Traffic and safety data analysis: from correlation to causation and policy support

George Yannis, Professor
www.nrso.ntua.gr/geyannis
Structure of the Lecture

1. From correlation to causation and policy support
2. Evidence-based road safety policies
3. Road safety analysis methods
4. Road safety measures assessment
5. Road user behaviour

The Traffic Engineering Laboratory of the NTUA
The National Technical University of Athens

• The National Technical University (NTUA) is the oldest and most prestigious educational institution of Greece in the field of technology, and has contributed unceasingly to the country's scientific, technical and economic development since its foundation in 1836.

• **Nine Schools**, eight being for the engineering sciences, including architecture, and one for the general sciences.

• The personnel of the nine Faculties include more than 550 people as academic staff, 4,000 scientific assistants and 800 administrative and technical staff. The total number of undergraduate students is about 8,500 and the graduate students 1,500.
All transport modes
road, rail, sea, air, intermodal

All transport types
- passenger and freight transport
- urban and interurban
- national and international
- terminals

All stages of transport project lifecycle
- Planning, Design study, Tender
- Construction, Delivery
- Operation, Maintenance
Traffic Engineering Laboratory

- 3 Faculty - 25 PhD students and researchers
- 5 Courses
- More than 200 Diploma Theses
- 12 Ph.D. Theses
- More than 80 Research Projects
- More than 650 Publications
- Equipment - Software - Systems
Scientific Fields

- **Traffic Flow Theory, Traffic Control and Traffic Management**
  - Traffic flow analysis and modeling in urban and highway road networks
  - Traffic capacity and level of service estimation in highways, arterials and intersections
  - Analysis and modeling of urban signalized networks
  - Design and implementation of traffic management systems
  - Analysis of pedestrian and cyclist traffic
  - Traffic counts and surveys
  - Intelligent transportation systems for traffic management and road safety
  - Statistical analysis, mathematical modeling and computational intelligence for traffic and safety

- **Traffic Safety**
  - Traffic safety analysis
  - Investigation of hazardous locations
  - Statistical analysis, mathematical modeling in traffic safety

- **Parking**
  - Design and operation of parking systems
  - Systems operation of parking areas

- **Logistics**
  - Analysis and estimation of signalization, delays, queuing and tollway systems
  - Design and implementation of traffic management systems
Courses

5 courses School of Civil Engineering Transportation Engineering cycle

- **Traffic Flow theory**  
  *Obligatory - 7th semester*
- **Urban Traffic Networks**  
  *Obligatory - 8th semester*
- **Traffic Management and Traffic Safety**  
  *Obligatory - 9th semester*
- **Special Topics of Traffic Engineering**  
  *Compulsory elective - 9th semester*
- **Evaluation & Impact of Transportation Infrastructure Projects**  
  *Compulsory elective - 9th semester*
Research Projects

- 76 research projects in total
  14 in the last 3 years
- 47 projects for the European Commission
  9 in the last 3 years
- 37 projects for Greek Authorities
  4 in the last 3 years
- 47 projects with international competitive procedures
  12 in the last 3 years
- 7.3 average partners per project
  5.5 in the last 3 years
- 12.9 million € total budget
  3.88 million € in the last 3 years
Scientific Publications

• 202 Publications in Scientific Journals
  77 in the last 3 years
• 388 Publications in Conference Proceedings
  112 in the last 3 years
• 38 Book Chapters
  14 in the last 3 years
• 7 Books
• 2.696 Citations in Google Scholar
  1441 in the last 3 years
• 1.305 Citations in Scopus
  915 in the last 3 years
Equipment - Driving Simulator

Foerst Driving Simulator FPF 1/4 cab

- Motion Base
  - 2 degrees of freedom
- Programming Software Tool
  - Programming driving scenarios in different conditions
- Investigation of driver’s behavior in extreme traffic conditions and conditions of difficult geometry.

- Driver Behaviour Data
  - Kinematic characteristics
  - speed, acceleration, headways, time-headways
  - Time To Collision
  - Track of the vehicle
  - Reaction Time
Equipment - Traffic Counts

- Recording systems based on video traffic
- Manual traffic counters
- Counters of turning templates traffic
- Automatic traffic counters sectional road (ADR)
- Radar speed detection (Laser)
- Device for measuring and analyzing traffic to junction
- System for recording and analyzing real-time traffic (Autoscope)
- Device road traffic noise levels
- GPS devices log position information
Software

• Traffic Flow Analysis
  *HCM, Synchro, TSIS*

• Macroscopic and Microscopic Traffic Flow Simulation
  *AimSun, Saturn, Contram, Simtraffic, Corsim*

• Statistical Analysis
  *SPSS, R, MLWIN, MATLAB, LIMDEP*

• Traffic Flow Forecast, Traffic Flow distribution network
  *CUBE*
Information Systems

• Data base with disaggregate road accident data - SANTRA
  *Disaggregate data for 525,000 road accidents and 1,205,000 casualties*

• International Traffic Accident Databases
  EC-CARE, OECD-IRTAD

• Traffic Databases, field surveys and driver behaviour experiments

• Road safety research results dissemination portal
  www.nrso.ntua.gr

• Electronic Traffic Safety Library
  More than 3,500 Technical Reports
Priority Scientific Fields

• Dynamic analysis and management of real-time traffic
• Intelligent systems in traffic management and road safety
• Driver behaviour experiments analysis: actual driving conditions in a driving simulator simulated traffic
• Analysis of data traffic and road accidents big data social networks data algorithms use data from multiple sources
EVIDENCE-BASED ROAD SAFETY POLICIES
Road safety - Open questions

• Which are the current and future challenges of traffic and safety data analysis?
• How critical are data and evidence based decision making?
• What is the role of high quality traffic and safety data, as well as of appropriate analysis methodologies?
• How critical is the efficiency assessment of measures?
• How to link statistical analyses (correlation) and the interpretation of their results (causation) to policy support?
• Is the analysis of road user behavior the key for the establishment of the links between accident causes and impacts?
The **high complexity** of the road environment and road user behavior makes road safety choices a very difficult task, attempting to balance conflicting social needs and economical restraints, especially during the economic crisis.

- Traffic Efficiency (Speed) **Versus** Traffic Safety
- Vehicles **Versus** Vulnerable Road Users
- Expensive but safe **Versus** Cheap but unsafe (vehicle, infrastructure, management)
- Priorities in policies, measures, research, etc.
Different road safety progress in different countries

Road Fatalities change 2001-2010 (source: CARE)

Urban Areas
Inside   Outside
North-Western countries -48.4% -50.0%
Southern countries -47.7% -42.4%
Eastern countries -22.6% -24.3%
Different road safety problems in different countries

Road Fatalities 2010 (source: CARE)

<table>
<thead>
<tr>
<th></th>
<th>Urban</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Power Two Wheelers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-Western countries</td>
<td>848</td>
<td>3.776</td>
<td>22%</td>
</tr>
<tr>
<td>Southern countries</td>
<td>1,091</td>
<td>3.399</td>
<td>32%</td>
</tr>
<tr>
<td>Eastern countries</td>
<td>434</td>
<td>4.183</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Cyclists</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-Western countries</td>
<td>472</td>
<td>3.776</td>
<td>13%</td>
</tr>
<tr>
<td>Southern countries</td>
<td>203</td>
<td>3.399</td>
<td>6%</td>
</tr>
<tr>
<td>Eastern countries</td>
<td>400</td>
<td>4.183</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Pedestrians</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North-Western countries</td>
<td>1,249</td>
<td>3.776</td>
<td>33%</td>
</tr>
<tr>
<td>Southern countries</td>
<td>1,066</td>
<td>3.399</td>
<td>31%</td>
</tr>
<tr>
<td>Eastern countries</td>
<td>1,888</td>
<td>4.183</td>
<td>45%</td>
</tr>
</tbody>
</table>
Road safety is ideal for spending money for nothing

If you cannot measure it, you cannot improve it
(Lord Kelvin)

Road Safety is a typical field with high risk of important investments not bringing results
The need for evidence based decision making

The policy making cycle
• Vision and strategy;
• Problem identification;
• Target Setting and priority setting;
• Development of measures;
• Establishing and implementing the programme;
• Monitoring of and evaluation of outcomes.

The use of high quality road safety data is involved in each stage

Necessity to:
• Consolidate and organise existing data
• Make data and information available
• Provide a complete tool-kit (analyses, methodologies, benchmarking tools)
• Support road safety decision making at all stages
The need for evidence based decision making

Lack of data
(accidents, injuries, exposure, performance indicators,...)

Data not comparable

Data incompatible

Insufficient data details

Low reliability of data

Correlations but not Causations

Lack of standard methodologies

Analyses not solution oriented
The ‘pyramid’ of road safety management systems

• The road safety management ‘footprint’ of a country at specific point in time

• **Level 1**: Structural and cultural characteristics (i.e. policy input).

• **Level 2**: Safety measures and programs (i.e. policy output).

• **Level 3**: The operational level of road safety in the country (road safety performance indicators)

• **Level 4**: Final outcomes (i.e. road casualties).

• **Level 5**: The total social costs of road crashes.
Road safety data to support evidence-based policies

- Fatalities and their evolution
- Exposure
- Safety Performance Indicators
- Causation (in-depth accident investigations)
- Health indicators
- Economic indicators
- Driver behavior, attitudes etc.
- Road safety rules and regulations
- Road safety measures assessment

Do we have the data we need?

Do we need the data we have?
The necessary exposure data and SPIs

Exposure indicators
- Vehicle- and person-kilometres of travel
- Time spent in traffic
- Number of trips
- Vehicle fleet
- Population

Safety Performance Indicators
- Road user behavior (e.g. speeding, alcohol)
- Vehicles (e.g. crashworthiness, fleet age etc.)
- Infrastructure (e.g. meeting design standards)

The most useful data are the least available
A comprehensive and integrated road safety information system with aggregate data and information consolidating, organising and making available existing data and information, necessary for the support of road safety decision making in Europe

Key road safety analyses and syntheses
• Basic Fact Sheets
• Annual Statistical Report
• Country Overviews
• Road Safety Management Country Profiles
• Forecast Fact Sheets
• Safety Issues Syntheses (knowledge webtexts)
Statistical and Knowledge Fact Sheets

- Seasonality
- Motorcycle and Mopeds
- Main Figures
- Motorways
- Pedestrians
- Car occupants
- Elderly
- Causation
- Youngsters (aged 15-17)
- Single Vehicle Accidents
- Heavy Goods Vehicles & Buses
- Junctions
- Children
- Urban Areas
- Cyclists
- Gender
- Roads outside urban areas
- Young People (aged 16-24)
Road safety country overviews and forecasts

All layers of the Road Safety Pyramid
- Structure & Culture
- Programs & measures
- Road Safety Performance Indicators
- Road Safety Outcomes
- Social Cost

Syntheses
- Safety position
- Scope of problem
- Recent progress
- Remarkable road safety policy issues

Forecasts
- based on time series models
- fatalities and exposure
- mobility scenaria
A European Road Safety Knowledge System

SafetyNet and DaCoTA research projects, NTUA 2004-2012

Road Safety Knowledge System

ASR
ROAD SAFETY DATA
BFS
COUNTRY OVERVIEWS
FORECAST FACT SHEETS
SAFETY SYNTHESSES
ROAD SAFETY MANAGEMENT PROFILE
ROAD SAFETY KNOWLEDGE
COMPOSITE INDEX
MASTER TABLES
ROAD SAFETY ANALYSIS METHODS
The use of rigorous analysis techniques

Analysis techniques

• Descriptive and qualitative analysis
• Linear regression
• Generalised Linear Models / Non-Linear Models
• Multilevel analysis
• Time series analysis
• Cost-Benefit Analysis

Conventional analysis techniques may not be appropriate for the complex and hierarchical nature of road safety outcomes.
Hierarchical structures in road safety data are receiving increasing attention in the literature:
- Geographic / spatial hierarchies
- Accident components hierarchy

ML models are regressions (linear or GLM) in which the parameters are assigned a probability model.

This “higher-level” (probability) model has parameters of its own (mean, variance).

They aim to capture dependencies among observations due to their hierarchical structure.
Basic Multilevel model formulation:

Observations \( i \) are nested within units \( j \)

\[
y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \epsilon_{ij}
\]

\[
\beta_{0j} = \beta_{0} + u_{0j} \quad u_{0j} \sim N [0, \sigma_{0j}^{2}]
\]

\[
\beta_{1j} = \beta_{1} + u_{1j} \quad u_{1j} \sim N [0, \sigma_{1j}^{2}]
\]

ML model formulations:
(i) allow improving the fit of the model to the data,
(ii) allow identifying and explaining random variation at specific levels of the hierarchy considered,
(iii) yield different (more correct) conclusions than single-level model formulations with respect to the significance of the parameter estimates.
Example: effects of enforcement on road accidents

Yannis et al. (2007), AAP Journal

- Geographical Multilevel model

- Data for the period 1998-2002 for the 52 regions of Greece
  - Road fatalities
  - Number of speed / alcohol controls
  - Number of speed / alcohol violations
  - Demographics and vehicle fleet

- The Multilevel model shows that:
  - The effect of enforcement becomes more significant when considering the geographical hierarchy of the observations
  - The highest effect of enforcement was observed in those counties that had a low level of motorization and high violations rate.
Example: risk and protection factors in fatal accidents

Yannis et al. 2013, Traffic Injury Prevention Journal

- Accident hierarchy ML model

- Disaggregate data for 1000 fatal accidents from in-depth investigation methods for 7 European countries

- The fatality risk largely depends on the vulnerability of the road user (‘baseline risk’) compared to the vulnerability of the collision opponent.

- Several road, traffic and individual factors interact in terms of injury severity

- The ‘vehicle’ random effect is necessary to better capture these effects
Time series analysis

Commeur et al., 2013, AAP Journal

• Road safety observations are often **serially correlated** (time series) – autocorrelation

• Although linear and GLM models may, under certain conditions, be used for time series analysis, only dedicated techniques may fully account for this type of dependencies:
  
  • ARIMA models
  
  • State-space models

• Time series analysis components: Trend, Seasonality
Example: Forecasting fatalities in European countries

Dupont et al., 2014, AAP journal

- Data for the period 1960-2010 were used to forecast fatalities 2011-2020 in 30 European countries by means of state-space models

- Explicitly examined the presence or not of correlation between fatality series and exposure series.

- SUTSE (Seemingly unrelated time series model)
- LLT (Local Linear Trend model)
- LRT (Latent risk model)

- If fatalities and exposure are correlated, then the forecasts can be made on the basis of mobility scenarios (e.g. 20% increase in mobility by 2020)
Example: Forecasting fatalities in Greece

Yannis & Antoniou, 2013, AAP Journal

- Data for the period 1960-2010
- Exposure (vehicle fleet) and fatalities found to be related
- LRT (Latent risk model)
- Intervention variables introduced to capture shocks / breaks in the series (e.g. 1986 financial crisis, 1991 old-car exchange scheme, 1996 new road fatality definition)
- An intervention variable to model the current economic recession, assuming it ends on 2013.
- Data for the period 2009-2011 used to validate the model
- Forecasts for 2012-2020: actual data for 2012, 2013 lie within the 95% confidence intervals of the forecasts
Example: Effects of GDP changes on road fatalities

Yannis et al., 2013, Safety Science Journal

- The economic recession appears to have affected road fatalities in all European countries, due to decrease in mobility and other factors.

- Modelling the annual change of fatality rate against annual change of GDP (short-term effects)
Example: Effects of GDP changes on road fatalities

- Statistically significant relationship between annual GDP increase and fatality rate increase was established, as well as a statistically significant relationship between annual GDP decrease and fatality rate decrease.

- Particularly in Northern / Western European countries, annual GDP decrease is associated with fatality rate decrease on the same year, as well as on one year later.

- Once the socioeconomic conditions improve, fatalities may temporarily increase, “correcting” for the effect of external factors (GDP change).
ASSESSMENT OF ROAD SAFETY MEASURES
The need for efficiency assessment

- Need to make sure that the limited funds available are used effectively.
- A synthesis of diverse evaluation results allows for more universal understanding and application of safety effectiveness measures.
- The narrower the efficiency assessment results distribution, the larger is the probability that policy decisions are correct.
- Efficiency assessment allows more rapid adoption and dissemination of new safety measures.
- Efficiency assessments are the basis for evidence based safety policies.
Efficiency assessment tools

- **US, Australia**: The NHTSA Highway Safety Manual, The CMF Clearinghouse, the Austroads Road Safety Engineering Toolkit, etc.

- **Europe**: The Handbook of Safety Measures, the ROSEBUD Handbook, the SUPREME Handbook, the CEDR report on cost-effective infrastructure investments, etc.

**Challenges for transferability:**
- Lack of a uniform understanding of the value, importance and usage of CMFs in road safety decision making.
- Need to assess the particularities of setting, context, and implementation features of a specific measure.
Examples of assessed measures in Greece

- Before-and-after analysis on the basis of the ‘odds-ratio’ methods
- Upgrade of main interurban roads to motorways
- Traffic calming schemes in urban areas
- Enforcement of speeding and alcohol while driving
- Removal of advertising signs and billboards along urban arterials.

In Greece, as in several other countries, measures evaluation is the weakest component of road safety management.
CEDR report on cost-effective infrastructure investments

- **Analysis and ranking** of more than 55 specific investments falling within 4 infrastructure areas
  - motorways
  - interurban roads
  - urban roads
  - junctions

- **Identification and in-depth analysis** of the most promising investments
  - Speed management
  - Roadside treatments
  - Road lighting
  - Traffic calming
  - Junctions treatment
  - Traffic control

- **Cost-Benefit and Cost-Effectiveness**
Towards a European inventory of road safety measures

The PRACT project is aimed at addressing the lack of framework for the efficiency assessment of road safety measures in Europe, by developing a practical guideline and a user friendly tool (open access repository) that will allow the different road administrations to:

• adapt the basic Accident Prediction Model to local conditions based on historical data;
• identify the safety effects (Crash Modification Factors) that could be relevant for the specific application;
• verify if the selected safety effects are transferable to the specific condition;
• apply the calibrated model to the specific location to be analysed.
Understanding road user behaviour

• Human factors are the basic causes of road accident in 65-95% of road accidents.

• Human factors include a large number of specific factors that may be considered as accident causes, including:

  - driver injudicious action (speeding, traffic violations etc.),
  - driver error or reaction (loss of control, failure to keep safe distances, sudden braking etc.),
  - behaviour or inexperience (aggressive driving, nervousness, uncertainty etc.),
  - driver distraction or impairment (alcohol, fatigue, mobile phone use etc.).
Monitoring road safety attitudes and perceptions


- A common questionnaire on national random samples of 1000 participants

- Mobility and travel motivations
- Risk perception
- Attitudes towards measures
- Driver self-assessment
- Self-reported behavior
  - Speeding
  - Alcohol
  - etc.

Drivers, and on 2012 also PTW and pedestrians
Understanding pedestrian behaviour

Papadimitriou et al., 2009, Transportation Research F

- Revisiting the hierarchical model of pedestrian behavior
- Pedestrians utility vs. cost
- Pedestrians risk perception vs. trip optimization
- Interaction with vehicles
- Interaction with other pedestrians
- Pedestrian behavior (i.e. road crossing) is affected by road, traffic and human factors
- Existing simulation models focus on crowd modeling while the interaction with vehicles is not adequately examined.
- Statistical models for local behavior (e.g. at a junction)
Example: modelling pedestrian road crossing choices

Papadimitriou, 2012, Transportation Research F

Two NTUA field surveys video-recording pedestrian behaviour:
- 2010: following 490 pedestrians (randomly selected)
- 2013: following and interviewing 75 pedestrians (selected panel)
- Sequential logit models of crossing choices
- Pedestrian behavior (i.e. road crossing) is affected by road, traffic and human factors:
  - Type of road
  - Traffic flow
  - Traffic signal
  - Walking speed
  - Risk-taking behavior
  - Walking motivations
A research method for the observation of everyday driving behaviour of road users

• **Advantages**
  - Large degree of control over the variables that affect driving behavior
  - Researchers study issues that cannot be investigated in a lab
  - Help support the external validity of research

• **Disadvantages**
  - Difficult to determine the exact cause of a behavior
  - The experimenter cannot control outside factors
  - Traffic incidents are very rare
Driver behaviour – Simulator studies

Driving simulators allow for the examination of a range of driving performance measures in a controlled, relatively realistic and safe driving environment

• **Advantages**
  - Collection of data which would be very difficult to collect under real traffic conditions
  - Exploration of any possible driving scenario
  - Driving conditions are identical for all drivers

• **Disadvantages**
  - Non totally realistic simulated road environment
  - Possibility of adopting a different driving behavior
  - Feeling of safety
  - Simulator sickness
Distracted driving

Driver distraction: “The diversion of attention away from activities critical for safe driving toward a competing activity”

• Distracted driving sources: in-vehicle or external

• Mechanism

- Distance from the central axis
- Reduced speed
- Increased reaction time
- Increased accident probability
Mobile phone use, driver speed and accident probability
Driving Simulator Experiment, NTUA, 2010

- Investigation of the interrelation between mobile phone use, driver speed and accident probability.
- The research focuses on the behaviour of 30 young drivers aged between 18 and 30 years old.
- A driving simulator experiment took place, in which participants drove in:
  - urban / interurban areas
  - good / rainy weather conditions
  - with / without the use of mobile phone
- Binary logistic regression methods were used to analyse the combined influence of mobile phone, driver speed and other parameters on the probability of an accident.
Mobile phone use leads to:
- Significant **decrease of mean speed** in urban and interurban environment
- **Increase of accident probability**
Mobile phone - Simple vs. complex conversation

*Driving Simulator Experiment, NTUA, 2010*

Participants: 48 drivers aged between 19 and 27 years, out of which 29 were males and 19 were females. The experiment included 3 simulated drives in a rural road environment during good weather conditions.

- Simple mobile phone conversation: basic questions on driver’s characteristics (age, name, job, hobbies, news etc.)
- Complex mobile phone conversation: questions requiring some concentration, as well as some logical and mathematical reasoning
- Listening to music

Simple and complex mobile phone conversations were associated with reduced speeds.

Listening to music was associated with increased speeds.

Only complex mobile phone conversation was associated with reduced reaction times and increased accident risk at unexpected incidents.
FROM CORRELATION TO CAUSATION AND POLICY SUPPORT
Is road safety management linked to safety outcomes?

Papadimitriou & Yannis, 2013, AAP journal

• A recent research examined the relationships between the layers of the pyramid by means of Poisson- and Beta-regression models on data for European countries

• Road safety management does not appear to directly affect road safety outcomes

• Road safety management appears to affect SPIs

• SPIs are in turn strongly related to the final outcomes (fatalities)

Confirming the structure suggested by the road safety pyramid
From correlation to causation

• The use of **rigorous methods for accident analysis** provides evidence of the relationships between road safety and various factors

• The use of **efficiency assessment** provides evidence for the effect of interventions on road safety

• The interpretation of these analyses alone may not always explain the underlying causes of the effects.

• The intermediate level of SPIs is required for understanding the road safety process

• The **monitoring and analysis of road user behaviour** allows the establishment of the underlying links between accident causes and impacts.
From data monitoring and analysis to policy support
Conclusion

• Despite the important improvements in road safety, the number of road accident casualties around the world is still unacceptable and there is need for intensification of efforts for further improvement.

• Effective road safety management systems need to be based on rigorous scientific evidence supporting the identification of accident causes and appropriate countermeasures.

• Open knowledge systems are key tools for evidence based policies (European Road Safety Observatory, European Road Safety Measures Repository).
Fundamental directions for road safety research

- Streamline road safety monitoring by exploiting also new technological advances (big data)
- Establish the links between accident causation and injury causation
- Establish the links:
  - between measures and behaviour
  - between behaviour and risk
- Focus on the most cost-effective measures
- Improve the transferability of methods and experiences
- Bridge the gaps between research and policy
The future road safety challenges
From correlation to causation and policy support

1. The Traffic Engineering Laboratory of the NTUA
2. Evidence-based road safety policies
3. Road safety analysis methods
4. Road safety measures assessment
5. Road user behaviour
6. From correlation to causation and policy support
Traffic and safety data analysis: from correlation to causation and policy support

George Yannis, Professor
www.nrso.ntua.gr/geyannis