SOUND LEVEL NOISE CORRECTIONS AND UNCERTAINTY

1. KEY CONCEPTS

Mean

The *mean* $\bar{x}$ is the central value of a finite set of numbers or *sample*. For an uncorrelated random variable x, where all observations N are equally probably, the mean is given by:

$$\bar{x} = \frac{1}{N}\sum_{i}^{N} x_i \tag{1.1}$$

Standard Deviation

The *standard deviation* is a measure of the spread of a set of values relative to their *mean*.

The *population standard deviation* σ is the spread of the values of an entire population. In statistics, the *population* refers to the subject of a statistical study, that is, a group of individuals with characteristics in common from which a statistical sample is drawn for study. The *population standard deviation* is used to calculate the spread of the observations of the entire population, or the spread of a sub-sample of the entire population and do not wish to extrapolate to the entire population. For a *population mean* μ and N observations, the population standard deviation is given by:

$$\sigma = \sqrt{\sum\nolimits_{i}^{N} (x_i - \mu)^2 / N} \tag{1.2}$$

The *sample standard deviation* S is the spread of the values of an entire population calculated from a small sample. In other words, S is an *unbiased estimator* of the population standard deviation σ based on a limited sample. If all you have is a sample, but you wish to make a statement about the population standard deviation from which the sample is drawn, you need to use the sample standard deviation. Do not confuse S with the standard deviation of the sample.

$$S = \sqrt{\sum\nolimits_{i}^{N} (x_i - \bar{x})^2 / (N-1)} \tag{1.3}$$

The standard deviation can be given as a relative metric. The *relative standard deviation* (RSD), also known as coefficient of variation (CV), is the ratio of standard deviation σ to the absolute value of the mean μ:

$$RSD = \frac{\sigma}{|\mu|} \approx \frac{S}{|\bar{x}|} \tag{1.4}$$

Standard Error

The *standard error* of a *statistic* is the standard deviation of its *sampling distribution*. A *statistic* is a statistical measure from a sample, that is, any quantity computed from values in a sample; the *mean* and *standard deviation* are examples of statistics. A *sampling distribution* is the probability distribution of a given statistic based on a random sample. The standard error is used as a measure of *uncertainty*.

The standard error can be calculated for any statistic. In this note, we are interested in the standard error of the mean and the standard deviation.

For a sample of M statistically independent and identically distributed (IID) observations $x = [x_1, x_2, \dots, x_N]$, the *standard error of the sample mean* (SEM) is calculated with the formula:

$$\mathrm{Std}(\bar{x}) = \frac{\sigma_x}{\sqrt{M}} \approx \frac{S_x}{\sqrt{M}} \tag{1.5}$$

where σ_x and S_x are the *population* and *sample* standard deviations of x, and $\bar{x}$ is the sample mean. The demonstration of this formula is straightforward and follows the expansion of the *variance of the sample mean* of a sample x of M statistically *independent and identically distributed* (IID) observations,

$$\mathrm{Var}(\bar{x}) = \frac{1}{M}\sum_{i}^{M} x_i \tag{1.6}$$

From *product* and *addition* properties of the variance $\mathrm{Var}(ax) = a^2\mathrm{Var}(x)$ and $\mathrm{Var}(x+y) = \mathrm{Var}(x) + \mathrm{Var}(y) + \mathrm{Cov}(x,y)$, and knowing that for IID observations $\mathrm{Cov}(x,y) = 0$ and all variances are σ^2, the variance of the mean becomes:

$$\mathrm{Var}(\bar{x}) = \frac{1}{M^2}\sum_{i}^{M} Var(x_i) = \frac{1}{M^2} M\sigma^2 = \frac{\sigma^2}{M} \tag{1.7}$$

Taking the square root of the result we arrive at the *standard deviation* (or standard error) *of the mean*. Note that the population standard deviation σ of each observation in x is the same as the population standard deviation of x, since all observation are statistically independent ($\sigma = \sigma_x$). It means that the standard error can be calculated directly from the observations in x.

Similarly, for a *normally-distributed* sample of M statistically independent and identically distributed (IID) observations $x = [x_1, x_2, \ldots, x_N]$, the *standard error of the sample standard deviation* (SES) is calculated with the formula:

$$\mathrm{Std}(S_x) = \sigma_x \sqrt{1 - \frac{2}{M-1} \cdot \left(\frac{\Gamma(M/2)}{\Gamma[(M-1)/2]}\right)^2} \tag{1.8}$$

where σ_x and S_x are the *population* and *sample* standard deviations of the standard deviation. A practical application of this formula will use S_x instead of σ_x. The demonstration of this formula is more elaborate and I will not include it here, but it follows the expansion of the *variance of the sample standard deviation of* a sample x of M statistically *independent and identically distributed* (IID) observations,

$$\mathrm{Std}(S_x) = \sqrt{E([E(S_x) - S_x]^2)} \tag{1.9}$$

An unbiased version of the formula above is:

$$\mathrm{Std}(S_x) = S_x \sqrt{\frac{M-1}{2} \cdot \left(\frac{\Gamma[(M-1)/2]}{\Gamma(M/2)}\right)^2 - 1} \tag{1.10}$$

Standard Uncertainty

The *standard uncertainty* $u(x)$ of a sample of M observations x is the estimated standard deviation of x. The *standard uncertainty* obtained by the statistical analysis of a series of observations is classified by the Guides for Measurement Uncertainty (GUM) as *type A*.

Individual observations differ in value because of random variations in the influence quantities or random effects. The best estimate of a type A standard uncertainty is given by the standard deviation of the sample mean, that is the, the *standard error of the mean*.

The *combined standard uncertainty* of a measurement accounts for all sources of error. It is calculated as the square root of the sum of variances or covariances of all contributing sources of error, weighted according to their individual effect on the measurement result.

Expanded Uncertainty

The *expanded uncertainty* is a quantity defining the interval about a measurement result $\hat{x}$ within which the value of the measurand x can be confidently asserted to lie. The expanded uncertainty U is obtained by multiplying the standard uncertainty u (or the combined standard uncertainty u_c) by a *coverage factor* k.

$$U = ku \tag{1.11}$$

It can be stated that the measurand x (i.e., the real value of the measured quantity) lies within $\hat{x} \pm U$ within a specific *confidence level*. The range of values $\hat{x} \pm U$ is known as the *confidence interval* (CI).

The *confidence level* is the percentage of times that the measurement is expected to fall within the confidence interval. In other words, the confidence level is the *area* of the probability distribution of x within which observations will likely lie, and the confidence interval are the top and bottom *limits* associated with that area.

The *coverage factor* k is chosen on the basis of the desired confidence level associated with the confidence interval. In a Gaussian distribution, coverage factors $k = 1, 2, 3$ correspond to $CI = 68\%, 95\%, 99\%$.

For a normal distribution $x \sim \mathcal{N}(\mu, \sigma^2)$ and confidence factor k, the confidence interval can be easily calculated using the *error function* $\mathrm{erf}(z')$, where z' is the *z-score* divided by $\sqrt{2}$. The *z-score* is the unbiased, normalised value of an observation x.

$$\mathrm{erf}(z') = \frac{2}{\sqrt{\pi}} \int_0^{z'} e^{-t^2} dt \tag{1.12}$$

$$z' = \frac{z}{\sqrt{2}} \tag{1.13}$$

$$z = \frac{x - \mu}{\sigma} \tag{1.14}$$

The formula above returns the *confidence interval ratio* (CRI). To obtain the confidence interval as a percentage (CI), multiply by 100.

The reason for using the error function for calculating the confidence interval is that $\mathrm{erf}(z')$ is a common algorithm widely available and efficiently implemented in most computer languages (see MATLAB erf.m). Its relation with the Gaussian probability function $g(x)$ can be easily demonstrated. The Gaussian function is:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2} \tag{1.15}$$

By replacing t with $z' = (x - \mu)/\sigma\sqrt{2}$ and dt with $dz' = x/\sigma\sqrt{2}$ in $\mathrm{erf}(z')$ we arrive at a more intuitive equation for the calculation of the confidence interval in a Gaussian distribution with coverage factor k.

$$CIR = \mathrm{erf}(z') = 2 \int_0^{z'} g(x) dx \tag{1.16}$$

But where is k in this equation? k is simply the value of x that will result in the coverage interval we are looking for. Therefore, for calculating CI the top integral limit becomes $z' = (k - \mu)/\sigma\sqrt{2}$.

In general, we start from a confidence interval for which we want to obtain the coverage factor. For that purpose, we apply the *inverse error function* erf^{-1} to the confidence interval ratio $CI/100$ and divide the result by $\sqrt{2}$.

$$k = \frac{\mathrm{erf}^{-1}(CIR)}{\sqrt{2}} \tag{1.17}$$

2. MEAN, STANDARD DEVIATION, AND UNCERTAINTY OF SOUND LEVELS

The following text is based on the analysis and conclusions from Taraldsen et al. (2015).

Mean

The mean of a random sample of sound level measurements L_i ($i = 1,2,3 \dots N$) is routinely estimated as the *mean of the observed energies* $L_{\overline{W}}$.

$$L_{\overline{W}} = 10 \log \frac{\overline{W}}{W_0} \quad \text{Energy Mean Level} \tag{2.1}$$

$$\overline{W} = \frac{1}{N} \sum_{i}^{N} W_i \quad \text{Energy Mean} \tag{2.2}$$

with W_0 the reference value. Note that W refers to any energy-scale metric, such as power, energy, intensity or exposure. $\overline{W}$ is referred to as the *energy mean* and $L_{\overline{W}}$ as the *energy mean level*.

Another expression for the mean of acoustic measurements is the *level mean* $\bar{L}$.

$$\bar{L} = \frac{1}{N} \sum_{i}^{N} L_i \quad \text{Level Mean} \tag{2.3}$$

The *Jensen inequality* shows that $\bar{L}$ will tend to underestimate $L_{\overline{W}}$ (i.e., $\bar{L} < L_{\overline{W}}$). The law of large numbers ensures that $\overline{W}$ asymptotically approaches a Gaussian distribution of mean μ_W and *standard error* $u(\overline{W})$. That is why $L_{\overline{W}}$ is preferred instead of $\bar{L}$ as an estimator of the mean decibel level. $\bar{L}$ is valid only for data with a small variance.

Uncertainty of the Mean

The *standard uncertainty of the mean* is a measure of the margin of error associated with a stated mean value. This section focuses on the *statistical part* of the measurement uncertainty, which is only one of all possible sources of error (e.g. accuracy of the measuring device, uncertainty in transducer position, meteorological conditions, etc). The uncertainty from all sources of error must be taken into account. There will be occasions when one source of error will dominate over the rest; even in those cases it is advisable to calculate the uncertainty of each contributor to justify the use of the dominant source of uncertainty.

The *standard uncertainty* u can be multiplied by a *coverage factor* k to obtain the *expanded uncertainty* U for a specific *confidence interval*. The common value of $k = 1.96$ corresponds to a 95% confidence interval.

When it comes to decibels, there are three methods that can be used for estimating the measurement uncertainty of the mean: *level uncertainty* $u(\bar{L})$, *GUM energy uncertainty level* $u(L_{\overline{W}})$, and *finite-difference energy uncertainty level* $u_f(L_{\overline{W}})$. The latter is preferred for most cases. The three options are described below.

Level Uncertainty $u(\bar{L})$

The *level uncertainty* $u(\bar{L})$ is the standard measurement uncertainty of the *level mean* $\bar{L}$ as described in the Guides for Measurement Uncertainty (GUM). This method of calculating the measurement uncertainty is not recommended except when the standard deviation is small ($\sigma_L \ll 1$ dB).

$$U(\bar{L}) = u(\bar{L}) \cdot k \quad \text{Level Uncertainty } (Expanded) \tag{2.4}$$

$$u(\bar{L}) = \frac{S_L}{\sqrt{N}} \quad \text{Level Uncertainty } (Standard) \tag{2.5}$$

$$S_L = \sqrt{\sum_{i}^{N} (L_i - \bar{L})^2 / (N-1)} \quad \text{Sample Standard Deviation of } L_i \tag{2.6}$$

where L_i are the measured levels, N the total number of observations, and k the coverage factor.

GUM Energy Uncertainty Level $u(L_{\overline{W}})$

The *energy uncertainty level* $u(L_{\overline{W}})$ is the standard measurement uncertainty of the *energy mean level* $L_{\overline{W}}$ calculated using the general approach recommended in the GUM.

$$U(L_{\overline{W}}) = u(L_{\overline{W}}) \cdot k \quad \text{Energy Uncertainty Level } (Expanded) \tag{2.7}$$

$$u(L_{\overline{W}}) = \frac{10}{\ln(10)} \frac{u(\overline{W})}{\overline{W}} \quad \text{Energy Uncertainty Level } (Standard) \tag{2.8}$$

$$u(\overline{W}) = S_W / \sqrt{N} \quad \text{Energy Uncertainty } (Standard) \tag{2.9}$$

$$S_W = \sqrt{\sum_i^N (W_i - \overline{W})^2 / (N-1)} \quad \text{Sample Standard Deviation of } W_i \tag{2.10}$$

where L_i are the measured energies, N the total number of observations, and k the coverage factor.

The demonstration for $eq.$ (2.8) follows the explanation of the GUM. A general smooth function $g(x)$ of a Gaussian variable x with mean x_0 and small standard deviation $\mathrm{Std}(x)$ can be expressed as its first order Taylor approximation:

$$g(x) \approx g(x_0) + g'(x_0)(x - x_0) \tag{2.11}$$

It follows that $g(x)$ is approximately Gaussian with mean $g(x_0)$ and standard deviation $\mathrm{Std}[g(x)]$. Knowing that $\mathrm{Std}[g(x)] = g(x) - g(x_0)$ and $\mathrm{Std}(x) = x - x_0$, we have that:

$$\mathrm{Std}[g(x)] \approx |g'(x_0)| \mathrm{Std}(x) \tag{2.12}$$

By substituting $x = \overline{W}$ and $g(x) = L_{\overline{W}}$ into $eq.$ (2.12) we have:

$$\mathrm{Std}(L_{\overline{W}}) \approx |L'_{\overline{W}}| \mathrm{Std}(\overline{W}) \tag{2.13}$$

$$L'_{\overline{W}} = \frac{dL_{\overline{W}}}{d\overline{W}} = \frac{10}{\ln 10} \frac{1}{\overline{W}} \tag{2.14}$$

Note that $L_{\overline{W}}$ is expressed as $\ln(\overline{W}/W_0) \cdot 10/\ln 10$ before calculating its derivative $L'_{\overline{W}}$. Hence:

$$\mathrm{Std}(L_{\overline{W}}) \approx \frac{10}{\ln 10} \frac{\mathrm{Std}(\overline{W})}{\overline{W}} \tag{2.15}$$

Note that $\mathrm{Std}(\overline{W})$ is the sample standard deviation of the energy S_W given by $eq.$ (2.10). Now if we apply the definition of standard uncertainty $u = \sigma/\sqrt{N}$ to $eq.$ (2.15) we arrive at the expression for the *energy uncertainty level* $u(L_{\overline{W}})$ from $eq.$ (2.8). It is interesting to note that this uncertainty, stated on a *decibel scale*, has no level metrics involved in its calculation.

The Taylor approximation in $eq.$ (2.11) is valid when the increment in the independent variable $(x - x_0)$ is small. Therefore, $eq.$ (2.8) is only valid for values of $u(\overline{W})/\overline{W} \ll 1$, or in terms of the energy uncertainty level, $u(L_{\overline{W}}) \ll 4.3$ dB.

Finite Difference Energy Uncertainty Level $u_f(L_{\overline{W}})$

The *energy uncertainty level* $u_f(L_{\overline{W}})$ is the standard measurement uncertainty of the *energy mean level* $L_{\overline{W}}$ calculated using the finite difference approximation of $eq.$ (2.12).

$$U_f(L_{\overline{W}}) = u_f(L_{\overline{W}}) \cdot k \quad \text{Energy Uncertainty Level } (Expanded) \tag{2.16}$$

$$u_f(L_{\overline{W}}) \approx \left|L_{\overline{W}+u(\overline{W})} - L_{\overline{W}}\right| = 10\log\left[1 + \frac{u(\overline{W})}{\overline{W}}\right] \quad \text{Energy Uncertainty Level } (\mathit{Standard}) \tag{2.17}$$

$$L_{\overline{W}+u(\overline{W})} = 10\log\frac{\overline{W} + u(\overline{W})}{W_0} \quad \text{Energy Level with Uncertainty } (\mathit{Standard}) \tag{2,18}$$

This formula is more elementary than $eq.\,(2.8)$ but in some cases it can be more accurate. In particular, $u_f(L_{\overline{W}})$ should be considered when the sample size N is small and the corresponding uncertainty $u_f(L_{\overline{W}})$ is large (comparable to or larger than 4 dB). Unlike $u(L_{\overline{W}})$, $u_f(L_{\overline{W}})$ will never give unrealistically large values.

The demonstration for $eq.\,(2.17)$ follows the application of a finite difference approximation to the derivative term in $eq.\,(2.11)$, which results in the expression below:

$$\mathrm{Std}_f[g(x)] \approx \left|\frac{g\big(x_0 + \mathrm{Std}(x)\big) - g(x_0)}{\mathrm{Std}(x)}\right| \mathrm{Std}(x) = \left|g\big(x_0 + \mathrm{Std}(x)\big) - g(x_0)\right| \tag{2.19}$$

By substituting $x = \overline{W}$ and $g(x) = L_{\overline{W}}$ into $eq.\,(2.19)$ we obtain:

$$\mathrm{Std}_f(L_{\overline{W}}) \approx \left|L_{\overline{W}+\mathrm{Std}(\overline{W})} - L_{\overline{W}}\right| = 10\log\left[1 + \frac{\mathrm{Std}(\overline{W})}{\overline{W}}\right] \quad \text{Energy Std. Dev. Level} \tag{2.20}$$

$$L_{\overline{W}+\mathrm{Std}(\overline{W})} = 10\log\frac{\overline{W} + \mathrm{Std}(\overline{W})}{W_0} \quad \text{Energy Level with Std. Dev.} \tag{2.21}$$

Now if we apply the definition of standard uncertainty $u = \sigma/\sqrt{N}$ to $eq.\,(2.20)$ we arrive at the expression for the energy uncertainty level $u_f(L_{\overline{W}})$ from $eq.\,(2.17)$.

The following formula relates the two energy uncertainty levels. It shows that the finite difference approximation gives more optimistic estimates as compared with the GUM method ($u_f(L_{\overline{W}}) < u(L_{\overline{W}})$).

$$u_f(L_{\overline{W}}) = 10\log\left[1 + \frac{\ln 10}{10}u(L_{\overline{W}})\right] \tag{2.22}$$

Standard Deviation

Level Uncertainty $\mathrm{Std}(\overline{L})$

See $Eq.\,(2.6)$.

GUM Energy Uncertainty Level $\mathrm{Std}(L_{\overline{W}})$

See $Eq.\,(2.15)$.

Finite Difference Energy Uncertainty Level $\mathrm{Std}_f(L_{\overline{W}})$

See $Eq.\,(2.20)$.

Uncertainty of the Standard Deviation

The *standard uncertainty of the standard deviation* is a measure of the margin of error associated with a stated standard deviation value. This section focuses on the *statistical part* of the measurement uncertainty, which is only one of all the possible sources of error.

The *mean* is typically provided as a sufficient statistic for describing a measurement sample, but in certain scenarios stating the standard deviation is also important. In those cases, calculating the uncertainty for the standard deviation is recommended. An example would be studying the likelihood that sound exposure measurements will exceed a specific threshold (e.g., in the context of long-term exposure in a music festival).

The standard uncertainty of the standard deviation follows a Chi-square distribution. Therefore, the *coverage factor* k will be different to the one used for calculating the expanded uncertainty of the mean, as the mean follows a Gaussian distribution. The Chi-square distribution is not symmetric and its uncertainty is typically given as a *two-sided confidence interval* (search for "Confidence Interval of a Standard Deviation"). Alternatively, since a standard deviation is positive value, its uncertainty can also be given as a *one-sided upper confidence interval*. The formulas for uncertainty of standard deviation included in this chapter refer to a single confidence interval (68%?).

When it comes to decibels, there is one main method for estimating the measurement uncertainty of the standard deviation: the *level uncertainty* $u(S_L)$. The GUM or finite difference methods available for the uncertainty of the mean are not applicable here. For this reason, providing the level mean $\bar{L}$ and its standard deviation S_L, with their corresponding uncertainties $u(\bar{L})$ and $u(S_L)$, is the recommended option when both statistical descriptors (mean and standard deviation) are necessary to provide a complete picture (e.g., for confirming threshold exceedance in the context of long-term sound exposure).

Level Uncertainty $u(\bar{L})$

The *level uncertainty* $u(S_L)$ is the standard measurement uncertainty of the *sample standard deviation of the observed levels* S_L as described in the Guides for Measurement Uncertainty (GUM).

For a normally-distributed population of levels, the level uncertainty is given by (Taraldsen et al, 2015):

$$u(S_L) = S_L\sqrt{1-\frac{2}{N-1}\cdot\left(\frac{\Gamma(N/2)}{\Gamma[(N-1)/2]}\right)^2} \quad \text{Level Uncertainty } (Standard) \tag{2.23}$$

$$\Gamma(N) = (N-1)! \quad \text{Gamma Function} \tag{2.24}$$

where N is the number of observations and ! the *factorial* symbol (see *gamma.m* and *factorial.m* in MATLAB).

An approximate formula for normally-distributed levels with similar results to $eq.$ (2.23) is:

$$u(S_L) \approx \frac{\sigma_L}{\sqrt{2N-2}} \tag{2.25}$$

where σ_L is the *population standard deviation*. The sample standard deviation S_L is an *unbiased estimator* of the population standard deviation, and as such it can replace σ_L with minimum error.

3. EXAMPLES

Maximum Noise level from a Passing Snowmobile

The pass-by maximum noise level was measured on several snowmobiles and under different driving conditions in Norway in March 2015. The measurement distance was 7.5 m from the centre of a 60 m straight driving path with a microphone at 1.2 m height. The resulting A-weighted maximum levels with fast time averaging for a particular snowmobile driving at 30 km/h were:

$$L = [66.97, 66.44, 67.10, 66.62, 66.96, 66.91]$$

These measurements can be summarised by the level mean $\bar{L}$ and level standard deviation S_L. It is important to state the standard deviation on measurements of maximum levels. The number in brackets is the standard uncertainty of the last digit. For example, $\bar{L} = 66.8(1)$ indicates that the estimated mean is $\bar{L} = 66.8$ dB and the standard uncertainty is $u(\bar{L}) = 0.1$ dB.

$$\bar{L} = 66.8(1)\ \mathrm{dB}$$

$$S_L = 0.25(8)\ \mathrm{dB}$$

In this case, the statistical part of the measurement uncertainty is negligible compared to the uncertainty from other factors (sound level metre, measurement geometry, meteorological conditions).

Noise Exposure Level from a Passing Snowmobile

In the experiment just described the noise exposure was also measured. The resulting A-weighted sound exposure levels for the same snowmobile passing by at 30 km/h were:

$$L = [77.97, 78.69, 78.39, 78.55, 78.08, 78.97]\ \mathrm{dB}$$

The results can be summarised by the mean. In this case, stating the standard deviation is not necessary and a better estimator of the mean, such as the *energy mean level* $L_{\overline{W}}$, can be used.

$$L_{\overline{W}} = 78.5(2)\ \mathrm{dB}$$

The corresponding uncertainties are $u(L_{\overline{W}}) = 0.1539$ and $u_F(L_{\overline{W}}) = 0.1539$; both are very close in this example ($\ll 4.3$ dB). Incidentally, $u_F(\bar{L})$ is also a good approximation in this case.

Sound Power of a Population of Road Vehicles

Measurements of pass-by noise exposure from passenger cars were made at Bratsberg in Trondheim in August 2000. Inversion of the sound propagation problem, including the ground effect, gives the following sound power level estimates in the 100 Hz third-octave frequency band for light vehicles driving at 55 km/h. The method for estimating sound power relied on an array of microphones.

$$L = [87.3, 85.5, 83.0, 116.7, 79.5, 84.2, 79.2, 79.9, 77.3, 78.1, \\ 78.3, 87.9, 77.7, 87.3, 87.6, 87.0, 79.3, 83.7, 85.5, 82.1, \\ 81.2, 80.9, 84.7, 86.2, 79.7, 84.2, 84.3, 84.6, 78.9, 81.2]\ \mathrm{dB}$$

The results can be summarised by the energy mean level $L_{\overline{W}}$.

$$L_{\overline{W}} = 102(3)\ \mathrm{dB}$$

The difference between this example and that of the snowmobile is that, in this case, the L values were sampled from a population of different vehicles, leading to a larger spread in the data. The standard *energy uncertainty level* is $u_F(L_{\overline{W}}) = 2.98$ dB, or its expanded value $U(L_{\overline{W}}) \approx 6$ dB for a 95% confidence interval. The statistical component of the uncertainty is significant and dominates over the other sources of uncertainty.

In this example, the other two types of uncertainty give unreasonable results. In particular, the energy uncertainty level is $u(L_{\overline{W}}) = 4.28$ dB and the level uncertainty is $u(\bar{L}) = 1.29$ dB. The assumption $u(L_{\overline{W}}) \ll 4$

dB is not satisfied; therefore, its value is expected to be inaccurate (overestimate). The level uncertainty also greatly underestimates the standard uncertainty.

Uncertainty of Sound Levels at Music Events

The World Health Organization (WHO) recommends that the A-weighted equivalent level over 4 h should not go above 100 dB at music events. The *equal energy hypothesis* assumed by the WHO states that an equal amount of energy always has the same damaging potential.

The Hove music festival lasted 5 days. Four persons with personal noise dose meters gave the following 17 maximal four-hour equivalent levels. Each level corresponds to a maximum 4-h equivalent level for one person during one day of concerts, but every person did not attend every festival day.

$$L = [\,93.2, 97.8, 98.7, 99.0, 93.4, 90.9, 93.7, 93.1, 95.1,\\ 96.9, 97.6, 98.9, 104.0, 103.0, 102.0, 101.0, 102.0]\ \mathrm{dB}$$

The results can be summarised by the level mean $\bar{L}$ and level standard deviation S_L.

$$\bar{L} = 97.6(9)\ \mathrm{dB}$$

$$S_L = 3.9(7)\ \mathrm{dB}$$

In this example, like in the first one, it is the distribution of levels itself that is of interest, and which gives for instance the probability of exceeding the 100 dB limit defined by the WHO for similar festivals. The *distribution* of measurements is completely specified by the mean and standard deviation assuming a Gaussian distribution.

4. CORRECTION OF BACKGROUND NOISE

Acoustic measurements are accompanied by a certain amount of undesired noise. Background noise is the combined result of self-noise from the electronics, electromagnetic interference and ambient noise in the acoustic medium. Any noise will affect the accuracy of our measurements. In this section, I explain the process to correct a measurement contaminated with noise, taking into account measurement uncertainties.

Let's start defining a few metrics:

<u>Signal to Noise Ratio (SNR)</u>: difference between the sound level of the signal (S) and the background noise level (N). This metric is suitable for those cases when the original signal is known. SNR lies within $-\infty$ and ∞.

$$SNR = S - N \tag{4.1}$$

<u>Signal-plus-Noise to Noise Ratio ($SNNR$)</u>: difference between the sound level of the noise-contaminated signal (SN) and the background noise level (N). This metric is an alternative to the SNR for measured signals. $SNNR$ lies within 0 and ∞.

$$SNNR = SN - N \tag{4.2}$$

<u>Corrected Signal Level (S_c)</u>: sound level of the signal after noise correction, calculated as ten times the base-10 logarithm of the energies difference between the measured (noise-contaminated) signal and the noise.

$$S_c = 10\log\left(10^{SN/10} - 10^{N/10}\right) \tag{4.3}$$

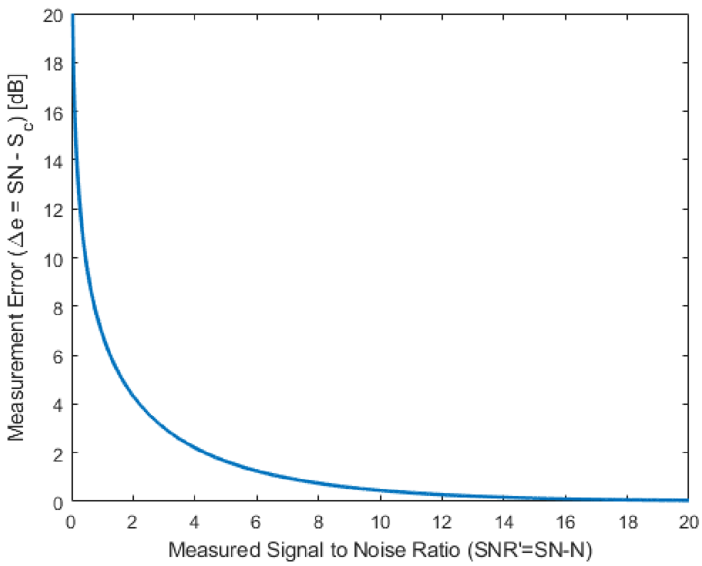
<u>Noise Correction Level (Δe)</u>: difference between the sound level of the measured (noise-contaminated) signal (SN) and sound level of the noise-corrected signal (S_c). It can also be defined as the difference between the signal-plus-noise to noise ratio ($SNNR$) and the signal to noise ratio (SNR). The closer the level of the original signal (S) is to the background noise (N), the further away the level of the corrected signal (S_c) will be from that of the original. For a signal completely buried in the noise, $SNNR = 0$, $S_c = -\infty$, and $\Delta e = -\infty$. This metric is a good indicator of the *uncertainty* of our measurements.

$$\Delta e = SN - S_c = SNNR - SNR = -10\log\left(1 - 10^{\frac{-SNNR}{10}}\right)^{*} \tag{4.4}$$

* Demonstration of how this equation has been obtained:

$$S_c = N + 10\log\left(10^{\frac{SNNR}{10}}\right)$$

$$\Delta e = SN - N - 10\log\left(10^{\frac{SNNR}{10}}\right) = SNNR - 10\log\left(10^{\frac{SNNR}{10}}\right) = -10\log\left(1 - 10^{\frac{-SNNR}{10}}\right)$$



SNR'	Δe
0.1	16.4
0.2	13.5
0.3	11.8
0.4	10.6
0.5	9.6
0.6	8.9
0.7	8.3
0.8	7.7
0.9	7.3
1.0	6.9
2.0	4.3
3.0	3.0
4.0	2.2
5.0	1.7
...	...

...	...
6.0	1.3
7.0	1.0
8.0	0.7
9.0	0.6
10.0	0.5
11.0	0.4
12.0	0.3
13.0	0.2
14.0	0.2
15.0	0.1
16.0	0.1
17.0	0.1
18.0	0.1
19.0	0.1
20.0	0.0

It becomes clear that the smaller $SNNR$, the larger the correction Δe we will need to apply to the measured signal. Since small values of $SNNR$ have a great effect on Δe, a moderate uncertainty in the measured levels SN and N can have a large impact on the noise-corrected level S_c.

Since the exact noise level that contaminated the measured signal is unknown, in order to get a reliable estimate of the original level of the signal S, it is advisable to calculate SN and N as the mean values of multiple observations ($\overline{SN}$, $\overline{N}$). The larger the number of observations, the smaller the uncertainty of the mean (see section about "Standard Error"). However, our certainty about $\overline{SN}$ and $\overline{N}$ will also depend on the variance of those values. In general, the uncertainties of $\overline{SN}$ and $\overline{N}$ will range from a fraction of a decibel (0.1 dB) to a few decibels (2 dB), meaning that a $SNNR$ of less than 1 dB will tend to result in an extremely large confidence interval for S_c. The mean ($\bar{S}_c$), top (S_{c1}) and bottom (S_{c2}) limits of the noise-corrected level can be calculated as follows:

$$\bar{S}_c = 10\log\left(10^{\overline{SN}/10} - 10^{\overline{N}/10}\right) \tag{4.5}$$

$$S_{c1} = 10\log\left(10^{\overline{SN}+\mathrm{u}(\overline{SN})/10} - 10^{\overline{N}-\mathrm{u}(\overline{N})/10}\right) \tag{4.6}$$

$$S_{c2} = 10\log\left(10^{\overline{SN}-\mathrm{u}(\overline{SN})/10} - 10^{\overline{N}+\mathrm{u}(\overline{N})/10}\right) \tag{4.7}$$

5. RELATIONSHIPS BETWEEN SNR, $SNNR$, AND Δe

Here I present various formulas to relate the *noise correction level* Δe, the *signal to noise ratio* SNR, and the *signal-plus-noise to noise ratio* $SNNR$.

$$\Delta e = -10\log\left(1 - 10^{\frac{-SNNR}{10}}\right) \tag{5.1}$$

$$\Delta e = 10\log\left(1 + 10^{\frac{-SNR}{10}}\right) \tag{5.2}$$

$$SNR = 10\log\left(10^{\frac{SNNR}{10}} - 1\right) \tag{5.3}$$