# ASSESSMENT OF THE ACCURACY AND CLASSIFICATION OF WEIGH-IN-MOTION SYSTEMS

# STATISTICAL BACKGROUND

**Bernard Jacob** [1]

*1. Department for Science and Technology, Laboratoire Central des Ponts et Chaussées, 58 Bd. Lefèbvre, F75732 Paris, Cedex 15, France; Tel: +33-1-40435312; Fax: +33-1-40435498; Email: jacob@lcpc.fr*

*International Journal of Vehicle Design - Heavy Vehicle Systems, Vol. 7, Nos. 2/3, 2000, 136-152*

.

**Abstract**

This is the first part of a two-part paper which addresses the issue of accuracy in weigh-in-motion (WIM) systems. This part describes and where necessary develops the background necessary for any system of accuracy classification applied to a WIM system. The second part describes a draft European specification for WIM of road vehicles, prepared by the COST 323 management committee. The system of accuracy classification used in the specification is based on the principles described in this part.

The common metrological definitions of accuracy are not applicable to WIM systems because the measured quantities are not traceable. Moreover, many WIM sensors may not be used to weigh static loads and/or cannot be checked with calibration masses. The WIM operations are not fully repeatable. Therefore particular definitions, sound reference values and statistical tools of which users should be reminded, are presented. Relevant formulas and methodology are developed and illustrated with numerical examples.

**Keywords**

Metrology, accuracy, precision, weigh-in-motion, WIM, WIM systems, calibration, acceptance test, statistics.

# 1. Scope and WIM accuracy related issues

## *1.1. Motivations and objectives*

The accuracy of Weigh-In-Motion (WIM) data is an issue of particular interest for both the suppliers of WIM systems and the customers or users. Different levels of accuracy are required by the users according to the application proposed, COST323 (1997-b). The suppliers often promote the performance of systems on the basis of accuracy and would welcome a common framework for its measurement. WIM data users also need to know the accuracy of the vehicle and axle load results that may be taken into account in their studies, and the sensitivity of the final results to data inaccuracy.

National and International specifications and standards are required to facilitate the relationships between customers and suppliers, to clarify the real system performance under different conditions of use, and to guarantee results. Any specification or standard must refer to accuracy levels, classes, etc., and to calibration and checking or testing procedures to verify the system performance claimed. It is therefore necessary to have precise definitions, mathematical, statistical and physical background, and harmonized procedures.

This part of a two-part paper on WIM system accuracy classification identifies the specific difficulties relating to WIM measurements and explains how they can be addressed. The second part describes an implementation of the principles developed here, in a draft European WIM specification (COST 323, 1997). In this part, precise definitions and concepts are either introduced or reviewed, some of them well known in the fields of Physics or Mathematics, but not by many WIM users or suppliers. Statistical tools, such as confidence interval estimation theory, will be reviewed in the context of WIM, to clarify the proposed background and to allow WIM experts and users to better understand the newly developed technical specifications and standards, or to improve them in the future. Numerical examples are given in this first part to illustrate these developments while a more comprehensive illustration is given in part II.

Thus this paper combines theoretical developments and practical recommendations.

## *1.2. Metrological accuracy of a weighing system*

Any measurement system may be characterized by its accuracy under specific conditions of use. Accuracy is defined by the International Organization of Legal Metrology (OIML) as the maximum permissible error (mpe) made on an individual measurement OIML (1992).

For example, a weighing scale with an accuracy of 1% in the weight range, 10 to 100 N, will provide a relative error $x$, such that, for any mass in the specified range:

$$|x| = \left| \frac{W_m - W_t}{W_t} \right| \leq 0.01 \qquad \forall W_t \in [10;100] \tag{1}$$

where $W_m$ is the measured weight and $W_t$ is the ‘true weight’, i.e., the mass multiplied by the acceleration due to gravity.

In such a case a load of 10 N will be weighed with an error $e$ within ± 0.1 N while a load of 100 N will be weighed with an error within ± 1 N.

If the accuracy is given as an absolute value such as '0.5 N', it means that:

$$|e| = |W_m - W_t| \leq 0.5 \text{ N} \qquad \forall W_t \in [10;100] \tag{2}$$

This term, $e$, is referred to as the absolute error.

The check of accuracy of the scales must be done with a calibration mass (or preferably with a set of calibration masses) with known 'true value(s)', $W_t$ ; but because the real 'true value' is never known, a calibration mass is provided with a reference value $W_r$ given with an acceptable mpe. The OIML (1992) recommends that the mpe of the reference value(s) must be less than 1/5 to 1/10 of the mpe of the scales to be checked. This means that a much more accurate scales must be used to weigh and certify the calibration mass(es). In such a way the masses are 'traceable', i.e. they may be compared step by step to the International calibration masses. The acceleration due to gravity is assumed to be constant for such weighing procedures, and therefore the weighing operation is assumed to be fully repeatable.

### *1.3. Difficulties with WIM*

The difficulties that arise with vehicle WIM stem from two principle sources:

**1.3.1.** A road sensor laid on or in the pavement, does not measure the weight $W$ (product of mass and acceleration due to gravity) of an axle or a vehicle when it is traveling at speed, but an instantaneous impact force $F$ which results from the various masses in the vehicle (suspended and not suspended, body, axle or bogie, tire, etc.), each of them being affected by their respective vertical accelerations. Such acceleration includes that due to gravity but also includes the effects of the pavement/vehicle interaction. The problem is rather complex because these accelerations depend on many factors such as the vehicle speed, the suspension parameters (damping, friction and dry friction, stiffness), the vehicle configuration and geometry and the pavement roughness adjacent to the sensor. Some other factors also play a part such as wind, driving torque or, particularly in the case of a liquid, the pay-load of the vehicle.

The accelerations are also time-dependent with the result that the applied force, $F$ (≥0) varies with distance along the pavement (or with the time), with a non-periodical (and non-stationary) behavior. Some studies have been carried out (Jacob and Dolcemascolo 1997) which show that the frequency spectrum of $F$ is quite complex with various individual frequencies $f$ corresponding to wheel, body and axle effects. It also contains typical wavelengths $\lambda$ ($\lambda = v/f$, where $v$ is velocity) due to pavement and wheel perimeter effects. Moreover the tendency for a vehicle to impose the same loading $F$ during different passes on a road surface having a given profile has been identified and termed 'spatial repeatability', Cole and Cebon (1992). It was shown, Cole et al. (1996), O'Connor et al. (1996), that this phenomenon may lead to repeatable discrepancies between $F$ and $W$ of up to 10% on a rather smooth road, and up to 25% on a rough road.

**1.3.2.** Another problem in WIM results of the sensor size. Both narrow-based (strip) sensors (such as piezo or capacitive strips), with widths in the traffic direction of a few centimeters, and larger bending plates or scales (with strain gauges or load cells) and mats (such as capacitive mats) with widths of less than 0.4 to 0.5 m, are all much smaller than a vehicle length. Accordingly the axle forces are measured (or the axle 'weighed') one by one as the vehicle travels over the sensor. WIM systems generally provide individual axle loads (equal to the measured impact forces), axle group loads (tandem and tridem), and gross vehicle weights where the latter two are calculated as the sum of the individual axle forces.

The procedure of weighing axle forces individually leads to two questions:

(i) The gross vehicle weight is computed as the sum of axle loads (forces) which are not measured at the same time. Because of pitching, the vehicle mass distribution on each axle may differ from time to time and, even if the acceleration at the center of gravity is equal to that due to gravity, this estimate of the gross weight is randomly biased. In fact, because of the accelerations mentioned in 1.3.1, the relationship between the axle forces measured at different times and the gross weight is not clear.

(ii) The definition of axle (or axle group) mass has no physical sense, because an axle or a bogie is an indivisible part of a whole vehicle. Thus, even if the vehicle is stopped, the weight of an axle cannot be defined such as the product of a mass by the acceleration due to gravity. Practically, a 'static axle load' is measured when the vehicle is stationary on a horizontal surface, by a scales placed under the axle or wheels. However, such a measurement is in fact always affected by some uncertainties (Myklebust 1995) and is not fully repeatable:

- a very small slope (such as a few millimeters of difference in level between two axles) can lead to a bias of 100 to 2000 N for the axle load, depending on the height of the center of gravity and on the vehicle dimensions;
- the dry friction in the vehicle suspension induces an imbalance in the distribution of load in a stationary vehicle after braking, depending on the braking conditions.

In summary, the difficulties in defining the accuracy of a WIM system are:

1. the parameter measured is not the weight (referred to here as the static weight or load) itself;
2. only the vehicle gross weight has a metrological definition (traceability) as it is only the entire vehicle that can be weighed at once on a large weigh-bridge, while the static axle load is non-traceable and is not a fully accurate reference.

Concerning the first difficulty, it would be possible to compare the WIM measurement with another (more accurate) measurement of the instantaneous axle impact force $F$ used as a reference if it were measured at the same time. However, it is quite difficult to measure accurately such an impact force, and particularly to repeat such a measurement. It is possible using instrumented vehicles, Leblanc et al. (1992), with a synchronization system with the road sensor, but experience has shown that such systems do not provide fully repeatable measurements because of the traveling conditions and are still affected by some uncertainties. Because of the required staff and cost of operating such vehicles, their use is at present limited to some research programs and it would appear to be too early to take them into account in standardized WIM system calibration or checking procedures.

Therefore in most cases the reference loads and weights considered for WIM accuracy investigations are the static ones. That requires some preliminary investigation about the accuracy of these static loads.

## 2. Accuracy of static weights and loads

### *2.1. Notation and definitions*

Let us consider a *p*-axle vehicle. We note and define:

- $Ws_t$ : the (static) gross weight of the vehicle, equal to its mass multiplied by the acceleration due to gravity; it may be determined accurately on a large weighing scales (weigh-bridge), approved with a given mpe (section 1.2) and capable of measuring the whole vehicle at once, which may in turn be calibrated or checked using calibration masses (traceability). The measured value $\widetilde{Ws}_t$ is such that: $\left|\widetilde{Ws}_t - Ws_t\right| \leq \text{mpe}$.
- $Ws_i$ ($1 \leq i \leq p$) are the 'ideal or reference static loads' for each axle *i*; as mentioned above; such loads are not defined as a mass multiplied by the acceleration due to gravity, but as the static force applied by each axle on a horizontal surface when the vehicle is stopped, the brakes released and the engine off. $Ws_i$ and $Ws_t$ are linked by:

$$Ws_t = \sum_{i=1}^{p} Ws_i \qquad (3)$$

- similarly, we may define the 'reference static axle group load', for any group of *k* axles (generally 2 or 3, with spacing less than about 2 m) by: $Ws_g = \sum_{i=1}^{k} Ws_i$

Now the vehicle is weighed axle by axle on static scales, either using as many scales as the number of wheels or axles to be weighed or only one scale. In the first case all axles are weighed at once, while in the second case they are weighed successively. We note:

- $\widetilde{Ws}_i$ ($1 \leq i \leq p$) the measured static axle loads, and
- $\widetilde{Ws}_g = \sum_{i=1}^{k} \widetilde{Ws}_i$ for a group of *k* axles, or preferably it is recommended to weigh all *k* axles at once while they are simultaneously on the scales, which provides $\widetilde{Ws}_g$ directly.

According to remark (ii) of section 1.3.2, it is not possible to define the accuracy of an axle or an axle group load measurement ($\widetilde{Ws}_i$ or $\widetilde{Ws}_g$) in accordance with equation (1) or (2) as axle weights are not traceable, because these measurements are not fully repeatable and because $Ws_i$ and $Ws_g$ cannot be quantified. Therefore the accuracy of axle load will be defined in the next section.

## *2.2. Statistical accuracy of static axle and axle group loads*

A (small) weighing scales - with a plate length, say, of between 0.3 and 2.5 m - may be calibrated and checked following traditional metrological procedures, with calibration masses. Its intrinsic accuracy (e.g., in the range of 0.5 to 200 kN) may be defined (section 1.2) by its maximum permissible error (mpe), denoted $x_{\max}$ (relative error) or $e_{\max}$ (absolute error). This accuracy only applies for (calibration) mass weighing purposes.

However, the whole accuracy definition of an axle load $\widetilde{W}s_i$ combines the inherent uncertainty of the measurement scales and the statistical uncertainties due to the vehicle stopping conditions, its suspension state and mass distribution according to the pavement or weighing platform profile. Hence this whole accuracy must be defined as a 'statistical accuracy'. It is assumed that the statistical uncertainties do not introduce any bias, i.e., the static axle load applied on the scales is a random variable centered on the unknown reference value $Ws_i$ . If not, the bias may be allocated to the scales (and its environment and using conditions).
Let us note $\sigma_0$ and $C_0$ , the standard deviation and the coefficient of variation respectively of a series of static measurements of the same axle under given weighing conditions (repeatability), if the scales error is neglected. Then it is possible to assess the whole statistical accuracy of this axle load measurement, by a centered confidence interval $[-\delta ; \delta]$ (relative error) or $[-\varepsilon ; \varepsilon]$ (absolute error) for a confidence level $\pi = (1-\beta)$. It means that:

$$\mathrm{P}\left[\left|\widetilde{W}s_i - Ws_i\right| < \varepsilon\right] = \mathrm{P}\left[\left(\left|\widetilde{W}s_i - Ws_i\right| / Ws_i\right) < \delta\right] \geq \pi = 1-\beta \qquad (4)$$

The following theorem provides this confidence interval, with acceptable hypotheses:

Theorem: for a given risk $\beta$ or confidence level $\pi$, and under the hypotheses:

(i) the intrinsic error of the scales is a random variable Z with a density h, distributed in the interval $[-e_{\max};e_{\max}]$;
(ii) the static axle load applied on the scales is a Normal variable, with a variance $\sigma_0^2$; Y is the centered Normal variable with a variance $\sigma_0^2$, and its density is $\varphi$;
(iii) Y and Z are independent (Z only depends on the scales while Y represents the variations in the weighing and vehicle stopping conditions).

then $\delta$ and $\varepsilon$ may be expressed by:

$$\delta = \rho x_{\max} + k_{\beta'/2} C_0 \quad \text{(relative accuracy)} \qquad (5)$$

where $\beta' < \beta$ and $\rho = (1-\beta)/(1-\beta')$ and $\beta'$ is chosen so as to minimize $\delta$,

and

$$\varepsilon = \rho e_{\max} + k_{\beta'/2} \sigma_0 \quad \text{(absolute accuracy)} \qquad (6)$$

where $\beta' < \beta$ and $\rho = (1-\beta)/(1-\beta')$ and $\beta'$ is chosen so as to minimize $\varepsilon$,
$k_{\beta/2}$ is the standard Normal distribution $(1-\beta/2)$-quartile (e.g. $k_{\beta/2} = 1.96$ for $\beta = 0.05$).

Proof of (6):

E = Y + Z is the random variable representing the total 'error' on the measured static axle load. Using (i), (ii) and (iii) we have: $(1-\beta) = \int_{-k_{\beta/2}\sigma_0}^{k_{\beta/2}\sigma_0} \varphi(y)dy \int_{-e_{\max}}^{e_{\max}} \mathrm{h}(z)dz$, where the second integral is equal to unity. Now for any $\beta' < \beta$, it is possible to choose $\rho < 1$ such that:

$$(1-\beta) = \int_{-k_{\beta'/2}\sigma_0}^{k_{\beta'/2}\sigma_0} \varphi(y)dy \int_{-\rho e_{\max}}^{\rho e_{\max}} \mathrm{h}(z)dz \text{, or } \int_{-\rho e_{\max}}^{\rho e_{\max}} \mathrm{h}(z)dz = \frac{(1-\beta)}{(1-\beta')} \quad \text{(a)}$$

then (iii) $\Rightarrow$ $(1-\beta) = \mathrm{P}\left[\left(|\mathrm{Y}| \le k_{\beta'/2}\sigma\right) \text{ and} \left(|\mathrm{Z}| \le \rho e_{\max}\right)\right] \le \mathrm{P}\left[|\mathrm{Y}+\mathrm{Z}| \le k_{\beta'/2}\sigma + \rho e_{\max}\right]$

If Z is assumed to have a uniform distribution, equation (a) leads to : $\rho = (1-\beta)/(1-\beta')$.
For any other distribution of Z, $\rho$ may be calculated from (a).
If the density h of Z does not have its maximum close to $-e_{\max}$ or $e_{\max}$, the value of $\rho$ will be smaller than $(1-\beta)/(1-\beta')$, which provides a smaller confidence interval than (6).

Equation (5) is derived directly from (6).

The statistical accuracy of an axle group static load may be defined in the same way.

A more empirical but practical approach is proposed by the OIML in the draft recommendation on WIM, OIML (1996), to assess the reference values $Ws_i$ . If $n$ ($\approx$10) static measurements of each axle load are performed, the results being $\widetilde{Ws}_{i,j}$ , $1 \le i \le p$ and $1 \le j \le n$, then the reference value $Ws_i$ may be estimated by:

$$Ws_i \cong \frac{1}{n} \sum_{j=1}^{n} \left( \widetilde{Ws}_{i,j} . Ws_t \Big/ \sum_{i=1}^{p} \widetilde{Ws}_{i,j} \right) \quad (7)$$

where $Ws_t$ is either measured on an approved weigh-bridge or estimated itself by the mean of the $n$ gross weights equal to the sum of the axle loads: $Ws_t \cong \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{p} \widetilde{Ws}_{i,j}$ .

In both cases, equation (3) is satisfied.

## 3. Accuracy of WIM measurements

### *3.1. Notation and definitions*

For a pass of a $p$-axle vehicle on a road sensor WIM system, we note and define:

- $Wd_i$ : the 'dynamic load' applied by axle $i$ ($1 \le i \le p$) on the road surface, recorded in motion by the WIM system, and $Wd_t = \sum_{i=1}^{p} Wd_i$ the 'dynamic gross weight' obtained as the sum of the instantaneous dynamic axle loads (recorded by the same sensor, generally at the same place but not at the same time);
- $Wd_g = \sum_{i=1}^{k} Wd_i$ the force applied by a $k$-axle group, called 'dynamic axle group load'.

It is also useful to define the following quantities, related to the dynamic behavior of a vehicle in motion on a road:

- $F_i$ : the 'true' impact force of the axle $i$ ($1 \le i \le p$) as it passes over the WIM system (sensor), and $F_t = \sum_{i=1}^{p} F_i$ the impact force of the vehicle, obtained as the sum of the instantaneous axle impact forces (recorded by the same sensor, generally at the same place but not at the same time); the ‘true’ impact forces are generally not known, because they are very difficult to measure. Nevertheless, with accurately instrumented vehicles or a suitable vehicle simulation model, it may be possible to record or calculate the wheel/axle impact forces continuously, or at a high frequency (such as 200 to 500 kHz), with an uncertainty less than a few percentage points. In such a case, these values may be taken as reference values for the impact forces;

- $IF_i = F_i / Ws_i$ (estimated by $F_i / \widetilde{Ws}_i$) the impact factor of the axle $i$ ($1 \leq i \leq p$) as it passes over the WIM system (sensor), and $IF_t = F_t / Ws_t$ the impact factor of the vehicle (again in this definition the axle impact forces are generally considered at the same place but not at the same time);

- $DF_i = (IF_i - 1)$ the dynamic increment of axle $i$ ($1 \leq i \leq p$) as it passes over the WIM system or sensor, and $DF_t = (IF_t - 1)$ the dynamic increment of the vehicle.

Physically it would be better to sum up the individual axle impact forces measured at the same time to provide the vehicle impact force, but it is generally not possible with the common WIM systems. Therefore the previous definitions are widely accepted by engineers.

### *3.2. Errors for an individual WIM measurement*

Definitions of the errors (and accuracy) of WIM measurements are now proposed, with respect to the static loads. As indicated above, it would be preferable to refer to the ‘true values’ of the instantaneous impact forces $F$. However in most cases it is impossible to determine these ‘true values’ and most WIM users are, in fact, more interested in the estimate of weight (static load).

Hence, the ‘error’ of a WIM measurement results both from the inherent WIM system uncertainty, and from all the effects of the vehicle and pavement interaction which lead to discrepancies between the static weights and the impact forces. Accordingly, the impact factors differ from unity. The error of an individual WIM measurement may be expressed as:

$$e_j = (Wd_j - Ws_j) = (Wd_j - F_j) + (F_j - Ws_j) \quad (8)$$

and

$$x_j = (Wd_j - Ws_j)/Ws_j \quad (9)$$

where $j = i$, $g$ or $t$ for an axle, an axle group or the gross weight respectively.

In (8), the first term $(Wd_j - F_j)$ is the measurement error due to the WIM system, and the second term $(F_j - Ws_j) = Ws_j . DF_j$ is the error due to dynamic effects. If the WIM system is calibrated or designed to provide an unbiased estimate of the static loads, as is assumed in the following, then $e_j$ is the error to be considered; if the WIM system is used for instantaneous impact force evaluation, then only the first term of equation (8) should be considered.

Because of the randomness of the pavement/vehicle interaction, of the vehicle characteristics, and of the intrinsic system uncertainties, the errors $e_j$ and $x_j$ are random. The accuracy must be again defined statistically, by a confidence interval for a confidence level $\pi = (1-\beta)$:

$$P\left[\left|e_j\right| < \varepsilon\right] \geq \pi = (1-\beta) \qquad \text{or} \qquad P\left[\left|x_j\right| < \delta\right] \geq \pi = (1-\beta) \tag{10}$$

As in section 2.2 for the static axle load accuracy, under some hypotheses about the distribution of $e_j$ or $x_j$, $\varepsilon$ and $\delta$ may be expressed as functions of $\beta$ or $\pi$ (see section 3.4). The parameters of these distributions depend on the samples considered, and particularly the repeatability or reproducibility conditions of measurements. These conditions will be detailed in the next section.

### *3.3. Repeatability and reproducibility conditions*

As referred to in 1.3, the difference between a static axle load measurement and an in-motion (dynamic) axle load measurement, is influenced by many vehicle and pavement effects. For a sample of *n* in-motion measurements, the following weighing conditions may be distinguished:

(r1) full repeatability: the same vehicle makes *n* passes on the WIM system, at the same speed, the same load and approximately the same traveling conditions (lateral position, driver, etc.);

(r2) extended repeatability: the same vehicle makes *n* passes on the WIM system, with different combinations of speed and load (e.g., empty, half loaded and fully loaded);

(R1) limited reproducibility: a small set of *k* vehicles (e.g. 2 to 10 lorries) with different silhouettes, loads and suspension types each make several ($n_i$ ) passes on the WIM system (such that $n = \Sigma\, n_i$ ) at different speeds;

(R2) full reproducibility: a large set of *n* vehicles (e.g., some tens or hundreds of lorries) taken from the traffic flow, pass once over the WIM system under normal traffic conditions.

The respective standard deviations of $x_j$ under each of these conditions are denoted $\sigma_j(r1)$, $\sigma_j(r2)$, $\sigma_j(R1)$ and $\sigma_j(R2)$. The higher the variability of the test conditions, the higher the scattering of the results. Hence we have:

$$\sigma_j(r1) < \sigma_j(r2) < \sigma_j(R1) < \sigma_j(R2) \tag{11a}$$

If the conditions are well specified, the standard deviation used will be denoted $\sigma_j$.

Moreover, for any given sample size and weighing conditions, the variability of the results also depend on the environmental conditions, i.e. the climatic conditions and the period of measurements. The three following conditions may be distinguished:

(I) environmental repeatability: the measurement period is limited to a couple of hours within a day, or spread over a few consecutive days, such as the temperature, climatic and environmental conditions do not vary significantly during the measurements;

(II) limited environmental reproducibility: the test time period extends at least over a full week or several days spread over a month, such that the temperature, climatic and environmental conditions vary during the measurements, but no seasonal effect has to be considered;

(III) full environmental reproducibility: the test time period extends over a whole year or more, or at least over several days spread all over a year, such that the temperature, climatic and environmental conditions vary during the measurements and all the site seasonal conditions are encountered.

For the same weighing conditions (r1) to (R2), if the respective standard deviations of $x_j$ under each of these environmental conditions are denoted $\sigma_j(I)$, $\sigma_j(II)$, and $\sigma_j(III)$, we have again:

$$\sigma_j(I) < \sigma_j(II) < \sigma_j(III) \tag{11b}$$

Finally, each sample of $n$ in-motion measurements should be affected by the corresponding measuring conditions, such as (I-r1), (I-R1), or (III-R2), etc.

### *3.4. Accuracy of a WIM system*

We may now express the statistical accuracy of a WIM system in its environment (road roughness, vehicle suspension and characteristics) and under specified traffic conditions (repeatability or reproducibility conditions). Equation (10) specifies that an individual measurement $Wd_j$ falls in the confidence interval $[Ws_j - \varepsilon ; Ws_j + \varepsilon]$ with a probability greater than $\pi$ (or outside this interval with a risk $\beta$), or reversely that, with a level of confidence $\pi$, the interval $[Wd_j - \varepsilon ; Wd_j + \varepsilon]$ contains the true static load $Ws_j$. The respective centered confidence intervals for $e_j$ and $x_j$ are $[-\varepsilon ; +\varepsilon]$, or $[-\delta ; +\delta]$. $\beta$ may be considered to be the risk deemed acceptable to the customer.

If we assume that the WIM measurements are unbiased with respect to the static loads, which may be obtained through a proper calibration, and if $\sigma_j$ is the standard deviation of $x_j$ under the specified repeatability or reproducibility conditions (r1 to R2, and I to III), the risk $\beta$ and the width of the confidence interval may be related by the Bienaimé-Tchebichev inequality (this upper bound is known to be a crude upper limit):

$$\beta \leq \sigma_j^2/\delta^2 \tag{12}$$

A better limit on this risk may be given if the distribution of $x_j$ is known, such as for a centered Normal distribution where:

$$\delta = k_{\beta/2} . \sigma_j \tag{13}$$

or, if the WIM measurements are biased, i.e., the WIM system is uncalibrated, with a relative bias $m$ with respect to the static loads:

$$\delta = m + k_{\beta/2}\, \sigma_j . \tag{14}$$

### *3.5. Practical considerations relating to the reference values of static loads*

According to the OIML requirements, OIML (1992), if a static weighing scales is used to measure the reference static loads or weights, its accuracy must be significantly better than the expected accuracy of the WIM system to be tested. A weigh-bridge (capable of weighing the whole vehicle at once) must have mpe's less than 1/10 of the expected $\varepsilon$ and $\delta$ values of the WIM system for gross weights. An axle scales must also have mpe's $x_{max}$ and $e_{max}$ less than 1/10 of the expected $\varepsilon$ and $\delta$ values of the WIM system for axle loads. Moreover the $\varepsilon$ and $\delta$ values of the static axle scales, according to equations (5) and (6) of section 2.2, should be less than 1/5 of those of the WIM systems.

In some cases it also may be possible to take the reference load values from a previously tested and much more accurate WIM system than the one to be checked. In such a case it is again recommended that the ratio of 1/5 would be obtained between the accuracy levels of both systems ($\varepsilon$ and $\delta$ values).

## 4. Procedure for assessing and checking the accuracy of WIM systems

This section briefly reviews the statistical background to be used and introduces the proposed procedure for accuracy classification and verification of a WIM system installed on a road. It also gives numerical values and examples from European WIM trials.

During a WIM session, the individual random errors, as defined in section 3.2, are assumed to be independent and to have a Normal distribution. This rather arbitrary assumption is nevertheless often verified by testing in measured samples, and is useful in that it leads to further statistical inferences. Non-Normal random errors may result from some malfunctioning of the WIM system.

### *4.1. Statistical framework*

In order to take into account the possibility that the static loads considered may be different (if different axles, axle groups or gross weights are taken into account in the samples), only the relative accuracy's and errors are considered: $X = (Wd - Ws)/Ws$. The parameter $X$ is considered to be a random variable while $Wd$ is also random. $Ws$ is unknown but deterministic, and is estimated by $Wd$ in a WIM session. However, to calibrate or check a WIM system, $Ws$ must be known.

In a WIM trial, either for calibration or checking purposes, the sample $(x_1, ....x_n)$ of relative errors for axles, axle groups or gross weights are considered. $(x_1, ....x_n)$ are the sample values of a series $(X_1, ....X_n)$ of independent identically distributed Normal variables, with unknown mean $\mu$ and standard deviation $\sigma$. $M$ and $S^2$ are the estimators of $\mu$ and $\sigma$ for a sample of size $n$, and are themselves random variables. $m$ and $s$ are the sample values of $M$ and $S$ for the sample considered $(x_1, ....x_n)$. Therefore $\mu$ is estimated by $m$ and $\sigma$ by $s$ in such a trial.

$M$ is a Normal variable with a mean $\mu$ and a variance $\sigma^2/n$, while $(n-1)S^2/\sigma^2$ has a $\chi^2$-distribution with $\nu = n-1$ degrees of freedom. Then the variable $T = \sqrt{n}(M - \mu)/S$ has a Student t-distribution with $\nu$ degrees of freedom.

The confidence interval of $\mu$, for a risk $\alpha$ taken equal to 0.05 in the numerical examples (commonly accepted value), is $[m - t_{\nu,1-\alpha/2}\, .s/n^{1/2}\ ;\ m + t_{\nu,1-\alpha/2}\, .s/n^{1/2}]$, from which it follows that:

$$P[- t_{\nu,1-\alpha/2} \le T \le t_{\nu,1-\alpha/2}] = 1-\alpha \qquad (15)$$

### *4.2. Confidence Interval for an individual WIM measurement*

Now let us consider a WIM system, either newly installed or in-service, hence uncalibrated or with a calibration done some time previously, which might have become suspect. Then, for a measured sample ($x_1$ ,....$x_n$), the true mean $\mu$ and standard deviation $\sigma$ of the variable $X$ are unknown, and the results of sections 3.2 and 3.4 may not be used. Of interest is the probability $\Pi$ that an individual measurement falls within a confidence interval centered on the static load :

$$\Pi = P[-\delta \leq X_i \leq \delta] \quad (16)$$

If Φ is the Normal standardized CDF (cumulative distribution function), the probability $\Pi$ becomes:

$$\Pi = P[(-\delta - \mu)/\sigma \leq (X_i - \mu)/\sigma \leq (\delta - \mu)/\sigma] = \Phi((\delta - \mu)/\sigma) - \Phi((-\delta - \mu)/\sigma) \quad (17)$$

For unknown $\mu$ and $\sigma$, using equation (15) a lower bound of $\Pi$ is found with the risk $\alpha$ :

$$\pi = \Psi((\delta - m)/s - t_{\nu,1-\alpha/2}/n^{1/2}) - \Psi((\delta - m)/s + t_{\nu,1-\alpha/2}/n^{1/2}) = \Psi(u_1) - \Psi(u_2) \quad (18)$$

where : $u_1 = (\delta - m)/s - t_{\nu,1-\alpha/2}/n^{1/2}$ , $u_2 = (-\delta - m)/s + t_{\nu,1-\alpha/2}/n^{1/2}$

and Ψ is the CDF for a Student t-variable with $\nu$ degrees of freedom.

For large $\nu$ (e.g., larger than 60), Ψ may be replaced by Φ in equation (18).

Then after testing, the bias of the WIM system may be estimated by $m$, while its accuracy may be again expressed by the centered confidence interval of the relative error [$-\delta$ ; $\delta$] for a confidence level greater than or equal to $\pi$. In this approach, $\delta$ is chosen first, and then $\pi$ is calculated.

By testing, it is possible to (re)calibrate the WIM system. The bias is removed, either by revising some amplification factor in the hardware or setting a calibration factor in the software. Then the system becomes unbiased and the parameters $u_1$ and $u_2$ of equation (18) are equal and opposite which leads to:

$$\pi = 2\,\Psi(u_1) - 1 = 2\,\Psi(\delta/s - t_{\nu,1-\alpha/2}/n^{1/2}) - 1 \quad (19)$$

Since $t_{\nu,1-\alpha/2}/n^{1/2}$ decreases as $n$ increases (see the tabulated values in Table 1), it can be inferred that the level of confidence increases with $n$. The implication is that, for larger sample sizes, there is less uncertainty in the WIM accuracy evaluation, but the testing cost increases.

### *4.3. WIM system accuracy*

The accuracy of a WIM system may also be defined in another way by testing. WIM users or standards generally specify a minimum acceptable level of confidence, say $\pi_0$. After testing, the statistics $m$, $s$ and $n$ of the measured sample ($x_1$ ,....$x_n$) are known, and equation (18) allows the calculation of the smallest $\delta$, denoted $\delta_{min}$ such that:
$\pi_0 = \Psi(u_1) - \Psi(u_2)$.

$\delta_{min}$ is called the statistical accuracy of the WIM system at the confidence level $\pi_0$.

### *4.4. Numerical examples*

For single axles, the standard deviation $\sigma$ (estimated by $s$) is assumed, under some repeatability or reproducibility conditions, to be equal to 0.07. For a mean static axle load of 70 kN, this corresponds roughly to a standard deviation of 5 kN.

For gross weights, the standard deviation $\sigma$ (estimated by $s$) is assumed, under the same conditions, to be equal to 0.045. For a mean static weight of 300 kN, this corresponds to a standard deviation of 13.5 kN.

We choose $\delta = 0.15$ for the axles and $\delta = 0.10$ for the gross weights, as required for Accuracy Class B in some specifications (see section 4.5) such as in METT-LCPC (1993) and COST323 (1997a). For a perfectly calibrated WIM system, assuming unbiased results, and for an accepted risk $\alpha = 0.05$, the lower bound $\pi$ of the confidence levels of the required confidence intervals is derived from equation (19) and increases with the sample size $n$ (see Table 2). When $n$ tends to infinity, $\pi$ tends to the upper limit $2\Phi(\delta/\sigma) - 1$, which is the confidence level of the interval $[-\delta ; \delta]$ for a centered Normal variable with a known standard deviation $\sigma$.

If a bias of $m$=0.035 is assumed for the axles, which leads to $m/s = 0.5$, the $\pi$-values are derived from equation (18). They are smaller than without bias, and for the axles they are given in Table 3.

The comparison between the values of the first and second lines of Tables 2 and 3 show that, for n>60, the error resulting from replacing $\Psi$ with $\Phi$ is less than 1%.

### *4.5. Specification and acceptance test of WIM systems*

Any standard or specification of WIM systems should propose classes of accuracy, which may be required by users for their particular applications. The accuracy class of a WIM system not only depends on the system itself, but also on the environment, and above all on the road profile. According to the previous developments, such classes should be defined by the statistical relative accuracy $\delta$ (see equation (16)). For clarity and traceability, $\delta$ should be independent of the conditions (measuring conditions in repeatability r1, r2 or reproducibility R1, R2, and environmental conditions I to III) of any acceptance test, but should only depend on the entity to be measured (criterion), i.e. gross weight, single axle load, group of axles or axle of a group taken individually. However, the test conditions (repeatability or reproducibility) affect the standard deviation $\sigma$, and the sample size, $n$ and hence $s$. Therefore the confidence level $\pi$ of a confidence interval $[-\delta ; \delta]$ depends on the test conditions, and should be known by users.

A standard or specification document should specify a minimum required value $\pi_0$ for the level of confidence, which depends on the test conditions performed to check the accuracy of a WIM system. If the customer has a particular requirement for $\pi_0$, it is their responsibility to perform or specify an acceptance (checking) test which complies with this value. It means that the repeatability/reproducibility conditions and the sample size *n* cannot be chosen arbitrarily. Moreover the test conditions should be representative of the conditions of use of the WIM system.

A typical standardized table of $\delta$ values is given here in Table 4, taken from the European specification (COST 323, 1997a). The classes are identified by letters and by the values of $\delta$ for the gross weight criterion; the latter allows the use of any class from 1 to 25 or more.

The $\pi_0$ values must be independent of the accuracy class, but may also be independent of the entity considered if the values of $\delta$ chosen are such that $\delta/\sigma$ remains constant (from one entity to another), for any given test conditions. That is the case in the European specification, which leads to the values given in Tables 5 to 7, under three environmental repeatability/reproducibility conditions and four weighing repeatability/reproducibility conditions. ***These $\pi_0$ values were derived from the assumption that $\delta/\sigma$=2.675, 2.36, 2.155 and 2, for r1, r2, R1 and R2 respectively, in environmental repeatability conditions (I).*** Such an assumption is based on the available tests results, but may be checked and possibly modified in the future if needed. For the other environmental conditions (II) and (III), according to Eq. (11b), it was assumed that the standard deviation increased such as: $\sigma_j(II)=1.05\sigma_j(I)$ and $\sigma_j(III)=1.1\sigma_j(I)$. Thus, we get the ratio ***$\delta/\sigma$=2.548, 2.248, 2.052 and 1.905, in environmental limited reproducibility conditions (II), and $\delta/\sigma$=2.432, 2.145, 1.959 and 1.818, in environmental full reproducibility conditions (III)***.

The procedure to check by testing if a WIM system complies with a given accuracy class $\delta$ is now described. For one entity, e.g., the gross weights, the sample statistics *m*, *s* and *n* of the relative errors with respect to the static reference values are calculated. Then two methods may be applied:

(i) the level of confidence $\pi$ is derived using equation (18), and the condition of acceptance is : $\pi \geq \pi_0$ for a given $\delta$, where $\pi_0$ is either taken from a specified table of a standard (e.g. Tables 5 to 7), or is specified by the user, generally between 0.90 and 0.99;
(ii) the required level of confidence $\pi_0$ is taken as the target value in equation (18), and $\delta$ is calculated to satisfy $\pi=\pi_0$. This (lowest) value, denoted $\delta_{min}$, is the best acceptable accuracy class for the entity considered.

For an initial verification, immediately following calibration, and if the mean bias is removed by using the same sample of data for calibration and checking, it is recommended to reduce the specified $\delta$ value by a factor $k$<1, i.e., to consider $k.\delta$ instead of $\delta$ for method (i), or to consider $\delta_{min}/k$ instead of $\delta_{min}$, for method (ii). *k* should be in the range of 0.5 to 0.9, and a value of 0.8 is proposed in the European specification. This is necessary to account for possible (small) drift and future limited bias in the WIM system.

### *4.6. Examples of acceptance tests and interpretation*

In order to illustrate the procedure described above, numerical examples are now given.

*4.6.1. Initial verification:* let us first consider a newly installed WIM system, which requires a calibration. We assume that this calibration is performed using (statically) pre- or post-weighed vehicle(s) making several passes over the system. The eventual bias $m$ on the gross weight is removed afterwards by software, and the system calibrated by fitting a calibration coefficient or the amplification of the signal to the data. Then the relative errors are centered and the sample estimator of the standard deviation, $s$, is computed from the sample data. According to the sample composition and to the test conditions, the repeatability/reproducibility conditions are determined, I, II or III, and r1, r2, R1 or R2.

For example, if the criterion of gross weight is considered, the environmental condition is (I) and the sampling condition is (r2) the sample size is $n = 20$, and the sample standard deviation is $s = 0.028$, then for $\delta$=0.10 and $k$=0.8, i.e. $k.\delta$=0.8, the probability given by equation (19) is $\pi$=0.973, which exceeds the target value of $\pi_0$=0.941 (Table 5). Therefore the system may be accepted in class B(10) for the criterion of gross weight. Moreover $\delta_{min}$ =0.069, $\delta_{min}/k$=0.086 and the system could be accepted in an intermediate class (9).

If, for the same sample conditions, the standard deviation $s$ is 0.045, then, for $k.\delta$=0.08, the probability is only $\pi$=0.794 and the system should not be accepted in class B(10) although it is clearly acceptable in class C(15) ($\pi$=0.96), or $\delta_{min}/k$=0.112/0.8=0.14.

After a calibration procedure and for an initial verification, a WIM system should be unbiased, at least on average (e.g. for the gross weights) and the tolerance $\delta$ is reduced. If the calibration is performed under repeatability conditions (r1 or r2), the level of confidence in the assumed accuracy class under reproducibility conditions (R1 or R2) remains unknown but it may be reasonably assumed to be less than the level of confidence calculated. For that reason, it is strongly recommended not to accept an accuracy class with a confidence level of less than 0.95 under repeatability conditions.

*4.6.2.. In-service verification*: Let us now consider a WIM system which has been in operation since its last calibration, some time previously. A test is performed to check the real accuracy, but no recalibration is authorized (at least in the first instance). Again the test is performed using a set of (statically) pre- or post-weighed vehicles, under repeatability or reproducibility conditions, with a total of $n$ passes. The sample statistics of the relative errors are $m$ (bias) and $s$ (standard deviation). Equation (18) gives the level of confidence $\pi$ for any proposed accuracy class $\delta$.

As an example, after a test carried out in conditions (I-R1) for gross weight, the sample statistics were:

(i) $m$=0.05, $s$=0.035, $n$=30, then $\pi$=0.85<0.925 for $\delta$=0.1, and the system may not be accepted in class B(10). It may however be accepted in class C(15), with $\pi$=0.99, and $\delta_{min}$ =0.115. If a recalibration is made in order to eliminate the bias, with a multiplicative calibration coefficient of 0.95, then $m$=0, and $\pi$=0.934 for $k.\delta$=0.8*0.1=0.08 and the system may be accepted in class B(10), while $\delta_{min}$ =0.078 and $\delta_{min}/k$=0.0975.

(ii) $m$=0.015, $s$=0.042, $n$=20, then $\pi$=0.914>0.908 ($\delta$=0.1), $\delta_{min}$ =0.098 and class B(10) is just acceptable, but with a risk of 8.6% that the specified relative error is exceeded. If the number of passes $n$ is increased to 60, and assuming that the same values for $m$ and $s$ are found, then $\pi$=0.951>0.942, $\delta_{min}$ =0.097 and class B(10) is still accepted with a reduced risk of 4.9%.

In case (i), the bias $m$ was substantial and therefore after recalibration the WIM system accuracy is upgraded by one class. In case (ii), the bias is low and does not significantly affect $\pi$ or $\delta_{min}$ ($\pi$=0.929 and $\delta_{min}$ =0.094 with $m$=0). Moreover, if this small bias is removed and the same sample re-used to perform an initial verification, then the system would be downgraded to class C(15), with $\delta_{min}/k$=0.118.

If the case (i) is again considered, but assuming that the results were collected over a whole year and with traffic flow vehicles passing only once on the WIM system, i.e. in conditions (III-R2), then the conclusions are modified: $\pi=\pi_0$=0.85 ($\delta$=0.1), $\delta_{min}$=0.10 and class B(10) is just accepted.

## 5. Conclusions

This paper establishes or presents the physical and statistical background for any WIM specification, but particularly for the European specification COST323, and introduces precise definitions for WIM system accuracy assessment.

The first issue addressed was the clear definition of reference values, that is not an obvious question for WIM. Although any WIM system measures instantaneous axle/wheel impact forces, the static loads are commonly required by users and agreed to be the reference for accuracy checks. However, these static axle loads cannot be defined by the product of a mass by the acceleration due to gravity, and are themselves affected, if measured on static scales, by experimental uncertainties. The metrological maximum permissible error (mpe) is not sufficient to give an account of these uncertainties, and a statistical accuracy must be considered.

For a WIM system, errors with respect to static loads clearly become random, with standard deviation highly dependent on the test measuring and environmental repeatability/reproducibility conditions. Therefore the accuracy of a WIM system is assessed using the classical theory of estimation by confidence intervals. A detailed procedure is proposed, which allows any WIM system to be classified into accuracy classes with respect to four criteria (gross weight, single axle load, group of axle loads and axle of group loads). Numerical examples are given as illustration. In addition to being the basis for the European WIM specification, this paper may be used for any future WIM specifications or standards. The numerical values proposed were analyzed and confirmed in the light of results from the European test program 1997-98 (COST323, 1996); they may need to be modified and adapted to specific conditions in other regions.

## 6. References

Cole, D.J. and Cebon, D. (1992), 'Spatial repeatability of dynamic tire forces generated by heavy vehicles', *Proc. Inst. Mech. Eng.*, vol 206, pp.17-27.

Cole, D.J., Collop, A.C., Potter, T.E.C., Cebon, D. (1996), 'Spatial repeatability of measured tire forces', *Proc. Inst. Mech. Eng.*, vol 210, pp.185-197.

COST 323 (1996), *European WIM Test Program 1996-1998 (CET & CMT)*, Technical report of the Management Committee, draft 2, COST323/WAVE, EC/DGVII, COST Transport, September, 34 pp.

COST 323 (1997a), *European Specification on Weigh-in-Motion of Road Vehicles*, draft 2.2, EUCO-COST/323/5/1997, Cost Transport, EC/DGVII, LCPC, Paris, June.

COST 323 (1997b), *Collection and Analysis of Requirements as Regards Weighing Vehicles in Motion*, EUCO-COST/323/2E/1997, Cost Transport, EC/DGVII, January.

Jacob, B. and Dolcemascolo, V. (1997), *Spatial repeatability*, OECD/DIVINE project, Final report of the Element 5, OECD/LCPC, Paris.

Leblanc, P.A., Woodrooffe, J.H.F., Papagianakis, T. (1992), *A comparison of the accuracy of two types of instrumentation for measuring vertical wheel load*, Heavy Vehicles and Road Technology, Safety and Policy, pp.86-94, T. Telford, London.

METT - LCPC (1993), *Recommendations for the choice of sites and the acceptance of WIM stations*, National WIM project report, fasc.2, November.

Myklebust, T. (1995), *Approval of automatic weighing instruments for road vehicles (WIM)*, OIML Seminar 'Weighing towards the Year 2000', OIML, Paris, September.

O'Connor, T.L., O'Brien, E.J., Jacob, B. (1996), 'An experimental investigation of spatial repeatability', *International Journal of Vehicle Design, Heavy Vehicle Systems*, in print.

OIML (1992), *Non automatic weighing instruments Part 1: Metrological and technical requirements - Tests*, International Recommendation R76-1.

OIML (1996), *In-motion road vehicle weighing instruments*, Second committee draft Recommendation, International Organisation for Legal Metrology, November.

## Tables

| $n$ | 10 | 20 | 30 | 40 | 60 | 120 | $\infty$ |
|---|---|---|---|---|---|---|---|
| $t_{\nu,1-\alpha/2}/n^{1/2}$ | 0.7153 | 0.4680 | 0.3734 | 0.3195 | 0.2582 | 0.1807 | 0 |

**Table 1 -** Tabulated values for $t_{\nu,1-\alpha/2}/n^{1/2}$ , with $\alpha = 0.05$ (for other values of $n$, interpolation is allowable)

| | $n$ | 10 | 20 | 30 | 60 | 120 | ∞ |
|---|---|---|---|---|---|---|---|
| **π-values for axle weights** | using ψ<br>using Φ | 0.813<br>0.847 | 0.890<br>0.906 | 0.913<br>0.923 | 0.936<br>0.941 | 0.948<br>0.950 | 0.968 |
| **π-values for gross weights** | using ψ<br>using Φ | 0.834<br>0.868 | 0.904<br>0.921 | 0.925<br>0.936 | 0.940<br>0.950 | 0.957<br>0.959 | 0.974 |

**Table 2 -** Example of lower bound on the confidence levels (unbiased system, $\alpha = 0.05$)

| | $n$ | 10 | 20 | 30 | 60 | 120 | ∞ |
|---|---|---|---|---|---|---|---|
| **π-values for axles** | using ψ<br>using Φ | 0.768<br>0.796 | 0.851<br>0.865 | 0.877<br>0.886 | 0.904<br>0.908 | 0.919<br>0.921 | 0.945 |

**Table 3 -** Example of lower bound on the confidence levels (biased system, $\alpha = 0.05$)

| **Accuracy classes**<br>**Entity** | **A (5)** | **B+ (7)** | **B (10)** | **C (15)** | **D+(20)** | **D (25)** | **E** |
|---|---|---|---|---|---|---|---|
| Gross weight | 0.05 | 0.07 | 0.10 | 0.15 | 0.20 | 0.25 | > 0.25 |
| Group of axle | 0.07 | 0.10 | 0.13 | 0.18 | 0.23 | 0.28 | > 0.28 |
| Single axle | 0.08 | 0.11 | 0.15 | 0.20 | 0.25 | 0.30 | > 0.30 |
| Axle of a group | 0.10 | 0.14 | 0.20 | 0.25 | 0.30 | 0.35 | > 0.35 |

**Table 4 -** Accuracy classes definition: values of $\delta$

| **Sample size ($n$)**<br>**Test (sampling) conditions** | **10** | **20** | **30** | **60** | **∞** |
|---|---|---|---|---|---|
| Full repeatability (r1) | 0.95 | 0.972 | 0.979 | 0.984 | 0.992 |
| Extended repeatability (r2) | 0.90 | 0.941 | 0.953 | 0.964 | 0.981 |
| Limited reproducibility (R1) | 0.85 | 0.908 | 0.925 | 0.942 | 0.970 |
| Full reproducibility (R2) | 0.80 | 0.874 | 0.896 | 0.918 | 0.954 |

**Table 5 -** Proposed $\pi_0$ values in "environmental repeatability" conditions (I)

| **Sample size ($n$)**<br>**Test (sampling) conditions** | **10** | **20** | **30** | **60** | **∞** |
|---|---|---|---|---|---|
| Full repeatability (r1) | 0.933 | 0.962 | 0.970 | 0.978 | 0.989 |
| Extended repeatability (r2) | 0.875 | 0.925 | 0.939 | 0.953 | 0.975 |
| Limited reproducibility (R1) | 0.819 | 0.887 | 0.907 | 0.927 | 0.960 |
| Full reproducibility (R2) | 0.766 | 0.849 | 0.874 | 0.900 | 0.943 |

**Table 6 -** Proposed $\pi_0$ values in "environmental limited reproducibility" conditions (II)

| **Sample size ($n$)**<br>**Test (sampling) conditions** | **10** | **20** | **30** | **60** | **∞** |
|---|---|---|---|---|---|
| Full repeatability (r1) | 0.914 | 0.950 | 0.960 | 0.970 | 0.985 |
| Extended repeatability (r2) | 0.847 | 0.907 | 0.924 | 0.941 | 0.968 |
| Limited reproducibility (R1) | 0.786 | 0.864 | 0.887 | 0.911 | 0.950 |
| Full reproducibility (R2) | 0.730 | 0.823 | 0.851 | 0.881 | 0.931 |

**Table 7 -** Proposed $\pi_0$ values in "environmental full reproducibility" conditions (III)