Received: 25 January 2024 | Revised: 12 March 2024 | Accepted: 25 March 2024

DOI: 10.1002/prep.202400022

Propellants, Explosives, Pyrotechnics

RESEARCH ARTICLE

Estimating sensitivity with the Bruceton method: Setting the record straight

Dennis Christensen[1, 2] | **Geir Petter Novik**[2, 3] | **Erik Unneberg**[2]

[1]Department of MathematicsUniversity of OsloOslo Norway

[2]Norwegian Defence Research Establishment (FFI)Kjeller Norway

[3]Faculty of Science and TechnologyUniversity of StavangerStavanger Norway

Correspondence
Dennis Christensen, Department of Mathematics University of Oslo Oslo Norway.
Email: dennis.christensen@ffi.no

Funding information
Norwegian Defence Research Establishment

Abstract

Accurate estimates of sensitivities of energetic materials are crucial for ensuring safe production, transport, usage and destruction of explosives. When estimating sensitivities, researchers most commonly follow the NATO standard guidelines (STANAGs), in which the Bruceton method is imposed. Introduced in 1948, this method contains (i) an experimental design for choosing which stimulus levels to measure at and (ii) a recipe for computing sensitivity estimates. Although the former experimental design is supported by both theory and simulations, few modern researchers are aware that the latter recipe was only intended as a pen-and-paper approximation of the maximum likelihood estimates, which are easy to compute today. The persistent use of this outdated approximation has led to many unfortunate misconceptions amongst users of the Bruceton method, including the rejection of many perfectly valid data sets and neglect of uncertainty assessments via confidence intervals. This is both dangerous and unnecessarily wasteful. This paper sets the record straight and explains how researchers should estimate sensitivity via maximum likelihood estimation and how to construct confidence intervals. It also shows explicitly how wasteful said approximation is via both simulations and with real data.

KEYWORDS
confidence intervals, maximum likelihood estimation, the Bruceton method

1 | INTRODUCTION

It is crucial to understand the risk associated with explosives being exposed to external stimuli such as mechanical stress, shock, friction, physical impact, electrostatic discharge or heat. The susceptibility of an explosive substance to react to these stimuli, resulting in either combustion or detonation, is referred to as its sensitivity [1]. This particular property might be a key factor in determining the practical applicability of a given explosive, and a proper understanding of the sensitivity thresholds is therefore crucial for ensuring safe handling, storage, transport and destruction of explosives, both in military and civil sector. This paper will focus primarily on impact sensitivity to make the examples concrete, but the reader should be aware that the analysis generalises to all above cases.

In addition to the usage and manufacturing of new explosives, sensitivity estimates are particularly important in the study of explosive remnants of war (ERW) and dumped ammunition, of which there exist millions of tonnes worldwide [2]. It has been recognised that

ERW represent an explosive threat if disturbed, and more recent research also proves that some ageing explosives in munitions become increasingly sensitive to external stress, and are even susceptible to detonate spontaneously [3,4]. Societal concerns about potential environmental implications of explosive contamination have also intensified recent years [5]. An increasing number of people have been affected by the humanitarian consequences from man-made disasters in the latest decades, and the Russian aggression against Ukraine, which started in 2014, is the most recent example of the horrors of mines, unexploded ordnance (UXO) and other ERW. In many cases, the ammunition used in this conflict may have a low reliability due to age and storage conditions, causing them to malfunction [6]. Based on the estimated failure rate of these munitions, the expenditure of artillery rounds alone could produce over 200,000 new UXO each passing month [7–9]. Additionally, there are also the UXO produced by other weapon systems, as well as ammunition that has yet to function as intended, such as mines, booby traps, and munitions stockpiles that may be abandoned and partially destroyed. According to a recent assessment (as of January 2023) by the Ukrainian Prime Minister Denys Shmyhal, about 250,000 square kilometres, or roughly 40% of the Ukrainian territory, is now contaminated with munitions [10].

There are multiple ways of estimating impact sensitivity, but arguably the most common is to employ an apparatus called a fallhammer, in which a weight of fixed mass is repeatedly dropped onto samples of the explosive of interest [11]. The researcher then observes which drops lead to explosions. Ideally, one would have preferred to estimate a material's impact sensitivity as a single energy level, where one would know for certain that any impact of energy below this value would never cause an explosion, and, vice versa, that any impact of energy above this level always would. However, when dropping the fallhammer weight from the same height repeatedly, it is rarely the case that all drops cause the same reaction. The task of estimating sensitivity is therefore a statistical problem, in which the best one can do is to estimate the probability of an explosion occurring given a certain impact energy. As one would expect, this probability of an explosion occurring increases with the drop height. Most commonly, the median value h_{50} is estimated, which represents the (log) height at which there is a 50% probability of an explosion occurring. However, for many practical applications, extreme quantiles like h_{01} or h_{99} (corresponding to the (log) heights at which there are a 1% and 99% probability of an explosion occurring, respectively) are more relevant.

Researchers, manufacturers and military personnel determining the sensitivity of energetic materials most commonly follow NATO's standard guidelines (STANAGs) when conducting their experiments. The STANAGs for measuring sensitivity to friction [12], shock [13] and impact [14] all impose the use of the Bruceton method. Introduced by Dixon and Mood [15] in 1948, this method consists of two parts:

- an experimental design for choosing the drop heights for the fallhammer experiments (henceforth referred to as the *Bruceton design*),
- an approximate estimate of h_{50} from the data obtained (henceforth referred to as the *1948 approximation*).

Although more sophisticated experimental designs have been introduced since [16–19], the Bruceton design is nevertheless still imposed in the aforementioned STANAGs, and provides a simple and model-independent recipe for choosing the impact levels. Furthermore, the long-term behaviour of up-and-down designs (which include Bruceton as a special case) has been extensively studied [20–25], and the construction of confidence intervals based on the large-sample properties of the Bruceton design (see Section 2.2) has been explicitly mathematically verified [23–25]. However, what is not well-known to many modern researchers is that the 1948 approximation was only intended as a pen-and-paper alternative to calculating the maximum likelihood estimators (MLEs) of the model for the sensitivity. Dixon and Mood [15] explain that ideally, one would calculate the MLEs directly, but that "a significant simplification can be made by neglecting a small part of the information in the sample." This was a useful simplification in 1948, before the advent of computers. Today, the MLEs in models for binary outcomes can easily be computed numerically in a fraction of a second with standard spreadsheet software like Microsoft Excel or virtually any general-purpose programming language like Python, MATLAB or R.

Even worse, a serious confusion has arisen amongst modern users of the Bruceton method from the fact that the 1948 approximation is not always valid. As Dixon and Mood [15] explain, the accuracy of their approximation depends on a certain numerical criterion being satisfied by the data, and it can result in inaccurate estimates if applied wrongly. However, many believe that when the approximation is not valid, then the data are not good enough and must therefore be discarded. That is, the researcher has obtained a perfectly valid data set from his or her experiments; the MLEs exist and can be computed numerically. Yet, when the 1948 approximation is employed, the data will be deemed "not

valid", simply because the approximation does not apply. However, this is of no relevance, as the MLEs can be computed directly, regardless of the validity of the approximation. Therefore, uncritical use of said approximation is both unnecessary and wasteful. Instead, impact sensitivities ought to be estimated nowadays using maximum likelihood theory directly, not using an obsolete pen-and-paper approximation from over 75 years ago.

Beyond acquiring an accurate point estimate of h_{50} from data, it is equally important to assess the uncertainty of this estimate. Without a proper quantitative evaluation of uncertainty, a point estimate is just as useful as an initial guess. In statistical inference, uncertainty is most commonly reported by means of confidence intervals (CIs). Unfortunately, the construction of CIs has almost been entirely neglected in the sensitivity testing literature. It seems like the primary focus has been on inventing new experimental designs which yield more efficient point estimates, without any regard to the construction of CIs [16,17,19]. NATO's standard guidelines for sensitivity tests do not require the researcher to report CIs, or even instruct them on how this should be done [12–14]. As a result, many databases on impact sensitivity data only contain a single estimate (namely $\hat{h}_{50}$) without any CIs or the original data from which CIs could be computed [26–29]. Thus, without the original data, CIs have to be redundantly re-estimated using regression analysis [30]. For binary response data, there are three main alternatives for constructing CIs, namely via the delta method, via Fieller's theorem or via the likelihood ratio test. Indeed, Dixon and Mood explain how to combine their approximation with the delta method to obtain CIs for h_{50} [15]. However, numerous simulation studies have since demonstrated that the delta method is consistently unfavourable for binary regression of this kind [31–35]. The last of these studies looked at the Bruceton design specifically, and found that Fieller's theorem yielded the most accurate confidence intervals, and that the delta method performed the worst in all cases.

This paper reports how to compute MLEs, estimate quantiles and construct confidence intervals when using the Bruceton design for estimating sensitivity. In particular, it is explained why the 1948 approximation is to be avoided, by comparing it to the MLEs in a simple simulation study. Then, five real datasets measuring the impact sensitivity of remnants of amatol from the Second World War are considered, all of which are deemed "not valid" by the 1948 approximation. This illustrates the point that a proper statistical analysis can yield a far better insight into the data obtained than uncritical usage of the 1948 approximation.

2 | STATISTICAL ESTIMATION OF IMPACT SENSITIVITY

This section covers the mathematical background required for statistical estimation of impact sensitivity. Suppose one has conducted n fallhammer drops, where, typically, $n = 30$. Letting $h_1, \ldots, h_n$ denote the heights from which the hammer was dropped, one observes the binary outcomes $y_i \in \{0, 1\}$, for $i = 1, \ldots, n$, where $y_i = 1$ if an explosion occurs at height h_i and $y_i = 0$ otherwise.

For many types of explosives, the probability that $y_i = 1$ (i.e. that an explosion occurs) can be modelled reasonably accurately using probit regression on a log scale,

$$\Pr(y_i = 1) = \Phi(\alpha + \beta x_i), \tag{1}$$

where $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2} \mathrm{d}t$ is the standard normal cumulative distribution function (cdf), $x_i = \log h_i$ for $i = 1, \ldots, n$ for $i = 1, \ldots, n$, and the constants α, β are the parameters of the model, to be estimated from data. Figure 1 provides a visual illustration of the model in Equation 1. The question of how the (log) heights $x_1, \ldots, x_n$ are chosen is discussed in detail in Section 3.

From Equation 1, the median h_{50} is given by

$$h_{50} = -\alpha/\beta, \tag{2}$$

and, more generally,

$$h_{100q} = \frac{z_q - \alpha}{\beta}, \tag{3}$$

where $z_q = \Phi^{-1}(q)$ is the q-quantile of the standard normal distribution. Note that $z_{1/2} = 0$, so Equation 2 can be recovered from Equation 3.

Another common parametrisation of Equation 1 is given by

$$\Pr(y_i = 1) = \Phi\left(\frac{x_i - \mu}{\sigma}\right), \tag{4}$$

from which it is clear that

$$\mu = -\alpha/\beta \quad \text{and} \quad \sigma = 1/\beta.$$

In this paper, the parametrisation of Equation 1 is used, since this is arguably more natural and insightful. Indeed, firstly, if one were to include another covariate x' to the model rather than just the (log) height x of the drop, one would extend linearly,

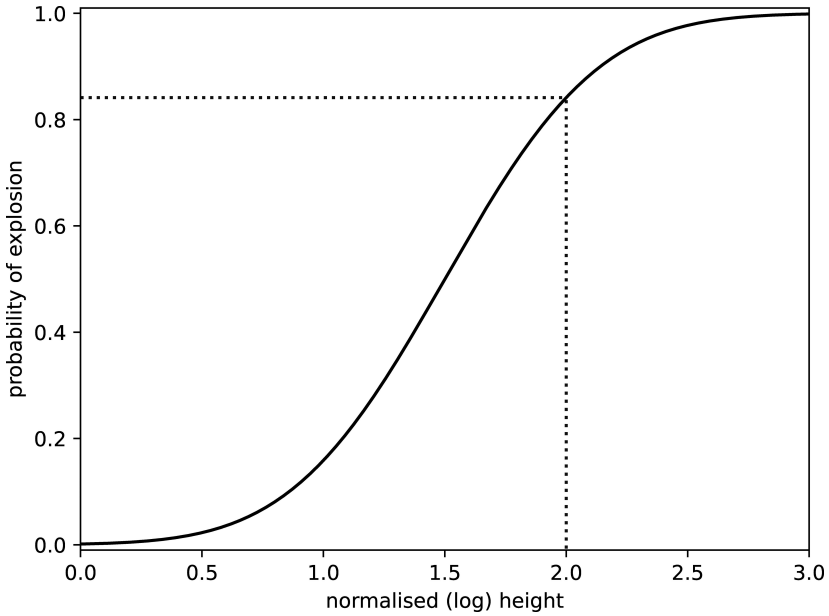

FIGURE 1 The probit curve governing the probability of an explosion occurring, with $\alpha = -3.0$ and $\beta = 2.0$. Here, on a normalised scale, the (log) height $x = 2.0$ yields a probability of 84% of an explosion occurring.

$$\Pr(y' = 1) = \Phi(\alpha + \beta x + \gamma x'),$$

where the parameters of the model are now α, β and γ. Secondly, it is more insightful to realise h_{50} as a ratio of two parameters (namely $-\alpha/\beta$) rather than its own separate parameter, as this explains why asymmetric confidence intervals naturally arise when estimating h_{50}. This point is elaborated further on in Section 2.2.

2.1 | Maximum likelihood estimation (MLE)

Given data $(x_1, y_1), \ldots, (x_n, y_n)$, the model in Equation 1 yields the likelihood

$$L(\alpha, \beta) = \prod_{i=1}^{n} \Phi(\alpha + \beta x_i)^{y_i} [1 - \Phi(\alpha + \beta x_i)]^{1-y_i} \tag{5}$$

and the log-likelihood

$$\begin{aligned} \ell(\alpha, \beta) = &\sum_{i=1}^{n} \{y_i \log \Phi(\alpha + \beta x_i) \\ &+ (1 - y_i) \log[1 - \Phi(\alpha + \beta x_i)]\}. \end{aligned} \tag{6}$$

Note that maximising the expressions in Equation 5 and Equation 6 is equivalent. This maximiser is called the maximum likelihood estimator (MLE), and is the most efficient estimator, given that the model in Equation 1 is correct (see e. g. Ferguson [36]). We write

$$\hat{\alpha}, \hat{\beta} = \arg \max_{\alpha, \beta} \ \ell(\alpha, \beta).$$

Although expressions for $\hat{\alpha}$ and $\hat{\beta}$ are not available in closed form, the log-likelihood in Equation 6 is globally concave, and so $\hat{\alpha}$ and $\hat{\beta}$ can easily be found numerically via standard optimisation techniques as the Newton-Raphson method. This functionality is built into all standard spreadsheet software, like Microsoft Excel, and virtually any general-purpose programming language, like Python, MATLAB or R.

If the parametrisation in Equation 4 is used, then the MLEs for μ and σ are easily obtained as

$$\hat{\mu} = -\hat{\alpha}/\hat{\beta} \quad \text{and} \quad \hat{\sigma} = 1/\hat{\beta}.$$

Hence, the estimate for h_{50} is given by $\hat{h}_{50} = -\hat{\alpha}/\hat{\beta}$. Similarly, the estimate for a general quantile h_{100q} is given by $\hat{h}_{100q} = (z_q - \hat{\alpha})/\hat{\beta}$.

For some datasets, the MLEs $\hat{\alpha}$ and $\hat{\beta}$ do not exist. Consider the dataset in Table 1, where there is a gap between all non-explosions and explosions. Optimising the log-likelihood for this dataset would only force $\beta \to \infty$ (or equivalently, $\sigma \to 0$), as the optimal solution is

TABLE 1 A dataset with perfect separation (no overlap).

Normalised height	1.0	1.5	2.0	2.5
Explosions	0	0	10	10
#Trials	10	10	10	10

obtained when $\Pr(y_i = 1)$ is a step function. Such datasets are usually said to satisfy *perfect separation*. In the sensitivity testing literature, datasets without perfect separation are sometimes said to contain an overlap [17,19]. In cases where the MLEs do not exist, more measurements must be made.

2.2 | Confidence intervals (CIs)

Once an estimate of a parameter has been obtained, it is crucial to address its uncertainty via CIs. As an example, a CI for h_{50} is an interval $[u, v]$ (which depends on the data) such that

$$\Pr(u \leq h_{50} \leq v) = 1 - \gamma,$$

where $1 - \gamma$ is the desired confidence level, the most standard choice being $\gamma = 0.05$, resulting in a 95 % CI. As u and v depend on the data, they are random variables, as opposed to h_{50}, which is a fixed, albeit unknown, parameter. The remainder of this section briefly explains how to obtain CIs from sensitivity data (for a more complete account, see Christensen et al. [35]). Since $\hat{\alpha}$ and $\hat{\beta}$ are functions of the outcomes $y_1, \ldots, y_n$, they are random variables with an associated probability distribution. It is via this distribution that confidence intervals are derived. However, the exact distribution of $\hat{\alpha}$ and $\hat{\beta}$ is rarely available in closed form, and so one relies on large-sample theory (i. e. what happens as $n \to \infty$) to obtain an approximation. One of the most fundamental results in mathematical statistics asserts that, under mild regularity conditions on the log-likelihood function, the MLEs are approximately normally distributed, where the accuracy of this approximation increases as $n \to \infty$. More specifically, let $\theta = (\alpha, \beta)^{\mathrm{T}}$ and $\hat{\theta} = (\hat{\alpha}, \hat{\beta})^{\mathrm{T}}$, and define the *Fisher information matrix* J by

$$J = -\mathbb{E}_\theta\left[\frac{\partial^2 \ell}{\partial\theta\partial\theta^T}\right] = \sum_{i=1}^{n} \frac{\phi^2(\eta_i)}{\Phi(\eta_i)[1 - \Phi(\eta_i)]} \begin{pmatrix} 1 & x_i \\ x_i & x_i^2 \end{pmatrix}, \quad (7)$$

where $\eta_i = \alpha + \beta x_i$ and $\phi(x) = (2\pi)^{-1/2} e^{-x^2}$ is the standard normal probability density function (PDF). Then

$$\hat{\theta} \approx \mathrm{N}(\theta, V), \quad (8)$$

where $V = J^{-1}$. We write $\hat{J}$ for the matrix obtained from substituting $\hat{\alpha}$ and $\hat{\beta}$ for α and β in Equation 7, and let $\hat{V} = \hat{J}^{-1}$. From Equation 8, Fieller's theorem [37] yields the following $100(1 - \gamma)\%$ level CI for h_{50},

$$\hat{h}_{50} + \frac{g}{1-g}\left(\hat{h}_{50} + \frac{\hat{V}_{12}}{\hat{V}_{22}}\right) \pm \frac{z_{\gamma/2}}{\hat{\beta}(1-g)}\{\hat{V}_{11} + 2\hat{h}_{50}\hat{V}_{12} + \hat{h}_{50}^2\hat{V}_{22} - g\left(\hat{V}_{11} - \frac{\hat{V}_{12}^2}{\hat{V}_{22}}\right)\}^{1/2},$$

where $g = z_{\gamma/2}^2 \hat{V}_{22}/\hat{\beta}$. Here, it is assumed that $g < 1$. Christensen et al. [35] also cover the case where $g \geq 1$.

A common misconception in the energetic materials literature, particularly in cases where the parametrisation in Equation 4 is employed, is that $\hat{\mu} \pm z_{\gamma/2}\hat{\sigma}$ serves as a $100(1 - \gamma)\%$ CI for h_{50}. By this logic, $[\hat{h}_{02.5}, \hat{h}_{97.5}]$ would be a 95% CI for h_{50}. However, this is incorrect, as it wrongly assumes that σ is the standard deviation of $\hat{h}_{50}$. Although the letter σ is commonly used to denote standard deviations in statistics, it is important to realise that in the context of Equation 4, $\sigma = 1/\beta$ is simply one of the two model parameters, associated to the horizontal stretching of the probit curve drawn in Figure 1. It is indeed possible to derive the asymptotic standard deviation of $\hat{h}_{50}$ from Equation 8 via the delta method (see e. g. Ferguson [36]). Doing so, one obtains

$$\mathrm{sd}(\hat{h}_{50}) \approx \frac{1}{\beta}\sqrt{V_{11} + 2h_{50}V_{12} + h_{50}^2 V_{22}},$$

from which CIs for h_{50} can be obtained directly. As pointed out in the introduction, however, these CIs do not perform as well as those created using Fieller's theorem in simulations [31–35].

3 | THE BRUCETON METHOD

The Bruceton method, which is the recommended procedure for obtaining and analysing sensitivity data [12–14,38], is now studied in more detail. Since the method comprises two parts, namely the Bruceton design and the 1948 approximation, these are analyses in turn.

3.1 | The Bruceton design

Also known as the 'up-and-down' design, the Bruceton design refers to the sequential experimental design presented by Dixon and Mood [15]. The design depends on two parameters, an initial (log) drop height x_1 and a step size d>0. After observing the first binary outcome

$y_1 \in \{0, 1\}$ from x_1, the subsequent heights $x_2, \ldots, x_n$ are decided by the rule

$$x_i = \begin{cases} x_{i-1} - d & \text{if } y_{i-1} = 1, \\ x_{i-1} + d & \text{if } y_{i-1} = 0. \end{cases} \tag{9}$$

That is, one goes up one step at the next test when observing no explosion in the current test, or down one step size when observing an explosion. In this way, the Bruceton design has a rubber banding effect built into it, such that the drop heights oscillate around the median h_{50}. It has been proved that this rubber banding behaviour guarantees asymptotic normality of the MLEs, and the approximation in Equation 8 is thus mathematically justified [23–25].

In practice, the amplitude of the oscillations of the Bruceton design depends on the choice of the step size d, and one needs to take care to make sure the step size is not too small or too large. If d is too small, then the sequence $x_1, \ldots, x_n$ will only explore a small region around the median h_{50}, assuming that the initial guess x_1 is chosen close to h_{50}. If x_1 is far away from h_{50}, then a small step size will increase the number of steps required to converge to h_{50} in the first place. Conversely, if d is too large, then the drops will only oscillate between extreme quantiles, increasing the likelihood of perfect (or nearly perfect) separation occurring. This will in turn make it less likely that the MLEs will exist, as explained in Section 2.1. Nevertheless, Christensen et al. [35] demonstrated via simulations that for a wide range of step sizes, the Bruceton design (combined with maximum likelihood estimation and confidence intervals via Fieller's theorem) yields satisfactory results for h_{50} given $n = 30$ measurements. For more extreme quantiles like h_{99}, a larger value of n is needed (Christensen et al. [35] recommend $n = 100$), but the Bruceton design still works consistently well.

3.2 | The 1948 approximation

Dixon and Mood [15] also provide a recipe for estimating h_{50} from the obtained data $(x_1, y_1), \ldots, (x_n, y_n)$. Unbeknownst to many, however, is that this recipe is only meant as an approximation to the MLEs, derived in Section 2.1. In Appendix A, Dixon and Mood [15] write the following:

The estimation of of μ and σ^2 is based on the principle of maximum likelihood. We shall not maximize [the likelihood] directly, however because a material simplification in the analysis can be made by neglecting a small part of the information in the sample.

In order to see how the approximation operates, let $N = \min\{\sum_{i=1}^n y_i, n - \sum_{i=1}^n y_i\}$, so that N is the number of occurrences of the least frequent binary outcome in the data. Also, let x' be the smallest (log) height on which this less frequent outcome occurred, and let n_j denote the number of times it occurred at the height $x' + dj$, for $j = 0, 1, \ldots, k$. Here, k is the maximum height at which the less frequent outcome occurred. Define the sums

$$A = \sum_{j=0}^{k} j n_j \quad \text{and} \quad B = \sum_{j=0}^{k} j^2 n_j.$$

The first simplification to make is to exploit that the number of explosions at any particular height, say x, will be at most one away from the number of non-explosions at height $x - d$, due to the nature of the experimental design in Equation 9. The likelihood in Equation 5 can therefore be simplified by treating these two numbers as the same. Maximising the simplified likelihood obtained then yields an equation involving the function

$$\alpha(x) = \frac{\phi(x - d/2\sigma)}{1 - \Phi(x - d/2\sigma)} - \frac{\phi(x + d/2\sigma)}{\Phi(x + d/2\sigma)}$$

which cannot be solved analytically. However, Dixon and Mood [15] point out that if $d < 2\sigma$, then $\alpha(x)$ can be very closely approximated by a linear function. Performing the linearisation and simplifying the likelihood further still results in the estimates

$$\mu^\star = y' + d\left(\frac{A}{N} \pm \frac{1}{2}\right),$$

$$\sigma^\star = 1.620d\left(\frac{NB - A^2}{N^2} + 0.029\right),$$

where 1/2 is added if $N = n - \sum_{i=1}^n y_i$ and subtracted if $N = \sum_{i=1}^n y_i$ in the equation for $\mu^\star$.

The above derivation shows where the distinction between *valid* and *not valid* estimates comes from in the 1948 approximation. If $d \geq 2\sigma$, then the linear approximation of $\alpha(x)$ is poor and the approximation should be avoided. Similarly, Dixon and Mood [15] give reasons for also avoiding cases where $\sigma \geq 2d$. Thus, if a particular data set yields an estimate $\sigma^\star$ for which $d \geq 2\sigma^\star$ or $\sigma^\star \geq 2d$, then the data are deemed not valid. Note, however, that this criterion has virtually no relation to the existence of the MLEs, or to their accuracy or stability.

4 | SIMULATIONS

In order to evaluate the performance of the 1948 approximation versus that of direct maximum likelihood estimation, we conducted a simple simulation experiment. For the sample sizes sizes $n = 30, 50, 100$ and for step sizes $d = 1.0, 1.25, 1.5, 1.75, 2.0$, we created $S = 100,000$ data sets via the Bruceton design with $x_1 = 2.0$ and true underlying parameters $\alpha = 0$ and $\beta = 1$. For each data set, we computed the MLEs and theestimates $\mu^\star$ and $\sigma^\star$ based on the 1948 approximation. The results are given in Table 2, where we see the proportion of iterations for which the MLEs exist and the 1948 approximation yielded valid estimates, respectively. For the iterations where both methods successfully yielded an estimate, we also report the average square difference between the estimates of α, β and $h_{50} = -\alpha/\beta$. That is, the column for α lists the average square difference $(\hat{\alpha} - \{-\mu^\star/\sigma^\star\})^2$, and so on. From our results we see that although the 1948 approximation is generally in good numerical agreement with the MLEs, it often yields non-valid estimates and discards way too many data sets where the MLEs exist and can be computed without problems. This is particularly the case for smaller step sizes, like for example the case where $n = 100$ and $d = 0.5$, where the 1948 approximation deemed almost half of the iterations to be not valid even though not a single one of these yielded non-existing MLEs.

5 | REAL DATA

In this section we consider the data reported by Novik and Christensen [4] to illustrate with a concrete example how direct maximum likelihood estimation yields better insight than the 1948 approximation. These data were gathered from fallhammer experiments (as described in Section 1) on five samples of amatol extracted from live ordnance originating from the Second World War. The samples are labelled as substance A_2, B, C, D and E. Table 3 summarises the use of the 1948 approximation on these data. Even though Novik and Christensen [4] successfully computed the MLEs and confidence intervals for all of these datasets, not a single one is deemed valid by the 1948 approximation. It is also worth pointing out that as NATO's standard guidelines for explosives and impact sensitivity tests [14] do not include a recipe for constructing CIs, Novik and Christensen [4] use Fieller's theorem to do so, as recommended by Christensen et al. [35]. This points to an improvement to

TABLE 2 Results from the simulations comparing the 1948 approximation with direct maximum likelihood estimation. The columns on the left report the mean square difference between the estimates yielded via maximum likelihood estimation and the 1948 approximation. The columns on the right report the proportion of iterations for which the MLEs converged and the 1948 approximation yielded valid estimates, respectively.

	$n = 30$					$n = 50$					$n = 100$				
	Mean square difference			Converged (%)		Mean square difference			Converged (%)		Mean square difference			Converged (%)	
d	α	β	h_{50}	MLE	1948	α	β	h_{50}	MLE	1948	α	β	h_{50}	MLE	1948
0.50	0.015	0.059	0.004	**96.24**	67.80	0.004	0.014	0.002	**99.97**	61.84	0.001	0.003	0.000	**100.0**	57.24
0.75	0.008	0.011	0.005	**97.71**	89.18	0.002	0.003	0.002	**99.88**	94.88	0.000	0.001	0.000	**100.0**	98.18
1.00	0.005	0.008	0.004	**92.39**	88.66	0.002	0.003	0.001	**98.59**	97.68	0.000	0.001	0.000	**100.0**	99.96
1.25	0.002	0.004	0.003	**92.09**	83.99	0.001	0.002	0.001	**98.87**	93.04	0.000	0.001	0.000	**99.98**	99.01
1.50	0.002	0.002	0.002	**81.45**	70.27	0.001	0.001	0.001	**95.13**	81.41	0.000	0.000	0.000	**99.82**	92.81
1.75	0.003	0.001	0.003	**51.21**	46.61	0.001	0.001	0.001	**70.89**	63.52	0.000	0.001	0.000	**91.62**	78.09
2.00	0.003	0.000	0.005	**28.62**	27.10	0.001	0.001	0.001	**43.62**	41.82	0.000	0.002	0.000	**67.50**	60.54

TABLE 3 The 1948 approximation applied to the amatol datasets.

Substance	Analysis based on	A	B	$m^\star$	$\sigma^\star$	$\sigma^\star/d$	Valid/Not valid
A_2	Non-explosions	85	323	1.367	0.224	4.484	Not valid
B	Non-explosions	31	119	1.216	0.235	4.706	Not valid
C	Explosions	36	86	1.803	0.107	2.136	Not valid
D	Non-explosions	155	937	1.467	0.474	9.474	Not valid
E	Non-explosions	148	712	1.186	0.202	4.035	Not valid

NATO's current recommendations for impact sensitivity testing.

6 | CONCLUSION

This paper has covered the topic of estimating the sensitivity of energetic materials, a proper understanding of which is crucial for ensuring safe handling, storage, transport, use and destruction of explosives. With particular focus on impact sensitivity, it was shown how to obtain accurate estimates of quantiles like h_{50}via maximum likelihood estimation, and how to construct confidence intervals via Fieller's theorem. In particular, it has been explained and demonstrated that the 1948 approximation, currently imposed in the NATO standard guidelines measuring explosives' sensitivity to friction, shock and impact [12–14], discards a substantial proportion of perfectly valid datasets. The approximation is therefore unnecessarily wasteful compared to direct maximum likelihood estimation. With modern computers being able to maximise the log-likelihood directly in a fraction of a second, there is no reason to still employ said approximation. Through an increase in accuracy of explosives testing methodology, the recommendations made in this paper could therefore prove to be advantageous in several phases of life cycle management, including engineering, design, manufacture, use, in-service surveillance and final disposal of explosive substances and munitions containing explosives. In particular, as studies show that risks related to explosive remnants of war could intensify rather than diminish over time, we must manage these risks in safe and effective ways. Such positive actions will extensively depend upon a thorough knowledge of the sensitivity thresholds of the various energetic materials.

ACKNOWLEDGMENTS

This research is funded by the Norwegian Defence Research Establishment. The funding source had no involvement in the research and/or preparation of this article.

CONFLICT OF INTEREST STATEMENT

The authors declare no potential conflict of interests.

DATA AVAILABILITY STATEMENT

The data analysed in this paper are available at Dryad: https://doi.org/10.5061/dryad.sqv9s4n91.

ORCID

Dennis Christensen http://orcid.org/0000-0002-7540-7695

REFERENCES

1. M. Sućeska, *Test Methods for Explosives*; Springer: New York, 1995.
2. G. P. Novik, Complex risk management in explosive-contaminated areas: Explosive remnants of war. PhD thesis, University of Stavanger, Norway, 2023.
3. G. P. Novik, *Sci. Total Environ.* **2022**, *842*.
4. G. P. Novik, D. Christensen, *Roy. Soc. Open Sci.* **2024**, *11*, 1110. 1098/rsos.231344.
5. D. Koske, Dumped munitions: Effects, metabolism and detection of explosive compounds and chemical warfare agent-related chemicals in fish from the Baltic Sea. PhD thesis, Christian-Albrechts-Universität zu Kiel, Germany, 2020.
6. E. C. Duncan, S. Skakun, A. Kariryaa, A. V. Prishchepov, *Science of Remote Sensing* **2023**, *7*.
7. J. Ataman, C. Sebastian, Ukraine is firing shells faster than can be supplied. Can Europe catch up? CNN, 2023, URL: https://edition.cnn.com/2023/09/17/europe/ukraine-shell-supplies-intl/index.html.
8. C. Caroll, Unexploded ordnance is scattered across Ukraine's front lines. UMD researchers are mapping hot spots with AI, Maryland Today, University of Maryland, 2023, URL: https://today.umd.edu/unexploded-ordnance-is-scattered-across-ukraines-front-lines-umd-researchers-are-mapping-hot-spots-with-ai.
9. P. K. Kerr, A.Feickert, Cluster munitions: Background and issues for Congress. Report No. RS22907, Congressional Research Service (CRS), Washington, DC, 2023.
10. dpa, Ukraine has largest minefield in the world, prime minister says, Deutsche Presse-Agentur, 2023, URL: https://www.rferl.org/a/ukraine-largest-minefield-shmyhal/32214163.html.
11. R. Meyer, J. Köhler, A. Homburg, *Explosives*, 6th Edition; John Wiley & Sons: Weinheim, 2007.
12. North Atlantic Treaty Organization (NATO), STANAG 4487 - Explosives, Friction sensitivity tests, 2009.
13. North Atlantic Treaty Organization (NATO), STANAG 4488 - Explosives, Shock sensitivity tests, 2009.
14. North Atlantic Treaty Organization (NATO), STANAG 4489 - Explosives, Impact sensitivity tests, 1999.
15. W. J. Dixon, A. M. Mood, *J. Am. Stat. Assoc.* **1948**, *43*, 109–126.
16. H. J. Langlie, A reliability test for "one-shot" items. Report No. U-1792, Aeronutronic Division, Ford Motor Company, Newport Beach, CA, 1962.
17. B. T. Neyer, *Technometrics* **1994**, *36*, 61–70.
18. V. R. Joseph, *Biometrika* **2004**, *91*, 461–470.
19. C. F. J. Wu, Y. Tian, *J. Stat. Plan. Infer.* **2014**, *149*, 1–15.
20. C. Derman, *Ann. Math. Stat.* **1957**, *28*, 795–798.
21. S. D. Durham, N. Flournoy, Random walks for quantile estimation. In: Gupta, S. S.; Berger, J. O. (Eds.) *Statistical Decision Theory and Related Topics*; Springer, 1994, pp. 457–476.
22. N. Flournoy, A. H. Up-and-down designs. In: El-Shaarawi, W. W. Piegorsch, (Eds.) *Encyclopedia of Environmetrics, Volume 4*; John Wiley & Sons, 2002, pp. 2297–2302.
23. R. K. Tsutakawa, Ann. Math. Stat. 1967, 38, 1822–1828, DOI: 10.1214/aoms/1177698615.
24. W. F. Rosenberger, N. Flournoy, S. D. Durham, J. Stat. Plan. Infer. 1997, 60, 69–76, DOI: 10.1016/S0378-3758(96)00120-6.

25. D. Christensen, E. A. Stoltenberg, N. L. Hjort. Theory for adaptive designs in regression. arXiv preprint, *arXiv:2312.13387,* 2024.
26. C. B. Storm, J. R. Stine, J. F. Kramer, Sensitivity Relationships in Energetic Materials. In: Bulusu, S. N. (Ed.) *Chemistry and Physics of Energetic Materials*; Springer: Dordrecht, 1990, pp. 605–639.
27. W. S. Wilson, D. E. Bliss, S. L. Christian, D. J. Knight, Explosive Properties of Polynitroaromatics. Report No. NWC TP 7073, Naval Weapons Center, China Lake, CA, 1990.
28. J. C. Oxley, A survey of the thermal stability of energetic materials. In: Politzer, P. A.; Murray, J. S. (Eds.) *Energetic Materials, Part 1: Decomposition, Crystal and Molecular Properties*; Elsevier, 2003, pp. 5–48.
29. E.-C. Koch, *High Explosives, Propellants, Pyrotechnics*; De Gruyter: Berlin, 2021.
30. T. L. Jensen, J. F. Moxnes, E. Unneberg, D. Christensen, *J. Mol. Model.* **2020**, *26*.
31. K. M. Abdelbasit, R. L. Plackett, *J. Am. Stat. Assoc.* **1983**, *78*, 90–98.
32. C. Cox, *Biometrics* **1990**, *46*, 709–718.
33. R. R. Sitter, C. F. J. Wu, *J. Am. Stat. Assoc.* **1993**, *88*, 1021–1025.
34. D. Faraggi, P. Izikson, B. Reiser, *Stat. Med.* **2003**, *22*, 1977–1988.
35. D. Christensen, E. Unneberg, E. Høyheim, T. L. Jensen, N. L. Hjort, Improved measurements of impact sensitivities of energetic materials. In: *Proceedings of the 25th International Seminar on New Trends in Research of Energetic Materials (NTREM)*, Institute of Energetic Materials, University of Pardubice, Czechia, 2023.
36. T. S. Ferguson, *A Course in Large Sample Theory*; Chapman & Hall: London, 1996.
37. E. F. Fieller, *J. Roy. Stat. Soc. B* **1954**, *16*, 175–185.
38. E. L. Baker, An overview of single event testing - methods and analysis. Report No. MSIAC O-216, Munitions Safety Information Analysis Center (MSIAC), 2021.

[Correction added on 25 September 2024, after first online publication: Supporting Information have been deleted in this version.]

How to cite this article: D. Christensen, G. P. Novik, E. Unneberg, *Propellants, Explos., Pyrotech.* **2024**, *49*, e202400022. https://doi.org/10.1002/prep.202400022