ABSTRACT

With increased numbers of vehicles on Irish roads, there is now a need to be able to scientifically assess the quantity of pollutant material to which populations are exposed. Traditionally, emissions have been determined using kinematic (vehicle speed) data but recent studies have identified that other parameters are of interest. The work in this Paper focuses on the development and testing of a purpose-built software system to extract on-board diagnostic data from a vehicle in order to derive a driving cycle and to use other engine characteristic data to better inform local pollutant and energy consumption models for Dublin. Comparisons with GPS data shows the system to be cost effective (price and computing overhead) and reliable.

INTRODUCTION

During the period 1990 to 2005, private vehicle numbers using Irish roads increased from about 796,000 to 1,662,000 (109%). During the same period, commercial vehicle numbers rose from 143,000 to 287,000 (100%) [1]. The average engine size also increased, with a distinct shift from small (< 1.4 l) engines to mid-sized (1.4 l – 2.0 l) engines. In 2000, only 12 % of the Irish private car fleet was diesel powered. By 2006, this had risen to 16 % [2], amounting to an absolute increase of nearly 70,000 vehicles.

With these shifts in the structure and composition of the vehicle fleet, concerns about air quality, fuel consumption and availability of supply, and the contributions of road transport to greenhouse gas (GHG) emissions have risen. While air quality in Ireland is generally considered to be good [3], there are concerns over the potential risks to human health, particularly in towns and cities. Vehicular emissions are currently considered to be the most important contributor to poor air quality in urban areas [4–7]. The combination of narrow streets [8], high traffic volumes and large populations [9] living and working in close proximity to major roads [10] means that both local and national authorities must be proactive in assessing the level of pollutant emissions and employing measures that reduce the negative impacts on human health. Recently released figures from the United Nations [11] suggest that by 2050, nearly 69% of the World's population (6,398 million people) will live in urban centers. For Ireland, the projected value is much higher, reaching 78% (4.82 million people) by 2050. Clearly, a considerable amount of work needs to be done in Ireland in order to fully quantify and limit the potential negative impacts of vehicular emissions on human health.

With such concerns in mind, this Paper describes a low-cost system used to extract data from vehicles using the standard on-board instrumentation to develop a series of representative driving cycles, such that emissions, based on vehicle activity within the transport network, can be predicted [12]. Aggregation of emissions from individual vehicles in given parts of the city will permit a clearer understanding of vehicular emissions in the urban area.

The Paper focuses on the software system designed to achieve this goal and on the comparison and validation of data extracted in this way. A comparison is made with global positioning data to show the effectiveness of the system for describing kinematic driving cycle data.

REVIEW OF DRIVING CYCLE DEVELOPMENT

RATIONALE – Traditionally, pollutant emissions are determined using standardized driving cycles, which are plots of vehicle speed against time [5, 13–15]. Broadly speaking, driving cycles can be divided into two categories, those used for legislative emissions testing, and those developed as a means to examine emissions and fuel consumption for a given location [15–17]. Legislative cycles, such as the New European Driving Cycle (NEDC) or the Federal Test Procedures (FTP) in the United States, are used by national authorities to ensure conformity of emissions to the appropriate standards for the vehicle under scrutiny.

DRIVING CYCLE DEVELOPMENT – The development of a driving cycle involves three distinct steps [18], namely the recording of driving condition under real operating conditions, the analysis of the data in order to fully characterize those conditions and the development of a composite driving cycle that adequately describes such conditions and parameters set out above.

Early attempts to construct driving cycles usually resulted in the development of a ‘synthetic’ cycle, which were descriptions of simplified modes based on durations and averaged values of acceleration,
deceleration and cruise speed [18]. These cycles were normally constructed by manually recording the durations, as in the case of the original Californian 7-Mode Cycle. Indeed, the European ECE-15 and NEDC both follow from such synthesized roots. Later, instrumented vehicles were used, making use of simple instrumentation, such as a fifth wheel such that speed could be measured continuously [18].

Such cycles are very useful for comparing vehicle operation for conformity or comparison purposes. However, they often do not adequately represent the real world and so are of questionable value in trying to estimate emissions from vehicles under genuine operating conditions. Thus a variety of tests have been carried out to more carefully describe local conditions for driving. Driving cycles have recently been developed for Athens [5], Edinburgh [19], Hong Kong [16] and for eleven cities in China [15], for example, as well as a pan-European project to describe a set of coherent driving cycles for Europe [13].

Such systems involve extensive instrumentation of vehicles and the use of external measurement devices, such as accelerometers, speed sensors and global positioning systems [15–17,21]. Early experiments in Hong Kong examined vehicle speed by means of an infrared photoelectric sensor pointed at the axle [16], while later experiments made use of a microwave speed sensor in the follow-car configuration [17]. While many of these systems can be put together quite effectively, they do have a serious limitation in that only a small number of vehicles can be examined at any one time and that such instrumentation tends to be expensive.

REAL WORLD EMISSIONS FROM VEHICLES – While traditional emissions tests have relied on standard, speed-based driving cycles, many studies have shown that the vehicle speed is only part of the issue and that engine operations and ambient conditions significantly impact the nature and quantity of pollutants emitted [12–15,20–24]. Thus, factors such as engine speed, engine load and engine thermal conditions should be considered to effectively estimate pollutant output in towns and cities as a function of vehicle activity. This is due to the fact that when a driver requires more ‘speed’ from the vehicle, the depressing of the throttle results in increased engine torque. That, in turn, derives from increased pressures and temperatures within the engine cylinder, frequently accompanied by changes in the air:fuel ratio from cycle to cycle. These changes in combustion behavior, and their interaction with any emissions control systems, are ultimately responsible for determining the nature and quantity of pollutants emitted from the vehicle. Examining vehicle speed alone, therefore, does not provide sufficient information for accurate prediction of pollutant emissions.

It is possible to account for these factors by exploiting the pre-existing sensor arrays within modern vehicles. Recent research papers have reported on using manufacturers’ proprietary software and interface systems in order to communicate with the test vehicles [6,19]. Such experimental arrangements have the advantage of being able to potentially communicate with a given make of vehicle, such as Ford or Renault as in the case of the cited studies. However, such systems are limited: they can only be used to examine those specific marks of vehicle.

ON-BOARD DIAGNOSTICS AND DRIVING CYCLE DESCRIPTIONS – An alternative approach is to make use of on-board diagnostic (OBD) data that is available for extraction from the vehicle's engine control unit (ECU) [12], in order to conveniently communicate with a large number of vehicles without the need to extensively instrument a vehicle. This paper examines the use of standard OBD data interfaces for the development of representative driving cycles, as well as synchronous evaluation of the operating conditions of the engine. Current OBD regulations require that vehicles support a minimum set diagnostic information, pertaining to emissions, to off-board equipment [25].

THE USE OF OBD SYSTEMS - The United States has been to the fore in implementing OBD. The California Air Resource Board (CARB) set out its first OBD requirements in 1988 [26]. This system required monitoring of emissions-related electrical components, and the storing of any generated fault codes within the on-board memory. Further refinement of the OBD requirements resulted in the monitoring not only of the emissions related components but also the components that are used in their monitoring. Thus, the scope of the OBD system has grown to allow for analysis of a wide variety of vehicle and engine parameters [25,26].

The European Union has been somewhat slower to adopt the various OBD protocols. All new petrol-fuelled Private Cars (PCs) and Light Duty Vehicles (LDVs) have been OBD compliant since 2000 [27]. Diesel vehicles, in both the PC and LDV categories, had to be OBD compliant by 2003 [26]. These regulations have been further extended to include diesel heavy duty vehicles (HDV) with effect from 2006 and 2008 [12,26].

A variety of OBD protocols exist [12]. Certain protocols are used almost exclusively in Europe and Japan (ISO 9141-2 for PCs and ISO 12430-4 for PCs and LDVs), while others are used in the United States (SAE J1850 for PCs and SAE J1708 for commercial vehicles). In the future, the Controller Area Network (CAN) bus will be used to handle communications between the vehicle and the tester [26].

EMPLOYING OBD SYSTEMS FOR DATA EXTRACTION – With a comprehensive suite of existing sensors, it is possible to extract driving cycle data and engine operation information without the need for multiple and costly measurement devices. An examination was made of the Standard [25] by researchers at UCD to determine which parameters are...
of most interest. Other researchers, using external measurement devices, have focused on vehicle speed and acceleration [16,17]. The ability to connect directly to the ECU has allowed other researchers [6,12,19] to examine certain engine operating characteristics such as engine speed, temperature and throttle position, as well as the parameters mentioned above. Connection to the OBD network of sensors also allows for determination of air- and fuel-flow parameters, in order to better estimate fuel consumption and emissions in near real-time [12].

**EXPERIMENTAL PROGRAMME**

The authors set out to examine the use of OBD systems (in conjunction with GPS data) to describe both kinematic and engine-based data for determination of driving cycles. Tests have been carried out on two principal test routes, one linking the City Centre to Dublin Airport along the M1 (The Northern Transport Corridor) and the other connecting Celbridge, in County Kildare, to the UCD Richview Campus in Clonskeagh (The Celbridge Transport Corridor). A number of ad hoc trips have also been recorded and normally correspond to trips that the drive had to make as a matter of course.

In total 16 hours of driving, covering some 380 kilometers, have been completed to date. Two test vehicles have been employed thus far, one a Volkswagen Polo (2006) and the other a Honda Civic (2003). These are both representative of vehicles within the Irish fleet [2]. Both vehicles have been operated in an ‘as-is’ condition, that is, no specific reconditioning has been carried out to facilitate this work.

During all periods of testing, regular visual inspections were made to ensure that the ‘virtual’ and ‘real’ instrumentation were in agreement. For speed inspections, each of the vehicle, the software and the GPS unit, were inspected. Good agreement was observed, except where canopyning affects or canopy layers interfered with GPS operations [12,20].

<table>
<thead>
<tr>
<th>Table 1 - Test vehicle data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Characteristics</strong></td>
</tr>
<tr>
<td><strong>Make &amp; Model</strong></td>
</tr>
<tr>
<td><strong>Engine Size</strong></td>
</tr>
<tr>
<td><strong>Model Year</strong></td>
</tr>
<tr>
<td><strong>Standard</strong></td>
</tr>
<tr>
<td><strong>Transmission</strong></td>
</tr>
<tr>
<td><strong>Odometer</strong></td>
</tr>
</tbody>
</table>

**EXPERIMENTAL EQUIPMENT**

In order to show that that this system can accurately collect, parse and log data, a minimal compliment of equipment is employed for each test comprising the OBD scan tool itself, a laptop and the GPS device, along with the appropriate cables. For each test, the OBD system has logged engine speed (rpm), coolant temperature (°C), vehicle speed (km h⁻¹, m s⁻¹), distance travelled (m) at 1 Hz. A number of problems with the load or throttle position interpretation meant that such values were not always extracted. Both the laptop and the GPS device were synchronized using the Microsoft ActiveSync software package to allow for future synchronization of the collected data such that a suitable comparison of the two speed traces can be made.

In practice, and as a form of validation process as well as a data collecting exercise, the OBD data extraction system is complemented by a Trimble GeoXT global positioning system (GPS). The GPS system is used to gather data as accurately as possible, rather than at high logging speeds, where spatial accuracy is not as good.

A specially developed piece of software has been used to extract four vehicle parameters from the OBD system: vehicle speed (PID 0D), in order to develop a traditional driving cycle; engine speed (PID 0C) and load (PID 04), since these parameters describe the engine condition and consequently the pollutant formation and outputs; and coolant temperature (PID 05) which allows for examination of cold start emissions contributions. Load values are only reported as a percentage but can be converted to absolute load values in conjunction with the appropriate torque maps [12].

Functionally, the program is divided into four parts: connect, request, parse, and log.

The first element establishes a serial connection between the vehicle and the off-board testing equipment, (a laptop). Having established the appropriate link with the vehicle, a queuing function generates a series of timed requests for data. The serial nature of the connection, and the fact that certain systems only allow for one line data access [26] means that a specified delay must be left between the requests to allow for the return of the response. A maximum of five parameters can be successfully extracted in 0.92 s [12]. Thus extraction delays are normally selected based on the number of sought parameters and the desired data logging rate.

Responses are then received by the program and interrogated by a parameter recognition sub-routine before they are stripped by the parsing algorithms and the hexadecimal byte data are converted to decimal values. These decimal values are then scaled according to the provisions of the Standard [25]. The final functional module is the data logging system, set at 1 Hz.

The GUI displays the collected data in real time and has been constructed to give the generalized appearance of a vehicle dashboard. Data on the current connection to the vehicle, data storage paths, date and time and test duration are all given, along with a series of dials and panels to display vehicle (vehicle speed) and engine (engine coolant temperature, engine speed and throttle position) operating conditions. The vehicle speed value,
returned in kilometers per hour [25], is further manipulated to give equivalent values in meters per second, which is then used to instantaneously generate a total distance travelled in meters. A further conversion scheme allows for the display of distance in kilometers and miles, since both are representative of the current Irish PC fleet.

A trace of each parameter is also shown to allow the operator to see any striking anomalies during testing. The ability to see such anomalies has led to significantly better data recording, since time need not be wasted on gathering data when the system is clearly malfunctioning, or if the OBD connection is lost.

The collected data have been analyzed using a combination of SQL and spreadsheet data manipulation techniques. Initial analysis has focused on the comparison of OBD and GPS data to examine the correlations and to investigate the suitability of the OBD to adequately represent driving patterns.

**COMPARISON OF OBD AND GPS DATA**

In keeping with current practice, analyses have examined the various aspects of vehicle kinematics. Additionally, engine load and speed, and coolant temperature, have been analyzed. Such engine data is of particular interest when examining acceleration profiles or vehicle idle conditions. Consider the kinematic data described in Figure 2. This is typical of the Northern Corridor Route, between Dublin Airport and Bolton Street, a total distance of just over 11 km. The average speed of the vehicle was just 8.2 m s$^{-1}$ (29.5 km h$^{-1}$). The average speed while driving, defined here as the cruise speed, was somewhat higher, standing at 12.24 m s$^{-1}$ (44.1 km h$^{-1}$). These data indicate that a significant portion (33%) of the journey duration is spent at idle – a total of 342 seconds - during which time fuel is used and pollutant and greenhouse gases are emitted with no benefit to the driver.

Figure 1 - Schematic of the OBD data extraction software showing the flow of data.

Figure 2 - (a) Kinematic and (b) engine data for vehicle testing on the Northern Corridor

Figure 3 shows, for short tests on both the Honda and the Volkswagen, the relative distribution of engine load and engine speed. The long idle durations, based on analysis of the time-speed plots, are reflected in the clustering of data point in the lower left hand corner of the graph. It is also possible to see that at low engine speeds, very high loads are requested under certain circumstances (93% engine load at 990 rpm. It should, of course, be noted that this is 97% of the maximum load for that engine speed and not the maximum rated torque). This is due to the driver wanting to achieve very high acceleration from stop. Clearly, such analysis of engine speeds and loads allows for a more robust description of aggressive driving without simply relying on acceleration. Again, this highlights the benefits of extracting simultaneous engine and vehicle data.
Figure 3 – Engine load and engine speed values as a function of speed for vehicle tests in the Honda Civic Test Vehicle. A significant degree of clustering can be observed at low engine speeds and low engine loads. The light grey dots refer to Volkswagen operation and the darker dots pertain to Honda tests.

With regard to analyzing the suitability of the OBD system to describe vehicle kinematics, it is useful to examine the correlation between GPS, OBD and Geographic Information Systems (GIS) data.

Consider the data shown in Figure 4. We can see significant correlations between speed values obtained by both GPS and OBD. This segment describes a journey along the M50 motorway from just south of the N7 (Naas Road) Junction to the Sandyford Road. As shown in Table 2, the overall correlation between average values in terms of speed and acceleration is remarkable. The discrepancies in acceleration and deceleration are due to a brief loss of signal using GPS, at about 3421 s, half way through the actual cycle shown, and the observed canyoning effects at the entrance to the urban area of Dundrum at the end of the segment.

The differences in idle durations are due to the fact that the GPS is sampling at 0.5 Hz. The system, although set to a sampling interval of 1 s, only samples after a delay of 1 s, effectively leading to timestamps of 2 s intervals.

Table 2 – Comparison of OBD and GPS data collected for the Celbridge Transport Corridor

<table>
<thead>
<tr>
<th>Durations Sequence</th>
<th>Idle (s)</th>
<th>Maximum Speed (m s(^{-1}))</th>
<th>Average Speed (m s(^{-1}))</th>
<th>Cruise Speed (m s(^{-1}))</th>
<th>Maximum Acceleration (m s(^{-2}))</th>
<th>Maximum Deceleration (m s(^{-2}))</th>
<th>Distance (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBD</td>
<td>1842</td>
<td>508</td>
<td>20.83</td>
<td>5.89</td>
<td>7.52</td>
<td>0.64</td>
<td>13792</td>
</tr>
<tr>
<td>GPS</td>
<td>1842</td>
<td>363</td>
<td>20.83</td>
<td>6.02</td>
<td>7.18</td>
<td>0.84</td>
<td>13785</td>
</tr>
</tbody>
</table>

For the entire test segment, the total distance, determined using a GIS model of the Dublin road network, is given as 706 m. Analysis of the OBD data, by means of integrating the speed data, yields a distance travelled of 706.21 m. The GPS information shows a significant over-estimation of distance. (This contrasts somewhat with the slight underestimation of distance described in Table 2). This discrepancy is mostly based on the method used to determine distance. The OBD speed data (m s\(^{-1}\)) is integrated over the cycle extracted and a value for distance is determined. Where the GPS signal is lost, this is not possible, since, in the case of Figure 5, for example, a straight line from 0 m s\(^{-1}\) at 200 s to 8 m s\(^{-1}\) at 370 s would result in the formation of a significant linear segment in the distance-time graph.
Figure 4 – OBD and GPS data covering a time segment of some 2500 s. The segment relates to motorway driving along the M50. There is some loss of GPS signal at just before 4500 s due to the vehicle entering the built up area of Sandyford and Dundrum.

Figure 5 – OBD and GPS data covering a time segment of 400 s in the Drumcondra area of Dublin. OBD signal remains clear while GPS signal is lost due to shadowing affects of the trees. In the early stages (0 – 200 s), both OBD and GPS correspond well.

Figure 6 - Ariel photograph of the Drumcondra area, showing waypoints for the segment described in Figure 4. Image provided by Google Earth, 2008 © DigitalGlobe.

Table 3 - Details of Drumcondra data sample. The GPS cycle distance (denoted by *) was calculated using GIS data for the Dublin Road network

<table>
<thead>
<tr>
<th></th>
<th>0–200 s</th>
<th>201–400 s</th>
<th>Cycle</th>
<th>0–200 s</th>
<th>201–400 s</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (m)</td>
<td>208.93</td>
<td>497.28</td>
<td>706.21</td>
<td>208.31</td>
<td>-</td>
<td>706.27*</td>
</tr>
<tr>
<td>Average Speed (m s(^{-1}))</td>
<td>1.04</td>
<td>2.49</td>
<td>1.76</td>
<td>0.98</td>
<td>-</td>
<td>1.31</td>
</tr>
<tr>
<td>Acceleration (m s(^{-2}))</td>
<td>1.06</td>
<td>0.87</td>
<td>0.94</td>
<td>1.01</td>
<td>-</td>
<td>1.24</td>
</tr>
<tr>
<td>Deceleration (m s(^{-2}))</td>
<td>-0.85</td>
<td>-0.55</td>
<td>-0.66</td>
<td>-0.87</td>
<td>-</td>
<td>-0.85</td>
</tr>
</tbody>
</table>
CONCLUSION

Traditional methods for the extraction of kinematic data from vehicle trips have relied on external devices and have considered only a limited number of parameters based on vehicle speed. Since vehicular emissions are really a function of the operation of the engine, an understanding of the engine operation is clearly required in order to fully quantify emission. To this end, two systems for data extraction have been evaluated, one following the strictly traditional approach (GPS) and the other combining both the traditional and the subsequently identified engine operating parameters required to develop a full picture of emissions. The systems considered are 'low-cost' both in terms of monetary outlay for equipment and in terms of processing overhead. The combination of kinematic data with synchronous engine operation information can significantly increase the level of understanding that researchers have of driving conditions in urban areas and the consequential release of pollutant and greenhouse gases and energy consumption.

Tests carried out using the UCD Driving Cycle Data Extraction System have been very positive. A useful quantity of engine and vehicle data has been extracted so far, and testing will continue to elucidate diurnal, weekly and monthly variations in drive-cycle characteristics. This information will allow for the development of a series of driving cycles which can then be applied to a range of scenarios, such that more informed decisions about traffic-based pollutant emissions, energy consumption and GHG emissions can be made, both now and in the future.

The successful use of existing sensors and simple data extraction systems has been demonstrated, meaning that any vehicle can effectively become an instrumented test vehicle. No modifications need to be made to the software or the hardware in order for transfer from one vehicle to another. Currently, very generic information is extracted but the potential to extend the parameters that can be extracted is enormous. The Standard [25] describes the extraction of fuel consumption rates, catalytic converter temperatures and lambda sensors, clearly allowing for a significant level of information to be gathered with regard to emissions and fuel economy in real time. However, not all of these parameters have been fully implemented by manufacturers. It should be pointed out that manufacturers are currently not obliged to make such parameters available Even so, the currently extracted data, namely engine speed, engine load, vehicle speed and coolant temperature is in keeping with current best practice and does extend the ability of automated test equipment for driving cycle data extraction by allowing for detailed (second-by-second) examination of both vehicle kinematics and the effect of engine operation on emissions.

Comparison of the OBD data with external data collections, such as GIS and GPS shows that the system is well able to adequately represent distances and speeds. As an extension of this fact, the authors are satisfied that the engine speed, engine load and other parameters of interest are also adequately represented. Other researchers have shown that OBD data extracted in Rome and compared to dynamometer tests have shown similar correlations [12]. Again, in-vehicle examination of the ‘real’ and ‘virtual’ dashboards would corroborate this conclusion. Ideally, the low-cost system will be deployed to many vehicles simultaneously to allow for more extensive data collection campaigns is the Dublin Area.

In carrying out the tests described above, the combination of the OBD and GPS data has been shown to be quite useful, despite the losses in signal in the GPS. While the goal of the data extraction campaign is to develop a series of kinematic sequences for the derivation of a driving cycle, other work has raised questions about the validity of using a sequence of kinematic sequences that is, speed taken between successive periods of idle, need not necessarily be representative of fluid vehicle movements [12,16]. Instead, the authors propose to examine the kinematics of individual links within the network, as a function of time or day and the consequential traffic flows. This work will is being carried out as part of the Urban Environment Project at UCD. The combination of traffic data and kinematic and engine operation information for specific routes and links should yield significantly better estimations of emissions, based on actual vehicle activity, in given locales and hence allow for more accurate determination of the effects of pollutant exposure on the populations in these areas.

ACKNOWLEDGMENTS

The authors would like to acknowledge the funding support offered by the Urban Environment Project through the EPA (Ireland) ERTDI Programme.

REFERENCES