Themes & trends

Key Focus: Driver & Technology

Key Targets: Safety, Energy, & Mobility

40,000/year × $9.6 M/human life

USD 384 billion/year + Injuries
A Simple Picture of CAVs

BIG DATA

Sensors in vehicle

Warnings
Automation & impacts

Work/residence locations, auto ownership & use (4%)
  - Shared Mobility

Trip Generation & Distribution
  - More trips (aging/underserved), Longer trips
  - Changing demand generators?

Mode choice
  - New shared ride modes (Uber, Lyft) (↑)
  - Drive alone?
  - Transit (↓)

Route assignment
  - Efficient routing/platooning
  - Undetermined cost functions
  - System optimum vs User equilibrium
  - Micro-behaviors

Vehicle & Infrastructure
  - Light weighting & Electrification/AFVs (Safety?)
  - Automated lanes, CMS/HAR, Pavement markings
Concept Development - Driving Volatility

Broadly defined *Driving Volatility*

Narrowly defined *Driving Volatility*

**Take-away concept**

*Driving Volatility:*
The extent of difference from the norm.

Sources:

Different measures of volatility

- Speed-Based Volatility
- Acceleration/Deceleration Based Volatility
  - Longitudinal
  - Lateral
- Vehicular Jerk-Based Volatility
  - Longitudinal
  - Lateral
Different measures of volatility

Driving volatility

- Speed-Based Volatility
- Acceleration/Deceleration Based Volatility
  - Longitudinal
  - Lateral
- Vehicular Jerk-Based Volatility
  - Longitudinal
  - Lateral

Statistical Measures of Volatility

- Std.dev
- Mean ± X*SD
- Mean absolute deviance
- Coefficient of variation
- …..

- Historical financial volatility

\[ SD(ln\left(\frac{x_i}{x_{i-1}}\right) * 100) \]
Volatility matrix

- Trip-based volatility
- Location-based volatility
- Event-based volatility
- Driver-based volatility
Volatility matrix

- Trip-based volatility
- Location-based volatility
- Event-based volatility
- Driver-based volatility

VI: Volatility index
Volatility matrix

Trip-based volatility

Location-based volatility

Event-based volatility

Driver-based volatility

Location “i”

A → B: VI:10%

C → D: VI:14%

E → F: VI: 4%
Volatility matrix

- Trip-based volatility
- Location-based volatility
- Event-based volatility
- Driver-based volatility

Location “i”

A → B: VI:10%
C → D: VI:14%
E → F: VI:4%
G →: VI:34%
Long-term Goal of our NSF Study*

- Reduce driving volatility ->> control assists
- Increase cooperation between proximate vehicles to respond to unexpected events
- Ideas...
  - Understand microscopic driving volatility using Bayesian and Dynamic Markov Switching Framework.
  - MDP, Inverse Reinforcement Learning, Gossip algorithms
- Understand drivers’ decisions at different speeds

CAV Data – Basic Safety Messages

**SAE J2735 BSM** standard:

- BSMs transmitted using DSRC-range ~1,000 ft

Red indicates High Priority Data Elements

Delivering improved alerts, warnings, and control assistance using basic safety messages transmitted between connected vehicles (Liu and Khattak)

Quantification of Driving Volatility

\[
\text{Volatility Score} \% = \frac{\text{Seconds of "Instantaneous Decision" > Threshold}}{\text{Seconds of Entire Trip}} \times 100 \tag{Equation 1}
\]

Where, \( \text{Threshold} = m \text{ (mean)} + \sigma \text{ (standard deviation)} \) of \text{Instantaneous Driving Decision} (acceleration or vehicular jerk) within a speed range \( k \) (and/or direction of decisions).

\[
\text{Vehicular Jerk}, j = \frac{da}{dt} = \frac{d^2v}{dt^2} = \frac{d^3r}{dt^3} \tag{Equation 2}
\]
Data Visualization

Surface Threshold for identifying critical driving moments

Gray dots: Volatile Seconds, outside of the Surface Threshold
Instantaneous Driving Profile on Road

Warning and control assist based on extreme driving seconds (0.5 seconds above 95th percentile)
Applications: Driver Feedback

For Drivers

Real time Info

Driving Profile

Daily driving performance summary (or monthly, yearly)

You drove 55 km today
Your average volatility is 22%
Average volatility for other drivers in your area is 12%

Analysis of volatility in driving regimes extracted from basic safety messages transmitted between connected vehicles (Khattak & Wali, 2017).

Citation: Khattak, A. J., & Wali, B. (2017). Analysis of volatility in driving regimes extracted from basic safety messages transmitted between connected vehicles. Transportation Research Part C: Emerging Technologies, 84, 48-73.
Analysis of volatility in driving regimes extracted from basic safety messages transmitted between connected vehicles

**Behavior Conceptualization**

Associated Factors: Instantaneous driving contexts

**Three-Regime Markov Switching Dynamic Regression Framework**

For short-term prediction...
Markov-Switching Dynamic (abrupt-change) Probabilistic Model

General Time-series Framework

Regime 1: \( y_t = \mu_1 + \phi y_{t-1} + \epsilon_t \)

Regime 2: \( y_t = \mu_2 + \phi y_{t-1} + \epsilon_t \)

If timing of switching is known, then:

\( y_t = s_t \mu_1 + (1 - s_t) \mu_2 + \phi y_{t-1} + \epsilon_t \)

However, \( s_t \) is not observed i.e. stochastic time series driving process. Formulation of MSDR:

\( y_t = \mu_{s_t} + X_t \alpha + Z_t \beta_{s_t} + \epsilon_t \)

Where:
\( y_t \) is the dependent variable \\
\( \mu_{s_t} \) is the regime-dependent intercept term \\
\( X_t \) is a vector of exogenous variables with regime-independent coefficients \\
\( Z_t \) is a vector of exogenous variables with regime-dependent coefficients \\
Both error terms \( \sim N(0, \sigma^2) \) and \( N(0, \sigma^2) \)

Markov Chains

Irreducible and aperiodic Markov chain from ergodic distribution:

\[
P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}
\]

\( s_t \) is not observed. \( s_{t+1} \) depends on current state \( s_t \) and not on previous history \( s_{t-1}, s_{t-2}, \ldots \)

\[
\text{Prob}(s_{t+1} = 1 | s_t = 2) = p_{21}
\]
\[
\text{Prob}(s_{t+1} = 2 | s_t = 1) = p_{12}
\]

Likelihood function with latent regimes

\[
\text{Pr}(s_t = i; y_t; \theta) = \frac{f(y_t | s_t = i, y_{t-1}; \theta) \text{Pr}(s_t = i; y_{t-1}; \theta)}{f(y_t | y_{t-1}; \theta)}
\]

\[
L(\theta) = \sum_{t=1}^T \log f(y_t | y_{t-1}; \theta)
\]

In simplest case i.e. constant-only model, \( \theta = [\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_{1-2}, p_{2-1}] \)

Transitional Probabilities

\[
p_{ij} = \frac{\exp(-q_{ij})}{1 + \exp(-q_{11}) + \cdots + \exp(-q_{ij})} \quad j \in (1, \ldots k - 1)i.e. 1 & 2
\]

\[
p_{ik} = \frac{1}{1 + \exp(-q_{i1}) + \cdots + \exp(-q_{ij})} \quad q_{ij} = -\log \left( \frac{p_{ij}}{p_{ik}} \right)
\]

Maximum a posteriori (MAP) parameterization is maximized through Expectation-Maximization algorithm

Data quality checked...
Results: Two-regime constant-only Models (summary of all trips)

• Two distinct driving regimes exist – empirical data strongly favor MSDR
• Category 1 & 2 trips, drivers decelerate at higher rate
• Deceleration is more volatile

Regime Durations

<table>
<thead>
<tr>
<th></th>
<th>Category 1 trips</th>
<th>Category 2 trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceleration</td>
<td>75</td>
<td>25</td>
</tr>
<tr>
<td>Deceleration</td>
<td>58</td>
<td>53</td>
</tr>
</tbody>
</table>

Transition Probabilities

<table>
<thead>
<tr>
<th>Category</th>
<th>Acceleration</th>
<th>Deceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1 trips</td>
<td>0.918</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>0.039</td>
<td>0.961</td>
</tr>
<tr>
<td>Category 2 trips</td>
<td>0.912</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>0.043</td>
<td>0.957</td>
</tr>
</tbody>
</table>
Summary of specified **three-regime models** (see paper for details)

**Table 11**

Three-regime Markov switching models – summary of direction of effects for all trips.

<table>
<thead>
<tr>
<th>Roadway Type</th>
<th>Driving Regimes</th>
<th>Variable</th>
<th>↑</th>
<th>↓</th>
<th>Not Significant at 95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeways &amp; State Routes (N = 13 trips)</td>
<td>High Rate Acceleration-Regime 1</td>
<td>Constant</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objects indicator</td>
<td>2 (15.4%)</td>
<td>6 (46.2%)</td>
<td>5 (38.5%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range</td>
<td>3 (23.1%)</td>
<td>7 (53.8%)</td>
<td>3 (23.1%)</td>
</tr>
<tr>
<td></td>
<td>High Rate Deceleration – Regime 2</td>
<td>Constant</td>
<td>0</td>
<td>13 (100%)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objects indicator</td>
<td>3 (23.1%)</td>
<td>6 (46.2%)</td>
<td>4 (30.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range</td>
<td>2 (15.4%)</td>
<td>8 (61.5%)</td>
<td>3 (23.1%)</td>
</tr>
<tr>
<td></td>
<td>Constant/Cruise around 0 – Regime 3</td>
<td>Constant</td>
<td>3 (23.1%)</td>
<td>6 (46.2%)</td>
<td>4 (30.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Objects indicator</td>
<td>4 (30.8%)</td>
<td>2 (15.4%)</td>
<td>7 (53.8%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range</td>
<td>8 (61.5%)</td>
<td>1 (7.7%)</td>
<td>4 (30.8%)</td>
</tr>
</tbody>
</table>
Short-term regime predictions

- Potential to predict driving regime at specific instant of time
- Regime - stochastic event - Description of probability law governing the change from regime 1 to regime 2 and vice versa
- Example: 25-minutes trip undertaken on freeway I-94 in Ann Arbor, Michigan
  Smoothed probabilities using all-sample information
Short-term regime predictions

Dynamics of Driving Regimes Extracted from Basic Safety Messages Transmitted Between Connected Vehicles

Acceleration associated with lower volatility
Location-Based Volatility

How is Driving Volatility Related to Intersection Safety in a Connected Vehicles Environment?


Motivation

- Location-based volatility >> leading indicators of crashes
- **Objective:** identify locations where *crashes are waiting to happen*

  **Methodological:** Frequentist vs Full-Bayesian Approach, unobserved heterogeneity
Descriptive Statistics

Traditionally, standard deviation/variance has been used to capture dispersion.

**TABLE 1: Descriptive Statistics of Key Variables**

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>All intersections</th>
<th>Signalized Intersections</th>
<th>Unsignalized Intersections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
<td>Min</td>
</tr>
<tr>
<td>Volatility related</td>
<td>Mean Speed</td>
<td>20.04</td>
<td>10.24</td>
<td>5.65</td>
</tr>
<tr>
<td></td>
<td>Standard deviation of speed</td>
<td>11.38</td>
<td>2.28</td>
<td>4.96</td>
</tr>
<tr>
<td></td>
<td>Minimum speed</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Maximum speed</td>
<td>57.78</td>
<td>10.07</td>
<td>37.80</td>
</tr>
<tr>
<td></td>
<td>Coefficient of variation of speed</td>
<td>0.76</td>
<td>0.40</td>
<td>0.13</td>
</tr>
</tbody>
</table>

**COV as a new safety surrogate...**
Accidents waiting to happen (visual illustration)

**Top Panel (A):** Crash frequency & Intersection-specific volatilities

**Bottom Panel (B):** focuses on subset of intersections which are highlighted in top panel. Solid circles refer to “known hot-spots” and dashed circles refer to intersections where “crashes may be waiting to happen”
## Table: Full Bayes Gibbs Sampler Random Parameter Estimation of All Intersections

<table>
<thead>
<tr>
<th>Variables</th>
<th>Poisson Model</th>
<th>Hierarchical Random Parameter Poisson Model</th>
<th>Hierarchical Random Parameter Poisson Log-Normal Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parameter Estimates (95% credible intervals)</td>
<td>Parameter Estimates (95% credible intervals)</td>
<td>Parameter Estimates (95% credible intervals)</td>
</tr>
<tr>
<td>COV of Speed</td>
<td>0.4244 (0.1789,0.678)</td>
<td>0.4951 (0.08915,0.8385)</td>
<td>0.5997 (0.1558,1.137)</td>
</tr>
<tr>
<td>Major road AADT (log form)</td>
<td>0.7152 (0.5571,0.8952)</td>
<td>0.6039 (0.416,0.8305)</td>
<td>0.5107 (0.3362,0.692)</td>
</tr>
<tr>
<td>Minor road AADT (log form)</td>
<td>0.2221 (0.09374,0.3678)</td>
<td>0.1857 (0.004587,0.3903)</td>
<td>0.2656 (-0.01793,0.5976)</td>
</tr>
<tr>
<td>Signalized intersection</td>
<td>0.7823 (0.5577,0.9989)</td>
<td>0.7838 (0.4507,1.065)</td>
<td>0.7901 (0.4667,1.142)</td>
</tr>
<tr>
<td>Four leg intersection</td>
<td>0.2595 (0.08564,0.4335)</td>
<td>0.2612 (0.03635,0.5226)</td>
<td>0.2096 (-0.08113,0.4904)</td>
</tr>
<tr>
<td>Total left turn lanes</td>
<td>0.08368 (0.02484,0.1431)</td>
<td>0.07753 (0.008434,0.1722)</td>
<td>0.07318 (0.02981,0.1806)</td>
</tr>
</tbody>
</table>
### Full Bayes Hierarchical Random Parameter Models

#### Table: Full Bayes Gibbs Sampler Random Parameter Estimation of All Intersections

<table>
<thead>
<tr>
<th>Unobserved effects</th>
<th>Estimate 1</th>
<th>Estimate 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation (COV of Speed)</td>
<td>0.04136</td>
<td>0.04267</td>
</tr>
<tr>
<td>Standard deviation (Major road AADT)</td>
<td>(0.01607, 0.1105)</td>
<td>(0.01638, 0.113)</td>
</tr>
<tr>
<td>Standard deviation (Signalized intersection)</td>
<td>0.03767</td>
<td>0.03431</td>
</tr>
<tr>
<td>Standard deviation (Four leg intersection)</td>
<td>(0.02585, 0.05049)</td>
<td>(0.01806, 0.05033)</td>
</tr>
<tr>
<td>Standard deviation (total left turn lanes)</td>
<td>0.04591</td>
<td>0.04355</td>
</tr>
<tr>
<td>Standard deviation (Four leg intersection)</td>
<td>(0.01651, 0.1294)</td>
<td>(0.01615, 0.1318)</td>
</tr>
<tr>
<td>Extra-Poisson variance</td>
<td>0.0463</td>
<td></td>
</tr>
<tr>
<td>Goodness of Fit</td>
<td>630.128</td>
<td>515.01</td>
</tr>
<tr>
<td>Dbar</td>
<td>622.91</td>
<td>432.174</td>
</tr>
<tr>
<td>pD</td>
<td>7.218</td>
<td>82.836</td>
</tr>
<tr>
<td>DIC</td>
<td>637.34</td>
<td>597.846</td>
</tr>
</tbody>
</table>

See paper for details on segmented models for signalized intersections
### Difference in Elasticity Effects Across Models

A one percent increase in volatility increases crashes by (after controlling for observed & unobserved factors):

- **0.37%** for all intersections
- **0.66%** for signalized intersections.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Poisson Model</th>
<th>HRPP Model</th>
<th>HRP-PLN Model</th>
<th>Poisson Model</th>
<th>HRPP Model</th>
<th>HRP-PLN Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection-specific volatility</td>
<td>0.32 (0.76)</td>
<td>0.37 (0.88)</td>
<td>0.45 (1.07)</td>
<td>0.55 (0.88)</td>
<td>0.66 (1.08)</td>
<td>0.67 (1.11)</td>
</tr>
<tr>
<td>Major road AADT (log form)</td>
<td>7.04 (7.66)</td>
<td>5.94 (6.47)</td>
<td>5.02 (5.47)</td>
<td>5.73 (6.16)</td>
<td>4.72 (5.08)</td>
<td>4.83 (5.20)</td>
</tr>
<tr>
<td>Minor road AADT (log form)</td>
<td>2.01 (2.26)</td>
<td>1.68 (1.89)</td>
<td>2.40 (2.71)</td>
<td>2.17 (2.44)</td>
<td>2.41 (2.72)</td>
<td>2.23 (2.51)</td>
</tr>
<tr>
<td>Signalized intersection</td>
<td>0.543</td>
<td>0.543</td>
<td>0.546</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Four leg intersection</td>
<td>0.229</td>
<td>0.230</td>
<td>0.189</td>
<td>0.255</td>
<td>0.279</td>
<td>0.272</td>
</tr>
<tr>
<td>Total left turn lanes</td>
<td>0.12 (0.50)</td>
<td>0.11 (0.46)</td>
<td>0.11 (0.43)</td>
<td>0.16 (0.44)</td>
<td>0.13 (0.34)</td>
<td>0.13 (0.34)</td>
</tr>
</tbody>
</table>

Notes: In parenthesis are the maximum elasticities (see text for explanation); HRPP is Hierarchical Random Parameter Poisson Model; and HRP-PLN is Hierarchical Random Parameter Poisson Log-Normal Model.

Note differences across Bayesian & heterogeneity-based Bayesian models.
Event-Based Volatility in Naturalistic Driving

Driving Volatility, Crash Propensity, & Injury Severity


Event volatility and crash propensity

**Key Idea:**
Use traditional & emerging data sources to understand links b/w driving volatility & crash propensity

**Event Types:**
1. Safety Critical Events:
   - Crash /Near-Crash
2. Baseline Events: Normal Driving Control Factors

**Driving Volatility Concept:**
Characterize driver performance prior to involvement in safety critical events

**Comparison Group:** Baseline events

**Different Performance Measures**

**Empirical toolbox:**
- ANOVA
- Discrete Choice Framework

**Unobserved heterogeneity & omitted variable bias**

**Impacts:**
- Proactive safety approach
- Behavioral countermeasures
- Alerts & warnings

**Large-Scale Data Analytics**

Volatility Indices
Event volatility and injury severity

Volatility measure based on entire 30 seconds vehicle kinematics data

Volatility based on 1st 10-sec bin (Intentional volatility)
Volatility based on 2nd 10-sec bin
Volatility based on 3rd 10-sec bin (Unintentional volatility)

Event Detail Table

Crash severity, given a crash

On-Board Data Acquisition Systems
Knowledge-Base: Spatial Distribution of Ped-Vehicle Crashes across the US

- Analysis of current pedestrian-vehicle fatal crashes and conflicts
- Development of future travel scenarios (CAV, behavior shifts)
- Assessment of future pedestrian-vehicle conflicts

Key “Fatal” Pre-Crash Behaviors-peds (80%)
- Dart-out/Dash (9.0%)
- Failure to yield right of way (21.6%)
  - Standing, Working, Lying, Playing
- Dark clothing/Not visible (10.2%)
- Inattention (1.2%)
  - Talking, Eating

2018 TRB paper: Analysis of Crashes Involving Pedestrians across the United States: Implications for Connected and Automated Vehicles
Knowledge-Base: Driving Volatility & Ped-Bike Safety Critical Events

- Analysis of SHRP-2 Naturalistic Driving Data
- How driving volatility differs across baseline, ped-bike, and no ped-bike involved events.

<table>
<thead>
<tr>
<th>Category</th>
<th>Volatility (Negative vehicular Jerk: longitudinal direction)</th>
<th>Volatility (Positive vehicular Jerk: longitudinal direction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedbike Involved Crashes</td>
<td>1.79</td>
<td>2.24</td>
</tr>
<tr>
<td>Vehicle Involved Crashes</td>
<td>1.66</td>
<td>1.97</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.82</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Key points
- Greater volatility increases the likelihood of both crash and near-crash events.
- Volatility in deceleration (negative vehicular jerk), both lateral and longitudinal, has more negative consequences than volatility in acceleration (positive vehicular jerk).

NDS Data: Event Data Table, 9543 events
Baseline sample = 7562
Vehicle (no ped-bike) involved crashes = 1851
Ped-bike involved crashes/near-misses = 74
Closure...

• Profound CAV impacts on mobility & safety by altering driving volatility
• Data to support micro-level decisions through control assists
• Customized driver feedback
• Hazard anticipation & notification systems → short term regime predictions
• Proactive safety → trip-based, location-based, & event-based volatility concepts
• Knowledge base for vulnerable road users
• Research needs/implications of our analysis
  ■ Predict changes in travel behavior
  ■ Future vehicles: How will their use be substantially different?
  ■ Network performance-partial automation-open question?
Thank YOU

Asad J. Khattak, Ph.D.
akhattak@utk.edu

Behram Wali
bwali@vols.utk.edu