3rd International Conference on Road Safety and Simulation
Session: MODERN SAFETY MANAGEMENT
Thursday, September 15

# DETECTING & MODELING SECONDARY ACCIDENTS ON TOLLWAYS

Matthew G. Karlaftis, Ph.D.
mgk@mail.ntua.gr

# OUTLINE

- Secondary Incidents-crashes
- Scope
- Application and Results
- Conclusions

# SECONDARY INCIDENTS

- **Crashes within a predefined spatiotemporal region of a primary incident.**
- **Importance**
  - Major source of freeway incidents → additional traffic delay
  - Much more severe than the primary incident
  - Longer clearance times
  - Affect freeway operations
- **Identification contributes to maintaining and increasing safety levels on freeways**

# PREVIOUS RESEARCH

- **Secondary Incident Detection**
  - Pre-defined spatiotemporal criteria
  - Dynamic thresholds based on reported queue data
  - Estimation of maximum queue length and traffic delay based on cumulative arrival and departure plots.
- **Influencing factors of Secondary Incident Likelihood**
  - Traffic (but, usually AADT, day and time, period of day).
  - Weather conditions not explicitly considered

# SCOPE OF WORK

- **Analytically define dynamic thresholds of the influence area of a primary incident**
  - Use of detailed real-time collected traffic data from upstream loop detectors
  - Accident occurring within the defined influence area → secondary

- **Modeling secondary accident likelihood**
  - Concentrate on real-time traffic and weather effects
  - Compare methodological approaches (Stats & AI)

# METHODOLOGY

ASDA Model

$$x_{up}^{(jam)}(t) = L_{i+1} - \int_{t_0^{(i+1)}}^{t} \frac{q_0^{(i)}(t) - q_{min}}{\rho_{max} - \left(q_0^{(i)}(t)/w_0^{(i)}(t)\right)} dt$$

$$x_{down}^{(jam)}(t) = L_j - \int_{t_1^{(j)}}^{t} \frac{q_{out}^{(j)(jam)}(t) - q_{min}}{\rho_{max} - \left(q_{out}^{(j)(jam)}(t)/w_{max}^{(j)}(t)\right)} dt$$

**Dynamic spatiotemporal boundaries of the influence area of the primary incident fully defined**

- Automatic tracking of the propagation of moving traffic jams at any time
- Accurate calculation of the positions of the upstream and downstream front of the bottleneck
- Calculation of the jam width, queue length propagation and queue duration

# METHODOLOGY

- **Secondary accident likelihood modeling**
  - Logit (Probit, Gompit)
  - Multilayer Perceptrons

- **Independent Variables**
  - Duration
  - Collision Type
  - Nr. Lanes
  - Nr. Vehicles
  - Heavy Vehicle
  - Travel Speed
  - Hourly volume
  - Rainfall
  - Alignment
  - Downstream Geometry
  - Upstream Geometry

# METHODOLOGY

## Neural Networks as Explanatory Models

- Mutual Information $$I(Y\mid\vec{X}) = \frac{H(Y)-H(Y\mid\vec{X})}{H(Y)}\times 100$$
- Partial Derivatives $$\frac{\partial y_j}{\partial x_i} = \sum_k \frac{\partial y_j}{\partial net_j}\frac{\partial net_j}{\partial h_k}\frac{\partial h_k}{\partial net_k}\frac{\partial net_k}{\partial x_i} = \sum_k f_k' w_{kj} f_k' w_{ik}$$

***Interpretation***

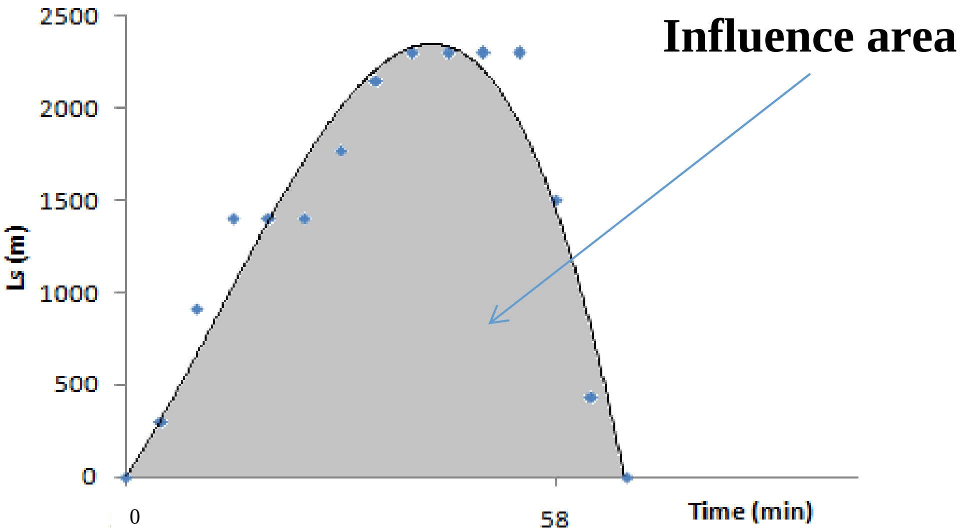
| | | Partial Derivative | |
|---|---|---|---|
| | | **High** | **Low** |
| **Mutual Information** | **High** | Small changes in the variable's value can have a high impact on the magnitude of the response. The variable has a high contribution to the response value. | Small changes in the variable's value have limited or no effect on the magnitude of the response, but its absence would have a large impact. |
| | **Low** | Although the variable contributes less to the magnitude of the response, changes to the variable's value might have an impact on the magnitude of the response. | The variable has low contribution to the response, and offer limited opportunity to effect change to the magnitude of the response. |

# DESCRIPTION OF DATA

| Variable | Type | Description |
|---|---|---|
| Duration | Continuous | The incident duration in minutes |
| Collision Type | Categorical | 0 to 4, from no history to delays |
| Nr. Lanes | Categorical | 1 to 3, 1:1 lane, 2: two, 3: more than 2 |
| Nr. Vehicles | Categorical | 1 to 3, 1:one vehicle, 2: two vehicles, 3: more than 2 vehicles involved |
| Heavy Vehicle | Categorical | 0 to 1(Heavy Vehicle involved) |
| Travel Speed | Continuous | Travel speed (km/h) at the occurrence of the incident |
| Hourly volume | Continuous | Hourly volume (veh/h/lane) at the occurrence of the incident |
| Rainfall | Continuous | Rainfall at the occurrence of the incidentin mm/10min |
| Alignment | Categorical | 0 to 1(curve) |
| Downstream Geometry | Categorical | 0 to 4, 0: no special geometry, 1: adjacent to tunnel, 2: adjacent to toll, 3: adjacent to entrance/exit, 4: more than one |
| Upstream Geometry | Categorical | 0 to 4, 0: no special geometry, 1: adjacent to tunnel, 2: adjacent to toll, 3: adjacent to entrance/exit, 4: more than one |

# RESULTS

- Traffic jam's width ($L_s$) versus time




- Accidents falling within the curve's boundaries are considered as secondary

# RESULTS

- **Identified secondary accidents:**
  - 3.5% of the total number of accidents based on our approach
  - 3.3% of the total number of accidents using fixed boundaries of 2 hours and 3.2 km
  - 1.6% of the total number of accidents using fixed boundaries of 15 minutes plus clearance time and 1.6 km

# RESULTS

***Variable contribution to the secondary accident risk with respect to the conditional mutual information.***

*High contribution to the response value*

Low contribution to the response value

| Variable | Normalized Mutual Information |
|---|---|
| Speed | 1.00 |
| Duration | 0.72 |
| Hourly volume | 0.69 |
| Rainfall | 0.67 |
| Nr. Vehicles | 0.51 |
| Upstream Geometry | 0.45 |
| Nr. Lanes | 0.45 |
| Collision Type | 0.39 |
| Alignment | 0.38 |
| Heavy Vehicle | 0.33 |
| Downstream Geometry | 0.27 |

# RESULTS

**Input variable significance[1] with respect to partial derivatives**

| Variable | Partial Derivative | |
|---|---|---|
| | MLP(11-12-2)[2] | Logit |
| Duration | 0.014 | 0.014 |
| Collision Type | -0.023* | -0.023* |
| Nr. Lanes | -0.065** | -0.065** |
| Nr. Vehicles | 0.063** | 0.062** |
| Heavy Vehicles | -0.098** | -0.097** |
| Speed | -0.091** | -0.091** |
| Lane Volume | 0.053** | 0.053** |
| Rainfall | 0.021** | 0.018** |
| Alignment | -0.063** | -0.066** |
| Downstream Geometry | 0.003 | 0.003 |
| Upstream Geometry | 0.034** | 0.035** |

*[1] Significance is calculated from bootstrapped distributions*
*[2] The numbers in parentheses signify the number of neurons of the input, hidden and output layer.*
**significance at 95% level, ** significance at 99%level*

# RESULTS

**Error Percentages for out-of-sample forecasting**

| Model[1] | False Positive | False Negative | Percent Correct |
|---|---|---|---|
| MLP(11-6-2) | 0.042 | 0.102 | 72 |
| MLP(11-12-2) | 0.050 | 0.089 | 63 |
| Logit | 0.044 | 0.102 | 70 |

*[1] The numbers in parentheses signify the number of neurons of the input, hidden and output layer of the MLP.*

# RESULTS

- **Evaluation of the two sensitivity measures**
  - Small changes in speed and volume have a high impact on secondary accident likelihood.
  - Upstream geometry complexity, # of vehicles involved and rainfall intensity have limited effect on secondary accident likelihood
  - Changes in the % heavy vehicles, alignment, total rainfall, and # of lanes have an impact on secondary accident likelihood.

# CONCLUSIONS

- **Identification of dynamic boundaries defined using real-time traffic data**
  - May work with historical data
  - Analytical calculation of the position of the upstream and downstream front of the moving jam
  - No demanding calculations
  - ‘Accurate’ determination of secondary crashes

# CONCLUSIONS

- **Influential factors: traffic conditions at the occurrence of an incident and rainfall.**
  - speed and volume, number of blocked lanes, heavy vehicles and alignment were found significant.
  - Changes in downstream geometry and **duration** of an incident were not significant.
- **Both models (NN, Stats) give similar results.**
- **Neural Network models for transportation applications with tractable explanatory power.**