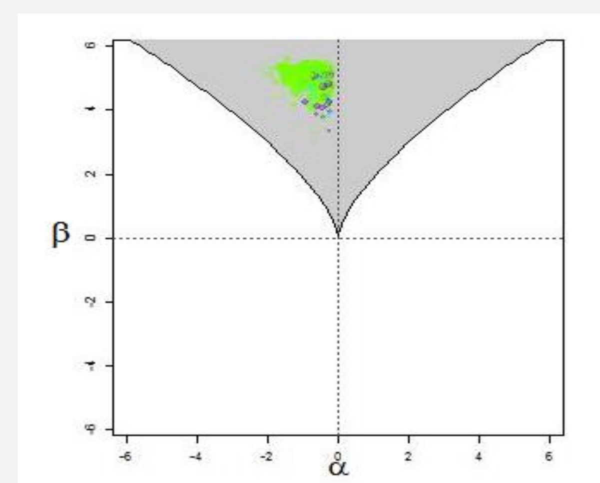
Severity_1 (Number Of Severely Injured and Killed By Total Number Of Persons Involved)

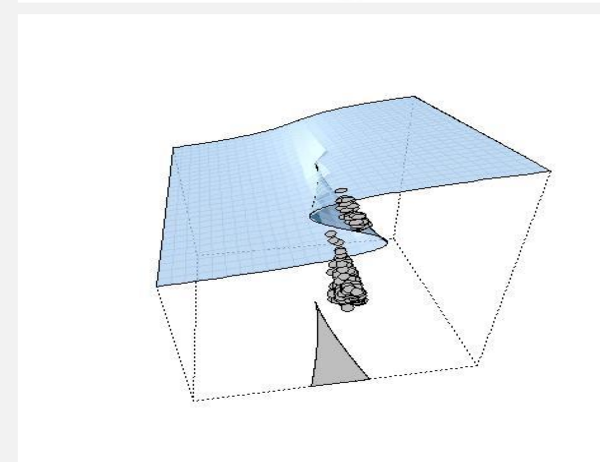
- **Cusp catastrophe**

Assymetry factor a	Variable	Coefficient	Std. error	p-value
a_0	Constant term	-0.200	0.093	0.032**
a_1	V_cv_1h_up	-0.720	0.372	0.052*
a_2	Acc.type1	-0.417	0.228	0.067*
a_3	Acc.type2	-0.280	0.144	0.052*
a_4	Acc.type3	-1.021	0.420	0.015**
a_5	Acc.type4	-0.493	0.286	0.084*
a_6	W.Sp_1h_max	-0.068	0.032	0.031**
Bifurcation factor β				
b0	Constant term	-	-	-
b1	log(Q_avg_1h_up)	0.755	0.034	0.000**
Dependent variable y				
w0	Constant term	-2.342	0.047	0.000**
w1	Severity_1	4.732	0.101	0.000**
McFadden R^2	0.789			
pseudo-R^2	0.845			
Logistic model R^2	0.095			

**=95% significance level

*=90% significance level





- **Censored regression**

Variable	Coefficient	Std.error	p-value
Constant term	7.078	2.943	0.016**
V_cv_1h_up	-1.975	1.627	0.225
Acc.Type1	-1.731	1.049	0.098*
Acc.Type2	-1.225	0.712	0.085*
Acc.Type3	-12.729	563.132	0.982
Acc.Type4	-1.763	1.075	0.101
W.Sp_1h_max	-0.302	0.169	0.074*
log(Q_avg_1h_up)	-1.312	0.498	0.008**
Madalla R^2	0.119		

**=95% significance level

*=90% significance level

Findings (2/2)

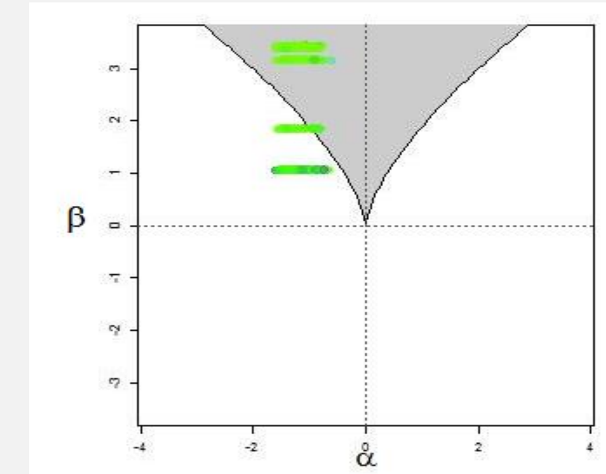
Severity_2 (Total Number Of Persons Involved By Total Number Of Vehicles Involved)

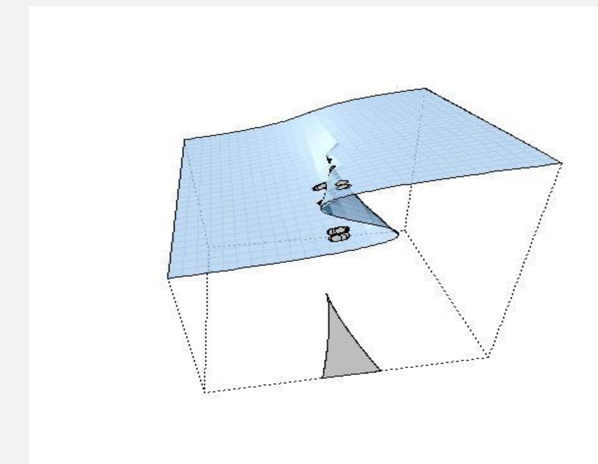
- **Cusp catastrophe**

Assymetry factor a	Variable	Coefficient	Std. error	p-value
a_0	Constant term	-0.251	0.226	0.267
a_1	V_cv_1h_up	0.172	0.037	0.004**
a_2	log(Q_avg_1h_up)	-0.101	0.047	0.031**
a_3	Sol_1h_max	-0.001	0.000	0.006**
Bifurcation factor β				
b_0	Constant term	1.048	0.082	0.000**
b_1	Acc.Type1	0.783	0.280	0.005**
b_2	Acc.Type2	2.115	0.077	0.000**
b_3	Acc.Type3	2.350	0.242	0.000**
b_4	Acc.Type4	2.392	0.276	0.000**
Dependent variable y				
w0	Constant term	-2.847	0.041	0.000**
w1	Severity_2	1.689	0.056	0.000**
McFadden R^2	0.251			
pseudo-R^2	0.303			
Logistic model R^2	0.313			

**=95% significance level

*=90% significance level





- **Censored regression**

Variable	Coefficient	Std.error	p-value
Constant term	1.484	0.291	0.000**
V_cv_1h_up	0.057	0.132	0.663
log(Q_avg_1h_up)	-0.041	0.000	0.367
Sol_1h_max	0.000	0.000	0.009**
Acc.Type1	-0.259	0.078	0.000**
Acc.Type2	-0.500	0.060	0.000**
Acc.Type3	-0.571	0.068	0.000**
Acc.Type4	-0.543	0.078	0.000**
Madalla R^2	0.300		

**=95% significance level

*=90% significance level

Conclusions - Discussion

- Promising fit of the cusp models.
- Possible evidence of presence of cusp and imply strong nonlinear relationships between crash injury severity and independent variables for Severity_1.
- The cusp catastrophe model is superior for Severity_1.
- It is found that small changes to crash type, maximum wind speed, traffic flow and variation of speed leads to sudden changes to crash severity.
- All crash cases lie inside the instability area.
- For Severity_2, the cusp model has a reasonable fit but is not superior to the traditional linear models.
- Although some traffic and weather variables have strong non-linear relationships with crash severity, the cusp model does not outperform the traditional linear model.
- When severity is expressed as the number of severely injured and killed divided by the total number of persons involved in a crash (Severity_1), the safety system is highly non-linear and sudden transitions from unsafe (high severity) to safe regions (low severity), due to small changes of traffic and weather parameters.
- When severity is expressed as the total number of persons involved by total number of vehicles involved in the crash (Severity_2), it seems that the linearity of the system is preserved.
- The findings indicate that the dynamic change in urban road safety levels expressed by crash severity is likely to be nonlinear in nature.

Contact information

- Email: atheofil@central.ntua.gr
- Paper ID: 17-02516