Road Safety Data Analysis
Conflicting Objectives - Modeling Implications

Matthew G. Karlaftis, Ph.D.
National Technical University of Athens
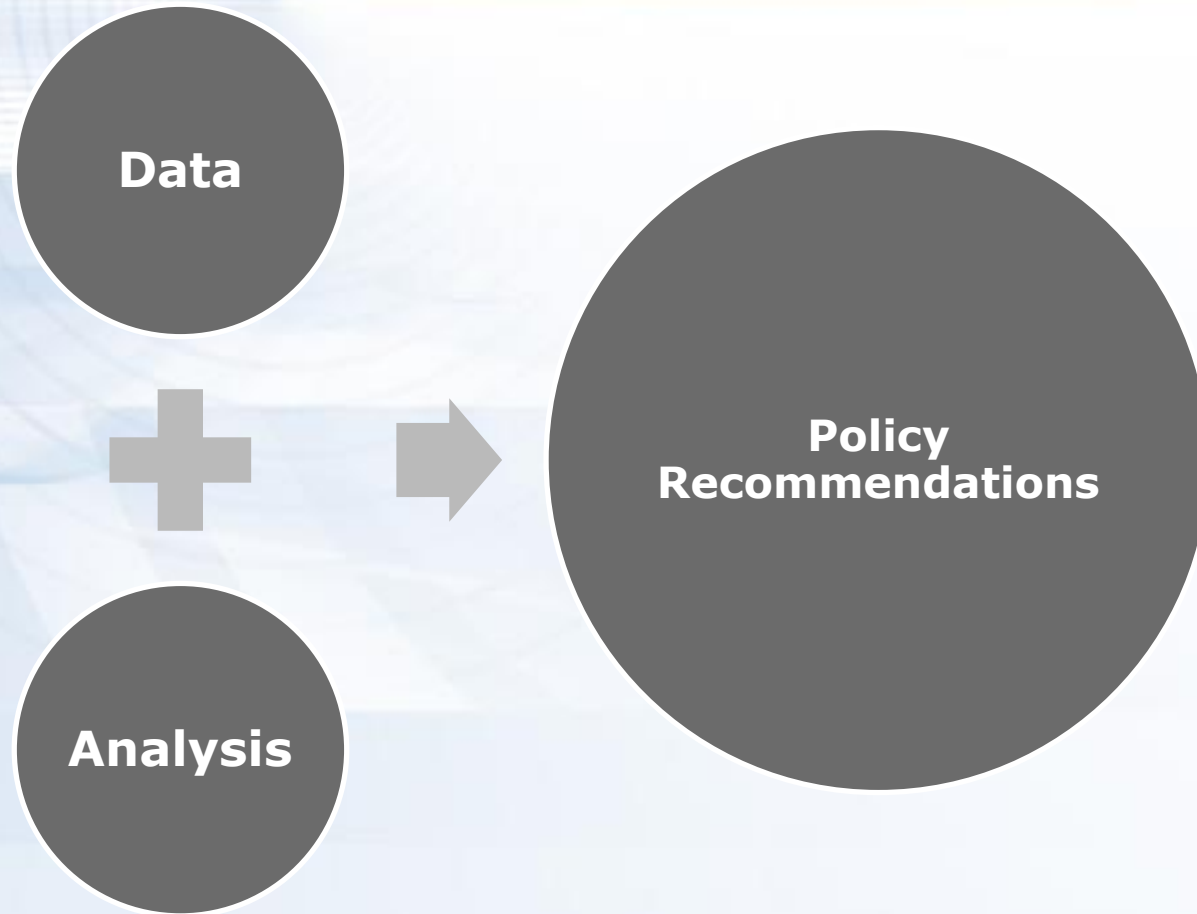http://users.civil.ntua.gr/mgk

RSS
2013

Outline

1. **How are decisions made?**

2. **New data in Safety**

3. **Emerging modeling challenges**

Data Analysis in Safety

Data

Analysis

Policy Recommendations

Does this relationship (always) work?

But…

❑ **Analyses may lead to "erroneous" policies**

1. Errors in data

2. Model misspecification

3. Complexity and hidden relationships

(1) Data Errors

❑ **Manual data collection**

  ✓ police reports problematic (severity assessment, location)

❑ **Incomplete datasets**

  ✓ Lack of exposure data – insufficient information

❑ **Spatial and temporal coverage**

  ✓ Pre-, during, and post- crash data

(2) Model Misspecification

❑ **Unreliable statistical analyses**

✓ Model selection

✓ Parsimonious vs fully specified models

- Spatial- and temporal correlations
- Unobserved heterogeneity
- Selectivity, bias/Endogeneity
- Under-reported phenomena (severe crashes)
- Risk compensation

(3) Complex Relationships

Safety

**Economics/
Social Equity**

**Environmental
Sustainability**
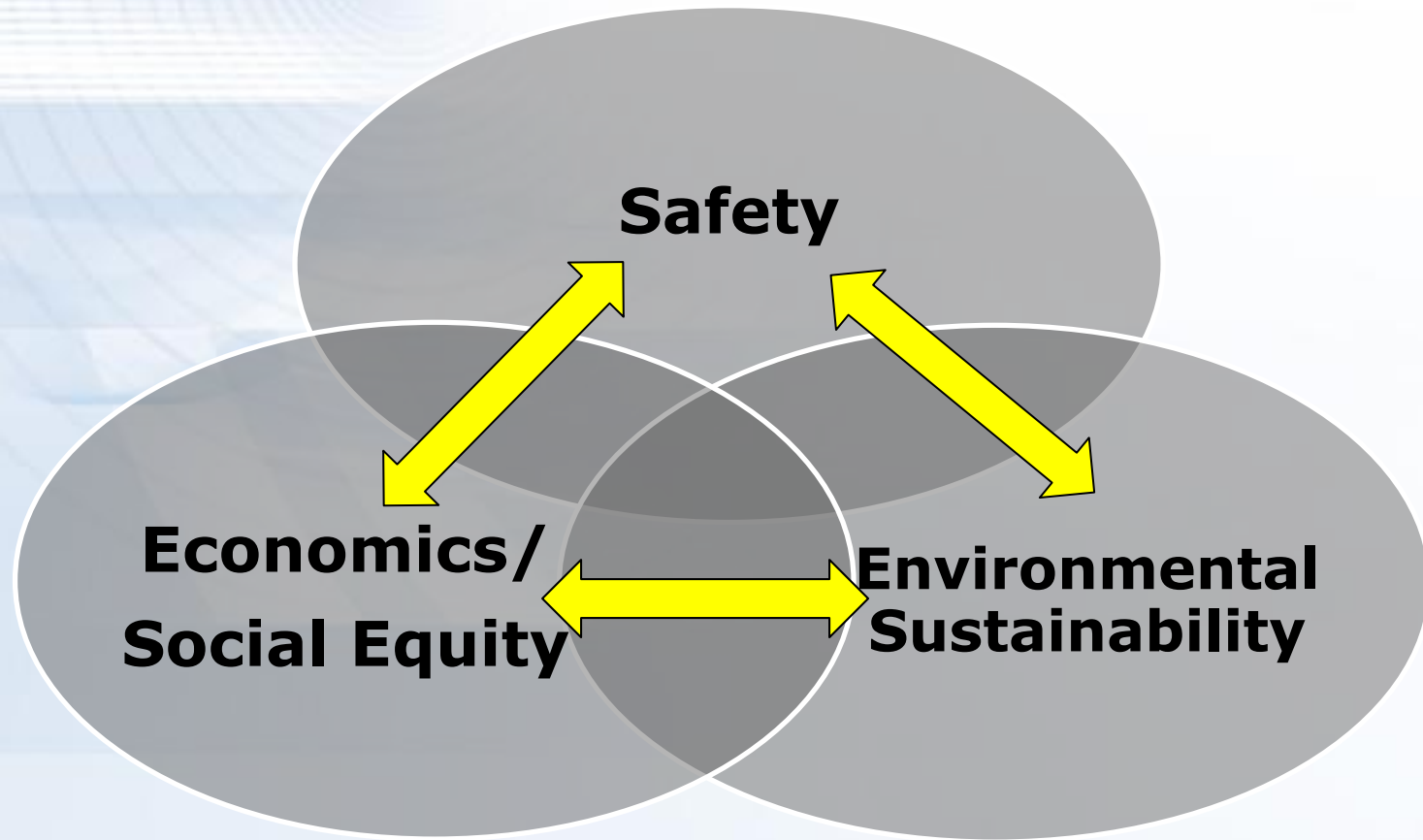
Conflicting objectives?

Conflicting Objectives

❏ **Environmental sustainability ➡ highway safety**

 ✓ Improve fuel efficiency and reduce GHG → smaller vehicles

 ✓ Climate change → crash frequencies and severities

❏ **Highway safety ➡ environmental sustainability**

 ✓ Life cycle costs of road safety countermeasures (pavement rehabilitation, friction enhancements, …)

Conflicting Objectives

❑ **Highway safety ➡ social equity**

  ✓ Increased costs associated with advanced safety features

  ✓ Vehicle safety features impose externalities on those without

❑ **Economics/social equity ➡ Highway safety**

  ✓ Demand for freight movement → increases truck traffic & speed

  ✓ Increased mobility → traffic speeds & volume

Understanding Complexities

❑ Many traditional tools are applicable (*efficient?*)

❑ Understanding of the process has been data-limited

❑ Driving Forces for Change

    1.   New Technologies and Data Sources

    2.   Novel Modeling Paradigms

… where Change lies, Challenges Emerge!
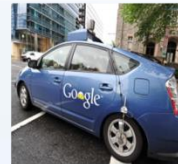
Emerging Data Sources

❑ Safety- Related

- ✓ More detailed crash and injury data
- ✓ Naturalistic driving data
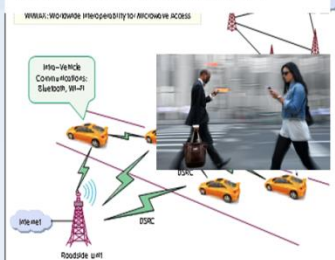- ✓ Vehicle Event Data Recorders (EDRs)

Mobile Internet

Autonomous vehicles

THE INTERNET OF THINGS

BIG Data

Wireless communications

Cloud computing

Energy Storage Technologies

Emerging Data Sources: Big Data

Policy decisions with big data is … different

❑ **Much more operational data**
  ✓ Volumes, speeds, travel time etc.

❑ **Much more unstructured data**
  ✓ social media data
 ✓ Web transactions
  ✓ Smart objects

Emerging Data Sources: Challenges

1. **Database size and high dimensionality**

2. **Overfitting and assessing statistical fit**

3. **Rapidly changing and imperfect data**

4. **Complexities and interactions**

Can traditional statistical modeling efficiently transform massive data into knowledge?

Big Data Analysis

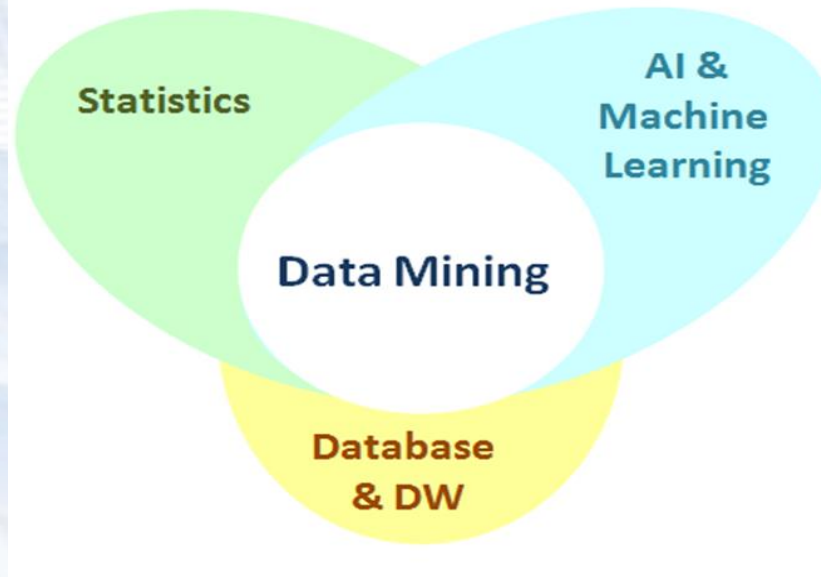**Debatable...**

Statistical Modeling

Novel Modeling Paradigms
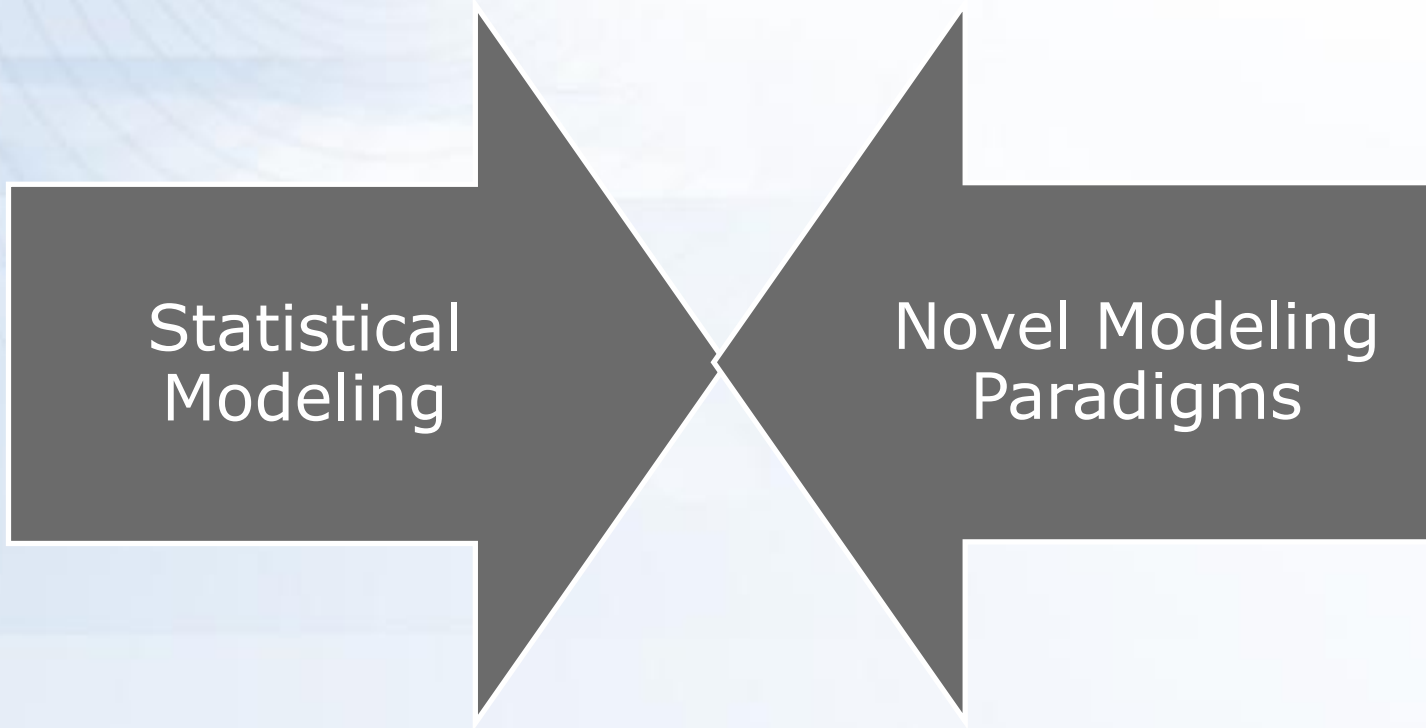
Analysis: Statistical Modeling

❑ **Classical models**

❑ **But we also have extensions...**

- ✓ Random Parameter models
- ✓ Finite mixture latent class
- ✓ Markov switching
- ✓ Multivariate models
- ✓ Psychometric (ICLV)
- ✓ …

Analysis: Data Mining

❑ **Classical techniques from statistics, machine learning and CI**

❑ **Faster analysis of growing data with errors and imperfections**

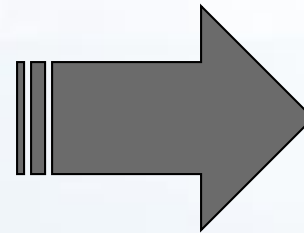❑ **Enables (easier) data fusion**

❑ **Robust to uncertainty**

Data Mining in Transportation

Three promising and rapidly growing research fields

1. ITS Applications

2. User Experience Mining

3. Behavioral Recognition

Significant implications to proactive safety policies

The road ahead contains well developed statistical models and intelligent data mining

….but with caution

Key Drivers

❑ **Demystifying the gap between modeling and policy making**

   ✓ Caveats

❑ **Unraveling the complex relationships**

   ✓ Emerging technologies and novel methodologies

   ✓ Means for success

      ▪ **Synergies of statistics and Data Mining**

      ▪ **Testbeds and Test data**

Key Drivers

Efficient policy recommendation will stem from understanding complexities

❑ **Disregard complexities**
   - ✓ We will be simply ignoring potentially important information moving blindly forward

❑ **Do something about them but poorly**
   - ✓ Will lead to erroneous and ineffective policies that will adversely impact safety, social equity and environmental sustainability

A New Era in Safety Analysis

❑ **Safety will always be the cornerstone of modern societies**

❑ **Emerging technologies**

❑ **New and exciting data**

❑ **We also must consider**
  - ✓ Modeling Efficiency
  - ✓ Informed Decision Making
  - ✓ Proactive Management

Some Thoughts

❑ **We need to expand our educational boundaries**

❑ **New Learning Paradigms**

    ✓ Formulating new problems - revisiting old

❑ **New tools for dealing with old and new problems**

❑ **Beware of the hype...**

Road Safety Data Analysis
Conflicting Objectives - Modeling Implications

Matthew G. Karlaftis, Ph.D.
National Technical University of Athens
http://users.civil.ntua.gr/mgk

RSS
2013