The influence of road grade and GPS signal reconstruction on real driving behaviour and emissions

L. Della Ragione, G. Meccariello

Abstract—In recent years in Italy the progress of the automobile industry, in the field of reduction of emissions values, is very remarkable. Nevertheless their evaluation and reduction is a key problem, especially in the cities, that account for more than 50% of world population. The aim of this study is to investigate the parameters influencing the real driving emission monitoring with particular attention towards the influence of road gradient. Several experimental campaigns are carried out with some cars instrumented for both the acquisition of kinematic data, polluting emissions in continuous, and GPS latitude, longitude and altitude data to obtain precise geolocalization and slope variation during a path. Moreover this paper deals with the definition of a quantitatively approach for the reconstruction of GPS coordinates and altitude, in the context of correlation study between driving cycles / emission / geographical location. Also cluster analysis of kinematic data is realized, keeping in mind new corrected GPS data, and a basic statistical analysis of exhaust emissions of CO, THC, NOx, CO2 acquired on road.

Keywords—Driving cycle, Emissions, GPS Signal reconstruction, Slope variation.

I. INTRODUCTION

Vehicle pollutant emissions and fuel consumption evaluation remain until now a major challenge for Italy. Positive results were obtained but in the case of Italy, some actors contributing negatively to this situation are present. So different aspects related to evaluation of vehicles pollutant emissions is still the subject of research studies. The dependence of emission not only from the average speed but also by many other parameters characterizing the kinematics of the vehicle is fully established [1], [2].

Furthermore, interaction with path road not only in terms of traffic experienced by the vehicle, but also in terms of orographic characteristics of the path has an overt acclaimed effect on the emission and consumption.

Various studies have been conducted to evaluate the effect of the traffic in terms of time of day and day of the week, to characterize the driving behavior.

The equally important feature that could be not neglect concerns the correlation between kinematic sequences / emissions / topography of the route runs from the vehicle in real use [3], [4], [5]. In fact kinematic sequences and their features represent driving behavior in its totally, so it could be important defining in the best way every factor. In particular, it could be interesting to suggest paths based not only on the minimum distance, but also on the minimization of fuel consumption as a function of the geomorphologic features of the territory. In our integrated approach, kinematic, emissions and geomorphologic data are statistically analyzed, aiming to locate precisely where a specific situation may occur more frequently. For this purpose, the activities will be aimed at integrating automatically, the commercial digital maps with geomorphological data relating to the real three-dimensional pattern of the road network. But to do in a correct manner this activity, data must be pre-treated and corrected by any form of error that can be amplified in subsequent statistical analysis.

The aim of this paper is to give a primarily answer to the problem of signal reconstruction of GPS signal especially for altimetry [6], [7], to better analyze and to outline the behavior of low environmental impact vehicles in city traffic situations and in a precise location [8], [9]. Also it could be desirable to perform a quantitative analysis of altimetry to evaluate the slope variability during a path, so this variable can contribute to correlate kinematic behavior with emissions [10].

II. GLOBAL APPROACH FOR GPS DATA

In this paper, experimental data coming from four repetitions of a specific path, with different gradient situations, are analyzed. The experimental path, with a length of about 8 km, is shown in Fig. 1. It is chosen taking into account a roads with numerous elevation changes. In particular the identified experimental path is well balanced in terms of slope with an average uphill of 6.9%, while downhill has an average gradient of 6.5%. In the fig. 1 the range of the elevation profile is about between 50m and 190m.
The following Fig. 2 illustrates the general layout of the overall approach.

The measured emissions, the kinematics and GPS data are filtered, synchronized and analyzed using statistical methods.

The central block refer to the reconstruction of the GPS signal with the aim to solve the frequently problems that can arise from the acquisition of this signal in real time [12], [13], [14], [15], [16]. The third block described the statistical approach to correlate the emission and consumption values in a specific road position [17].

The vehicles are equipped with a GPS instrument that returns the values of position and elevation of the path with a 10 Hz frequency. During the review of the data of the GPS signal we found that the data for altimetry varies in a range of +/- 10 ms, this change makes it not completely accurate analysis of the slope on the way.

Such errors, for what concerns the coordinates, have occurred in the form of:
- position values sometimes missing;
- misplacement of the vehicle, for example, in areas that are not followed during the journey;
- drift of the position values with the vehicle stationary.

For what concerns instead the elevation errors, they were found:
- changes in short strokes, with slope values are incompatible with reality;
- discrepancies in absolute terms, by comparison with ordinary topographic maps and Google Maps.

So the logical process defined to individuate an algorithm to perform in a mathematical and in an automatic way the signal reconstruction consists of the steps described below. On the first phase, through the Google API (Application Program Interface), we can get the correct map data DTM (Digital Terrain Model) of the theoretical path, on which we performed the real tests. A second phase is to reconstruct accurate geolocation and slope of the vehicle respect to the GPS signal recorded in real use. To this end, a software solution has been implemented in C# .NET, which should facilitate the correction and the filling of the data set of coordinates.

The following part describes the implementation of the approach for the analysis of the identified path as shown in

- a Semtech gas analyzer produced by Sensor to measure at 1Hz CO, NOx and CO2 emissions. This analyzer uses NDIR cell (Non-Dispersive Infrared) for CO and CO2 measurements, NDUV cell for NO/NO2 and separate electrochemical sensor for oxygen. The analyzer is calibrated on a regular basis and zeroes itself on start-up using outside air.
- a Semtech-EFM (Exhaust Flow Meter) by Sensor.
- an OBD interface and logging computer running propriory software (EDS) to acquire engine operating parameters (speed, rpm, engine air flow).
- a GPS receiver by Racelogic Ltd to acquire the spatial position (for distance: resolution = 1cm; accuracy = 0.05% - for altitude: 10 m 95% Circle of Error Probable).
- a video camera to record traffic situations.

The first block concerns the design of experiments, the experimental activity and the sampling of all the kinematic quantities, vehicle position and emission factors.

Two vehicle have been instrumented with a complete Portable Emission Measurement System (PEMs) [11] comprised of gas analyzers, to measure the concentrations of regulated gaseous pollutants in the exhaust gas, an exhaust mass flow meter, a Global Positioning System, sensors to measure the ambient temperature and pressure and a connection with the vehicle ECU.

The main components of the PEM system used in this work are:

- a Semtech gas analyzer produced by Sensor to measure at 1Hz CO, NOx and CO2 emissions. This analyzer uses NDIR cell (Non-Dispersive Infrared) for CO and CO2 measurements, NDUV cell for NO/NO2 and separate electrochemical sensor for oxygen. The analyzer is calibrated on a regular basis and zeroes itself on start-up using outside air.
- a Semtech-EFM (Exhaust Flow Meter) by Sensor.
- an OBD interface and logging computer running propriory software (EDS) to acquire engine operating parameters (speed, rpm, engine air flow).
- a GPS receiver by Racelogic Ltd to acquire the spatial position (for distance: resolution = 1cm; accuracy = 0.05% - for altitude: 10 m 95% Circle of Error Probable).
- a video camera to record traffic situations.
Fig. 1. Figures 3 and 4 show the graph of Experimental Path (EP) and the Model Path (MP) imported through the realized software. In particular, figure 3 shows the EP on a latitude and longitude graph. In the analysis of this path problems can be arise due to continuous changes of elevation, signal drift and the presence of tall buildings leading to the loss of the GPS signal.

Fig. 3 Lat. vs Long. of the Experimental Path

Fig. 4 shows the graph of the corresponding Model Path used to match and aligns the GPS data of the EP.

Fig. 4 Lat. vs Long. of the Model Path

Fig. 5 shows the elevation profile of the MP obtained through the Google API.

Fig. 5 Elevation profile of the Model Path

We have to take in mind that the elevation profile obtained by Google API, shown in Fig. 5, is speed independent and in the same way, consequently the calculated distance. Instead, the distance obtained from the real profile is obviously dependent on the recorded speed. Moreover, in the reconstruction phases we must consider that the recorded distance is different from the real one, due to the error of the recorded velocity signal especially in the speed range close to zero. In the data processing, we must keep in mind that gap. In fact as shown in Fig. 6 the cumulative distance of EP (red line) is lower than the MP (green line).

Fig. 6 Elevation profile of the Model Path

Applying the algorithm identified on the experimental data registered in different days, we can reconstruct automatically, for each trip, all the profiles of the elevations perfectly aligned with the kinematic.

After data processing, which globally consists in several steps based on the identification of the beginning and the end of the EP and MP, scale and realign the EP in relation to the MP model, you can get the correct EP elevation profile (Fig. 7). It is aligned with the kinematics and the real distance traveled by the vehicle (Fig. 8). The proper alignment of the data processing on the Path is especially noticeable, comparing Figs. 7 and 8, when the vehicle is stationary around the second 600 and 1200.

Fig. 7 Correct elevation vs time of the Path
This allowed us to make a proper assessment of the effect of the slope on the guide profiles acquired and consequently the emission factors. Figures 9 and 10 show the elevation profiles of two cycles carried out on the road, with the elevation obtained from the acquired GPS vs. that obtained by the process of realignment and correction developed by Meccariello et al. 2014. The following graphs show the difference between the altitude values obtained from the GPS instrument (Altitude) and those calculated through the numerical model of the reallocation coordinated (Altitude Dev.).

Only two days of acquisitions among those developed in the analysis are shown because in the other two trips, even though longitude and latitude values are present, for the 60% of the trip the elevation value was in full scale value of 500 m.

The figures show a strongly deviation between the two signals, and consequently a strong bias on slope variation.

On the other hand the correct reconstruction of the signal altimetry affects the measurement of emissions through the classification of kinematic sequences.

In fact, the emission values depend primarily on the average speed, equally strongly from the distribution of accelerations, from the idle time and other kinematic parameters. Moreover, these parameters are correlated with the realization of experimental paths with positive gradient, slope, or negative gradient, downhill. For this reason also the distribution of the percentage of time with positive or negative slope, must be used in the statistical analysis of the kinematic sequences; mostly when experimental campaign will be realized on a path made of routes on high slope. The use of automatic values altimetry then can have a misleading effect on the statistical analysis, especially in a perspective of automation of the process of identification of the driver on board recommended path for emission reductions. In fact, following this approach, latitude and longitude are used qualitatively, while altimetry will be used quantitatively.

Among the GPS signals, the altimetry is used in a quantitative way in the cluster analysis of the cinematic sequences. This leads to an influence of a sequence in belonging to a cluster rather than to another, and then the correlation of an average value of each pollutant with a level of traffic/driving behavior rather than with another.

### III. Statistical Analysis of Kinematic and Slope Data

Each acquired velocity profile is splitted into a succession of kinematic sequences, constituted the elementary statistical unit. So the entire set of kinematic sequences are analyzed by multivariate statistical analysis. Methodologies applied are principal component analysis, cluster analysis and discriminant analysis. Each of them realizes a specific aim in our approach. Basically principal component analysis performs a dimensionality reduction in the entire range of variables defining to characterize the set of kinematic sequences. Moreover cluster analysis subdivided them in groups that are internal homogeneous and very good differentiated each other. Discriminant analysis is applied to outline features and reciprocal differences of clusters. Canonical Discriminant Analysis is used to determine which
variables discriminate between clusters (groups) of multivariate observations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mv (km/h)</td>
<td>Mean of running speed (v&gt;0)</td>
</tr>
<tr>
<td>mv2 (km²/h²)</td>
<td>Mean of square speed (v&gt;0)</td>
</tr>
<tr>
<td>mv3 (km³/h³)</td>
<td>Mean of cube speed (v&gt;0)</td>
</tr>
<tr>
<td>Tral (s)</td>
<td>Idling time v=0 in second</td>
</tr>
<tr>
<td>Trunning (s)</td>
<td>Total running time (v&gt;0) in second</td>
</tr>
<tr>
<td>Dist (m)</td>
<td>Distance covered</td>
</tr>
<tr>
<td>DS1 (%)</td>
<td>% time with delta slope &lt; -0.70 meters (m)</td>
</tr>
<tr>
<td>DS2 (%)</td>
<td>% time -0.70 &lt;= delta slope &lt; -0.20 meters (m)</td>
</tr>
<tr>
<td>DS3 (%)</td>
<td>% time -0.20 &lt;= delta slope &lt; 0 meters (m)</td>
</tr>
<tr>
<td>DS4 (%)</td>
<td>% time 0 &lt;= delta slope &lt; 0.20 meters (m)</td>
</tr>
<tr>
<td>DS5 (%)</td>
<td>% time speed &gt; 60</td>
</tr>
<tr>
<td>DS6 (%)</td>
<td>% time with delta slope &gt;= 0.70 meters (m)</td>
</tr>
<tr>
<td>Time (s)</td>
<td>Total duration of the sequence (s)</td>
</tr>
<tr>
<td>(m_vap_{os} (m^2/s^3))</td>
<td>Mean of instantaneous values of product ((a(t) \cdot v(t))) when (v(t) &gt; 0) and (a(t) &gt; 0)</td>
</tr>
</tbody>
</table>

Table I. Acronyms of variables characterizing kinematic sequence.

Some optimal combinations (functions) of variables are automatically determined so that the first function provides the overall discrimination between groups; the second provides the second most, and so on. Functions are denoted as canonical variables (called in the paper CAN1, CAN2...). Each canonical variable is characterized by the percent amount of explained variance respect to the total variance. Plots of the first two canonical variables, generally explaining a consistent percent of variance, give a good two-dimensional representation of observations. Results are illustrated by cluster representation in the Can 1 sequence, Can 2 sequence scatter plot (Fig. 11) [18].

Cluster mean values of some original variables (mean velocity, sequence duration, idling time and distance covered) are reported. Cluster 1 represent sequences along the road path with a short distance realized and with 57% of the time in flat road part. Clusters 2 and 3 have about the same mean speed. They are essentially different because the cluster 2 is mostly downhill (only 20% uphill) while Cluster 3 has about 70% uphill. Cluster 4 are formed essentially by downhill sequences (15% uphill).

In fig. 12 an overlay of representative sequences profile of each cluster is shown. The kinematic characteristics are consistent with the values of the Table II.

Cluster mean values of sequences kinematic variables.

Table II. Cluster mean values of sequences kinematic variables.

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>29</td>
<td>15</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Tral (s)</td>
<td>10.93</td>
<td>3.73</td>
<td>7.38</td>
<td>41.33</td>
</tr>
<tr>
<td>Dist (m)</td>
<td>101.59</td>
<td>679.09</td>
<td>813.81</td>
<td>463.64</td>
</tr>
<tr>
<td>mv (km/h)</td>
<td>10.52</td>
<td>20.71</td>
<td>20.79</td>
<td>15.91</td>
</tr>
<tr>
<td>mv2 (km²/h²)</td>
<td>168.18</td>
<td>490.77</td>
<td>498.23</td>
<td>339.40</td>
</tr>
<tr>
<td>mv3 (km³/h³)</td>
<td>3396.59</td>
<td>12650.96</td>
<td>12956.17</td>
<td>8050.23</td>
</tr>
<tr>
<td>(m_vap_{os} (m^2/s^3))</td>
<td>0.80</td>
<td>0.59</td>
<td>0.56</td>
<td>0.47</td>
</tr>
<tr>
<td>Time (s)</td>
<td>37.31</td>
<td>123.67</td>
<td>154.25</td>
<td>127.50</td>
</tr>
<tr>
<td>DS1 (%)</td>
<td>0.49</td>
<td>9.71</td>
<td>0.53</td>
<td>76.90</td>
</tr>
<tr>
<td>DS2 (%)</td>
<td>6.19</td>
<td>37.70</td>
<td>6.24</td>
<td>5.31</td>
</tr>
<tr>
<td>DS3 (%)</td>
<td>14.90</td>
<td>26.53</td>
<td>5.64</td>
<td>2.73</td>
</tr>
<tr>
<td>DS4 (%)</td>
<td>57.46</td>
<td>17.52</td>
<td>25.29</td>
<td>12.02</td>
</tr>
<tr>
<td>DS5 (%)</td>
<td>20.21</td>
<td>8.39</td>
<td>51.36</td>
<td>3.05</td>
</tr>
<tr>
<td>DS6 (%)</td>
<td>0.14</td>
<td>0.27</td>
<td>15.25</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The following Table II presents the results of Cluster Analysis. The four road tests are subdivided in 58 sequences, which are grouped in four clusters. For each cluster the table presents mean values of variables most representative, so it is possible to point out fundamental differences in the kinematic features.

Fig. 11. Cluster representation of sequences.

Fig. 12. Speed profile overlay of representative sequence for each cluster.

Correlation analysis between a kinematic sequences and the corresponding emission sequence is performed in the same traffic situation (determined by cluster) and same GPS location. A trend analysis is performed via the visualization of instantaneous emission profiles and kinematic sequences for each cluster. Figures 13 and 14 report a GPS profile, while figures 15 and 16 report the elevation profile of the GPS instrument and those calculated through the numerical model (Dev). In particular, the road path characteristics are well related to cluster characteristics reported in previous table, which are an expression of uphill road, flat road and downhill road.
With particular regard to figures 8 and 9, it is possible to note how the cluster distribution is well identified in both the uphill and the downhill faces. Moreover, comparing the figures 8a vs. 8b, and 9a vs. 9b, it is possible to note as the reconstruction process has properly filled the missing parts. Furthermore in the reconstruction phases we must consider that the recorded distance is different from the real one, due to the error of the recorded velocity signal especially in the speed range close to zero. In the data processing, we have to keep in mind that gap.
IV. EMISSION RESULTS AND DISCUSSION

The time series of each regulated emission acquired during each trip is split into sequences. CO₂ and fuel consumption (FC) are expressed in g/s, CO and NO in mg/s. Correlation analysis between a kinematic sequences and the corresponding emission sequence is performed in the same traffic situation (determined by cluster) and same GPS location. A trend analysis is performed via the visualization of emission instantaneous profile and kinematic sequences for each cluster.

Fig. 17 reports a GPS (a) and elevation profile (b) of a trip with color related to the identify cluster, that identifies uphill road, flat road and downhill road. In particular, the road path characteristics are well related to cluster characteristics reported in previous table.

Furthermore in Fig. 18 pollutant emissions (a, b, c) and FC (d) profiles relative to the same trip are shown. Time series of CO₂, NO, CO, and FC are colored according to cluster sequences road characteristics. We have to consider that this result is not a qualitative one but it is obtained by statistical cluster analysis of succession of sequences. Here it is possible to note that at similar kinematic conditions (i.e. mean speed) in all sections of road we have different emission profile. This confirms the importance of the variable slope in emission evaluation in real use analysis, such as in the case analyzed in which the identified experimental path is well-balanced in terms of slope, with an average uphill of 6.9%, and downhill with an average gradient of 6.5%.

Fig. 16. Elevation profile of day trip 29/06 coloured according to cluster.

Fig. 17 (a) GPS profile of a trip colored according to cluster; (b) Elevation profile of the same trip.
In Table III, to better underline and quantify difference due to cluster of sequences and their characteristics, mean velocity (mv), mean of each emission and FC (fuel consumption) are summarized for each cluster. In addition, the maximum emission values are recognized in cluster 3, probably due to kinematics sequences high percentage of uphill phase. Although it shows the same mean velocity, a strong reduction effect for each emission could be evidenced in cluster 4, formed essentially by downhill sequences. Emission reduction are stronger for NO emission, but also clearly stated for CO₂.

Cluster 1 represents sequences along the road path mostly in uphill with a short distance and it shows medium values for each emission and for fuel consumption.

V. CONCLUSION

An experimental campaign is carried out in the city of Naples, in two areas different for the topography of the streets, but that present very similar vehicular traffic, typical of busy streets. The aim of this activity is to compare fuel consumption and emissions on road during real world experimental tests, in order to identify some characteristics of road that strongly influence emission production. Moreover, this paper seeks to give a contribution to on-board measurements, in different geographical areas, with PEMS using a statistical methodology. A procedure to implement the reconstruction of the GPS signal during a real experimental campaign is completely formalized. A more precise GPS signal and altimetry value help to better perform a quantitative analysis of driving behaviour. These signals, in particular altimetry value, may be statistically evaluate to produce variables that describe the variation of the slope and the distance travelled during a trip with greater precision.
Experimental data are subdivided in sequences, which are grouped in cluster, so it is possible to point out fundamental differences in the kinematic features. Particularly attention is given to the construction of appropriate variables to characterize the slope variation along the road path after the reconstruction of the GPS signal. These variables are among the most significant in kinematic sequences classification above all on a journey in path with many uphill and downhill phases. In fact the introduction of these variables in discriminant and cluster analysis of sequences produce different group with different kinematic characteristics.

The approach followed in this paper allowed us to evaluate the emissions trend with two cars gasoline fuelled on a particular path road characterized by uphill and downhill gradient variability. Results show an influence of this street feature on emissions and fuel consumption.

In the future, we will aim to conduct a comparative emissions analysis on mean emission values with and without these slope variables.

ACKNOWLEDGMENT

This activity research is realized within a national scientific project PON01_00744 DRIVE IN2, partially supported by the Italian Ministry for University and Research. It was carried out by the Istituto Motore of the National Research Council.

REFERENCES


[7] Cătălin Popescu, Augustin Mitu, Daniela Uta; (2012) "Design and implementation of methodologies for transport (CIVITAS SUCCESS) project(s) impact measurements"; Wseas Transactions On Environment And Development; E-ISSN: 2224-3496; Issue 1, Volume 8.


L. Della Ragione graduated in 1988 in mathematics from the University of Naples “Federico II” (Italy). Presently, she is a researcher of CNR-IM Her main research activities concern the environmental impact of vehicles. Her principal areas of interest include multivariate statistical analysis and time series analysis. The main research subjects are vehicle mission profile in urban areas and the development of methods for the evaluation of exhaust emissions as a function of vehicle operating conditions in the traffic. In this context, she has developed some statistical codes applied to determine driving cycles by vehicle speed records in the traffic obtained by instrumented cars performing planned. In addition, she is involved in the development of models for vehicular emission factors determination based on measurements of pollutant emissions. Dr. Della Ragione is responsible for research group working on “Statistical methods for driving cycle modelling and emission factors development”.

G. Meccariello graduated in 1999 in economics from the University of Naples “Federico II” (Italy). He received his PhD in 2005 from the Department of Engineering “Luigi Tocchetti” of the University of Naples “Federico II” (Italy). He is a researcher of vehicle performance and emissions for sustainable transportation at Istituto Motore of National Research Council of Italy. His research field of interest includes statistical and soft computing methods for modeling driving cycles and emissions of vehicles in real-use, with special concern for vehicle mission profile in urban areas and exhaust emissions evaluation. He is involved in the development of models for vehicular emission factors determination based on measurements of pollutant emissions. He is a developer in #net Framework. He is also, working with SAS System, particularly in the field of multivariate statistical analysis with SAS/STAT and SAS/GRAPH.