HOW MUCH DRIVING DATA DO WE NEED TO ASSESS DRIVER BEHAVIOR?

Extended Abstract

Anna-Maria Stavrakaki*
Civil & Transportation Engineer
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1367; Fax: +30 210 772 1454
e-mail: annamariastavrakaki@gmail.com

Dimitrios I. Tselentis, Ph.D. (ORCID: 0000-0001-5939-0136)
Research Associate
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 2210; Fax: +30 210 772 1454
e-mail: dtsel@central.ntua.gr

Emmanouil N. Barmpounakis, Ph.D. (ORCID: 0000-0001-5455-6643)
Research Associate
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1367; Fax: +30 210 772 1454
e-mail: manosbar@central.ntua.gr

Eleni I. Vlahogianni, Ph.D. (ORCID: 0000-0002-2423-5475)
Assistant Professor
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1369; Fax: +30 210 772 1454
e-mail: elenivl@central.ntua.gr

George Yannis, Ph.D. (ORCID: 0000-0002-2196-2335)
Professor
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 1326; Fax: +30 210 772 1454
e-mail: geyannis@central.ntua.gr

Word count: 1646 words text + 2 tables + 2 figures

ACKNOWLEDGEMENT

The authors would like to thank OSeven Telematics, for providing all necessary data to accomplish this study.

Submitted for presentation to the 98th Annual Meeting of the Transportation Research Board

*Corresponding author

INTRODUCTION

The aim of this study is to provide a methodological framework for estimating the amount of driving data that should be collected for each driver in order to acquire a clear picture regarding driving behavior. This amount is defined as the total driving duration and/or the number of trips that need to be recorded for each driver in order to draw a solid conclusion regarding where the rate of driving behavioral characteristics (e.g., per kilometer or per minute) has converged to a fixed point.

Several studies have taken advantage of new technologies such as In-Vehicle Data Recorders and smartphones for the evaluation of driving behavior *(1-5)*. However, the exact amount of driving data that need to be collected and evaluated to assess driving behavior with sufficient precision has not yet been determined. Both small and large data samples are likely to lead to questionable results by acquiring a sample either biased or computationally expensive to analyze, and thus, it is important to investigate the amount of driving data that should be recorded by each participant in the experiment.

In this study, the driving metrics used to identify driving behavior stabilization are the number of harsh acceleration and braking events, the time of mobile phone usage and the time driving above the speed limit (speeding), which are the main human factors used in literature as well *(6-9)*. Through the exploitation of cumulative sums, moving averages and Shewhart control charts, the driver's aggression and volatility is estimated, and consequently, the time point at which driving behavior converges is determined. The analysis indicated that for a certain driving characteristic, convergence depends largely on the aggressiveness and stability of the overall driver's behavior as well as the average duration of the trips being studied. The results of the analysis performed could be exploited both by the private and public sector, providing multiple social and economic benefits.

METHODOLOGY

The basis of this framework is an innovative data collection system developed by OSeven Telematics that continually records real-time driving behavior data of each participant using smartphone sensors *(10)*. Driving behavior is monitored and analyzed to determine the minimum observation time for each driver and the potential to group drivers based on their driving aggressiveness. The database used consisted of 21,610 trips that took place by 68 drivers, which were chronologically classified to observe the change in the magnitude of driving behavior characteristics over time. All data were anonymized before provided by OSeven Telematics and therefore there is no personal information for the driving sample studied.

As stated in the introduction, the driving metrics used to identify driving behavior convergence are the number of harsh acceleration and braking events, the time of mobile phone usage and the time of speeding. Cumulative sums of those metrics (per kilometer for harsh events and as percentage of driving duration for mobile phone usage and speeding) are used to reveal the time point at which driving characteristics stabilize or fluctuate around a fixed value over time. This trend is also captured in FIGURE 1 provided below.

The analysis is conducted on a trip basis, and three distinct trip duration categories are used (5, 10 and 20 minute trips). The variability of the above metrics is then examined to observe driving behavior evolution over time. For this purpose, the measures of simple moving average and volatility are used along with statistical limits (Shewhart charts) and conditions that need to be met, to identify convergence. For each of the sub-databases originating from the initial database of the 68 drivers, it is examined whether and when all of the following conditions are met simultaneously:

- The moving average is within the range Mean ± 1*Standard Deviation.

- For five consecutive trips the percent change (in absolute terms) between successive values of the moving average is less than or equal to 1.5%.
- The value of the moving average in the corresponding trip is a local extreme (this criterion ensures that the neighboring values of the moving average are smaller or larger than the selected one, and therefore it does not belong to a sequence of points that have a particular trend e.g. ascending or descending).

These criteria are separately applied on the cumulative sum measures and to their volatility measures. For each driver, each time step is iteratively examined to ensure when the above criteria are met. The first trip, for which all of the above conditions are met, is assigned to the drivers' database as the first time point at which the particular driving attribute stabilizes. At the same time, the values at which the cumulative sum metrics and their volatility converge, are also recorded.

FIGURE 1 is indicatively provided to illustrate the convergence point of a random driver regarding the harsh acceleration events for average trip duration of 5 minutes. The temporal change in the driving characteristic (HA) and its volatility as well as the time points at which driving behavior is converged can be noticed (at 44th and 169th trip respectively).

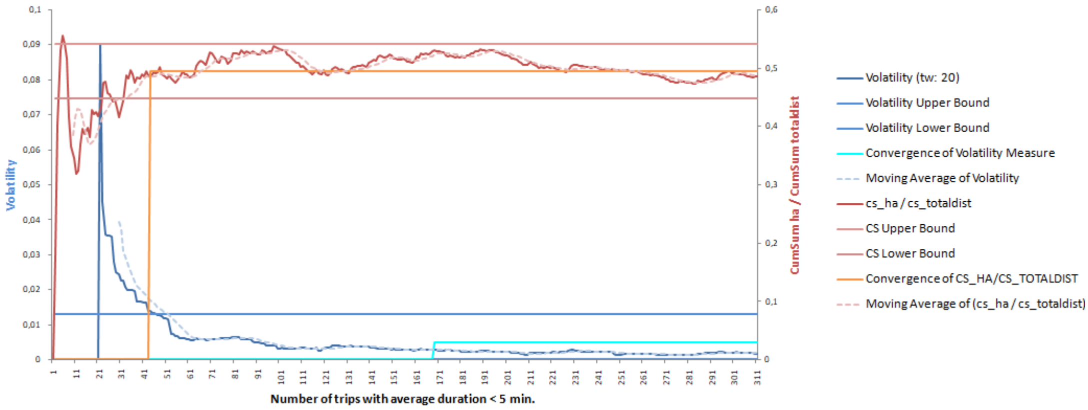

FIGURE 1 Convergence Plot of the Harsh Acceleration Events Rate for Randomly Chosen Driver

FINDINGS

The procedure described above is applied to the initial database of 68 drivers, separately for trips with an average duration of 5 minutes, 10 minutes and 20 minutes. It is not applied to trips with duration over 25 minutes though since the number of these trips is significantly lower, resulting to a very low number of trips for all drivers. Therefore, no duration category above 20 minutes is analyzed since this would probably lead to statistically insignificant and uncertain results. The final analysis performed included data from 29 drivers who were used to obtain the results illustrated in TABLE 1 Aggregated Table.

TABLE 1 Aggregated Table presents the descriptive statistics of the minimum number of trips after which convergence of the driving behavior metrics is reached for the above 29 drivers. Results of TABLE 1 Aggregated Tableare grouped by trip duration category and drivers' aggressiveness level i.e. the number of harsh acceleration (HA) / braking (HB) events per 100 km driven, the percentage of mobile usage (MU) and the duration of speeding (SP) while driving.

As observed in TABLE 1 Aggregated Table, the time point at which driving behavior stabilizes is not common for all drivers and/or all driving behavior metrics. This is an expected finding since driving aggression differs and therefore the analysis of the driving aggression profile should be preceded. Furthermore, the analysis indicated that the most aggressive drivers tend to converge at a faster rate than the more cautious drivers, confirming the results of the literature *(11)*. On average, more cautious drivers tend to converge (for all driving metrics and their volatility) at around 95 trips, while more aggressive drivers at around 80 and 65 trips for driving metrics and their volatility respectively.

TABLE 1 Aggregated Table of Minimum Number of trips required for Convergence

Trip duration	Metric limits	Metric					Volatility					No of drivers
		min	max	Average	Median	StDev	min	max	Average	Median	StDev	
5	**HA ≤ 15**	63	112	92	92	17	49	169	95	81	35	27
	HA > 15	52	136	86	85	27	36	97	65	70	19	
	HB ≤ 5	60	196	110	109	44	50	271	85	70	58	
	HB > 5	56	157	97	94	31	52	103	81	85	19	
	MU ≤ 10%	76	167	102	94	26	43	112	76	75	18	
	MU > 10%	52	104	76	73	17	38	187	73	67	38	
	SP ≤ 3,5%	69	145	104	104	29	41	157	79	70	34	
	SP > 3,5%	64	138	86	76	23	34	172	65	50	38	
10	**HA ≤ 15**	58	109	84	84	14	74	235	115	103	40	29
	HA > 15	49	134	80	75	26	43	119	67	62	22	
	HB ≤ 6	71	213	118	97	50	62	251	102	90	47	
	HB > 6	65	135	90	77	22	41	96	69	66	18	
	MU ≤ 7%	41	291	110	98	61	58	203	86	79	35	
	MU > 7%	67	134	95	87	21	46	105	64	63	16	
	SP ≤ 5%	18	154	89	88	32	62	201	99	83	46	
	SP > 5%	53	123	85	85	23	41	99	68	71	19	
20	**HA ≤ 12**	14	103	61	69	35	61	188	117	102	44	16
	HA > 12	29	81	59	63	17	42	50	46	46	6	
	HB ≤ 5	84	102	94	97	9	60	184	102	87	40	
	HB > 5	51	109	69	65	17	-	-	-	-	-	
	MU ≤ 10%	72	156	106	96	31	34	118	73	65	30	
	MU > 10%	58	103	80	80	19	38	116	65	41	44	
	SP ≤ 10%	56	126	87	88	27	40	166	85	83	40	
	SP > 10%	36	106	71	74	26	46	52	49	49	4	

Apart from the aggression, the number of trips a driver is required to be monitored also varies in terms of the average duration of the trips being studied. For example, it is clear from TABLE 1 Aggregated Tablethat the minimum number of trips required for convergence is generally smaller for trips with average duration of 20 minutes than the corresponding one for shorter trips (e.g. 5 or 10 minute-trips). FIGURE 2 is provided to better illustrate this finding.

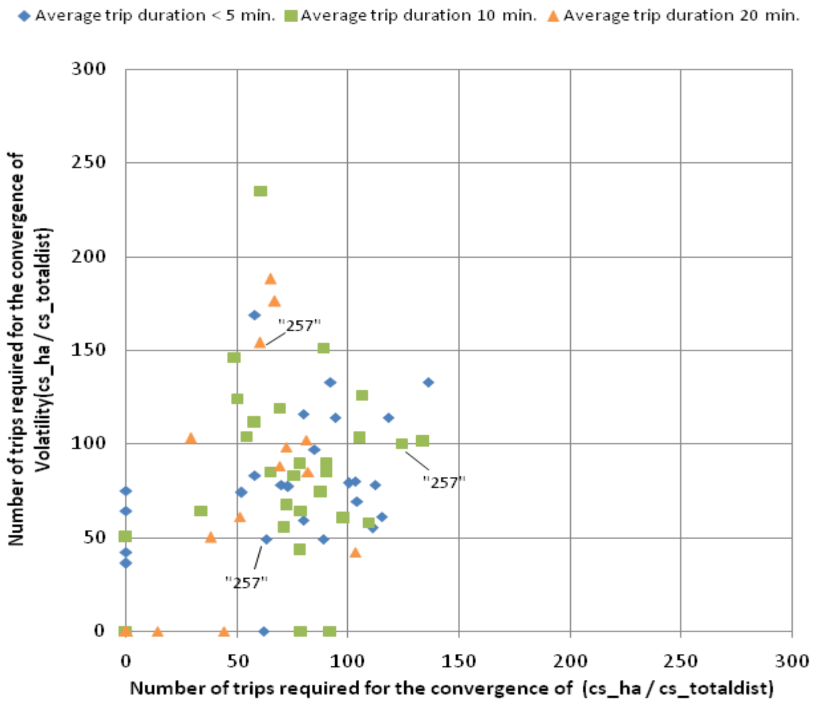

FIGURE 2 Minimum Number of Trips Required for the number of Harsh Acceleration Events per Km Rate to Converge

The driving behavior metric that converges later for each driver is the critical driving characteristic that determines the minimum number of trips that need to be collected to obtain a clear picture for his driving behavior. For the majority of drivers (~ 37%) the critical characteristic is the volatility of the number of harsh acceleration events per km as well as the percentage of time of mobile phone usage while driving (~ 34%).

TABLE 2 summarizes the results of the analysis performed on the convergence rates of the four driving metrics examined, which are categorized as fast or slow based on the minimum number of trips required to be collected. It also illustrates the aggressiveness and volatility limits noticed in each convergence rate group.

TABLE 2 Aggressiveness, Volatility Limits and Convergence Rate of Driving Behavioral Characteristics

	Minimum Required Number of trips		Average Conversion Rate of Driving Characteristics and Volatility			
	Fast Convergence	**Slow Convergence**	**Cautious**	**Aggressive**	**Stable**	**Volatile**
Harsh Acceleration events per km	< 50 (24.14%)	> 120 (10.34%)	< 0.11 (33.33%)	> 0.23 (17.24%)	-	-
Harsh Braking events per km	< 60 (13.79%)	> 140 (20.69%)	< 0.01 (5.75%)	> 0.12 (9.20%)	-	-
Percentage (%) of Time Mobile Usage	< 50 (17.24%)	> 120 (27.59%)	< 0.04 (32.18%)	> 0.16 (21.84%)	-	-
Percentage (%) of time Speeding	< 50 (24.14%)	> 120 (24.14%)	< 0.02 (12.64%)	> 0.14 (9.20%)	-	-
Volatility	< 60 (42.24%)	> 120 (21.55%)	-	-	< 0.005 (35.63%)	> 0.05 (23.75%)

The aggressiveness and volatility of drivers are determined from the average values at which driving behavior rates and their volatility converge. For example, if a driver's average number of harsh acceleration events/volatility is high, the driver is considered aggressive/volatile respectively. A driver may be cautious regarding the feature being studied, but at the same time exhibiting significant variations/fluctuations in his travel-related, and vice versa.

CONCLUSIONS

Data analysis indicated that for a certain driving characteristic, the amount of data required to be collected depends largely on the aggressiveness and stability of the overall driver's behavior as well as the average duration of the trips being studied. Particularly, more aggressive drivers require less monitoring than cautious drivers do. It is inferred that further investigation of the aggression level of drivers and the driving environment should be preceded. Aggressive drivers are those with a high number of harsh events and high percentages of time driving over the speed limit.

Apart from aggression, another driving characteristic that influences the time of convergence is the stability or volatility of driving behavior. Knowledge of drivers' behavioral volatility is of paramount importance when studying driving behavior as it provides important insights into their overall experience and the difference in behavior between trips.

The duration of the trips analyzed is also found to affect the point of convergence of a driver's behavior. It is particularly shown that the same driver may exhibit significant differences in the amount of data required to be collected with respect to a particular driving characteristic when considering trips of different average driving duration.

The results of the analysis performed could be exploited either for providing feedback to drivers on how to improve their driving behavior or to improve the services provided by usage-based insurance companies and car industries. Such schemes would bring multiple and significant benefits to the society, since the overall driving behavior of the population would be improved, leading to long-term accident reduction and even improvement of environmental conditions through limiting fuel consumption and emissions to the environment.

REFERENCES

1. Vlahogianni, E. I., and E. N. Barmpounakis. Driving Analytics Using Smartphones: Algorithms, Comparisons and Challenges. Transportation Research Part C: Emerging Technologies, Vol. 79, 2017, pp. 196–206. https://doi.org/10.1016/j.trc.2017.03.014.
2. Jensen, M., J. Wagner, and K. Alexander. Analysis of in-Vehicle Driver Behaviour Data for Improved Safety. International Journal of Vehicle Safety, Vol. 5, No. 3, 2011, p. 197. https://doi.org/10.1504/IJVS.2011.042850.
3. Händel, P., J. Ohlsson, and M. Ohlsson. Smartphone Based Measurement Systems for Road Vehicle Traffic Monitoring and Usage Based Insurance. IEEE Systems Journal, 2013. https://doi.org/10.1109/JSYST.2013.2292721.
4. Boquete, L., J. M. Rodriguez-Ascariz, R. Barea, J. Cantos, J. M. Miguel-Jimenez, and S. Ortega. Data Acquisition, Analysis and Transmission Platform for a Pay-As-You-Drive System. Sensors, 2010. https://doi.org/10.3390/s100605395.
5. Iqbal, M., and S. Lim. Privacy Implications of Automated GPS Tracking and Profiling. 2010.
6. Haque, M. M., and S. Washington. The Impact of Mobile Phone Distraction on the Braking Behaviour of Young Drivers: A Hazard-Based Duration Model. Transportation Research Part C: Emerging Technologies, 2015. https://doi.org/10.1016/j.trc.2014.07.011.
7. Saifuzzaman, M., M. M. Haque, Z. Zheng, and S. Washington. Impact of Mobile Phone Use on Car-Following Behaviour of Young Drivers. Accident Analysis and Prevention, 2015. https://doi.org/10.1016/j.aap.2015.05.001.
8. Speed management. Paris, OECD/ECMT Transport Research Centre (JTRC), 2006
9. Bonsall, P., R. Liu, and W. Young. Modelling Safety-Related Driving Behaviour - Impact of Parameter Values. Transportation Research Part A: Policy and Practice, 2005. https://doi.org/10.1016/j.tra.2005.02.002.
10. https://www.oseven.io/
11. Mantouka, E. G., E. N. Barmpounakis, and E. I. Vlahogianni. Mobile Sensing and Machine Learning for Identifying Driving Safety Profiles. 2018.