Statistical Analysis on Real Kinematic Urban Driving Cycles by preliminary use of VSP variable

LIVIA DELLA RAGIONE, GIOVANNI MECCARIELLO
Istituto Motori - CNR National Research Council,
Via G. Marconi, 80125 Napoles (NA), ITALY

Abstract. Still today true level emissions assess coming from transport activities is a key element in the evaluation of any transport policy both in Europe and the world context. In the next years, further reducing vehicle emissions will be one of the most important development of sustainable mobility. Increasingly stringent emission regulations require continuous optimization that must be validated in new test cycles. Therefore, several aspects related to the assessment of vehicle pollutant emissions are still the subject of research studies. some preliminary considerations will be developed trying to evaluate how it is possible to integrate the Vehicle Specific Power (VSP) in a statistical approach of vehicle kinematic. This kind of activity globally aims at valuation of traffic situation and emissions in real use. Driving data and emissions are acquired during an experimental campaign through four instrumented vehicles by a portable emissions measurement system (PEMS) for the simultaneous acquisition of emissions, kinematic variables and global positioning system (GPS) data. Statistical analysis results are relative to Principal Component Analysis (PCA) and Cluster Analysis (CA) performed on kinematic data to better characterize road tests and driving behaviour. In this paper we study the way to integrate VSP index in a statistical approach, since this index is widely recognized in the literature to synthesize the contribution of kinematics along with road gradient.

Keywords: Vehicle Specific Power, Driving behaviour analysis,Statistical Kinematic analysis

1 Introduction

The environmental impact assessment have to be treated from many angles, often integrated with one another and it appears necessary to reach a synergy of activities and research groups with different expertise in the field of pollution evaluation and impact. For example, the effects on human health, on cultural heritage and on social and economic aspects must be identified and assessed. Each of these different angles is assessed by defining specific impact indicators.

Specifically, in the context of traffic planning and mobility support to select and implement intervention policies, emission modelling is a key step in environmental impact assessment. Today most interventions applied by decision-makers are based both on traffic restriction in central areas and at peak times, and on the payment of tolls for access to urban centres. A Decision Support System (DSS) approach could also be useful both to set up a scheme of the subsequent steps required to solve problems, and consider the decision-makers' points of view. In a DSS approach, a knowledge base for modelling and predicting vehicle real-world emissions is function of driving behaviour kinematics [1]. Moreover, an emission model must be implemented in order to calculate emission factors based on real driving cycles that, in turn, represent different traffic levels. However, other interventions could be based on real-time evaluation of pollutant levels by establishing pollution data from real acquisitions to indicate alternative paths. In this last context, several studies are needed to take into account critical operating conditions affecting the real pollutant emissions and fuel consumption of in-use vehicles. More recently, in-use emissions testing with PEMS, became one of the solutions proposed to ensure that emissions are well controlled in real use of passenger cars. In particular, studies based on portable emissions measurement systems (PEMS) have shown that average on-road emission factors (EFs) have been estimated to be about 6−7 times higher than the regulated Euro 6 limit , [2-3-4]. Moreover, a new type approval procedure is going to be introduced by the European Commission, to take into account critical operating conditions affecting the real pollutant emissions and fuel consumption of in-use vehicles and not considered during the laboratory tests. The new procedure involves the monitoring of real driving emissions (RDE) on the road with PEMS. In addition, it is fundamental to keep in mind that nowadays, the new WLTP and WLTC cycle reproduce a situation closer to the reality with respect to the EUDC/NEDC driving cycle [5, 6]. Emission measurements are performed on the test bench and with (PEMS) [7, 8]. During real driving over pre-selected routes, results show that there can be substantial differences for some pollutants measured as 'real driving emissions' (RDE) using PEMS equipment, compared to the test bench cycles in the laboratory.

Furthermore, in literature there are discussion about some criticism of WLTC cycles. In particular, although an improvement over the NEDC, the WLTC cycles are still unrealistically slow in some acceleration phases. For example, the most rapid 0-30 mph (50 km/h) time is 15 seconds and this could be considered slow. Most drivers in western Europe will accelerate from rest to 30 mph in 5 to 10 seconds. There is also no hill climbing in the cycle. Modest gradients will increase engine loads by 2 to 3 times with subsequent increase in pollutants [9]. Results from this study can be considered indicative of emission patterns of modern technology vehicles. [10, 11]. During the development of ARTEMIS project proposed, in Istituto Motori, a new statistical approach capable to consider more attributes than the simple speed to characterize driving behavior, not only in the determination of driving cycles but also in the emission modeling [12, 13]. Results show that if we consider a specific road, kinematic performances change and consequently driving cycles. In this paper, some preliminary considerations will be developed trying to evaluate how it is possible to integrate the Vehicle Specific Power (VSP) in a statistical approach of vehicle kinematic. This kind of activity globally aims at valuation of traffic situation and emissions in real use. So globally we study the way to integrate VSP index in a statistical approach, since this index is widely recognized in the literature to synthesize the contribution of kinematics along with road gradient [14, 15].

2 Experimental Activity

An experimental activity was carried out with four gasoline vehicles from different manufacturers and with a wide variety in terms of mass, power, engine displacement and type approval technology.

Table 1. Vehicles characteristics.

Vehicles	Euro	Engine displacement (cm^3)
Fiat Panda Bifuel	4	1200
Fiat Panda Twin Air	4	900
Fiat 500	5	900
Fiat 500L	6	1400

The experimental campaign and data acquisition during the execution of the on-road tests are carried out in two areas different for the topography of the streets and for different road gradient.
The first pattern, named Path A, is predominantly flat and a length of about 22 km. The round trip Path A was divided in two sub-paths named respectively Path Ao (green color) from Istituto Motori to the Central Station Garibaldi and Path Ar (blue color), return path from Central Station Garibaldi to Istituto Motori. Both routes are situated in the center of Napoli city, making them greatly influenced by road traffic. The second, named Path B, on the hilly area of Naples with varied terrain and sudden changes of slope of about 6 km, includes positive (+2.9%) and negative (-3.6%) road gradient, as shown in figure 1. Exhaust emissions of carbon monoxide (CO), total hydrocarbons (THC), (NOx), CO2 were acquired on-road by using a PEMS connected also to the Engine Control Unit for saving the main engine parameters and to the GPS for the geographical coordinates and altitude.

Fig. 1. Experimental path: Path Ao: Outward (Istituto Motori-Central Station Garibaldi) and return Path Ar (Central Station Garibaldi - Istituto Motori); Path B: the hilly area of Naples.

3 Experimental setup

Vehicles have been equipped for on-road tests, as shown in the following figure 2. The main components of the monitoring system used in this work are:

- A PEMS Semtech-DS gas analyser produced by Sensor (RMC), in 2006, to measure at 1Hz CO, NOx and CO_2 emissions. This analyzer uses NDIR cell (Non-Dispersive Infrared) for CO and CO_2 measurements, NDUV cell for nitric oxide (NO)/nitrogen dioxide (NO_2) and a separate electrochemical sensor for oxygen. The analyzer is calibrated on a regular basis and zeroes itself on start-up using outside air.
- an EFM (Exhaust Flow Meter) by Sensor.
- an On-Board Diagnostics (OBD) interface and logging computer running proprietary software Engine Data Scan (EDS) to acquire engine operating parameters (speed, rpm, engine air flow…).
- a GPS receiver by Racelogic Ltd to acquire the spatial position.
- a video camera to record traffic situations.

The signals from all devices have been synchronized by using the same information obtained from different sources (i.e. speed from GPS and OBD). The emission measurements acquired by PEMS and the kinematics and GPS data are filtered, synchronized and analyzed using statistical methods.

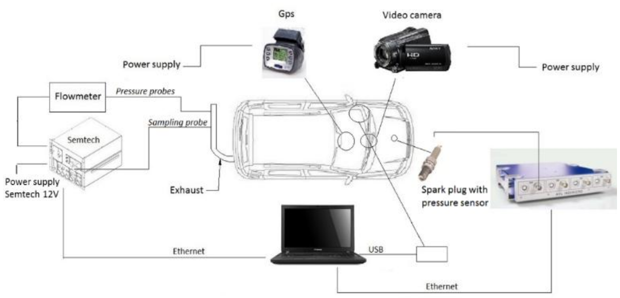

Fig. 2. On-board installation.

3 Driving behaviour analysis

The approach proposed in this paper for the determination of driving cycles is based on the analysis of road data by statistical methods to determine typical and statistically representative groups of driving cycles, engine operating conditions and emission correlated profiles. Driving behaviors are relative to different road networks, traffic conditions and specific features of each geographical area, with vehicles of different segment and technology. They constitute global sets from which it is possible to choose representative pieces to build a driving cycle going in the same direction of WLTP framework.

Following the general approach, the OBD speed profile is segmented in a succession of sequences, so that a sequence is the part of the motion of a vehicle between two successive stops. Therefore, the velocity profile of a trip was divided into blocks of sequences.

Driving cycles have been determined without any conditioning of data with respect to the road network, but keeping the information detected on road, in terms of GPS coordinates latitude, longitude and altitude. The method utilized to determine the driving cycles is based on sequence characterization. To characterize sequences pattern, the identified variables are partially related to the dynamic vehicle equation, plus idling time to consider standstill phase emission production and partially to slope variability. These variables are identified by considering emission variation as explained by the variation in exhaust mass (a function of energy spent by the vehicle in a driving cycle), and the frequency of acceleration events at different speeds (1). Moreover, the variables were identified considering two potential causes of variability in emissions for a driving cycle: energy expenditure from the vehicle in the cycle and the acceleration events at different speeds [1, 10].

$$M \approx \int P(t)dt \approx \int_t (a_0 + a_1 {*} v(t) + a_2 {*} v^2(t) + M_{va} {*} (a(t)) {*} v(t)dt \qquad (1)$$

where M is the total mass of exhaust, P(t) the engine power, v(t) the speed, a(t) the acceleration and Mva the vehicle effective mass.

Thus, variables describing each sequence are relative to speed and acceleration attributes, sequence time duration and length, idling and driving time, and relative frequency distribution. In table 2, variables characterizing driving behavior are reported.

Table 2. Variables characterizing kinematic DC.

Variable	Description
mv (km/h)	Mean of running speed (v>0)]
mv2(km^2/h^2)	Mean of square speed (v>0)
mv3(km^3/h^3)	Mean of cube speed (v>0)
Tral (s)	idling time v=0 in second
Trunning (s)	total running time (v>0) in second
Dist (m)	distance covered
Time(s)	Total duration of the sequence (s)
m_pos_acc (m^2/s^3)	Mean of instantaneous values of product (a(t)•v(t)) when v(t)>0 and a(t)>0
V20 (%)	%time speed <20
V30 (%)	%time 20<speed<30
V40 (%)	%time 30<speed<40
V60 (%)	%time 40<speed<60
V100 (%	%time speed>60
Paccel1 (%)	%time with acceleration in range [-∞;-1.4] m/s^2
Paccel2 (%)	%time with acceleration in range [-1.4;-0.6] m/s^2
Paccel3 (%)	%time with acceleration in range [-0.6;-0.2] m/s^2
Paccel4 (%)	%time with acceleration in range [-0.2;+0.2] m/s^2
Paccel5 (%)	%time with acceleration in range [+0.2;+0.6] m/s^2
Paccel6 (%)	%time with acceleration in range [+0.6;+1.4] m/s^2
Paccel7 (%)	%time with acceleration in range [+1.4;+∞] m/s^2

Consequently, observations (sequences) must be analyzed utilizing multivariate statistical methods. Since a sequence is represented by a considerable number of variables, such as distance, mean speed, idling time and running time, and these are mutually correlated, a Principal Component Analysis (PCA) is performed. Principal components (PC) are latent variables function of variables of Table 2, calculated by the matrix of observed X(*i, j*), where *i*=1, *k* and *j*=1, *NS*, *NS* being the total number of sequences. Each PC tends to characterize different typical features of sequences by a group of correlated variables, separating for example sequences with high mean speed and long running time and low acceleration, from sequences with low mean speed, high idling time and acceleration. Observed sequences are classified into homogeneous groups (clusters) by applying a clustering method, utilizing principal components calculated for each sequence, as variables characterizing the sequence. Kinematic sequences belonging to a cluster have similar patterns but display within-cluster variability, which is evaluated by statistical criteria. Moreover, to determine the sequence pattern most representing of each cluster, a multidimensional normal distribution is fitted to sequence PC data and its density function is estimated.

Sequences are ranked by density: those closest to the maximum density (the mode of distribution) are taken as the most representative. Then a Canonical Discriminant Analysis (CDA) is applied to outline features and reciprocal differences of clusters. CDA is used to determine which variables discriminate between clusters (groups) of multivariate observations. Clustering of sequences, by multivariate statistical analysis, gives the basic information to automatically split driving cycles from the real velocity profile detected on the road.

In the following, the whole experimental dataset is divided into two subsets, relating to the two paths. Statistical analysis results presented are relative to the Path A.

The road tests on Path A are subdivided in 595 sequences on which Principal Component Analysis (PCA) and Cluster Analysis (CA) are performed [16], to better characterize sequences and classify them into homogeneous groups. Principal component analysis of a data matrix extracts the dominant patterns in the dataset making use of a complementary set of score and loading.

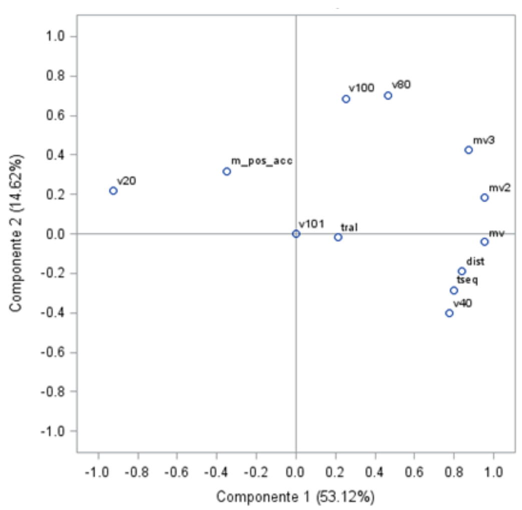

Fig. 3. First two Principal Components.

In particular, on the first axis, sequences which have a very low average speed (v20) with a strong acceleration phase (m_pos_acc) compared to those with a constant average speed of around 40km/h (V40) and with a greater distance (dist) are well distinct. In fact, they are located on the opposite side of the first axis. Instead, on the second axis, the few sequences emerge with a very high average speed (v80, v100), typical of street/path not congested. Detailed statistical approach was described in the previous paper [6, 10].

4 Results and discussion

4.1 Statistical Kinematic analysis

Driving cycles are clustered into groups to determine typical and statistical car performance in different roads with different experienced traffic conditions. Cluster analysis performed on cycle variables, some of them reported on table 3, groups them into three well-defined clusters.

Table 3. Cluster cycle mean characteristics.

CLUSTER	1	2	3
N cycles	94	82	4
mv (km/h)	16,05	25,48	43,14
Tral (s)	44,71	41,99	21,75
Dist (m)	1071,12	1294,25	1218,05
m_pos_acc (m^2/s^3)	0,66	0,57	0,63
tcyc (s)	234,05	211,68	130,25

In such a way, the different behavior and traffic situations that have occurred are synthesized and then the different driving styles are evidenced. For this reason, on each driving cycle (statistical unit), the average value of kinematic variables are calculated. For synthesis reasons, only the cluster analysis on Path A is described in detail. Significant variables are: cycle mean velocity (*mv*), cycle total idling time (*Tral*), total distance covered (*Dist*) and cycle total time (*tcyc*).
Results are illustrated by cluster representation in the Can 1, Can 2 scatter plot (figure 4), where each point is the statistical representation of a driving cycle. In the following figure, Can 1 values are correlated with variables that differentiate the cluster's cycles in terms of mean speed. In fact, it differentiate them from slow (mv≅16-25 km/h) to fast (mv≅43 km/h). The Cluster 1-3 cycle distances are about the same (dist≅ 1000-1300m).
Can 2 values differentiate them from slow to fast in terms of speed class velocity. Moreover, they are sorted in decreasing mode for time duration and idling.

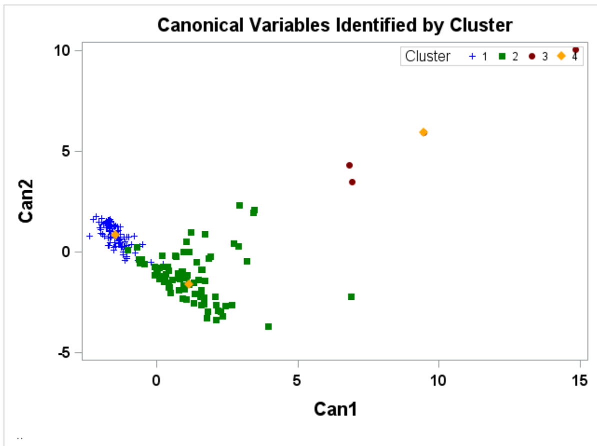

Fig. 4. Cluster representation of Path A cycles

From each group, the most representative cycle is determined by discriminant analysis. Therefore, in figure 4, the cluster named 4 in yellow, identifies the representative elements of each cluster (1, 2 and 3). They are calculated fitting a multidimensional normal distribution to the full driving cycles set, estimating its density function and finally keeping the mode of each group that has the maximum value of density function.

The previous Table 3 shows, for each cluster, the mean values of variables most representative, so it is possible to point out fundamental differences in the kinematic features. In this way, results coming from Can1-Can2 plot are confirmed and more simply understood in terms of physical phenomena. Finally, the global cycle could be built by retaining successively the most representative from each cluster, ordering according to the average speed, or by starting from cluster 1 to follow. This way could be considered as statistical criteria, but obviously, it has to be posed in a specific context, in the sense that it is necessary to visualize where it is realized over the path. However, it is specific for different kinematic conditions and summarizes information related to road/driver/traffic conditions more frequently experienced in real use.

The emissivity of the different cars is maintained in the different clusters, reflecting the kinematic differences encountered during the real use.

Results obtained by classifying the CO2 emissions of different cars, subdivided into the three clusters, show that the emission values in cluster two are more similar for the two Panda vehicles, perhaps more similar for engine technologies and vehicle characteristics. This feature is not evidence in cluster one, probably due to a higher prevalence of kinematic factors with respect to different engine technologies.

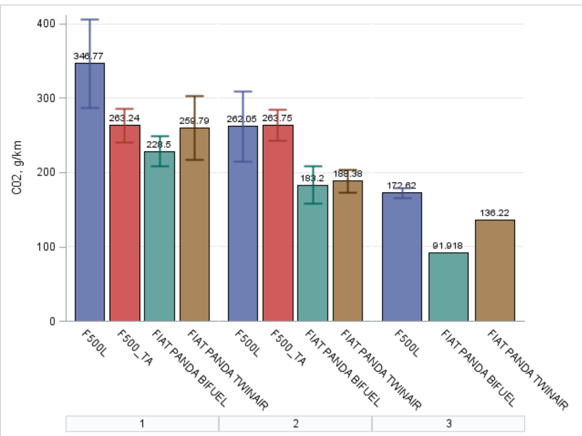

Fig. 5. Vehicle CO_2 emission vs cluster cycle for Path A.

4.2 VSP analysis

Vehicle Specific Power (VSP) is conventionally defined to represent the instantaneous vehicle engine power [17]. It has been widely utilized to reveal the impact of vehicle operating conditions on emission and energy consumption estimates that are dependent upon speed, roadway grade and acceleration or deceleration based on the second-by-second vehicle operation. The VSP of the test vehicle is calculated for each second of test data, from the 1 Hz vehicle speed data, recorded by OBD measurement and the 1 Hz road grade estimate generated by the GPS receiver.
In literature, VSP is commonly divided from 14 bins [14] to 59 [15] or more. In this research, we identify 24 bin to preliminary analyse the frequency distribution in the road trip and the correlation with the emissions.
As we can see in the figure 6 the CO_2 emission is strongly correlated with the highest mode were is present higher acceleration with high road grade. Moreover, the mode follow a normal distribution.

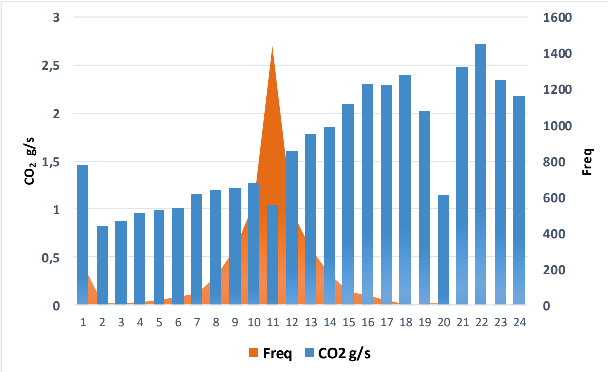

Fig. 6. CO_2 and bin frequency distribution

In Fig. 7 and 8, VSP values calculated for Fiat 500, are respectively reported as a function of acceleration and road grade.

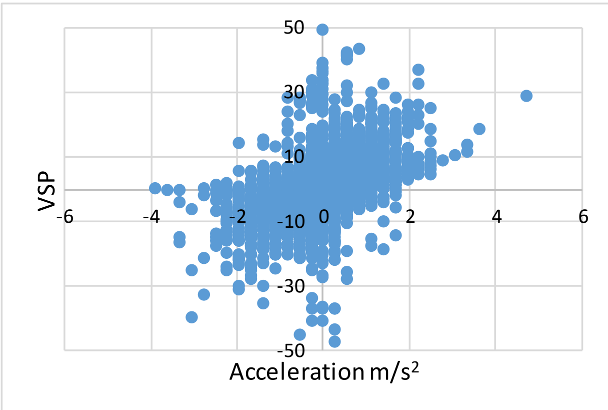

Fig. 7. VSP vs Acceleration m/s^2

During acceleration and in correspondence of positive slopes, VSP is mainly positive; on the contrary, the values become negative in deceleration and downhill.

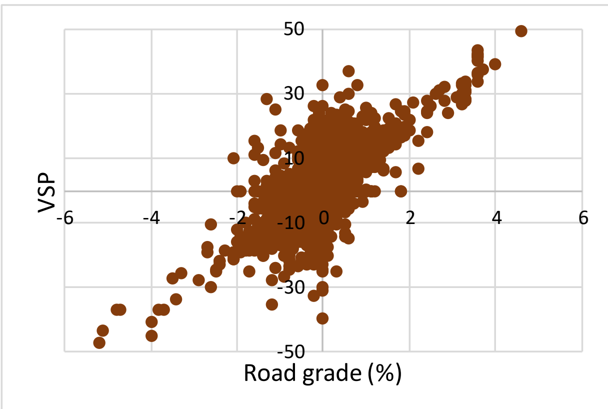

Fig. 8. VSP vs road grade

Study and experimentation on the road could help policy makers to apply different sustainable measurement in reducing urban traffic in certain areas or with particular morphological features of the territory.

5 Conclusion

In a DSS approach it is possible to integrate different aspects of emission environmental impact. In Istituto Motori particularly was developed a modeling approach of real-world emission aiming to develop a predicting tool, both for research and for political decision-makers. Some preliminary results from the statistical evaluation of real emission and kinematic data could contribute to produce various emission indexes that help policy makers in a context of sustainable mobility. In this context different experimental campaigns were realized in the roads of Naples city, by different cars equipped with kinematic data acquisition, GPS recording and PEMS (Portable Emission Measurement Systems) equipment. So, instrumented vehicles for the simultaneous acquisition of kinematic variables, localization and emissions data have been used. It is also to keep in account that the WLTC cycles, improvement of the NEDC cycle, are still unrealistically slow in some acceleration phases. Moreover, there is no hill climbing in the cycle with gradient limitations. So globally driving cycles must be as more representative of real use as possible.

In this paper, some preliminary considerations will be developed trying to evaluate how it is possible to integrate the Vehicle Specific Power (VSP) in a statistical approach of vehicle kinematic. This kind of activity globally aims at valuation of traffic situation and emissions in real use. So globally we study the way to integrate VSP index in a statistical approach, since this index is widely recognized in the literature to synthesize the contribution of kinematics along with road gradient.

Results of statistical analysis results presented are relative to the Path A, for which typical driving cycles was built summarizing different kinematic conditions related to road/driver/traffic conditions more frequently experienced in real use. Following this approach, three cluster of cycles were identified showing different kinematic characteristics in terms of average speeds, length and distance traveled, idling time, running time, distributions of acceleration/deceleration use. Results obtained by classifying the CO2 emissions of different cars, subdivided into the three clusters, show that the emission values in cluster two are more similar for the two Panda vehicles, perhaps more similar for engine technologies and vehicle characteristics. This feature is not evidence in cluster one, probably due to a higher prevalence of kinematic factors with respect to different engine technologies. Regarding VSP analysis, values calculated for a single car, as example, are respectively reported as a function of acceleration and road grade in order to highlight its correlation with this variables. During acceleration and in correspondence of positive slopes, VSP is mainly positive; on the contrary, the values become negative in deceleration and downhill. So a study on correlation between VSP values and driving kinematic mode could be very promising for a statistical evaluation and use of VSP variable.

In conclusion this activity obviously are a preliminary indication but also a strong starting point for reflection on the real representation of the cycles currently used for type approval, fuel consumption and emissions evaluation. Furthermore, they could be useful to both policy makers and vehicle manufacturers in developing future emission policy/technology strategies.

References

1. Rapone M., Della Ragione L., Meccariello G., An integrated knowledge base for modelling and predicting vehicle real-world emissions as a function of driving behaviour kinematics, 2008 December, CIT Journal of Computing and Information Technology 2008; 16(4): 279-286 Multivariate Regression Model, Environmental Modelling and Assessment, Springer Netherlands, 2008; 13:383-392
2. Yang L. , Franco V., Mock P., Kolke S. Zhang R., Wu Y., and German J., “Experimental Assessment of NOx Emissions from 73 Euro 6 Diesel Passenger Cars”, Environ. Sci. Technol., 2015, 49 (24), pp 14409–14415, DOI: 10.1021/acs.est.5b04242A
3. Franco, V.; Posada, F.; German, J.; Mock, P. Real-World Exhaust Emissions from Modern Diesel Cars: A meta-analysis of PEMS emissions data from EU (EURO 6) and US (Tier 2 Bin 5/ULEV II) diesel passenger cars; White paper, part 1: Aggregated Results; International Council on Clean Transportation, Washington, DC, 2014; http:// www.theicct.org/sites/default/files/publications/ICCT_PEMS-study_ diesel-cars_20141013.pdf.Mecke, I. Lee, J.R. Baker jr., M.M. Banaszak Holl, B.G. Orr, Eur. Phys. J. E **14**, 7 (2004)
4. Kadijk, G.; van Mensch, P.; Spreen, J. Detailed investigations and real-world emission performance of Euro 6 diesel passenger cars; TNO Report R10702; Netherlands Organization for Applied Scientific Research, Delft, The Netherlands, 2015; http://publications.tno.nl/publication/34616868/ a1Ug1a/TNO-2015-R10702.pdf.
5. Peter Mock, Jörg Kühlwein, Uwe Tietge, Vicente Franco, Anup Bandivadekar, John German “The WLTP: How a new test procedure for cars will affect fuel consumption values in the EU”, ICCT, October 2014.
6. Meccariello, G. and Della Ragione, L., "Statistical Determination of Local Driving Cycles Based on Experimental Campaign as WLTC Real Approach," SAE Technical Paper 2017-24-0138, 2017, https://doi.org/10.4271/2017-24-0138.
7. May J., Bosteels D., Favre C., "An Assessment of Emissions from Light-Duty Vehicles using PEMS and Chassis Dynamometer Testing," SAE Int. J. Engines 7(3):2014, doi: 10.4271/2014-01-1581.

8. Della Ragione L., Meccariello G., Prati M.V., "Statistical Investigation of In Use Emissions and Fuel Consumption Measured by PEM on Different Gasoline Cars", SAE Technical Paper 2013-01-1511, 2013, doi: 10.4271/2013-01-1511.
9. Wikipedia.https://en.wikipedia.org/wiki/Worldwide_harmonized_Light_vehicles_Test_Procedure.
10. Della Ragione L., Meccariello G, "The Influence of Road Gradient in an Integrated Approach of Real Driving Cycles and Emissions Factors Model", Transportation Research Procedia Volume 14, 2016, Pages 3179–3188, Transport Research Arena TRA2016, http://dx.doi.org/10.1016/j.trpro.2016.05.258.
11. Meccariello G., Della Ragione L., Prati M., Costagliola M. et al., "Real Time Emissive Behaviour of a Bi-Fuel Euro 4 SI Car in Naples Urban Area," SAE Int. J. Fuels Lubr. 6(3):959-967, 2013, doi: 10.4271/2013-24-0173.
12. Joumard, R., Andre, J.M., Rapone M., et al., 2007. Emission factor modelling and database for light vehicles Artemis deliverable 3.
13. André M., Joumard R., Vidon R., Tassel P., Perret P., "Real-world European driving cycles, for measuring pollutant emissions from high- and low-powered cars", Atmospheric Environment 40(31): 5944–5953, 2006.
14. Frey, H., Unal, A., Chen, J., Li, S., Xuan, C., "Methodology for developing modal emission rates for EPA's multi-scale motor vehicle & equipment emission system," State Univ. US EPA, Ann Arbor, MI, no. October, pp. 18-20, 2002.
15. "Characterizing Emission Variations due to Driving behaviour from On-Road Vehicles", International Vehicle Emissions Model; Attachment C. http://www.issrc.org/ive/;
16. Bouroche JM, Saporta G. (1983). L'Analisi dei Dati, CLU Napoli; 1983. ISBN-2-13-036505-1.
17. Yao, Z., Wei, H., Liu, H., Li, Z., Statistical Vehicle Specific Power Profiling for Urban Freeways, Procedia - Social and Behavioral Sciences, Volume 96, 6 November 2013, Pages 2927-2938 https://doi.org/10.1016/j.sbspro.2013.08.324