STANDARD QUALITY CHECKS FOR WEIGH-IN-MOTION DATA

Mr. van Loo has been involved in the field of Weigh-In-Motion technology and applications for more than 20 years. Since 2013 he has been working as an independent expert and consultant for both users and vendors of WIM systems.

Mr. Lees joined TDC Systems in 2010, prior to this he worked for 35 years with the UK Governments Department for Transport where he was Head of the automatic traffic and weigh in motion statistics branch. He has been involved in the field of Weigh-In-Motion for more than 20 years.

Hans van Loo
Corner Stone International Sagl.
Switzerland

Andy Lees
Q-Free TDC
United Kingdom

Abstract

This paper describes the development of a simple standardised test to make a first assessment of the quality of the data measured by Weigh-in-Motion (WIM) systems as part of the European EcoVehicle project ∑!7219. First the background, reasons and requirements for such a test are explained. These are followed by an explanation of the starting points and the approach used in the development of the test and its criteria. The individual quality checks and criteria were evaluated using actual WIM data from sites in four different European countries. An overview is given of the current procedures for the management of the quality of data in different European countries. Finally, a summary is provided presenting the main conclusions and recommendations for the future application of these tests and their potential to serve as a basis for an international standard data quality assessment tool.

Keywords: WIM, Footprint, EcoVehicle, Data, Quality, Checks, Validation

Résumé

Cet article décrit le développement d'un test standardisé simple pour faire une première évaluation de la qualité des données mesurées par des systèmes de pesage en mouvement (WIM) dans le cadre de la EcoVehicle européenne projet ∑! 7219. Tout d'abord, le contexte, les motifs et les exigences pour un tel test sont expliqués. Elles sont suivies par une explication des points de départ et l'approche utilisée dans le développement du test et ses critères. Les contrôles de la qualité individuelle et les critères ont été évalués à l'aide de données réelles de WIM de sites dans quatre pays européens. Un aperçu est donné des procédures en vigueur pour la gestion de la qualité des données dans différents pays européens. Enfin, on trouvera un résumé présentant les principales conclusions et recommandations pour l'application future de ces tests et leur capacité à servir de base pour un outil d'évaluation de qualité des données standard international.

Mots-clés: WIM, Footprint, EcoVehicle, données, qualité, vérifications, validation

1. Background

One of the objectives of the Eureka Footprint project ∑!2486 (Mayer, 2009) was the development of a method to identify vehicles by means of their "Environmental Footprint". The damage caused by the dynamic loading of a heavy vehicle on the road – and rail – infrastructure is one of the aspects of a vehicle’s environmental footprint. The quality (accuracy, reliability and stability) of the measurement data used in any study directly determines the quality of the results and conclusions of the study. The WIM data from different European countries used in the Footprint project (Poulikakos, 2009) have shown variability that could not be explained by mere differences in the national loading regulations alone. The differences in the data may have originated from variations in the local traffic flow, the environmental conditions or from differences in performance of the WIM systems, e.g. the type of WIM technology used or, possibly, structural measurement errors.

For a realistic comparison of the environmental impact of different vehicles – and in fact any other study on the impact of heavy truck traffic - the quality of the WIM data must be verified. This is especially important when comparing the effects in different European countries since the measurement data will come from different WIM systems based on different technologies, operating under different conditions and owned by different users. At present there is no uniform European standard procedure to make an assessment of the quality of WIM data from different systems. As a result many studies may be based on WIM data with little - if any - idea of the quality of the data and as a consequence some conclusions may be based on erroneous data.

A full guarantee of the quality of WIM data can only be given after an extensive evaluation of the performance of the WIM system, the traffic and environmental conditions over a long period of time (e.g. 1 year). In most cases, such an extensive evaluation is too time consuming, too expensive to carry out and also too complicated since it requires an in depth knowledge of WIM systems and sensor behaviour. A limited and simplified evaluation could fill the gap between an extensive and expensive test and no test at all, allowing for a quick assessment of the quality of the WIM data.

2. The EcoVehicle ∑!7219 Project

The goal of this European cooperative project is defining road and rail vehicles with a low environmental footprint (Lees, 2014). The principal tasks include: analysing data from real time measurements, defining limit values for environmental friendly vehicles and defining a combined environmental index for vehicles. An important EU objective is to reduce the environmental impact of transport. Characterising the environmental impact of individual vehicles enables the polluter pays principle to be applied to land transport. One of the parts of the EcoVehicle project focussed on the dynamic loading of heavy vehicles and included:

a. the development of a limited and simplified evaluation for a quick assessment of the quality of the WIM data;
b. the first international benchmark on the data quality management, procedures and criteria used by different users of WIM systems in Europe.

It is hoped, this project could lead, in time, to the direction of a harmonised European criteria and procedures for Data Quality Management for WIM systems. This paper will describe the results from both parts.

3. Data Quality Assessment

3.1 Objective

No WIM system can produce perfect data, even with high quality equipment and ideal site conditions. Data files are more than likely going to contain some invalid data. Regardless of the minimum data quality requirements are, any WIM system should be monitored and maintained to produce the best possible data given the system's potential. The key is to keep bad data to a minimum and to quickly recognise, identify, isolate and correct the cause of erroneous data. (FHWA, 2009)

Therefore, the objective of this part of the project was to develop a basic set of tests and criteria that will allow the user to make a quick verification of the quality of the data from any WIM system in Europe. These tests could then be used to compare the relative quality of different WIM sites (the quality of the data from site A is better than that of site B) and, if possible, to give an indication of the absolute quality of the data of a particular site (the data from site C has a quality that is sufficient). It is important to realise that these quality tests will not be able to distinguish between variations in the measurements by the WIM system and variations in the truck traffic at a certain site. This means that in case the test results would produce an "insufficient" verdict on the quality of data because of large variations in the WIM data, the reason for this could be explained by variations in the traffic flow and not because of the WIM system. In this case, the results of the tests should be interpreted as: "Do not use this data without additional checks on the quality of the data."

In general, the tests will look at the stability of certain elements or characteristics of the measured data. These test will provide an idea of the relative quality of the WIM data however may still contain a stable – and possibly significant – measurement error. The selection of the characteristics was based on an evaluation of international literature on WIM data quality management and the practical experience from the authors:

- The United States, the Long-Term Pavement Performance (LTPP) program initiated by the Transportation Research Board, the Federal Highway Administration and the American Association of State Highway and Transportation Officials (FHWA, 2010);
- The Netherlands, the WIM-NL network (currently consisting of 20 systems) developed by Rijkswaterstaat and used by the Transport Inspectorate for their weight enforcement program (Telman, 2013);
- South Africa, the South African National Roads Agency Ltd (SANRAL) has developed statistical methods for the calibration and quality assessment of the data from their network of about 50 WIM systems (De Wet, 2010).

3.2 Starting Points

In the development of the checks, the following starting points were used:

- the tests should give a first indication of the quality of the data measured by a certain WIM system;
- the tests should be easy to perform by anybody irrespective of whether they are specialists in Weigh-In-Motion or statistics or not;
- the calculations required for the tests should be available – or be easy to implement - in standard software like Excel, Access (or similar);

- it should be possible to do the tests on all measurement data from all different WIM systems currently operational in Europe;

the tests will be done on a limited sample from the WIM data only, e.g. one week representative of normal operational conditions. The test sample should be large enough to include possible variations over a few days and be small enough to be handled in Excel. All papers must be submitted for assessment in English. Authors from French, Portuguese and Spanish speaking countries and those who may get translations are encouraged to also submit their final papers (after revisions) in another language. Such papers will be reviewed and published on the ISWIM web site, and the best ones may be submitted to a French, Portuguese or Spanish journal with the authors' agreement.

3.3 Quality Checks and Criteria

Determination of the tests and criteria to assess the quality of the data. In other words this means finding characteristics of certain types of vehicles that show a very small variation in daily practice and are commonly found throughout Europe. This can either be caused by international regulations for heavy goods vehicles (examples 1 and 2) or by standards in vehicle design (examples 3 and 4). The following examples of such characteristics were used in the quality checks:

1. The vehicle length of Truck+Trailer combinations and that of Tractor+Semi-trailer (articulated) combinations. For most EU member states the maximum allowable lengths for these combination are respectively 18.75m and 16.50m;
2. The Gross Vehicle Weight (GVW) of 3 axle Trucks and that of 5 axle Tractor + Semi-trailer (articulated) combinations. For most EU member states the maximum allowable GVW's for these combination are respectively 26ton and 40/44ton;
3. The axle load of the first (steering) axle of – fully loaded - 5 and 6 axle articulated vehicles. International experience has shown that the load on this axle lies normally in a narrow bandwidth between 6.5 and 7.0 tonnes;
4. The axle distance between the 2nd and 3rd (driven) axles of 6 axle Tractor + Semi-trailer combinations. International experience has shown that the distance between these axles is very stable at 1.30m as this allows the highest axle loads;

3.4 Sample Data

The objective was to collect a sample of WIM data (one week of data in case of a WIM-site with high traffic volumes) from different users, different countries and different technologies. The aim was to try to collect more data from different countries, if possible based on different technologies and if possible data from 'good' and 'bad' quality sites to be able to see if the criteria are able to detect the bad data.

For the project, measurement data from two sites in each of four different European countries have been collected and evaluated. We have also deliberately included a ninth site which we knew was not working correctly and therefore providing erroneous data. We have included this site to highlight the value of identifying faulty sites. This site will appear as Site 9 in the graphs and tables.

It should be stressed that in all cases, the exact location of the WIM sites, the type of equipment deployed and the manufacturers have been kept anonymous to ensure any unintentional bias cannot be applied to the results. It was also felt that anonymity of the equipment should be maintained to avoid any unnecessary comparisons between technologies and vendors since this is not the objective of this project.

4. Data Analysis

4.1 Gross Vehicle Weight

The first analysis was carried out on the type of articulated goods vehicle that is probably the most frequently encountered vehicle on European roads; the two axle tractor and three axle semi-trailer unit.

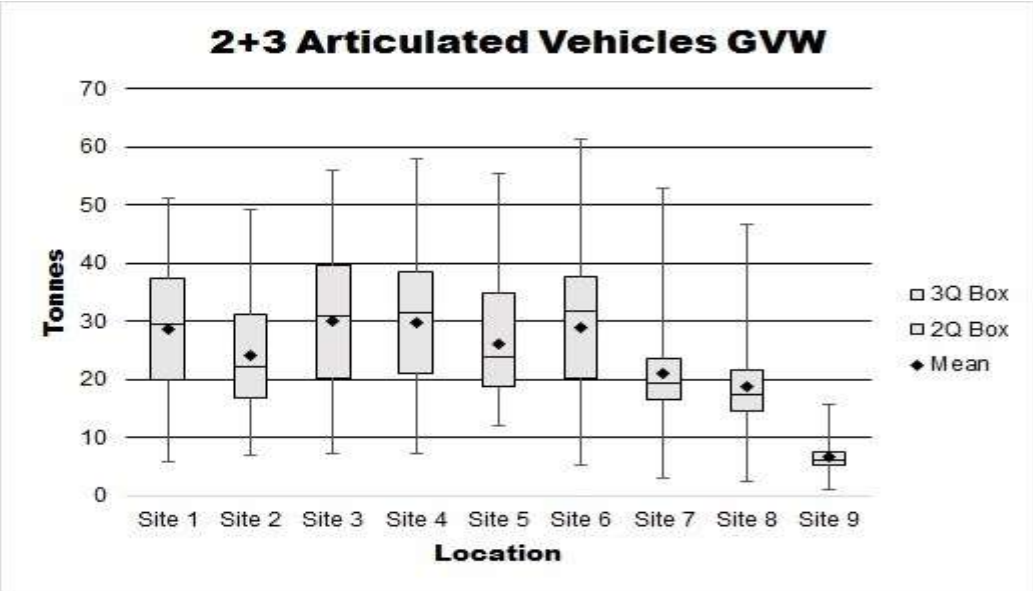

Figure 1 – Analysis of 5 axle articulated vehicles

Table 1 – Overview of 5 axle articulated vehicles

Site	1	2	3	4	5	6	7	8	9	Average
Mean	28.62	24.09	30.21	29.79	26.02	29.04	20.99	18.78	6.60	28.21
St. Dev	9.18	8.42	10.19	9.53	8.95	9.55	6.64	5.85	1.94	10.38

The average GVW for five axle articulated trucks is expected to be between the maximum permissable weight (40-44 tonnes) for international transports and the weight of empty trucks (around 20 tonnes). Hence an average value somewhere between 25 and 30 tonnes with a variation of ± 10 tonnes could be expected. When looking at the chart in Figure 1 it is clear that there is a discrepancy using sites 7 and 8 in any further analysis as the average weights are significantly lower than those seen elsewhere. Whether this is down to a measurement error in the data or local traffic conditions, it is unsure without further inspection. For site 9 the average GVW is extremely low and it was known that the data was faulty.

4.2 Steering Axle Load

For the next test, the steering axle weight of two axle tractor + three axle trailer articulated combinations was examined. More specific the axle loads of the first steering axle of these vehicles when fully laden were examined, i.e. in excess of 30 tonnes gross vehicle weight

(GVW). Obviously this was reliant on the "accuracy" of the test data to determine whether 30 tonnes GVW was met but this limit is actually not very strict and the results were rather consistent. For these next two tests it is clear from the previous results that Site 9 would not be able to be tested as none of the records for this vehicle class was measured in excess of 30 tonnes.

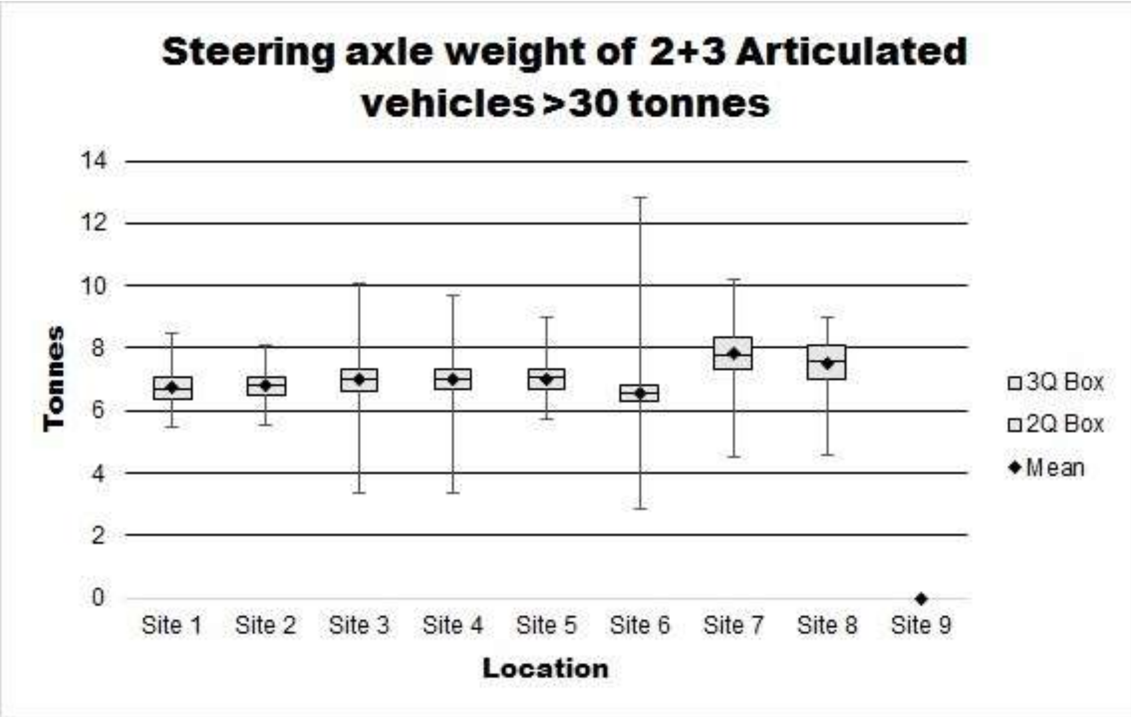

Figure 2 – Analysis of 1st axle load of five axle articulated vehicles

Table 2 – Overview of 1st axle load of 5 axle articulated vehicles

Site	1	2	3	4	5	6	7	8	9	Average
Mean	6.73	6.79	6.99	7.03	7.03	6.54	7.84	7.55	0	6.97
St. Dev	0.55	0.44	0.54	0.50	0.52	0.46	0.82	0.74	0	0.55

If site 6 is removed from the analysis, it has some extremes, the mean weight of the steering axle falls within 1 tonne of each other and there appears to be consistency in the 2nd and 3rd quartiles. Based on international experience the expected value for the first axle load is between 6.5 and 7.0 tonnes with a small variation. The first six sites follow these expectations, while sites 7 and 8 do appear slightly out of line with a higher average axle load and a larger variation. This should then alert the user to perhaps reconsider using the data from those sites in any analyses. Especially when this is combined with the lower average GVW for this vehicle class at these two sites.

Again the reason for these difference could also originate from the characteristics of the truck traffic at the site, e.g. a high percentage of light – partially loaded - vehicles that obviously have a lower GVW but tend to have a slightly higher axle load on the first axle because of a different distribution of the loads.

4.3 Vehicle Length

In addition to the above two weighing related tests we checked vehicle length, a parameter that is not reliant on the WIM sensors but the inductive loops. This time, again using the same >30 tonne articulated two axle tractor/three axle trailer combination.

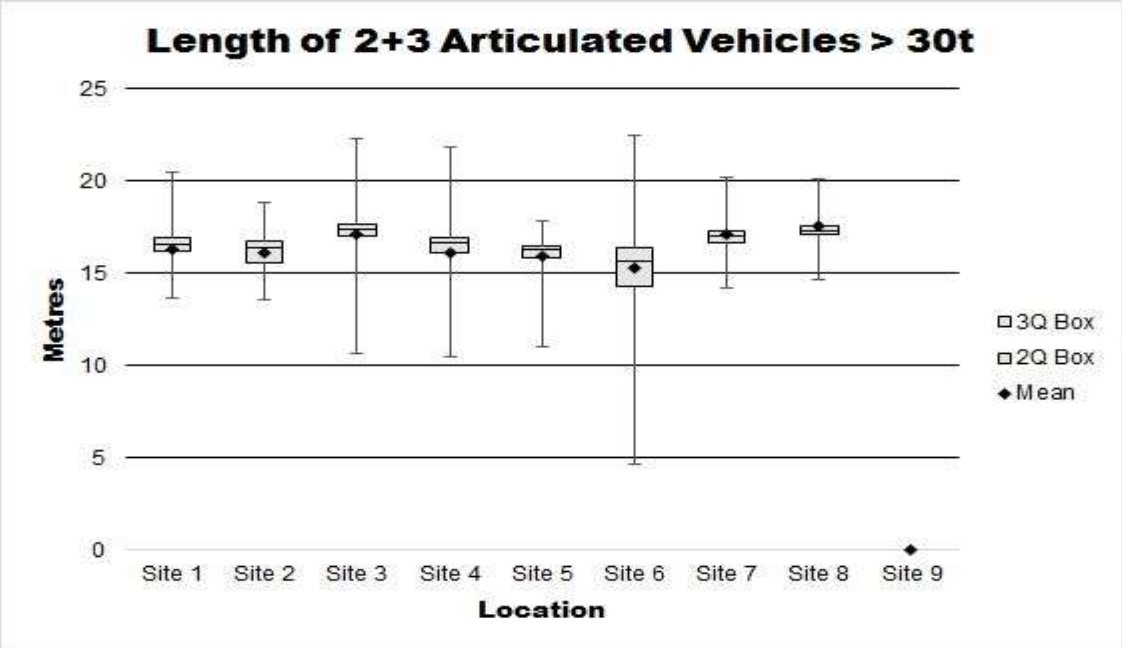

Figure 3 –Analysis of length of 5 axle articulated vehicles

Table 3 – Overview of length of 5 axle articulated vehicles

Site	1	2	3	4	5	6	7	8	9	Average
Mean	16.29	16.13	17.15	16.14	15.95	15.33	17.09	17.56	0	16.39
St. Dev	1.03	1.02	0.97	1.40	0.99	1.45	1.03	0.99	0	1.40

The legally permissible maximum length for this type of trucks is 16.5m in almost all EU-member states including the four countries considered in this test. Since transport companies and vehicle manufacturers seek to optimise their vehicles within the legal boundaries it is expected that the average vehicle length will be close to 16.5m with a small variation. The mean lengths of the vehicles across all eight sites is within 1.5 metres which is encouraging and would allow the user to have some confidence in using the data for length and classification purposes.

Although individual vehicles may exceed this maximum value it is unlikely that the average value is higher than the maximum limit. This indicates that the length measurement of all sites is 0.11m less than expected yet it has a small variation. It is interesting to note that sites 7 and 8 exceed the maximum length for this type of vehicle but it is known that data from these two sites were likely to have anomalies due to lack of site maintenance and recent calibration. In all cases, this kind of structural measurement error can easily be compensated through calibration. Only site 6 shows a slightly higher variation but this could originate from the local traffic conditions.

4.4 Axle Distance

The tests on the axle distance between the 2nd and 3rd (driven) axles of 6 axle Tractor + Semi-trailer combinations has not been implemented due to difficulties with the limited detail in the vehicle classification in the data from a few of the sites. In other words it has not been possible to filter out this specific vehicle class needed for the test.

5. Procedures for Data Quality Management

5.1 Questionnaire

It is possible that the way users maintained and checked their systems may have an effect on the quality of the data. Therefore, it was decided to question the users about how their sites were used and how they were maintained. A questionnaire was developed containing 18 questions within 4 main headings, these were;

- General usage; what is the main use and specifications for your WIM data?
- Site Maintenance; what are the procedures for maintenance and calibration of the systems?
- Data Checks; what are the procedures for checking the collection of measurement data?
- Data Quality Control; what are the procedures for validation on the quality of the data?

The idea behind these questions was to try and ascertain the extent to which users maintain their sites and, in particular, the frequency in which they analyse and check the quality of the data obtained, i.e. procedures for quality control of their systems and data.

5.2 Results

In terms of general usage, there is a diverse spread between statistics, pre-selection and tolling. However, nearly all of the sites are maintained to the specifications laid down in COST323 although the strictness of individual users' interpretation of the specified procedures are highly variable. This is particularly noticeable when responses are analysed for the calibration methods and frequency.

From the responses received regarding site maintenance there appeared to be no direct correlation between site maintenance and data quality. As an example, the user who only calibrated every two years (the longest period) was not the worst performer in terms of quality whilst one of the users' who calibrated every six months had the worst performing sites. In fact one of the users' didn't have a regular site maintenance routine, although they calibrated every six months, and their data seemed acceptable from the tests carried out.

Regarding data checks; all of the respondents carried out data checks at regular intervals. All had routines whether manual or automatic that carried out these tests and were able to identify faults at reasonably short notice. The frequency at which the checks were performed varied from daily to monthly, surprisingly, the owner of the worst performing sites carried out their quality controls on a daily basis.

It is was interesting to see that all of the users employ robust, regular maintenance and quality control of their data but it has shown the tests we have developed can bring into question the reliability of some of the quality controls employed. However, it should be stressed that the sites chosen for this paper were random and may well have been known to the users that some of them had reliability issues prior to our examinations of the data.

6. Conclusions

- A set of tests and criteria were developed and will allow users to make a quick verification of the quality of the data from any WIM system in Europe;
- The tests look at the stability of different characteristics of the data measured by the WIM system;
- These tests can be used to compare the relative quality of different WIM sites (the quality of the data from site A is better than that of site B);
- These tests can be used to give an indication of the absolute quality of the data of a particular site (the data from site C has a quality that is sufficient);
- For this, criteria were developed to assess the absolute quality derived from the maximum legal limits for international goods transport and values common for certain types of trucks;
- The criteria are:

Criterion	Min. Value	Max. Value
Av. GVW of 3 axle rigid	15t	20t
Av. GVW of 5 axle articulated	25t	40t
Av. Steering Axle Load	6.5t	7.0t
Av. Vehicle Length	15.5m	17.5m
Av. Axle Distance	-	-
Variation in # of registrations	-	-
Percentage of unclassified	-	5%
Percentage of meas. errors	-	5%
# hours without registrations	-	5 per week

- The results from a questionnaire on procedures for data quality management by various users did not show a clear correlation between the procedures used and the quality of the data.

7. Recommendations

- Since the outcome of the test is sensitive to the choice of what week of data is used. The selected weeks should represent normal operational conditions. Weeks with known variations due to holidays, road works or extreme weather conditions should be avoided.
- In case of a negative result of these tests this should be interpreted as: “Do not use this data without additional checks on the quality of the data.”
- In case of a positive result this should be interpreted as: “There are no reasons to suspect the quality of this data however this is not a guarantee”;
- By repeating the tests on data of one system from a number of different weeks from different periods over a year, the results of the will give a more reliable indication of the actual performance of the system.
- A further investigation is needed on the relation between site maintenance, data quality procedures and the quality of the measured data based on a larger set of different WIM systems.

8. References

1. G. de Wet, (2010), “Post-calibration and quality management of weigh-in-motion traffic data, (full thesis)“, Stellenbosch University, South Africa;
2. Jacob, B., O’Brien, E., Jeheas, S. (2002), Weigh-In-Motion of Road Vehicles- Final Report of the COST-323 Action, LCPC, Paris, France;
3. FHWA, (Federal Highway Administration), (2009), FHWA-IF-09-038, WIM Data Analysts Manual
4. FHWA, (Federal Highway Administration), (2010), Long Term Pavement Performance (LTPP) Programme; Information Management System, IMS Quality Control Checks, Virginia, USA;
5. Lees, A., Van Loo, H., (2014), Project plan WIM Quality International Benchmark, internal document as part of the EcoVehicle Project;
6. Mayer, R, Poulikakos L, Lees A, (2009), Impacts of vehicles with infrastructure and environment as measured by Footprint measuring systems, Eureka-Empa Report;
7. Poulikakos, L., Lees, A., Heutschi, K., Anderegg P., Comparisons of the environmental footprint of heavy vehicles, Transportation Research, Elsevier;
8. Telman, J., Hordijk, J., (2013), Monitoring prestaties WIM systemen, Rijkswaterstaat Dienst Verkeer en Scheepvaart, Delft, The Netherlands (in Dutch);
9. Van Loo H. (2001), ‘WIM-Hand Project, 1st Interim Report’, DWW-Publication: IB-R-01-09, Ministry of Transport, Public Works and Water Management