How to measure community tolerance levels for noise

Gunnar Taraldsen,[a)] Femke B. Gelderblom, and Truls T. Gjestland
The Norwegian University of Science and Technology and SINTEF, 7465 Trondheim, Norway

Relationships between noise exposure and transportation noise induced annoyance have been studied extensively for several decades. The annoyance due to aircraft noise exposure is in the present paper assumed to be influenced by the day–night yearly average sound level (DNL). It has long been recognized that the annoyance also depends on non-DNL factors, but this is complicated—resulting in a variety of different modeling strategies. Motivated by this, the community tolerance level (CTL) was introduced in 2011 for a loudness-based psychometric function. It is a single parameter that accounts for the aggregate influence of other factors. This paper suggests and investigates different methods for the measurement of the CTL. The methods are illustrated on data found in the literature, on recent surveys around two Norwegian airports, and on simulated data. The results from the presented methods differ significantly. An elementary method is shown to give a measurement of the CTL with smaller uncertainty, and is recommended as a replacement for the originally suggested least-squares method. Methods for evaluating the measurement uncertainty are also presented. © 2016 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4959134]

I. INTRODUCTION

A fundamental assumption is that there exists a relation between the noise annoyance experienced by an individual and the noise level the individual is exposed to. It is common to present a model for a population of individuals instead of a model for each individual. Historically, this has been done in terms of establishing dose-response relations that give the probability p of being highly annoyed as a function of the noise dose L_{dn}. Explicit examples of dose-response relations are given by

$$\text{Schultz (1978): } p = (0.8553L_{dn} - 0.0401L_{dn}^2 + 0.00047L_{dn}^3)\%,$$

$$\text{FICON (1992): } p = 1/(1 + \exp(11.13 - 0.14L_{dn})),$$

$$\text{Miedema and Oudshoorn (2001): } p = 1 - \Phi(4.542 - 0.0607L_{dn}),$$

$$\text{Fidell } et\ al.\ (2011): p = \exp(-\exp(4.6969 - 0.0691L_{dn})),$$

where Φ denotes the cumulative distribution function of a Gaussian variable with mean 0 and standard deviation 1. All of the above formulas can be found or derived by inspection of the indicated references. The Schultz relation was proposed as the best currently available estimate of public annoyance due to transportation noise of all kinds. The other relations have been derived by analysis of observations of annoyance due to aircraft noise.

Figure 1 shows the dose-response curves given by the above equations and also gives the resulting community tolerance level (CTL). The CTL is defined as the dose level L_{ct} at which the dose-response curve crosses $p = 50\%$. It should be noted that all curves have a similar shape and slope even though they have been obtained by different methods applied on different data.

The CTL L_{ct} is in this paper defined to be equal to the noise dose level where the probability of being highly annoyed equals 50%. This is as in the original paper by Fidell *et al.* (2011), but they considered only a particular class of loudness-based dose-response curve. This is discussed in more detail in Sec. III. We recommend that this
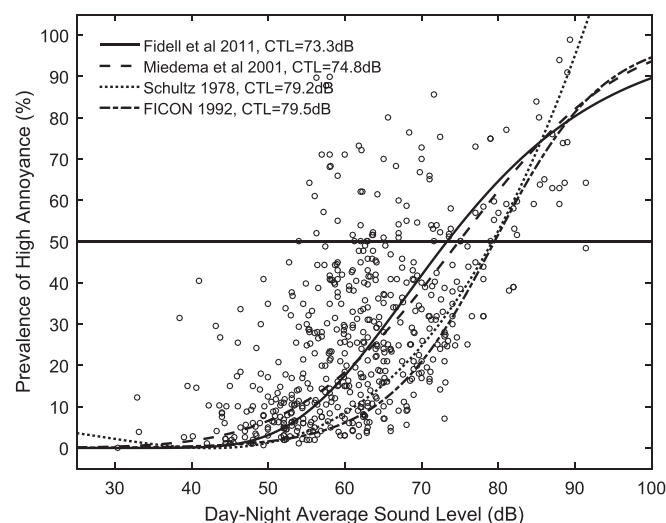

FIG. 1. Dose-response curves for noise and the corresponding CTLs. The 545 data points are the basis for the left-most dose-response curve (Fidell *et al.*, 2011, Table II).

a)Electronic mail: Gunnar.Taraldsen@ntnu.no

class should be used. For the presentation here it is, however, convenient to define the L_{ct} more generally as stated for any given dose-response curve. The identity $L_{ct} = \mathrm{CTL} = \mathrm{CTL}50$ indicates alternative use of symbols which is sometimes convenient.

Figure 1 also shows the observed (day–night yearly average sound level, %HA) $= (L_{dn}, p)$ pairs corresponding to the 545 interview sites from 43 aircraft noise surveys from Table II presented by Fidell *et al.* (2011). This will be referred to as the *Fidell dataset* in the following. The slope of the majority of the curves is visually consistent with the data points. It is, however, also noteworthy that visually the data points are shifted to the left when compared to the dose-response curves.

The curve obtained by Fidell *et al.* (2011) is given by the left-most curve with CTL = 73.3(1.1) dB. Throughout this paper we follow the convention that the number in parentheses is the numerical value of the standard uncertainty referred to the corresponding last digits of the quoted result (ISO/IEC, 2008, p. 25). The standard uncertainty of 1.1 dB is as stated by the original authors. The spread in the data points is large, but a visual fit based on the data points alone could seem to indicate a substantially lower CTL estimate somewhere in the range = 65–70 dB. The uncertainty of a visual fit is however large. This motivates the main question to be considered in this paper: How should the CTL be estimated based on observations?

The initial dose-response curve derived by Schultz (1978) can be seen simply as a polynomial fit to the observations from the studies available at that time. His Fig. 6 shows data points from 161 interview sites, and a visual fit gives a CTL close to 80 dB. This corresponds well to his polynomial fit, which gives CTL = 79.2 dB. It is noteworthy both that the data points in this case cluster around his curve, and that the majority of his data points around $p = 50\%$ are shifted approximately 15 dB to the right as compared with Fig. 1.

The purely visual fits can seem to indicate that people from the Schultz (1978) surveys tolerated approximately 15 dB more noise before being equally annoyed as the people in the more recent Fidell *et al.* (2011) surveys. Schultz (1978) considered, however, other transportation noise sources. Schomer *et al.* (2012) show that a CTL for road sources can be 5 dB higher than for air sources, and rail sources can have an even higher CTL if the source does not cause vibration and rattle. This can partly explain the 15 dB difference observed above: Noise from air sources are the most annoying sources. Others [Federal Interagency Committee on Noise (FICON, 1992)] have, however, presented results that indicate that the original Schultz (1978) curve is also a good fit for air sources as indicated by the FICON curve in Fig. 1.

This paper introduces and investigates methods for estimation of the CTL. We consider each method, which also includes a way to collect and organize the data, to constitute a method for measurement of the CTL. The measurand (ISO/IEC, 2008) we focus on is the CTL. The question is then: Which measurement method is best? In general we will prefer a method with small measurement uncertainty,

but other considerations such as simplicity and ease of interpretation are also important.

Section II presents a formalization of the visual method, and introduces the anchor method. The anchor method is simple, but is a main result in this paper. Both methods are illustrated by use on data found in the literature. The anchor method is derived in Sec. III based on a model motivated by the Stevens's (1957) power law on the apparent loudness of simple sound. Section III also explains that the CTL can be seen as a natural decibel valued quantity that replaces usage of both the community specific threshold A introduced by Fidell *et al.* (1988) and the alternative decibel valued quantity $D^* = (10/\beta)\ \mathrm{lg}(A)$ introduced by Green and Fidell (1991). Section IV generalizes and summarizes the methods presented in the foregoing, and presents several additional methods. Methods for calculating the uncertainty of the estimates are also indicated. Section V presents methods for calculation of summarized data from individual data and compares the resulting estimation methods by use on two recent Norwegian surveys, and by simulation. Section VI presents a discussion of the methods as applied to the examples and Sec. VII concludes.

II. THE VISUAL AND THE ANCHOR METHODS APPLIED ON HISTORIC DATA

Assume that L_{dn} is the day–night noise level and that p is the resulting probability of being highly annoyed. We define here the **visual method** by considering only the observations with $40\% < p < 60\%$. Each of the resulting L_{dn} observations can be considered to be an estimate of L_{ct}. This procedure gives a reduction to 99 observations when applied to the *Fidell dataset*, and the result is then the estimate $L_{ct} = 67.3(9)$ dB. The corresponding 95% confidence interval is (65.5, 69.1) dB. The estimate here is comparable with the purely visual estimate stated earlier, and is significantly different from the original estimate 73.3(1.1) dB.

A strength of the visual method is that it does not depend on any assumptions regarding the shape of the dose-response curve. A drawback is that it relies on the existence of a sufficient number of observations with $40\% < p < 60\%$. This is fulfilled for the complete *Fidell dataset* and for the data considered by Schultz, but for many surveys this requirement is not fulfilled. The two Norwegian surveys considered in Sec. V exemplify this since the majority of data points correspond to low noise levels. For these cases the visual method is impossible to use. A solution is given by establishing a dose-response curve based on more fundamental modeling assumptions. This approach was chosen by Fidell *et al.* (2011), and the method presented next relies on their assumptions.

It will be shown in Sec. III that the pair (L_{dn}, p) determines the corresponding CTL to be

$$L_{ct} = L_{dn} + \frac{10}{\beta}\,\mathrm{lg}\!\left(\frac{\mathrm{lg}(p)}{\mathrm{lg}(p_{ct})}\right), \tag{1}$$

where $\beta = 0.3$ is a model parameter related to the apparent loudness of simple sound exposure, and $p_{ct} = 50\%$ is as used

J. Acoust. Soc. Am. **140** (1), July 2016

Taraldsen *et al.* 693

in the definition of the CTL. Definitions and further explanation of the significance of these quantities are given in Sec. III. This includes in particular, the assumptions necessary for arriving at Eq. (1). The simple intuitive idea is that knowledge of one point (L_{dn}, p) on the dose-response curve is sufficient to anchor the curve. When the curve is known, the corresponding L_{ct} is also known and given by Eq. (1).

The CTL was introduced by Fidell *et al.* (2011). The observations in the *Fidell dataset* include uncertainties, as do most experimental results, and the observed pairs are hence only estimates of the true values. Equation (1) can be used to calculate the corresponding estimated CTL values. These estimates will be referred to here as the anchor estimates of the CTL. The **anchor method** referred to in the Abstract and in the title of this section is given by taking the arithmetic average of these anchor estimates as the estimate for the CTL. A more general and precise definition including weights is presented in Sec. IV. Section V shows how the anchor method can be applied on data collected according to the standardized ICBEN instrument (ICBEN, 2001), and this is the procedure we recommend.

Point number 15 in the SWI-534 survey in the *Fidell dataset* is an exception. The reported $p = 0$ for the corresponding $L_{dn} = 30.3$ dB. Equation (1) cannot be used for cases where $p = 0$ or $p = 1$, and the anchor estimates are hence only defined for $0 < p < 1$.

Figure 2 shows a histogram of the resulting anchor CTL values for the *Fidell dataset*. The mean 71.2(4) dB is an estimate of the CTL for the large community corresponding to all of the 43 aircraft noise surveys. This estimate differs significantly from the estimate 67.3(9) dB found from the visual method on p. 9, but differs also from the original estimate 73.3(1.1) dB in Fig. 1 (Fidell *et al.*, 2011).

It is noteworthy that the standard uncertainty 0.4 dB can be calculated by the usual procedure from the standard deviation and the number of observations: 9.3 dB$/\sqrt{544} = 0.4$ dB. This is a valid procedure since all the anchor estimates are independent. It is remarkable that the result is approximately

one-half of the standard uncertainty 1.1 dB found for the least-squares method by the original authors. This one-half result for the standard uncertainties coincides roughly with the finding for completely different simulated data in Sec. V. The least-squares method has a larger uncertainty than the anchor estimate method despite being more complicated to calculate.

The large spread in the CTL values in the histogram in Fig. 2 can be seen as a measure of the large spread between the dose-response curves found in the literature. Different communities have different CTLs.

The spread in the CTL values is, however, also due to the spread between the estimates for interview sites within each aircraft noise survey. Figure 3 shows a histogram of the resulting anchor CTL values for the Australian A/C noise survey from 1980. This noise survey is one of the 43 aircraft noise surveys found in Table II presented by Fidell *et al.* (2011). This subset of the *Fidell dataset* will be referred to here as the *AUL-210 dataset* in accordance with Fidell *et al.* (2011). The mean 78.6(8) dB is an estimate of the CTL for this community. This estimate is comparable to the estimate 79.0 dB found by the least-squares method used by Fidell *et al.* (2011).

One advantage of the method based on Eq. (1) compared to the least-squares method is its simplicity. It also directly provides the standard uncertainty 0.8 dB with a corresponding 95% confidence interval [77 dB, 80.2 dB]. The validity of this follows since the resulting anchor estimate is approximately normally distributed from Eq. (1). The approximate normality is also demonstrated by simulations in Sec. V. This additional information on the estimation uncertainty is not so easily available when using the least-squares method. The standard uncertainty of the estimate 79.0 dB found by Fidell *et al.* (2011) is not stated by the authors, and they do not give a method for calculation of the uncertainty.

The method just presented based on Eq. (1), including the possibility of plotting histograms and computing a
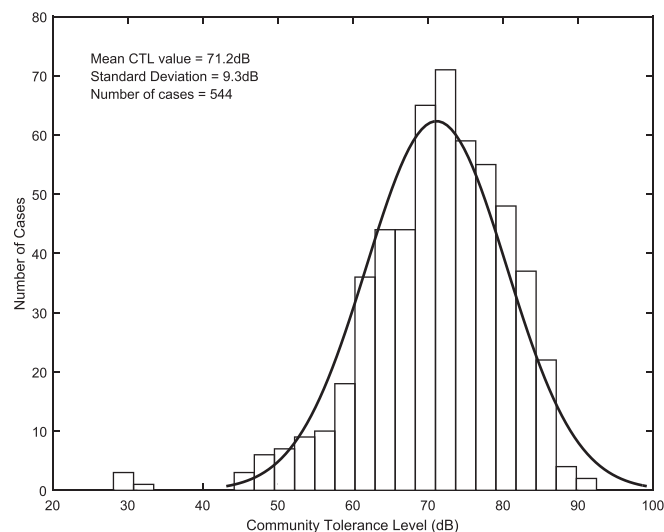

FIG. 2. Distribution of anchor CTL values from the *Fidell dataset* compared to Gaussian distribution of similar mean and variance.
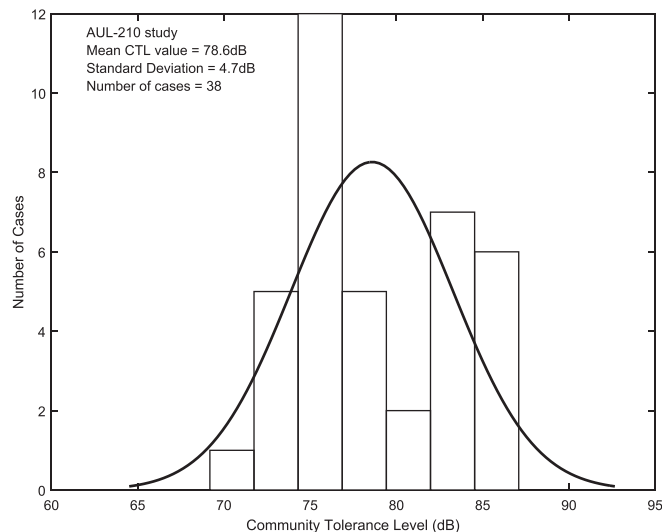

FIG. 3. Distribution of anchor CTL values from the *AUL-210 dataset* compared to Gaussian distribution of similar mean and variance.

Taraldsen *et al.*

standard uncertainty can be seen as the main result in this paper. This proposed simple method consists of (a) calculate the CTL for individual data points, (b) calculate the empirical mean and standard deviation of the CTLs, and (c) calculate the standard uncertainty for the mean CTL.

The purpose of the remainder of this paper is to further investigate this method and some alternatives. The somewhat surprising conclusion will be that this elementary method based on the anchor CTL estimates—after some refinement—is also the method with the smallest uncertainty in terms of precision and trueness.

III. THE CTL AND A COMMUNITY-SPECIFIC THRESHOLD A

The basic assumption of Fidell et al. (2011) is that an effective noise dose m defined by

$$m = 10^{\beta L_{dn}/10}, \tag{2}$$

gives the probability of being highly annoyed as

$$p = P(X_a > A) = \exp(-A/m), \tag{3}$$

where the random variable X_a is the annoyance of a random individual (Fidell et al., 1988, p. 2111), A is a community-specific threshold that defines the highly annoyed state, and $\beta = 0.3$. Originally, A was assumed to be a non-acoustic factor, but this is not assumed here. It may additionally depend on the source type, or other characteristics of the sound.

The choice of the exponential function in Eq. (3) is motivated by simplicity (Fidell et al., 2011, p. 793). It is, however, equivalent to assuming that the annoyance X_a is a random variable with an exponential probability distribution with mean m. Stevens's (1957) law states that the apparent loudness of simple sound exposure is proportional to the 0.3 root of acoustic energy, and Fidell et al. (1988, p. 2111) use this to conclude that $\beta = 0.3$ is a reasonable choice in Eq. (2).

The model given by Eqs. (2) and (3) contains only one unknown parameter given by the community-specific constant A. Different values of A give different dose-response curves as illustrated by Fidell et al. (1988, Fig. 2, p. 2112), and A can be estimated by minimizing the root-mean-square (rms) distance between the observations and the dose-response curve. Fidell et al. (1988, p. 2112) finds in particular, that the rms-distance between the curve given by $A = \exp(\ln(10)22.5/10) \approx 177.83$ and the Schultz (1978) curve is less than 0.035 for $40 < L_{dn} < 85$.

Fidell et al. (2011, p. 793) define A as above, but introduce additionally the CTL as an alternative and more convenient parameter. They define the L_{ct} from the midpoint of the dose-response curve, or more specifically by $\exp(-A/m_{ct}) = p_{ct} = 50\%$, where $10 \lg(m_{ct}) = \beta L_{ct}$ defines m_{ct} in accordance with Eq. (2). This can be solved for L_{ct}, and gives

$$L_{ct} = \frac{10}{\beta} \lg\left(\frac{A}{-\ln(p_{ct})}\right), \tag{4}$$

which is a one-one correspondence between A and L_{ct}. If a pair (L_{dn}, p) is known, then Eq. (3) gives $A = -m \ln(p)$ with m determined by Eq. (2). Insertion into Eq. (4) gives Eq. (1) as stated in Sec. II.

The above fit to the Schultz curve with $A = 177.83$ gives $L_{ct} \approx 5.31 + 33.33 \times 22.5/10 \approx 80.30$ dB which is close to the result $L_{ct} \approx 79.2$ dB found for the Schultz curve in Fig. 1. Green and Fidell (1991, p. 236) use the same modeling strategy, but introduce

$$D^* = \frac{10}{\beta} \lg(A) \tag{5}$$

as a convenient decibel valued quantity. Based on the Schultz data and an alternative estimation procedure they find the estimate $D^* = 73$ dB which gives $L_{ct} \approx 78.31$ dB.

It follows from the above that

$$L_{ct} = D^* - (10/\beta)\lg(-\ln(p_{ct})) \approx D^* + 5.31. \tag{6}$$

At first sight it could seem that the difference between L_{ct} and D^* is trivial, but this is not so. The quantity D^* is defined and restricted to the case where the dose-response curve is of a particular form. The CTL on the other hand can be defined more generally. It is well defined regardless of the particular form of the dose-response curve, and has a direct interpretation in terms of the observations.

With this definition it was shown above that the curve obtained by Schultz (1978) gives $L_{ct} = 79.2$ dB. As also explained above, Fidell et al. (1988) and Green and Fidell (1991) obtain estimates for the dose-response curve that corresponds to, respectively, $L_{ct} = 80.3$ dB and $L_{ct} = 78.3$ dB for the same data. Different estimation procedures give different results. Fidell et al. (1988) use a least-squares method, and this is also the method suggested by Fidell et al. (2011). Green and Fidell (1991) use a simpler method that does not require numerical optimization. Generalizations of both methods together with a novel third method are presented next.

IV. ESTIMATION OF CTL

It has been explained and illustrated that Eq. (1) can be used to estimate L_{ct}. Assume more generally that the dose-response curve is on the form

$$p = \psi(L_{dn} - L_{ct}), \tag{7}$$

where $\psi(0) = p_{ct}$. This is completely general. It is, furthermore, reasonable to assume that p increases with increasing noise L_{dn}. Assuming monotonicity gives

$$L_{ct} = L_{dn} - \psi^{-1}(p), \tag{8}$$

which is a generalization of Eq. (1). All dose-response curves given in Sec. I are of this form, and explicit alternative formulas corresponding to Eq. (1) can be derived.

Except for the Schultz polynomial, all curves can be expressed so that ψ is the cumulative distribution function of a random variable, and ψ^{-1} is the inverse cumulative

J. Acoust. Soc. Am. **140** (1), July 2016

Taraldsen et al.   695

distribution function. This means also that each of these models can be seen as models based on an underlying random effective annoyance score. The difference in the models can be seen as arising from a different probability distribution for the effective annoyance score.

The assumption $\beta = 0.3$ can be seen as an assumption on the slope $\psi'(0)$, or equivalently as an assumption on the value of the probability density at 0. It is approximately satisfied by all curves in Fig. 1. It is noteworthy that the data therefore seem to support the fundamental modeling assumption $\beta = 0.3$ in this sense.

Assume now that (L_j, p_j) are estimates of (L_{dn}, p) for each j in a finite index set J. The *Fidell dataset* contained in Table II presented by Fidell *et al.* (2011) have data in this format, where the index set J correspond to the $|J| = 545$ interview sites. Each interview site j is typically defined by a geographical area close to an airport. The p is estimated by the highly annoyed proportion p_j of n_j interviewed people and L_j is chosen as an representative L_{dn} for the interview site. The choice of method for collection of data is an important part of the resulting measurement method. Other possibilities for collection of data are also possible as will be illustrated in Sec. V.

Define a normalized weight $w_j = n_j / \sum_j n_j$, and use this weight to define averages $\bar{L} = \sum_j \omega_j L_j$, and $\overline{\psi^{-1}(p)} = \sum_j \omega_j \psi^{-1}(p_j)$. The anchor CTL estimate $\bar{L}_{ct}$ is then defined as

$$\bar{L}_{ct} = \bar{L} - \overline{\psi^{-1}(p)}, \tag{9}$$

in accordance with Eq. (8).

Unfortunately, the number n_j of interviews on each site is not contained in Table II presented by Fidell *et al.* (2011). The anchor CTL estimates corresponding to Figs. 2 and 3 have hence been calculated as if n_j is constant for all interview sites.

Equation (9) is very convenient for calculation since it is decomposed into a term separately for the effect of the observed noise and the observed annoyance state. The effect of different assumptions for ψ is also easily analyzed. Consider for instance the typical case where most of the observations p_j are smaller than p_{ct}. A dose-response curve ψ with a smaller slope will then tend to give a higher estimate for the CTL when applied to a fixed set of data. This is exemplified by allowing β to be smaller than the assumed 0.3.

The standard uncertainty $u(\bar{L}_{ct})$ is given by the weighted unbiased

$$u(\bar{L}_{ct})^2 = \frac{1}{|J| - 1} \sum_j \left(L_{ctj} - \bar{L}_{ct}\right)^2 \omega_j, \tag{10}$$

estimator for the variance. It is truly unbiased if the weight w_j is proportional the inverse of the variance of $L_{ctj} = L_j - \psi^{-1}(p_j)$. This choice minimizes the variance of $\bar{L}_{ct}$, and is in this sense optimal. The assumption on the weights is approximately true for the given choice of weights, and explains this choice.

Let $\phi_j = \phi(L_j) = \psi(L_j - L_{ct})$. The least-squares estimate $\tilde{L}_{ct}$ is here defined as the minimizer of the least-squares functional

$$\tilde{\mathcal{F}} = \overline{(p - \phi)^2} = \sum_j \omega_j (p_j - \phi_j)^2. \tag{11}$$

Similarly, the maximum-likelihood estimate $\hat{L}_{ct}$ is here defined as the minimizer of the negative log-likelihood functional

$$\hat{\mathcal{F}} = -\overline{\ln(\phi^p (1 - \phi)^{1-p})} = -\sum_j \omega_j \ln(\phi_j^{p_j}(1 - \phi_j)^{1-p_j}). \tag{12}$$

Both the least-squares and the maximum-likelihood estimates can be found by numerical optimization methods. In both cases the anchor estimate $\bar{L}_{ct}$ provides a good initial value for the search for a minimum.

In the generality presented, the least-squares and maximum-likelihood estimators are novelties here. It is also possible to introduce the possibility of allowing ψ to be an additional unknown parameter in a class of allowable functions. Minimization of the log-likelihood defines then the maximum-likelihood estimate $(\hat{L}_{ct}, \hat{\psi})$, and similarly for the least-squares estimate. It is possible, based on the previous and work by Taraldsen and Lindqvist (2013), to develop Bayesian and fiducial procedures for improved estimation of the CTL, but this is left for the future.

V. TWO RECENT NORWEGIAN SURVEYS

Telephone interviews according to the standardized ICBEN instrument, ICBEN (2001), were conducted during Spring 2014 for persons living near the airports of the Norwegian cities Bodø (BOO) and Trondheim (TRD). The top three points on the 11-point ICBEN scale were used to define the highly annoyed state. The corresponding noise dose was obtained by estimation of the yearly L_{dn} for each persons' home address by use of a noise mapping software as explained together with a more detailed description of these investigations by Gelderblom *et al.* (2014).

The observations for each location are given by pairs $(L_1, p_1),...,(L_n, p_n)$, where the vectors L and p give, respectively, the dose and response for each of the n respondents. The response p_i is an indicator variable which takes the value 1 if individual i is highly annoyed, and the value 0 otherwise. The resulting *BOO14 dataset* and *TRD14 dataset* have a sample size equal to $n = 302$ and $n = 300$, respectively.

The observations here are on the form assumed in Sec. IV, but the p_i values take only the extreme values 0 and 1. The maximum-likelihood estimator $\hat{L}_{ct}$ is well defined directly in this case by Eq. (12). It is well known that it is asymptotically optimal, and is hence the canonical choice in many different applications.

The competitors given by the mean $\bar{L}_{ct}$ and the least-squares $\tilde{L}_{ct}$ estimators are undefined, or intuitively unreasonable, for the case where the p_i values take only the extreme values 0 and 1. One solution to this problem is to transform the individual data to summarized data. This is exemplified by the *Fidell dataset*, which is summarized data based on individual data. The summarization has been done based on grouping by geographically defined interview sites as explained in Sec. IV. Some alternative procedures for

producing summarized data from individual data are described next. They do not depend on geographical data. This can be seen both as an advantage due to simplicity, but also as a disadvantage in that the location information is lost.

A common solution is to calculate a triple $(\bar{L}_j, \bar{p}_j, n_j)$ for each 5 dB interval centered at …, 45 dB, 50 dB, 55 dB,… where the number of observations n_j is larger than zero. The resulting index set J is here distinct and smaller than the corresponding index set I corresponding to the individual persons. Each $\bar{p}_j$ is calculated as the percentage being highly annoyed among the n_j observations in each interval $j \in J$. It equals also the arithmetic average of the observed p_i values in the interval as the notation indicates.

Gelderblom *et al.* (2014) defined $\bar{L}_j$ to be equal to the center of the interval, and this will be denoted here as **procedure 0.** The resulting least-squares estimate is denoted here as $\tilde{L}_{ct}$, but the notation $\tilde{L}_{ct0}$ can also be used. For the survey in TRD this results in the 7 datapoints (40 dB, 0.00%), (45 dB, 0.00%), (50 dB, 0.94%), (55 dB, 2.88%), (60 dB, 3.85%), (65 dB, 8.70%), (70 dB, 0%) shown as circles in Fig. 4. Similarly the survey in BOO gives the 7 datapoints (45 dB, 6.25%), (50 dB, 4.92%), (55 dB, 8.89%), (60 dB, 6.60%), (65 dB, 6.38%), (70 dB, 20.00%), (75 dB, 100.00%) shown as circles in Fig. 5.

An alternative is given by defining $\bar{L}_j$ to be equal to the arithmetic average of the observed levels in interval j. An advantage of this choice is that the resulting overall average $\bar{L}$ as needed in Eq. (9) equals the arithmetic average of all observed levels. The resulting transformation of the original (L_i, p_i) into the resulting $(\bar{L}_j, \bar{p}_j, n_j)$ will be denoted here as **procedure 1.** The resulting anchor estimate $\bar{L}_{ct1}$ will be referred to here as the quick anchor estimate. For the survey in TRD this results in the 7 datapoints (39.02 dB, 0.00%), (45.50 dB, 0%), (51.02 dB, 0.94%), (54.92 dB, 2.88%), (59.03 dB, 3.85%), (65.07 dB, 8.70%), (67.56 dB, 0%) which gives the quick anchor estimate in Fig. 4. Similarly the survey in BOO gives the 7 datapoints (45.67 dB, 6.25%), (50.24 dB, 4.92%), (54.69 dB, 8.89%), (59.89 dB, 6.60%), (64.70 dB, 6.38%), (68.78 dB, 20.00%), (75.80 dB, 100.00%) which gives the quick anchor estimate in Fig. 5.

A third approach is to keep the original L_i values, and define p_i equal to p_j obtained above, where j is the interval that contains L_i. An advantage of procedure 2 is that it is not necessary to calculate a weight w_j from n_j, and the two procedures gives identical results for the $\bar{L}_{ct}$.

A fourth, and final, approach is to replace the previously calculated $\bar{p}_j$ values by the values $\bar{p}_i$ obtained from an interval centered at each L_i. The resulting transformation of the original (L_i, p_i) into the resulting $(L_i, \bar{p}_i)$ will be denoted here as **procedure 3.** This defines the moving average anchor estimate $\bar{L}_{ct3}$. It is the mean of the resulting individual anchor estimates L_{cti}. It will be denoted here simply as the anchor estimate $\bar{L}_{ct}$ since it will be our recommended estimate. Procedure 3 can be seen as a natural refinement of procedure 2, but it requires some more work for calculating the moving average. For the survey in BOO this results in the datapoints shown as filled circles in Fig. 5.

Figures 4 and 5 show the estimated dose-response curves from the quick anchor estimate $\bar{L}_{ct1}$ from procedure 1, the anchor estimate $\bar{L}_{ct}$ from procedure 3, the maximum-likelihood estimate $\hat{L}_{ct}$ from individual data, and the least-squares estimate $\tilde{L}_{ct}$ from procedure 0 for the surveys at TRD and BOO, respectively. The circles represent data points from procedure 0 and the filled circles represent the moving average data points from procedure 3. The least-squares estimate deviates from the other three. Both anchor estimates are surprisingly close to the maximum-likelihood estimate. The standard error is calculated only for the quick anchor method since an explicit formula is available only for this method in these cases.

The previous procedure, and all the others, have corresponding analogues where the interval grouping is replaced by some choice of geographical interview site grouping. This will not be investigated here, but it is mentioned since the geographical location can be a factor of some interest.
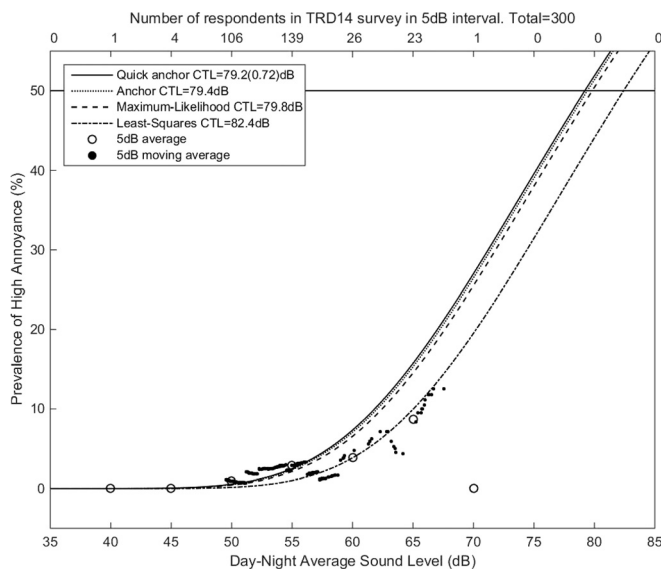

FIG. 4. Four estimated dose-response curves and the corresponding CTLs. The data points are from the *TRD14 dataset*.
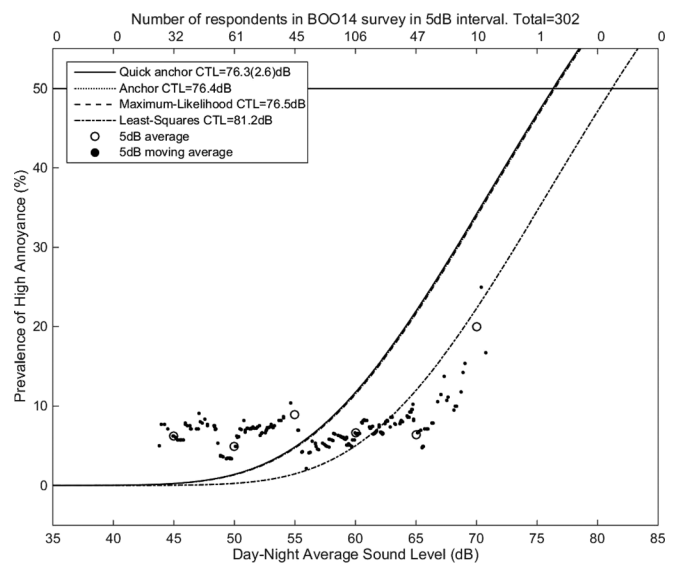

FIG. 5. Four estimated dose-response curves and the corresponding CTLs. The data points are from the *BOO14 dataset*.

There are also a variety of other possibilities given by for instance a moving smooth window replacing the rectangular interval window. We leave investigation of this and other alternatives for future work.

Figures 6 and 7 show a histogram of the resulting individual anchor CTL estimates obtained for the surveys at TRD and BOO, respectively, by procedure 3. The sample size in both histograms is smaller than the original sample size due to the requirement $0 < p_i < 1$.

Both histograms seem to be localized at two instead of one value, and this is also seen for the *AUL-210 dataset* as shown in Fig. 3. Inspection of the moving average points in Figs. 4 and 5 shows a jump at approximately 55 dB. The moving average points below 55 dB correspond to a lower estimate for the CTL than the moving average points above approximately 55 dB. This explains the two centers seen in the histograms. The explanation seems then not to be due to two different communities as could be speculated based only on the histograms. We find this interesting, but will not discuss this further here.

It is not possible to decide which estimate of the CTL is best in each specific case. The true CTL value remains unknown. It is, however, possible to compare the performance of the three methods by use on simulated survey data. In this way we can decide on which method is the best.

For this purpose the 300 observed L_{dn} values in the *TRD14 dataset* will be used. Computer simulated individual responses p_i are obtained by assuming that $p_i = 1$ with probability $\phi(L_i)$. The dose-response curve ϕ from Fidell *et al.* (2011) with the choice $L_{ct} = 80$ dB has been used for generation of the data presented next. We simulate 100 000 datasets for the given fixed parameter $L_{ct} = 80$ dB and the given fixed 300 observed L_{dn} values. Each method results then in 100 000 simulated measurement results in a case where the true value $L_{ct} = 80$ dB is known, and a comparison of the performance can be done.

A more complete simulation study would consider simulations for different parameter values, but this is not done here. The case given by the *TRD14 dataset* is similar to the
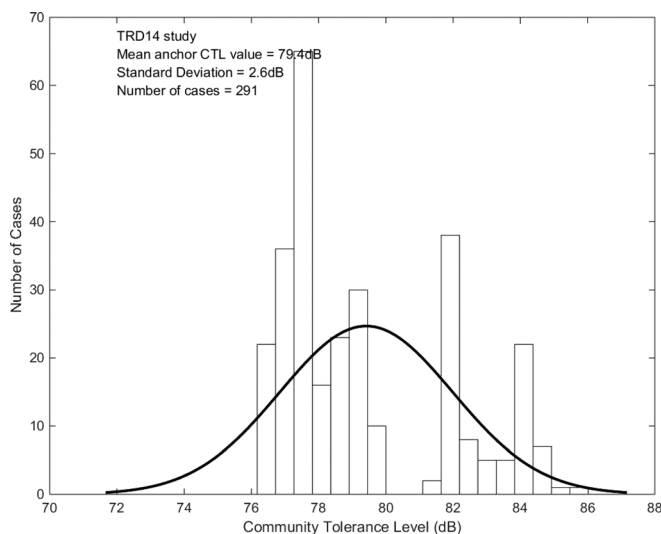
case given by the *BOO14 dataset* and represents an important and difficult case where large differences between the methods can be expected. Geometrically this can be explained by the small slope of the dose-response curve for low doses. A particular result of the simulation study is a Monte Carlo estimate of the uncertainty of each method for the *TRD14 dataset*.

Figures 8–11 give the resulting distribution of the 4 estimates for 100 000 simulated survey results. The least-squares estimator gives systematically on average a 0.8 dB too large estimate. The corresponding 0.2, −0.2, and −0.5 dB bias of the maximum-likelihood and the anchor estimators, respectively, are smaller. The least-squares estimator is hence inferior in terms of *trueness*.

The standard uncertainty of the least-squares estimator is 3.1 dB which is more than twice the standard uncertainty 1.4 dB of the anchor estimator. Somewhat surprisingly, the

FIG. 7. Distribution of anchor CTL values from the *BOO14 dataset* compared to Gaussian distribution of similar mean and variance.

FIG. 6. Distribution of anchor CTL values from the *TRD14 dataset* compared to Gaussian distribution of similar mean and variance.

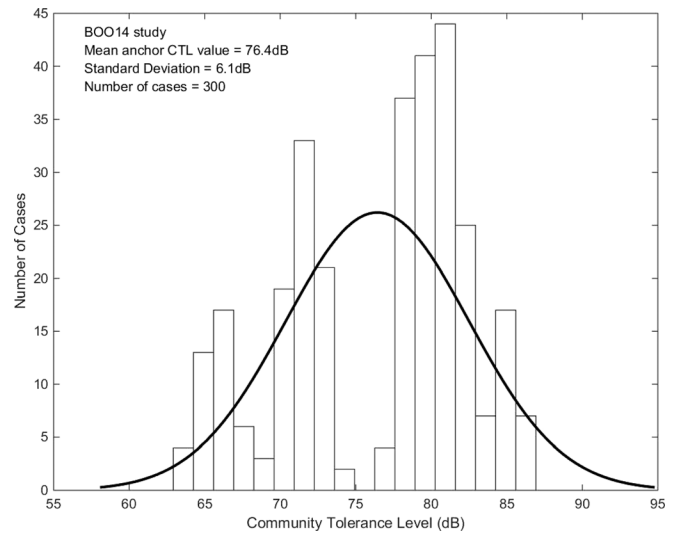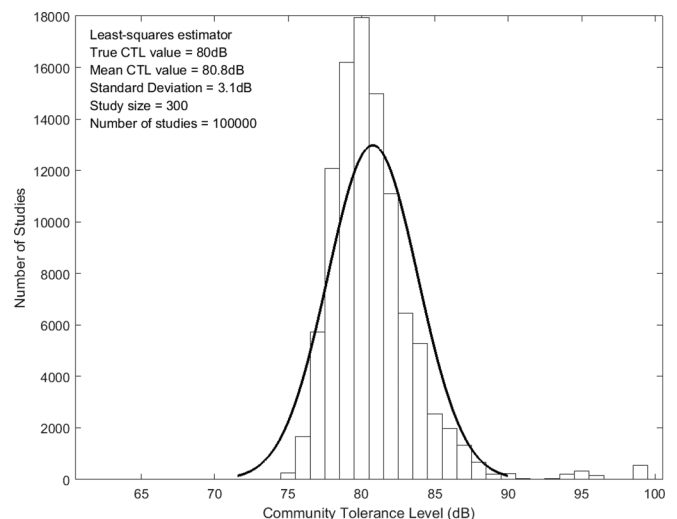FIG. 8. Distribution of least-squares estimated CTL values for simulated survey results compared to Gaussian distribution of similar mean and variance.
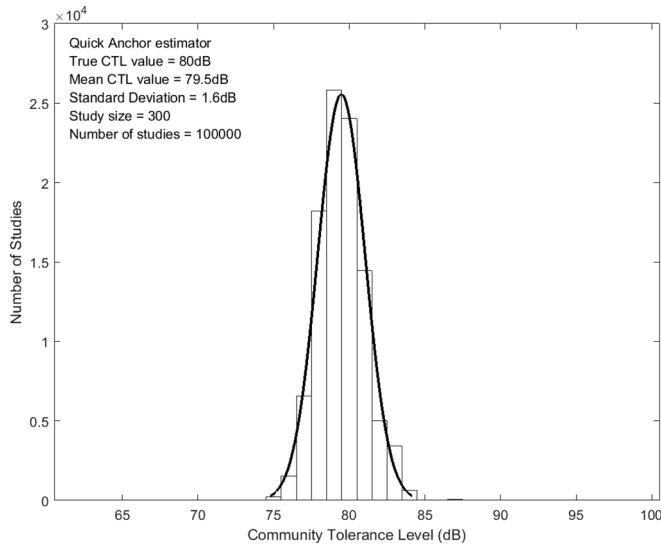
Taraldsen *et al.*

FIG. 9. Distribution of interval anchor estimated CTL values for simulated survey results compared to Gaussian distribution of similar mean and variance.

FIG. 11. Distribution of individual anchor estimated CTL values for simulated survey results compared to Gaussian distribution of similar mean and variance.

standard uncertainty 1.7 dB of the maximum-likelihood estimator is also larger than the uncertainty of both anchor estimators. Both anchor estimators are hence preferable also compared to the maximum-likelihood in terms of *precision*.

VI. DISCUSSION

Which CTL measurement method is best? A measurement consists of data collection for noise annoyance, data collection for noise levels, and processing the data to obtain an estimate. The latter is the main theme in this article and is discussed further below. For noise annoyance we recommend to use the standard ICBEN (2001) instrument. What instrument should be used for measurement of the noise levels?
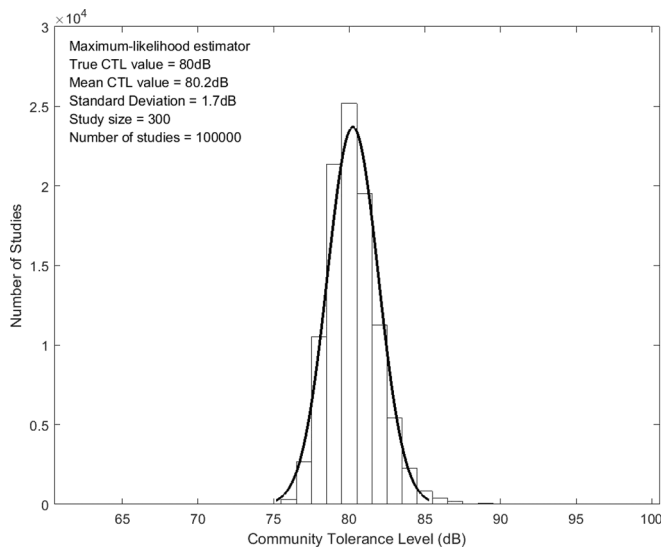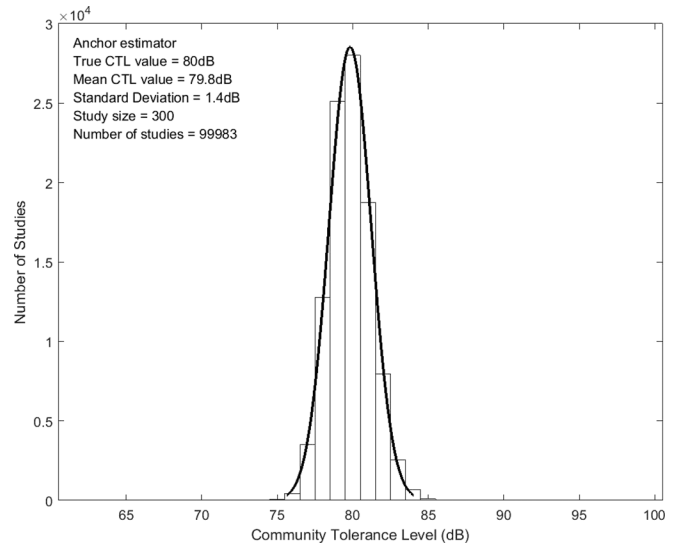

FIG. 10. Distribution of maximum-likelihood estimated CTL values for simulated survey results compared to Gaussian distribution of similar mean and variance.

The noise levels could in principle be measured by noise dose meters, but this is impracticable or impossible since the noise level is a weighted average for a complete and representative year. We recommend using a standardized noise mapping method for measurement of the noise levels. The instrument for the noise levels is then given by noise mapping software and their input in the form of source characterization and other data necessary for sound propagation calculations. Unfortunately, the choice of an instrument here is more difficult as all existing standardized methods are based on very simplified sound propagation models and source characterizations. A positive development (Jonsson *et al.*, 2008) is given by the methods developed in the Nord2000 project (Kragh *et al.*, 2002) and the HARMONOISE project (Nota *et al.*, 2004; van Maercke *et al.*, 2004). The uncertainty of the resulting noise level measurement instrument can and should be reduced by further development of standardized noise mapping methods. Today, the noise level instruments in use give a dominating and unnecessary contribution to the overall uncertainty of the measurement of the CTL.

It has been demonstrated on existing and simulated data in Secs. II, III, and V that different procedures for the estimation of the CTL give different results, and sometimes substantially different results. This is summarized and discussed next.

Section III shows that methods developed by, respectively, Schultz (1978), Fidell *et al.* (1988), and Green and Fidell (1991), give the CTL estimates 79.2, 80.3, and 78.3 dB based on identical data. The method presented by Fidell *et al.* (1988) is the least-squares method which we do not recommend to use, and we take this as an argument against the estimate 80.3 dB. The method presented by Green and Fidell (1991) is the method with closest resemblance with the anchor method, and we take this as an argument in favor of the estimate 78.3 dB. This moves the Schultz curve closer to the other dose-response curves presented in Fig. 1.

Section II shows that the visual method applied on the *Fidell dataset* gives $L_{ct} = 67.3(9)$ dB. Assuming a loudness-based dose-response curve as assumed by Fidell *et al.* (2011) is however recommended in favor of the visual method to reduce the large uncertainty in the estimate. The anchor method with this assumption applied on the *Fidell dataset* gives the estimate $L_{ct} = 71.2(4)$ dB as shown in Fig. 2. The measurement uncertainty of the anchor method is smaller than the measurement uncertainty of the original least-squares method and we take this as an argument in favor of the estimate $L_{ct} = 71.2(4)$ dB when compared with the original (Fidell *et al.*, 2011) estimate 73.3(1.1) dB in Fig. 1.

Section II also applies the anchor method on the data from the Australian A/C noise survey from 1980. The resulting CTL is 78.6(8) dB for this community. This estimate is comparable to the estimate 79.0 dB found by the least-squares method used by Fidell *et al.* (2011). Application of the anchor method is recommended in favor of the least-squares method for this and similar datasets since this method also gives an estimate of the measurement uncertainty and the possibility of visualization as in Fig. 3.

Section V considers two recent Norwegian surveys characterized by data dominated by low noise levels and corresponding resulting large uncertainty in the measurement of the CTL. It is in particular, for such cases that the choice of measurement method is important. The least-squares method applied initially led Gelderblom *et al.* (2014, p. 8) to conclude that the respondents tolerated approximately 8 to 10 dB more noise, before being equally annoyed as predicted by the Miedema curve. This dramatic and somewhat surprising solution gives motivation for exploring the uncertainty of the least-squares method and alternative methods as done in this paper.

The simulations show, however, that the standard uncertainty of the least-squares method is approximately 3.0 dB for these two surveys, and it is reduced to approximately 1.5 dB by using any of the anchor methods or the maximum-likelihood method. This can be used to conclude that the anchor results 76.4(1.5) dB and 79.4(1.5) dB should be used instead of the least-squares results 81.2(3.0) dB and 82.4(3.0) dB. The resulting dose-response curves for the Norwegian respondents are then also more in harmony with the dose-response curves presented in Fig. 1.

Many possible measurement procedures have been indicated, but the comparisons here have been restricted to the original least-squares method, the maximum-likelihood method, the quick anchor method, and the anchor method. The least-squares method has inferior performance, and the maximum-likelihood and the anchor methods have a similar performance. The anchor method is the recommended choice based on the following list of advantages:

(1) It is simple to implement and requires no optimization routines.
(2) It is the method with the smallest uncertainty in terms of both precision and trueness.
(3) It gives the possibility of visualizing the individual estimates in histograms and dose-response diagrams.
(4) Explicit formulas for the standard uncertainty and the corresponding confidence interval are available.

In conclusion, the least-squares estimator with its large variance and skewed distribution as shown in Fig. 8 cannot be recommended. The maximum-likelihood and the anchor estimators have comparable performance, but the anchor estimators can be calculated by explicit formulas without any optimization routine. The quick anchor estimator has surprisingly good performance, and has the advantage that Eq. (10) can be used to estimate the measurement uncertainty.

An improved estimate and characterization of the measurement uncertainty can be obtained by Monte Carlo simulation. This has been exemplified here for the TRD survey, and the results show that the least-squares method should not be used due to its large uncertainty. The anchor method even outperforms the maximum-likelihood method and should be used.

Additionally, there is uncertainty contributions (ISO/IEC, 2008) due to the procedures used for data collection. For the TRD example we judge that the respondents have been selected so that they are a representative sample for the community and the uncertainty from this is ignored. Based on experience with noise mapping calculations we subjectively estimate the uncertainty contribution from the noise level instrument to be 2.5 dB. The total measurement uncertainty for the TRD example is then $\sqrt{2.5^2 + 0.2^2 + 1.4^2}$ dB $= 2.9$ dB, and the measurement result is $L_{ct} = 79.4(2.9)$ dB. Given the overall uncertainty it is reasonable to state the result as $L_{ct} = 79(3)$ dB.

VII. CONCLUSION

How should the CTL for noise be measured?

The **first step** is to collect data by selecting respondents that are representative for the community of interest. It is strongly recommended to measure the degree of annoyance for each respondent according to the standardized (ICBEN, 2001) instrument. The noise level for each respondent should also be measured by a standardized instrument, but unfortunately the choice here is more difficult as discussed in Sec. VI.

The **second step** is to estimate the CTL based on the collected data. We recommend using the anchor method due to its simplicity, ease of interpretation, and small uncertainty as compared with the maximum-likelihood and least-squares methods. The maximum-likelihood method is a good alternative with comparable uncertainty.

The **third step** is to estimate the measurement uncertainty (ISO/IEC, 2008). The most refined approach for estimating this is by Monte Carlo simulation as exemplified for the measurements done in TRD, but the methods given in Sec. IV can be used as alternatives. Additionally, there is uncertainty due to the instruments used for data collection. The most important contribution to the uncertainty is given by the noise level instrument when the anchor method is used. If the least-squares method is used, then this will contribute with an uncertainty larger or comparable to the contribution from the noise level instrument.

ACKNOWLEDGMENTS

700   J. Acoust. Soc. Am. **140** (1), July 2016

Taraldsen *et al.*

Federal Interagency Committee on Noise (FICON). (**1992**). "Federal agency review of selected airport noise analysis issues," Technical report, U.S. Department of Defence, Washington, DC.

Fidell, S., Mestre, V., Schomer, P., Berry, B., Gjestland, T., Vallet, M., and Reid, T. (**2011**). "A first-principles model for estimating the prevalence of annoyance with aircraft noise exposure," J. Acoust. Soc. Am. **130**(2), 791–806.

Fidell, S., Schultz, T., and Green, D. M. (**1988**). "A theoretical interpretation of the prevalence rate of noise-induced annoyance in residential populations," J. Acoust. Soc. Am. **84**(6), 2109–2113.

Gelderblom, F. B., Gjestland, T. T., Granøien, I. L. N., and Taraldsen, G. (**2014**). "The impact of civil versus military aircraft noise on noise annoyance," in *Internoise. The 43rd International Congress and Exposition on Noise Control Engineering*, Melbourne, Australia (November 16–19).

Green, D. M., and Fidell, S. (**1991**). "Variability in the criterion for reporting annoyance in community noise surveys," J. Acoust. Soc. Am. **89**(1), 234–243.

ICBEN (International Commission on the Biological Effects of Noise). (**2001**). "Standardized general-purpose noise reaction questions for community noise surveys: Research and a recommendation," J. Sound Vib. **242**(4), 641–679.

ISO/IEC (**2008**). Guide 98-3: "Uncertainty of measurement—Part 3: Guide to the expression of uncertainty in measurement," Technical report, Geneva, Switzerland, pp. 1–120.

Jonsson, G. B., and Jacobsen, F. (**2008**). "A comparison of two engineering models for outdoor sound propagation: Harmonoise and Nord2000," Acta Acust. Acust. **94**(2), 282–289.

Kragh, J., Plovsing, B., Storeheier, S. Å., Taraldsen, G., and Jonasson, H. G. (**2002**). Nordic Environmental Noise Prediction Methods, Nord2000 Summary Report, General Nordic Sound Propagation Model and Applications in Source-Related Prediction Methods, DELTA Acoustics & Vibration Report 1719/01.

Miedema, H., and Oudshoorn, C. (**2001**). "Annoyance from transportation noise: Relationships with exposure metrics DNL and DENL and their confidence intervals," Environ. Health Perspect. **109**(4), 409–416.

Nota, R., van Leeuwen, J. J. A., Barelds, R., Beuving, M., Cremezi, C., Defrance, J., Jonasson, H. G., van Maercke, D., Talotte, C., Taraldsen, G., Watts, G., and Witte, J. (**2004**). "Harmonoise: Simplification of comprehensive source and propagation models into an accurate and practicable engineering method," in *Internoise. The 33rd International Congress and Exposition on Noise Control Engineering*.

Schomer, P., Mestre, V., Fidell, S., Berry, B., Gjestland, T., Vallet, M., and Reid, T. (**2012**). "Role of community tolerance level (CTL) in predicting the prevalence of the annoyance of road and rail noise," J. Acoust. Soc. Am. **131**(4), 2772–2786.

Schultz, T. J. (**1978**). "Synthesis of social surveys on noise annoyance," J. Acoust. Soc. Am. **64**(2), 377–405.

Stevens, S. S. (**1957**). "On the psychophysical law," Psychol. Rev. **64**(3), 153–181.

Taraldsen, G., and Lindqvist, B. H. (**2013**). "Fiducial theory and optimal inference," Annals Stat. **41**(1), 323–341.

van Maercke, D., Nota, R., van Leeuwen, J. J. A., Barelds, R., Beuving, M., Cremezi, C., Defrance, J., Jonasson, H. G., Talotte, C., Taraldsen, G., Watts, G., and Witte, J. (**2004**). "Fast and accurate prediction of outdoor noise propagation in the Harmonoise engineering model," in *Internoise. The 33rd International Congress and Exposition on Noise Control Engineering*.

J. Acoust. Soc. Am. **140** (1), July 2016

Taraldsen *et al.*   701