Transmitted by the experts of IWG MU

Informal Document **GRBP-74-11**
(74th GRBP, 15-17 January 2021,
agenda item 3)

DRAFT DOCUMENT OF REFERENCE:
A GENERAL APPROACH TO ESTIMATE MEASUREMENT UNCERTAINTIES

1. Background

In all kind of testing of objects according to standards, there is a certain measurement uncertainty. This is also the case of the measurement of sound levels of vehicles and tyres, for example during type approval of these objects. In standards used for such measurements (ISO, ANSI, CEN, etc.) a separate chapter on measurement uncertainty is mandatory. However, this is not the case in UN ECE regulations.

The focus on In-use compliance checking of vehicles is increasing, as the introduction of the Regulation (EU) 2018/858[1] (Marked surveillance) is showing. In the US., such testing has been in place for decades for emissions and safety (not noise).

These kinds of tests will then be performed by institutions not involved in the original type-approval test ("third party"). Therefore, uncertainties connected to such market surveillance tests will be of uttermost importance, as a failure they could withdraw any previous given type-approval to the vehicle/object.

Such third-party testing is not within the scope of UN ECE, however measurement uncertainties have also an important role in general for Conformity of Production (COP), which is part of UN ECE regulations for vehicles and tyres.

GRBP has therefore been asked to establish an Informal Working Group on measurement uncertainties to work on the following topics:

- Improvements of test methods
- Compensation, if possible (systematic errors)
- Remaining uncertainties (random errors)

This Draft rapport outline the general approach to measurement uncertainty, based on both ISO 5725[2] and ISO/IEC Guide 98-3 (GUM)[3]. However, the steps of defining the uncertainty of a measurement based on ISO 5725 do not differ significant from the GUM. Therefore, the statistical method described in this report will mainly focus on the GUM.

2. General considerations

Measurement procedures are always affected by factors causing disturbances leading to variation in the results observed by the same subject. The source and nature of these perturbations are not completely known and can sometimes affect the end-result in a non-predictable way.

A measured result shall be understood as an approximation to the true result, which by itself is unknown.

- Two measurements are deemed to provide the same result if their test results are within a given uncertainty.

Thus, the knowledge of the measurement uncertainty is important as it provides information about the precision and repeatability of measurements.

It is important to minimize the uncertainties, e.g. by narrowing ambient and test conditions or by corrections. Any residual uncertainty shall be covered by tolerances.

Sources of errors in measurements come from limitations in the sensitivity of the instruments or from imperfections in experimental design or measurement techniques. Errors are classified as rando or systematic.

Random errors, which cannot be compensated for

They are always present and are from operator approximating a reading and changes in the experimental conditions. There is equal probability that the reading will be too high or too low. To minimize random error, repeated measurements are taken, and the average or mean is calculated. If the same operator gets the same results, the results are said to be reproducible, see figure 2.1. Recording the precision or uncertainty is one way of representing random error:

measurement ± random error

Systematic errors, which can be compensated for

These are typically present and are from limitations in instruments, technology and operator skill. To minimize systematic errors, carefully calibration of the instruments can be done, and the operator uses the best techniques. Systematic errors lead to bias, moving the measurements away from the true value in one direction or the other, see figure 2.1. Recording the bias can be represented as:

measurement + systematic error or *measurement – systematic error*

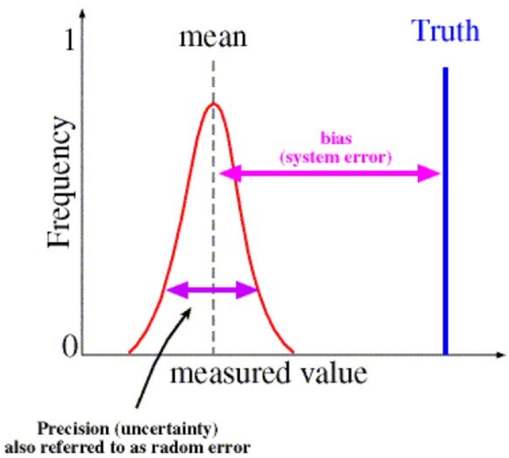

Figure 2.1 Graphical representation of random and systematic error.

In figure 2.1, the bias is *measurement – systematic error*. Normally, only the uncertainty is reported. Systematic errors are dealt with only if the true value is known and then the % error can be calculated and discussed.

Precision and accuracy in measurements

Precision reflects how reproducible the measurements are while accuracy reflects how close the measurements are to the true value. Ideally, we aim for both precision (smaller random errors) and accuracy (systematic error). The target analogy works well, as shown in figure 2.2.

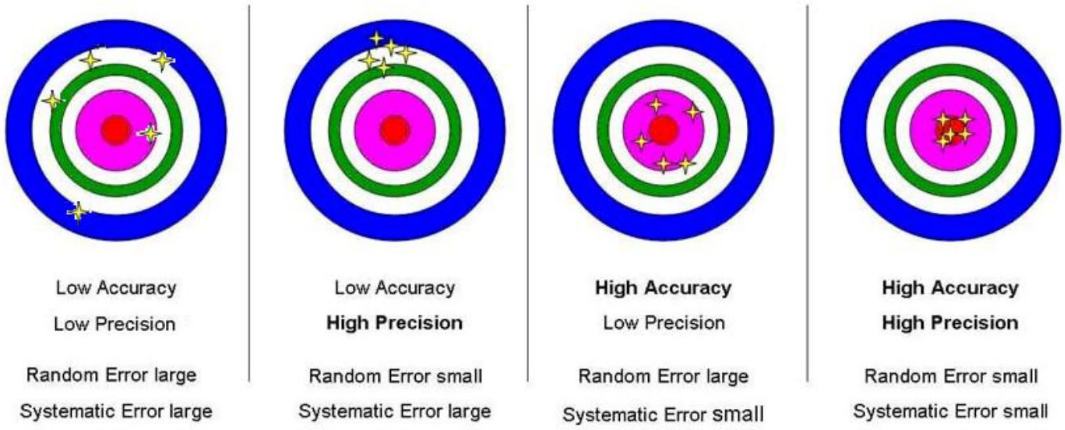

Figure 2.2 Precision and accuracy. Notice that random error is related to precision while systematic errors are related to accuracy.

A graphical representation is shown in figure 2.3.

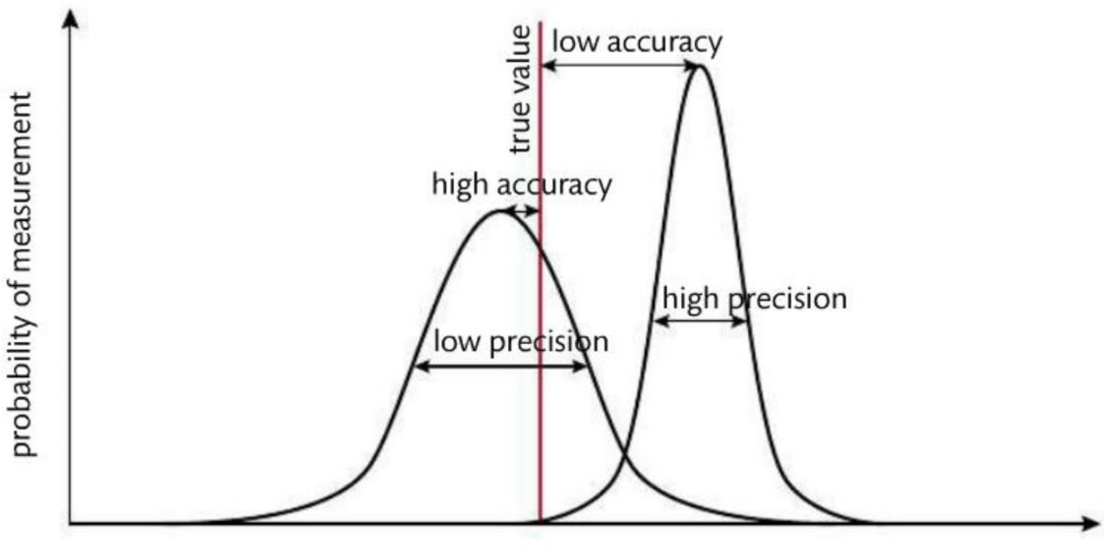

Figure 2.3 Graphical representation for precision and accuracy. Note that the set of readings to the left represent high accuracy and low precision. Those on the right indicate the values have high precision and low accuracy.

It should be noted that the approach to define the uncertainty contribution as given by the GUM do not distinguish between random or systematic errors.

3. How to handle measurement uncertainty

To reduce measurement uncertainty, the following approach is recommended[4]:

A. Avoidance of uncertainties

Normally, a regulation/measuring method defines certain tolerances within which the measurements can be performed. It is important to understand the possibilities to reduce uncertainty by limiting boundary conditions.

As an example, measurements according to UN ECE R117 on rolling sound can be performed within a test track surface temperature between +5 to + 50 °C. The measured sound level at a certain surface temperature shall then be corrected to a reference temperature of + 20 °C, based on a defined correlation correction between road surface temperature and sound level in this regulation. If the measurements can be made as close as possible to this reference temperature, the measurement uncertainty related to the influence of temperature can be reduced.

B. Use of compensations (reducing systematic errors)

Staying with the UN ECE R117 example, the measured sound level at a certain surface temperature shall then be corrected to a reference temperature of + 20 °C, based on a defined correction between road surface temperature and measured sound level. The correction does not eliminate measurement uncertainty, but it does reduce the measurement uncertainty. The lowest possible uncertainty is if all measurements are performed at + 20 °C.

C. Use of an uncertainty model

As there is never a "true" value for the final result, there is a need to use an uncertainty model to define the tolerances (as expected variance) of the measured value. Such uncertainty models are defined in ISO 5725 and in the ISO/IEC Guide 98-3.

D. Repetition of measurements

In a regulation/measuring method, a certain number of repetitions of a test condition can be defined, as a means to reduce uncertainties. Therefore, by repeating measurements under equal boundary conditions, using the mathematical mean of the measurements minimizes the uncertainty, as the influence of random errors will be reduced. An example of this practice is the use of four measurement runs in UN ECE R51.03 which are then mathematically averaged.

This approach is shown in figure 3.1

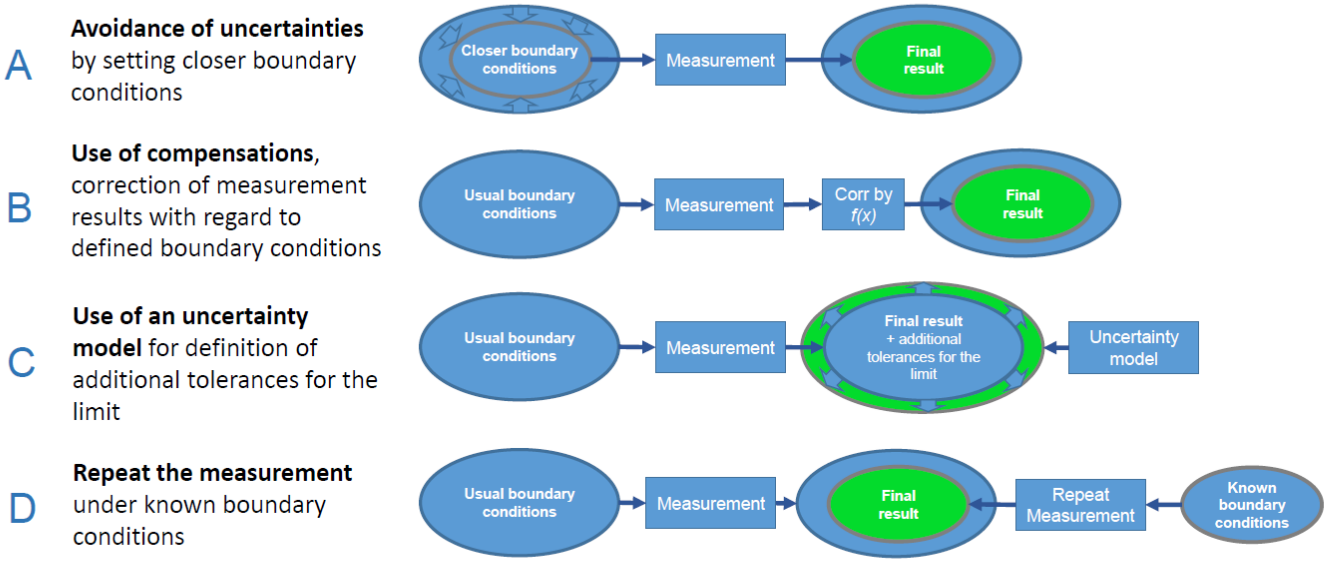

Figure 3.1 Approach to reduce measurement uncertainty[4]

4. Stages of uncertainty evaluation

There are in principle two stages to consider:

1. The formulation stage:

a) defining the output quantity *Y* (the measurand)
b) identifying the input quantities on which *Y* depends
c) developing a measurement model relating *Y* to the input quantities
d) on the basis of available knowledge, assigning probability distributions – Gaussian, rectangular, etc - to the input quantities (or a joint probability distribution to these quantities that are not independent)

2. The calculation stage consists of propagating distributions for the input quantities through the measurement model to obtain the probability distributions for the output quantity *Y* and summarizing by using distribution to obtain:

a) the expectation of *Y*, taken as an estimate *y* of *Y*
b) the standard deviation of *Y*, taken as the standard uncertainty of $\mu(y)$ associated with *y*
c) the coverage interval containing *Y* with a specific coverage probability.

5. ISO/IEC 98-3 (GUM) approach

The GUM uncertainty framework uses:

a) the best estimates of x_i of the input quantities X_i
b) the standard uncertainties $\mu(x_i)$ associated with x_i
c) the sensitivity coefficients c_i

to form an estimate *y* of the output quantity *Y* and the associated standard uncertainty $\mu(y)$.

An input quantity to the uncertainty model is never exact, so an assessment must be done.

In most cases, a measurand Y is not measured directly, but is determined from N other quantities $X_1, X_2,\ldots, X_N$ through a functional relationship:

$$Y = f\ (X_1, X_2,\ldots, X_N) \qquad (1)$$

A general expression for a measurement model is:

$$h(Y, X_1 \ldots\ldots X_N) = 0 \qquad (2)$$

It is taken that a procedure exists for calculating Y given $X_1\ldots..X_N$ in equation (2) and that Y is uniquely defined by this equation.

The input quantities $X_1, X_2,\ldots, X_N$ upon where the output quantity Y depends, may themselves be viewed as measurands and may themselves depend on other quantities, including corrections and correction factors for systematic errors.

An estimate of the measurand Y denoted by y, is obtained from Equation (1) using input estimates $x_1, x_2,\ldots.,x_N$ for the values of N quantities $X_1, X_2,\ldots,X_N$. Thus, the output estimate y, which is the results of the measurements, is given by:

3. $y = f\ (x_1, x_2,\ldots,x_N)$ (3)

If the input quantity can lie on both sides of the true value and the probability is higher if it is closer to the true value than further away from it, one can assume a normal ("gaussian") distribution as a good approximation. Figure 4.1 show such a normal distribution, where μ is the mean value of the variance V of the quantity and σ is the standard deviation ($V = \sigma^2$).

.

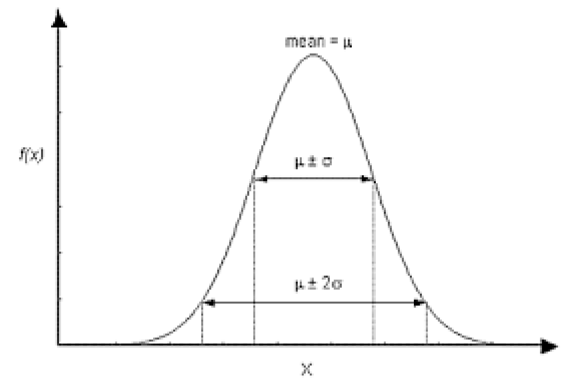

Figure 4.1 Normal ("gaussian") distribution

If all values of the input quantity are equally likely within a given interval, the distribution is rectangular, as shown in figure 4.2.

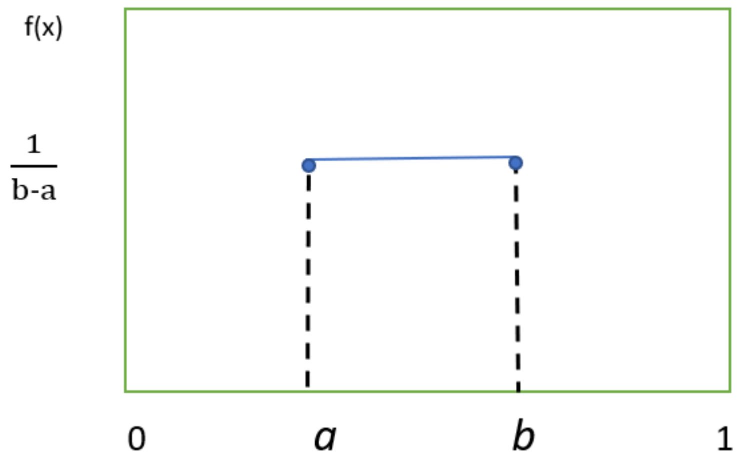

Figure 4.2 Rectangular probability distribution within the interval a to b.

In some cases, the input quantity can only lie above or below a fixed value, and in that case, one has a single-sided distribution. In some cases, a half-normal distribution (single-sided) can also be a good approximation, if for example, the input quantity is more likely to lie close to a limit value, than further away.

Knowledge about an input quantity X_i is established from repeated indication values (Type A evaluation of uncertainty) or scientific judgement or other information concerning the possible values of the quantity (Type B evaluation of uncertainty).

In **Type A** evaluation of measurement uncertainty, the assumption is often made that the distribution best describing an input quantity X given repeated indication values of it (obtained independently) is a Gaussian distribution (figure 4.1). X then has expectation equal to the average indication value and standard deviation equal to the standard deviation of the average.

When the uncertainty is evaluated from a small number of indication values (regarded as instances of an indication quantity characterized by a Gaussian distribution), the corresponding distribution can be taken as a t-distribution. Other considerations apply when the indication values are not obtained independently. (See ISO/IEC GUIDE 98-3 Annex G)

For a **Type B** evaluation of uncertainty, often the only available information is that X lies in a specified interval [a, b]. In such a case, knowledge of the quantity can be characterized by a rectangular probability distribution with limits a and b (figure 4.2). If different information were available, a probability distribution consistent with that information would be used.

Estimation of type B uncertainties is often based on calculations, experience, calibration, etc.

The final resulting value consists of the measured value + the input quantity for the uncertainty factor, δ_1 to δ_i

$$Y_{final} = Y_{meas} + \delta_1 + \delta_2 + \delta_3 + \delta_4 + \cdots\cdots + \delta_i \qquad (4)$$

The uncertainty contribution on the measurand due to the input quantity δ_i is $c_i\ \mu_i$, where c_i is the sensitivity coefficient and μ_i the uncertainty.

Sensitivity coefficients $c_1 \ldots\ldots c_N$ describe how the estimate y of Y would be influenced by small changes in the estimates $x_1 \ldots\ldots x_N$ of the input quantities $X_1 \ldots\ldots X_N$. For the measurement function (1), c_i equals the partial derivative of the first order of f with respect of X_i, evaluated at $X_1 = x_1$, $X_2 = x_2$, etc. For the linear measurement function

$$Y = c_1X_1 + \ldots\ldots\ldots + c_NX_N \qquad (5)$$

with $X_1 \ldots\ldots X_N$ independent, a change in x_i equal to $\mu_i(x_i)$ would give a change $c_i\,\mu_i(x_i)$ in y. This statement would generally be approximate for the models (1) and (2). The relative magnitudes of the terms $|c_i|\,\mu_i(x_i)$ are useful in assessing the respective contributions from the input quantities to the standard uncertainty $\mu(y)$ associated with y.

The sensitivity coefficients show how the variables in (3) will influence the magnitude of the result of y,
They function as a multiplier used to convert the uncertainty components to the right units and magnitude for the uncertainty analysis.

If there is no need for a sensitivity coefficient, for example if the input quantities or uncertainty contributors are all reported in the same unit of measure. In such cases, the sensitivity coefficient can be set to 1.

The combined standard uncertainty $\mu_c(y)$ will then be the positive square roots of the combined variances:

$$\mu_c(y) = \sqrt{\sum \mu^2} \qquad (6)$$

The combined standard uncertainty is expressed as the standard deviation of the measurand.

The expanded standard uncertainty, U, is calculated by multiplying the combined standard uncertainty, $\mu_c(y)$, with a coverage factor, k, for the chosen coverage probability:

$$U = k \cdot \mu_c(y) \qquad (7)$$

The coverage factor can be chosen such that the result U can be interpreted as the width of a certain confidence interval (although the GUM states that this is statistically not totally true).

Normally, the k factor lies between 2 and 3, which correspond to a level of confidence of approximately 95 % or 99 %. However, in other cases k can also be less than 2.

The result of the measurement is then conveniently expressed as:

$$Y = y \pm U \qquad (8)$$

For practical reasons, a table with an uncertainty budget should be set up, where all relevant quantities are defined. An example of such table is shown, below, taken from an ISO standard to measure the stationary sound pressure level from road vehicles[5].

Table 4.1 Uncertainty budget for determination of reported sound pressure level[5]

Quantity	**Estimate dB**	**Standard uncertainty, μ_i, dB**	**Probability distribution**	**Sensitivity coefficient, c_i**	**Uncertainty contribution, $c_i \mu_i$, dB**
$L_{\text{Ameas, i}}$	$L_{\text{Ameas, i}}$	-	-	-	-
δ_1	-	-	-	-	-
δ_2	-	-	-	-	-
δ_3	-	-	-	-	-
δ_4	-	-	-	-	-
δ_5	-	-	-	-	-
δ_6	-	-	-	-	-

6. ISO 5725 approach

ISO 5725 – Accuracy (trueness and precision) of measurement methods and results[2].

The method consists of 6 parts:
Part 1: General principles and definitions
Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method
Part 3: Intermediate measures of the precision of a standard measurement method
Part 4: Basic methods for the determination of the trueness of a standard measurement method
Part 5: Alternative methods for the determination of the precision of a standard measurement method
Part 6: Use in practice of accuracy values

This standard is primarily suited for inter- or intra-laboratory comparisons of results.

The following is a basic summary of the statistical model given in Part 1 of the standard and from UTAC[6]:

For estimation of the accuracy (trueness and precision) of a measurement method, one can assume that every test result, *Y,* is the sum of three components:

$$Y_{ij} = m + L_i + \varepsilon_{ij} \qquad (9)$$

where:
Y_{ij} is the jth test result from laboratory i
m is the general mean (expectation);
L_i is the laboratory effect i, I = 1 to p, with variance σ_L^2;
ε_{ij} is the residue (random error) on the jth result from laboratory I, j to n, with variances:

$$\text{var}(L) = \sigma_L^2 \qquad (10)$$
$$\text{var}(\varepsilon) = \sigma_\varepsilon^2 \qquad (11)$$

Methods are given in Part 3 for measuring the size of some of the random components of *L.* In general, *L,* can be considered as the sum of both random and systematic errors.

Within a single laboratory, its variance under repeatable conditions is called the within-laboratory variance and is expressed as:

$\sigma_L^2 = var(\varepsilon) = \sigma_W^2$ (12)

This arithmetic mean is taken over all those laboratories taking part in the accuracy experiment which remain after outliers have been removed.

When this basic model is adopted, the repeatability variance is measured directly as the variance of the error term ε, but the reproducibility variance depends on the sum of the repeatability variance and the between-laboratory variance in (10).

For precision evaluations:
- Repeatability standard deviation: $\sigma_r = \sigma_\varepsilon$
- Reproducibility standard deviation: $\sigma_R^2 = \sigma_L^2 + \sigma_r^2$

Variance component estimation:
- Repeatability: $s_r = s_\varepsilon$
- Reproducibility: $s_R^2 = s_\varepsilon^2 + s_L^2$

For trueness evaluations:

$$\delta = \mathrm{m} - \mu \qquad (13)$$

where μ is the reference value if it exists
Estimated by:

$$\hat{\delta} = \hat{\mathrm{m}} - \mu \qquad (14)$$

The combined uncertainty $\mu_c(y)$ comes from the values of precision:

- in conditions of repeatability: $\mu_c(y) = s_\varepsilon$
- in conditions of reproducibility: $\mu_c(y) = s_R$

The expanded uncertainty: $U = k \cdot \mu_c(y)$ (15)

Where k is the chosen coverage factor.

7. Example of estimation of calculation of expanded uncertainty – UNECE Reg.51.03 and ISO 362-1.

In UNECE Reg.51.03, the test method (Annex 3) is based on two driving conditions; a constant speed test, L_{crs}, and a wide-open throttle acceleration test, L_{wot}, to determine the final type-approval level, L_{urban}.

In the table[7] below, the impact of the different quantities on these indicators has been estimated for the Run-to-run, Day-to-day, Site-to-site and Vehicle-to-vehicle situations. Some of the different impacts are based on calculations from tolerances in the regulations, while others are based on experiences. Based on the probability distribution, the variance and the standard deviation is calculated. For each of the quantities, their contribution (in %) has been calculated and the colour scheme makes it easy to understand the influence of the quantity to the total uncertainty. Some of these quantities can be compensated for, like the influence of temperature and test track variations, while other is of random type, like instrumentation accuracy and cannot be compensated. In the example shown below, the estimated total expanded uncertainty has been calculated to ± 3.45 dB for a coverage factor of *k* = 2 (95 % level of confidence).

Table 6.1 Example of calculation of uncertainties for UNECE Reg.51.03[7]

Situation	Input Quantity	estimated deviations of the meas. result (peak-peak)		Impact on Lurb	Probability Distribution	Variance	Standard deviation	Share [%]	Combined standard uncertainty
		Lwot	Lcrs						
Run to Run	Micro climate wind effect	1.60	1.50	**1.57**	gaussian	0.15	0.392	5.1%	**0.53**
	Deviation from centered driving	0.50	0.50	**0.50**	rectangular	0.02	0.144	0.7%	
	Start of acceleration	0.60	0.00	**0.40**	rectangular	0.01	0.114	0.4%	
	Speed variations of +/- 1km/h	0.50	0.50	**0.50**	rectangular	0.02	0.144	0.7%	
	Load variations during cruising	0.00	1.00	**0.34**	gaussian	0.01	0.085	0.2%	
	Varying background noise	0.40	0.40	**0.40**	rectangular	0.01	0.115	0.4%	
	Variation on operating temperature of engine and tyres	0.80	0.80	**0.80**	rectangular	0.05	0.231	1.8%	
Day to Day	Barometric pressure (Weather +/-30 hPa)	0.70	0.00	**0.46**	gaussian	0.01	0.116	0.4%	**0.92**
	Air temperature effect on tyre noise (5-10°C)	0.00	0.00	**0.00**	rectangular	0.00	0.000	0.0%	
	Air temperature effect on tyre noise (0-40°C)	2.20	3.60	**2.67**	rectangular	0.60	0.772	20.0%	
	Varying background noise during measuremnt	0.00	0.00	**0.00**	rectangular	0.00	0.000	0.0%	
	Air intake temperatuire variation	1.60	0.00	**1.06**	rectangular	0.09	0.305	3.1%	
	Residual humidity on test track surface	0.90	2.10	**1.31**	rectangular	0.14	0.377	4.8%	
Site to Site	Altitude (Location of Test Track) 100 hPa/1000m	0.70	0.00	**0.46**	rectangular	0.02	0.134	0.6%	**1.24**
	Test Track Surface	3.40	5.50	**4.11**	rectangular	1.41	1.187	47.3%	
	Microphone Class 1 IEC 61672	1.00	1.00	**1.00**	gaussian	0.06	0.250	2.1%	
	Sound calibrator IEC 60942	0.50	0.50	**0.50**	gaussian	0.02	0.125	0.5%	
	Speed measuring equipment continuous at PP	0.10	0.10	**0.10**	rectangular	0.00	0.029	0.0%	
	Acceleration calculation from vehicle speed measurement	0.50	0.50	**0.50**	rectangular	0.02	0.144	0.7%	
Vehicle to Vehicle	Production Variation Tyre and aging of tyres	0.80	1.50	**1.04**	gaussian	0.07	0.259	2.3%	**0.57**
	Production Variation in Power	0.40	0.40	**0.40**	rectangular	0.01	0.115	0.4%	
	Battery state of charge for HEVs	0.00	0.00	0.00	rectangular	0.00	0.000	0.0%	
	Production Variability of Sound Reduction Components	1.10	0.00	**0.73**	gaussian	0.03	0.182	1.1%	
	Impact of variation of vehicle mass	1.60	1.60	**1.60**	rectangular	0.21	0.462	7.2%	
						2.98	1.73	100.0%	

Coverage Factor	Overall Combined Uncertainty +/-	Expanded uncertainty (95%) +/-
k=2 (95%)	1.73	3.45

In ISO 362-1, the appendix dealing with the measurement uncertainty has recently been updated in the ongoing revision.
In table 6.2, the uncertainty budget for the parameters influencing the total expanded uncertainty is listed.

Table 6.2 Uncertainty budget for determination of urban sound pressure level

Situation	Quantity	Peak to peak estimation L_{wot}	Peak to peak estimation L_{crs}	Impact on L_{urban}	Probability distribution	Standard uncertainty	95% Uncertainty
		dB	dB	dB		+/- dB	+/- dB
Run to Run	Micro climate wind effect	0,5	0,5	0,50	gaussian	0,13	0,6
	DRIVER #1: Deviation from centered driving	0,5	0,5	0,50	rectangular	0,14	
	DRIVER #2: Start of acceleration	0,5	0,5	0,50	rectangular	0,14	
	DRIVER #3: Speed variations of +/- 1km/h	0,3	0,3	0,30	rectangular	0,09	
	DRIVER #4: Load variations during cruising	0,3	0,5	0,37	gaussian	0,09	
	Varying background noise	0,1	0,1	0,10	rectangular	0,03	
	Variation on operating temperature of engine (WOT) and tyres (WOT&CRS) ==> See ISO 362-1 NOTE	0.25	0,25	0,25	rectangular	0,07	
Day to Day	Barometric pressure (Weather +/- 30 hPa)	1,0	0,0	0,66	gaussian	0,17	1,7
	Air temperature effect on tyre noise (5-10°C)	0,0	0,0	0,00	rectangular	0,00	
	Air temperature effect on tyre noise (10-40°C)	2,0	2,0	2,00	rectangular	0,58	
	Varying background noise during measurement	1,0	1,0	1,00	rectangular	0,29	
	Air intake temperature variation	1,5	0,0	0,99	rectangular	0,29	
	Residual humidity on test track surface	1,0	1,0	1,00	rectangular	0,29	
Site to Site	Altitude (Location of Test Track) -100 hPa/1000m (from 1015 to 915 hPa)	1,00	0,0	0,66	rectangular	0,19	2,7
	Test Track Surface	3,5	5,0	4,01	rectangular	1,00	
	Microphone Class 1 IEC 61672	0,6	0,6	0,60	gaussian	0,15	
	Sound calibrator IEC 60942	0,8	0,8	0,80	gaussian	0,20	
	Speed measuring equipment continuous at PP	0,1	0,1	0,10	rectangular	0,03	
	Acceleration calculation from vehicle speed measurement	0,5	0,0	0,33	rectangular	0,10	

Situation	Quantity	Peak to peak estimation L_{acc}	Peak to peak estimation L_{crs}	Impact on L_{urban}	Probability distribution	Standard uncertainty	95% Uncertainty
		dB	dB	dB		+/-dB	+/- dB
Vehicle to Vehicle	Production variation of tires, including aging of tires until 1 year	0,80	1,50	1,04	gaussian	0,259	**0,57**
	Production variation in engine power output	0,40	0,40	0,40	rectangular	0,115	
	Production variability of sound reduction components	1,1	0,0	0,73	rectangular	0,182	
	Vehicle mass variation from mass in running order	1,6	1,6	1,60	rectangular	0,462	

References

[1] Regulation (EU) 2018/858 *on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles.*

[2] ISO 5725:1994. *Accuracy (trueness and precision) of measurement methods and results – Part 1 to Part 6.*

[3] ISO/IEC Guide 98-3:2008. *Uncertainty of measurements. Part 3 – Guide to the expression of uncertainty in measurements* (GUM:1995).

[4] GRBP TFMU-02-04. *How to handle measurement uncertainties.* OICA, TF MU, Brussels, Nov.2019.

[5] ISO 5130:2019. *Acoustics – Measurements of sound pressure level emitted by stationary road vehicles.* Geneva, Switzerland.

[6] GRBP TFMU-01-05. *Experimental approach for evaluating uncertainties associated to stationary vehicle noise according to ISO 5725.* UTAC, TF MU, Brussels, May 2019.

[7] GRBP TFMU-02-06. (OICA) MU Calculation Sheet rev7 public.xlsx, Brussels, Nov.2019