Information Content of Aggregate Speed Measurements: A Distribution Analysis and Multimodality of Highway Speeds

Ilker Karaca, PhD
Instructor

Dept. Of Civil, Construction, and Environmental Engineering
Iowa State University, Ames, IA, 50011

2019 Mid-Continent Transportation Research Symposium
Iowa State University Institute for Transportation

August 21-22, 2019

Session 4D
Technology - Traffic

Outline

- Introduction and background
- Data description
- Methodology
- Main findings
- Conclusion

Summary

- This research examines the multimodality of driving speeds by using a data sample speed collected by the Iowa DOT in urban city highways.
- The findings suggest that different underlying subpopulations with unique distributional characteristics could be driving the observed multimodality in the sample.
- Both sets of findings are statistically significant and robust to a range of highway speed limits from 55 to 70 mph.
- The results are expected to be relevant to the ongoing policy debate on setting speed limits in a world where the shares of autonomous and connected vehicles are expected to climb steadily.
- Since the resulting heterogeneity in the underlying traffic subpopulations is only expected to become more palpable in the near future, these findings are expected to contribute to a better understanding of the current multimodality in speed density functions to manage potential shifts in highway traffic characteristics.

Background

- DOTs routinely gather vast amounts of vehicle speed data.
- Yet the data largely remain underutilized to investigate the potential shifts in today's traffic speed profiles as autonomous, automated, and connected vehicles become more mainstream.
- Thus such trends suggest potential changes to aggregate traffic patterns and the aggregate traffic flow patterns where AVs make up a sizeable share of the overall traffic remain largely untested
  - shifts in average hourly speeds
  - dispersion in acceleration and deceleration rates
  - changes in vehicle queuing and speed shockwave patterns.

Background

- Before such changes in aggregate traffic flow patterns can be identified, however, a reassessment and consolidation of the existing body of knowledge on naturalistic speed data would significantly contribute to future efforts to identify potential shifts in traffic flow characteristics.
- This study, thus, seeks to identify the main features of aggregate speed data that are expected to vary under increasingly autonomous driving conditions.

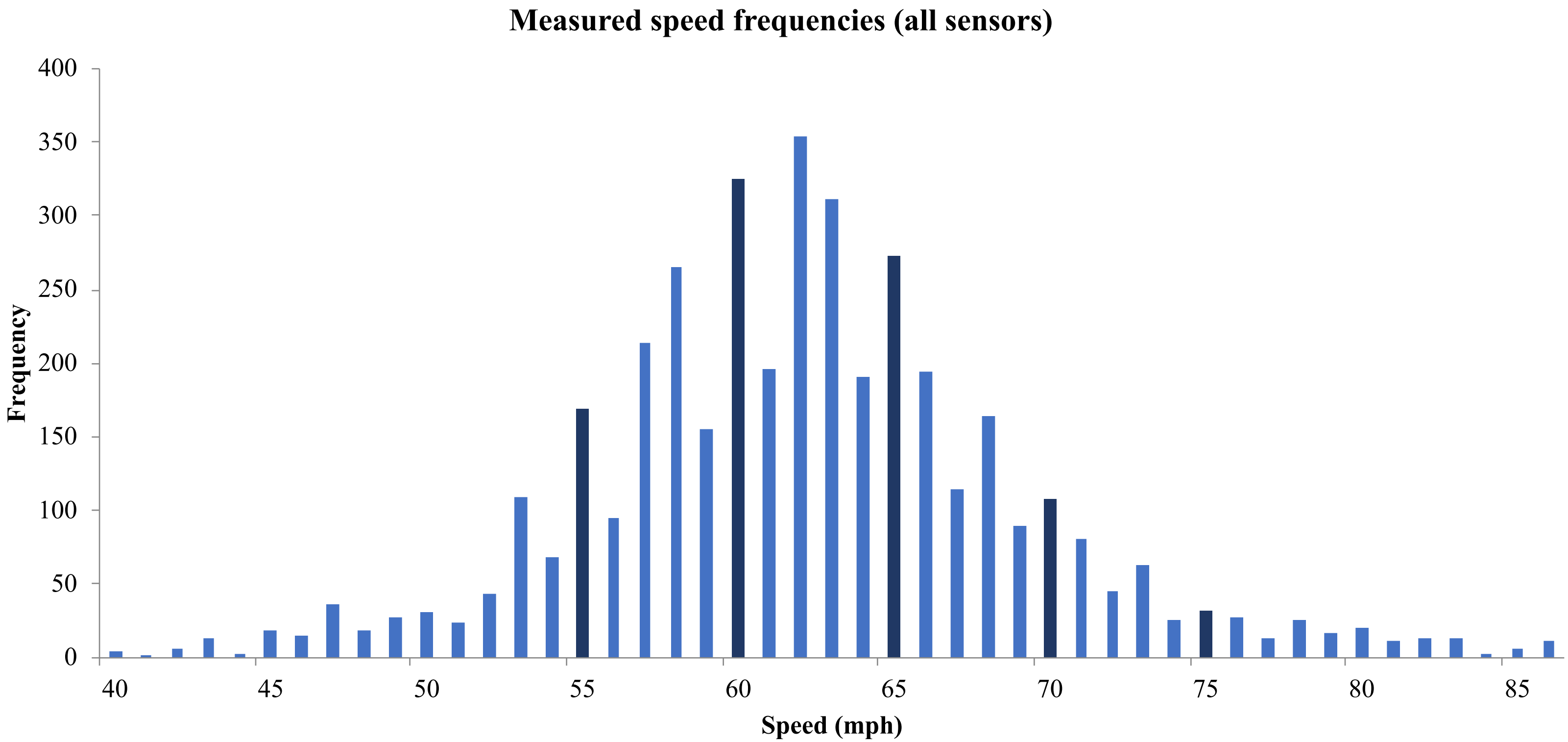
Data Summary

- The study uses a broad dataset collected by the Iowa Department of Transportation across the state.
- A subsample (those sensors with a posted speed limit of 60 mph) of the full dataset provides an illustration of the study's main results, which are representative of the findings from the full sample.
- The dataset sample includes 3,818 driving speed observations recorded by 31 speed sensors (Wavetronix SmartSensor™).
- Large and small-sized vehicles such as trucks and motorcycles were excluded from the sample.
- Only observations indicating free-flow traffic included.
    - Measurements were randomly taken in intervals of 20 seconds.
    - Data measured by sensors at exit and onramps were also excluded.

Summary Statistics

Sensor ID	No. of Obs.	Average	Std. Dev.	Min	Max
CRDS04-NB	78	74.2	6.5	59	84
CRDS04-SB	90	71.3	4.4	61	80
CRDS05-NB	104	73.5	6.9	60	95
CRDS05-SB	94	65.7	7.2	45	96
CRDS06-NB	103	69.5	5.7	58	84
CRDS06-SB	81	70.1	5.4	58	93
CRDS09-NB	44	61.8	3.3	52	69
CRDS09-SB	68	65.2	3.6	53	72
CRDS10-NB	225	62.4	4.7	48	75
CRDS10-SB	261	59.2	4.7	47	76
CRDS11-NB	242	64.0	4.8	50	80
CRDS11-SB	238	61.4	4.9	47	76
CRDS25-NB	205	60.1	4.2	50	77
CRDS25-SB	207	60.3	4.8	42	73
CRDS28-NB	187	62.5	4.2	51	75
CRDS28-SB	203	61.3	4.8	43	88
CRDS29-NB	270	59.1	5.2	42	83
CRDS29-SB	182	60.1	4.2	43	73
CRDS30-SB	203	62.6	4.8	48	77
CRDS32-NB	21	67.7	2.9	62	72
CRDS33-NB	31	65.1	4.3	54	69
CRDS33-SB	99	63.9	4.9	52	75
DMDS04-EB	26	73.7	7.8	50	86
DMDS04-WB	11	81.0	5.8	70	86
DMDS05-EB	67	45.8	2.0	38	50
DMDS05-WB	46	54.0	7.7	45	74
DMDS21-EB	159	59.1	5.7	47	79
DMDS21-WB	102	64.7	4.4	56	75
DMDS25-WB	45	62.2	6.8	51	79
DMDS26-EB	118	62.3	2.6	55	67
DMDS26-WB	8	43.9	5.5	39	54
Total	**3,818**	**62.53**	**6.87**	**38**	**96**

Spot speed measurements



- Common speed thresholds emerge as local modes consistently over a full range of posted highway speed limits

  (distribution shown for sensors with 60 mph posted speed limits)

Methodology

- In testing the multimodality hypothesis, the analysis that follows adopts a two-pronged approach.
  1. The first method uses a rank measure for each speed bin (in increments of 1 mph) given a sensor location.
     - The recorded driving speed frequencies are ranked by at each of the sensor locations.
     - The ranking of speed frequencies offers a proxy measure to examine the relative frequency of each speed bin across all locations.
     - The speed bin ranks for a given sensor *i* and a speed interval *j* are calculated by ranking the speed bins in a descending order by speed frequency counts.
     - The resulting rank ordering provides the proxy mode measure for each speed sensor.

Methodology

- **Hartigans' Dip Test**
  - Commonly used to rule out the unimodality of a sample density distribution provides further evidence for the significance of multimodality in the sample.

- Although other alternative statistical methods, such as the bimodality coefficient, can be used to detect multimodality, these methods lack the power necessary to detect multiple modes when the subpopulation means are relatively close.

Methodology

2. The second statistical method used is a finite mixture model (FMM), which provide the ability to test whether the observed density distributions can be decomposed into two separate speed distributions, each representing a different traffic regime at statistically significant levels.

Finite Mixture Models

- A finite mixture model is a probabilistic model that combines two or more density functions with weights estimated through an expectation-maximization procedure.
- The empirical observations $\boldsymbol{y}$ are assumed to be a composite of g subpopulations $f_1, f_2, \dots, f_g$ in proportions $\pi_1, \pi_2, \dots, \pi_g$. The probability density function can be expressed as in Eq. 1.

$$f(y) = \sum_{i=1}^{g} \pi_i f_i(\boldsymbol{y} \backslash \boldsymbol{x}' \boldsymbol{\beta}_i) \quad (1)$$

where π_i is the weight for the *i*th group and $f_i(.)$ is the conditional probability density function. The weights are subject to usual constraints: $0 \le \pi_1 \le 1$ and $\sum \pi_i = 1$.

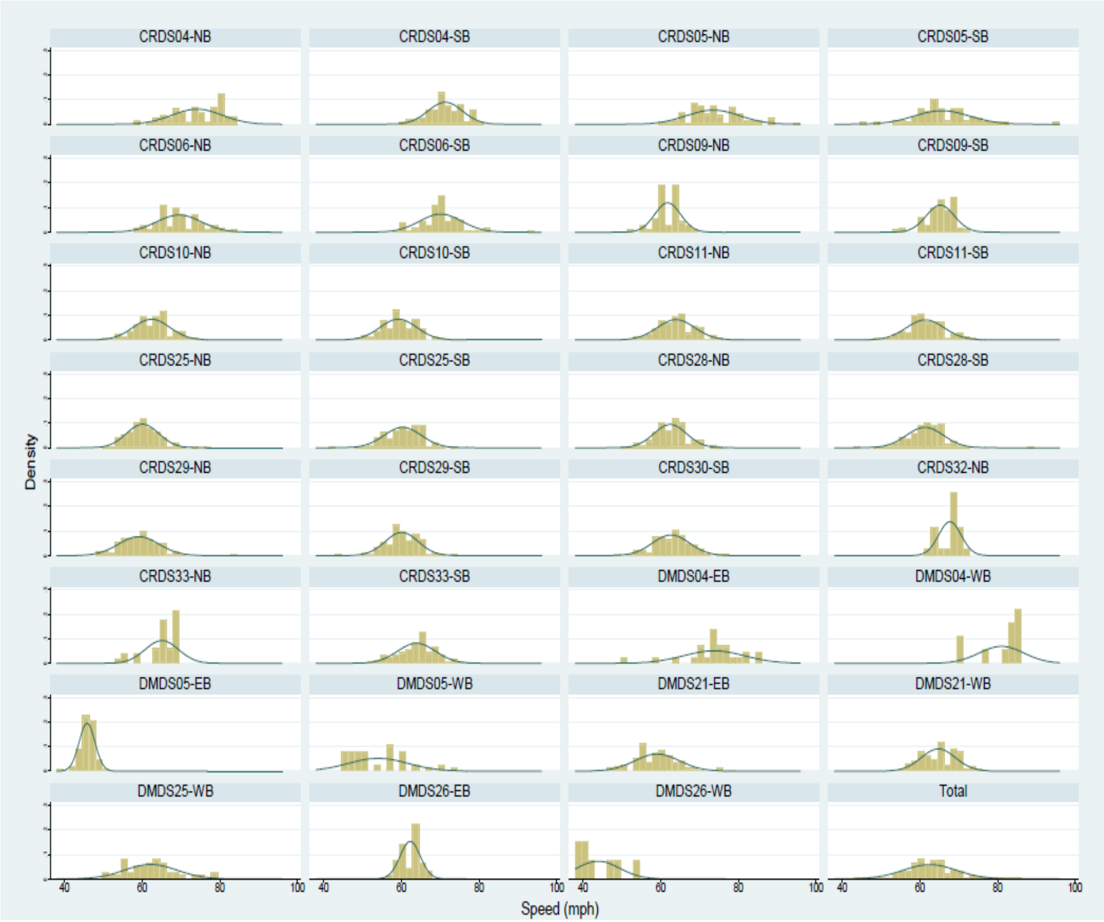
- The probabilities of each latent subpopulation are given by the multinomial logistic distribution, as shown in Eq. 2.

$$\pi_i = \frac{\exp(\gamma_i)}{\sum_{j=1}^{g} \exp(\gamma_j)} \quad (2)$$

where γ_i is an estimated value for the ith latent type through the FMM procedure.

- Since the first latent type is defined as the base level, $\gamma_i = 0$ and $\exp(\gamma_i) = 1$.

Speed Measurements by Sensor Location

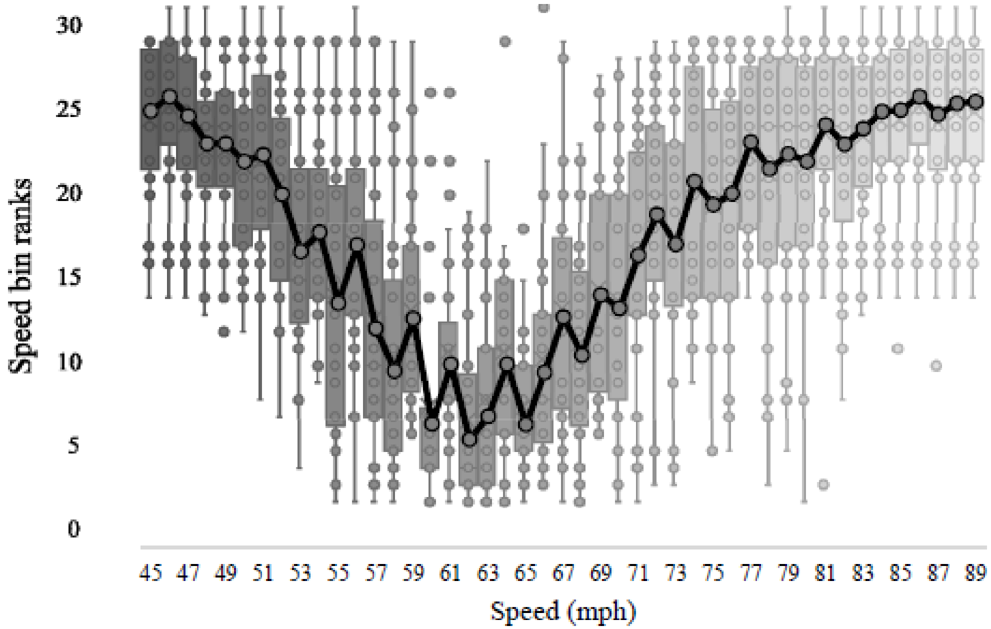


- The data show multimodality even as granularity increases when data are grouped by sensor locations.
- Almost all distributions offer examples of local clustering of speeds and sharp drop-offs following the usual speed thresholds, which further provide supporting evidence for the multimodal nature of driving speeds.

Clustering at Posted Speed Limits

- Given location and a posted speed limit, clustering behavior at *other* common posted speed limits suggests a “sticky” driving speed behavior
  - In free-flowing traffic, drivers may tend to maintain the same driving speed over long stretches of roadway.
  - This type of behavior can be observed either due to the ambiguity of the speed limits or the conscious driver decision to maintain a certain driving speed over long stretches a highway segment.
  - Many drivers may, for example, set the cruise control to an initial driver speed and choose to maintain the same speed even as the posted speed limits vary over the trip duration.

Box Plot of Speed Bin Rankings for all Sensors



- The dark line indicates the average rankings for each speed bin (45 mph to 89 mph).
- The average ranking data exhibits several local spikes, due to the clustering of travel speeds at common speed thresholds.
- Moreover, these deviations are statistically different from one another as indicated by the two-sample t-statistics reported in the following slide.

Two-sample *t*-tests for Speed Bin Ranks at Common Speed Thresholds

	55 mph	56 mph	60 mph	61 mph	65 mph	66 mph	70 mph	71 mph
Mean Ranking	11.57	13.95	6.92	8.97	6.73	8.68	10.97	13.16
Std. Dev.	7.44	7.21	6.12	5.46	4.64	6.29	6.89	7.30
Observations	37	37	37	37	37	37	37	37
Pearson Corr. Coeff.	0.81		0.68		0.34		0.71	
t-stat	-3.181		-2.671		-1.844		-2.459	
p-value (two-tail)	0.003***		0.011**		0.073*		0.019**	

*p<.05 ** p<.01 *** p<.001

Hartigans' Dip Test Results by Sensor

Sensor ID	No. of observations	Avg. Speed	Hartigans' Dip Statistic	p-value
CRDS04-NB	78	80.0	0.049	0.188
CRDS04-SB	90	71.0	0.061	0.013**
CRDS05-NB	104	68.0	0.053	0.028**
CRDS05-SB	94	63.0	0.053	0.046**
CRDS06-NB	103	68.0	0.058	0.009***
CRDS06-SB	81	70.5	0.074	0.001***
CRDS09-NB	44	64.0	0.114	0.000***
CRDS09-SB	68	67.0	0.066	0.024**
CRDS10-NB	225	62.5	0.058	0.000***
CRDS10-SB	261	58.0	0.059	0.000***
CRDS11-NB	242	62.0	0.046	0.000***
CRDS11-SB	238	60.0	0.059	0.000***
CRDS25-NB	205	60.0	0.066	0.000***
CRDS25-SB	207	63.0	0.058	0.000***
CRDS28-NB	187	63.0	0.059	0.000***
CRDS28-SB	203	62.0	0.062	0.000***
CRDS29-NB	270	57.0	0.057	0.000***
CRDS29-SB	182	59.0	0.060	0.000***
CRDS30-SB	203	63.0	0.052	0.000***
CRDS32-NB	21	68.0	0.095	0.102
CRDS33-NB	31	68.0	0.097	0.012**
CRDS33-SB	99	65.0	0.046	0.133
DMDS04-EB	26	74.0	0.058	0.731
DMDS04-WB	11	83.0	0.121	0.124
DMDS05-EB	67	47.0	0.112	0.000***
DMDS05-WB	46	47.9	0.087	0.006***
DMDS21-EB	159	56.0	0.054	0.002***
DMDS21-WB	102	65.5	0.059	0.008***
DMDS25-WB	45	64.0	0.056	0.358
DMDS26-EB	118	64.0	0.068	0.000**
DMDS26-WB	8	39.0	0.089	0.826

*p<.05 ** p<.01 *** p<.001

- 24 out of a total of 31 exhibit bimodality at the 5% confidence level, broadly supporting the multimodal nature of the driving speeds.

Hartigans' Dip Test Results by Sensor

Sensor ID	No. of observ.	Coef.	Std. Err	p-value	Class1 Wght. Est.	Class2 Wght. Est.	Class1 mean (mph)	Class1 Std. Dev.	Class2 mean (mph)	Class2 Std. Dev.
DMDS25-WB	45	-2.53	0.68	0.000***	0.93	0.07	60.99	28.22	77.51	3.61
DMDS26-EB	118	0.17	0.44	0.69	0.46	0.54	60.03	3.45	64.13	1.9
DMDS26-WB	8	-0.49	0.74	0.504	0.62	0.38	40.19	1.35	49.91	9.33
DMDS05-EB	67	-0.46	0.41	0.264	0.61	0.39	45.23	5.79	46.68	0.22
DMDS05-WB	46	0.15	0.43	0.727	0.46	0.54	47.83	4.78	59.36	41
DMDS04-EB	26	2.22	0.81	0.006***	0.1	0.9	56.08	28.89	75.65	24.72
DMDS04-WB	11	0.98	0.68	0.15	0.27	0.73	72.7	9.83	84.13	2.11
DMDS21-EB	159	-3.28	0.45	0.000***	0.96	0.04	58.48	23.16	75.69	3.44
DMDS21-WB	102	0.39	0.47	0.410***	0.4	0.6	60.5	5.28	67.61	8.32
CRDS04-NB	78	-0.49	0.3	0.104	0.62	0.38	70.24	24.77	80.53	2.41
CRDS04-SB	90	3.29	0.75	0.000***	0.04	0.96	61.83	0.19	71.63	16.72
CRDS05-NB	104	0.04	1.13	0.969	0.49	0.51	69.2	16.96	77.65	40.92
CRDS05-SB	94	-1.89	1.46	0.195	0.87	0.13	65.29	31.21	68.09	176.61
CRDS06-NB	103	-1.05	1.26	0.403	0.74	0.26	67.08	15.95	76.36	15.12
CRDS06-SB	81	-0.02	0.76	0.976	0.51	0.49	70.03	5.93	70.19	52.86
CRDS09-NB	44	2.7	0.82	0.001***	0.06	0.94	54.11	2.86	62.31	7.23
CRDS09-SB	68	0.89	1.06	0.404	0.29	0.71	61.9	17.97	66.61	4.47
CRDS10-NB	225	3.29	0.64	0.000***	0.04	0.96	52.02	3.48	62.77	18.88
CRDS10-SB	261	-1.98	5.35	0.712	0.88	0.12	58.86	18.2	61.79	42.64
CRDS11-NB	242	0.49	4.13	0.905	0.38	0.62	63.05	12.45	64.5	29.08
CRDS11-SB	238	0.73	0.68	0.283	0.33	0.67	60.24	6.1	61.92	31.29
CRDS25-NB	205	-4.21	0.72	0.000***	0.99	0.01	59.91	14.51	74.07	5.17
CRDS25-SB	207	-0.89	0.59	0.128	0.71	0.29	59.11	25.06	63.33	5.21
CRDS28-NB	187	0.68	1.14	0.55	0.34	0.66	62.3	6.46	62.53	23.4
CRDS28-SB	203	3.27	1.02	0.001***	0.04	0.96	60.53	185.18	61.31	16.73
CRDS29-NB	270	-3.08	0.87	0.000***	0.96	0.04	59.02	20.45	61.27	163.26
CRDS29-SB	182	3.11	2.3	0.175	0.04	0.96	57.16	65.32	60.19	15.32
CRDS30-SB	203	-0.24	1.31	0.852	0.56	0.44	61.97	30.61	63.49	11.18
CRDS32-NB	21	1.17	0.52	0.024**	0.24	0.76	63.4	0.67	69.05	2.77
CRDS33-NB	31	1.65	0.49	0.001***	0.16	0.84	56.2	2.56	66.77	2.95
CRDS33-SB	99	1.79	0.46	0.000***	0.14	0.86	55.96	3.84	65.22	14.45

- Each subsample is fit into two Gaussian distributions through an expectation-maximization algorithm.
- Observed speed densities decomposed into two types of distributions (referred to as classes in the table) in 13 sensor locations at the 5% significance level.

Conclusion

- Even though the AVs are not required to strictly follow the posted speed limits currently, the share of vehicles driving at constant speed can be expected to climb, which, in turn, can have a significant effect on the heterogeneity of existing traffic and speed patterns.
- Consequently, as the adoption of AVs and various cruise control technologies increase, there will likely be considerable changes in the density profiles of highway driving speeds.
- One such potential change is an increase in the bimodality of speed distributions since higher shares of autonomous and adaptive navigation algorithms can be expected to amplify the existing clustering of travel speeds.
- Monitoring the bimodality present in current speed distributions, thus, holds significant potential for building predictive models on the evolution of new traffic regimes, in which connected and autonomous vehicles will increasingly rely on various cruise control technologies.

Conclusion

- Different traffic regimes are at play and contribute to the empirical density functions, the findings suggest negative skewness and clustering at the posted and other common speed limit thresholds
- A framework to compare the basic speed characteristics of prevailing traffic flow under both urban and rural roadway conditions as the share of vehicles using autonomous driving technologies increases holds significant promise
- Illustrates that not only there seems to be a distinct pattern of speed restraining behavior but there are distinct traffic subpopulations centered on other common speed thresholds, putting in question whether the underlying traffic dynamics can be associated with dual cognitive processes.

Conclusion

- The analysis demonstrates the multimodality of driving speeds due to the clustering of data at several speed limit values (e.g., 55, 60, 65 and 70 mph), despite the uniqueness of the posted speed limit given a highway segment.
- This observation suggests that drivers' speed choices, which are commonly assumed to be largely a random variable process that centered on free-flow highway speeds, may be more complex and multi-faceted than previously recognized.
- Illustrates that not only there seems to be a distinct pattern of speed restraining behavior but there are distinct traffic subpopulations centered on other common speed thresholds, putting in question whether the underlying traffic dynamics can be associated with dual cognitive processes.
- Thus, different traffic regimes may exhibit bias against changing speeds in free-flowing traffic conditions presumably due to ingrained driving patterns or temporal lags in converging to posted speed limits.

Conclusion

- The study demonstrates the multimodality of driving speeds due to the clustering of data at several speed limit values (e.g., 55, 60, 65 and 70 mph), despite the uniqueness of the posted speed limit given a highway segment.
- This observation suggests that drivers' speed choices, which are commonly assumed to be largely a random variable process that centered on free-flow highway speeds, may be more complex and multi-faceted than previously recognized.
- There seems to be not only a distinct pattern of speed restraining behavior but distinct traffic subpopulations centered on other common speed thresholds, putting in question whether the underlying traffic dynamics can be associated with dual cognitive processes.
- Thus, different traffic regimes may exhibit bias against changing speeds in free-flowing traffic conditions presumably due to ingrained driving patterns or temporal lags in converging to posted speed limits.

Upcoming work

- A number of traffic flow attributes remain unexplored
  - average vehicle headway, average travelling speed,
  - mean deceleration and acceleration rates
  - Ultimately would aid agencies in calibrating speed and traffic data collection.
  - DOTs could benefit from an active approach that positions their data measurement programs to capture such anticipated shifts in aggregate traffic speed profiles.

Implications for Future Research

- The decomposition of aggregate speed distributions further raises interesting research questions, especially in relating the observed dual traffic regimes to the long and established body of research in cognitive psychology.
- Given the observed multimodality of driving speed profiles under free-flowing traffic conditions, to what extent are speed choices driven by dual cognitive processes?
- Despite the relatively limited attention dedicated to multimodality in highway speed profiles, elsewhere in applied sciences, interest in detecting multimodality has been growing.
- Briefly, the dual-process accounts demonstrate that human behavior is largely driven by two parallel cognitive processes: a relatively fast, automatic and nonconscious process, and another that is relatively slow, deliberate and conscious
    1. A deliberate and slow decision-making processes, which tend to draw upon attention and memory resources of the brain, could be associated with a subpopulation of drivers who tend to pay closer attention to speed limits and restrict driving speed accordingly.
    2. A nonconscious and fast cognitive mechanism could be related to driver behavior with a general lack of unawareness or indifference to posted speed limits.

Implications for Future Research

- The insights gained from distribution analyses of speed profiles will undoubtedly be critical in aiding policy decisions to review and maintain speed limits relevant as the share of autonomous, connected/automated and other innovative navigation technologies arise.
- The set of results presented here provides an example of how transportation agencies can leverage spot speed measurements both to gain a deeper understanding of and to monitor the ongoing evolution of the underlying traffic regimes that make up the observed speed density functions.
- Transportation agencies, in particular, could benefit from establishing baseline speed profiles that position their data measurement programs to capture such anticipated shifts in aggregate traffic characteristics.

The author would like to acknowledge the generous assistance provided by the Iowa Department of Transportation for the collection and analysis of the sample dataset used in this study.