Talanta 102 (2012) 79–87

Contents lists available at SciVerse ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Review: The approaches for estimation of limit of detection for ICP-MS trace analysis of arsenic

Ljubinka V. Rajaković [a,*], Dana D. Marković [b], Vladana N. Rajaković-Ognjanović [c], Davor Z. Antanasijević [a]

[a] Faculty of Technology and Metallurgy, University of Belgrade, Karnegijeva 4, 11000 Belgrade, Serbia
[b] Vinča Institute of Nuclear Sciences, University of Belgrade, P. O. Box 522, 11001 Belgrade, Serbia
[c] Faculty of Civil Engineering, University of Belgrade, Bulevar Kralja Aleksandra 73, Belgrade, Serbia

ARTICLE INFO

Available online 21 August 2012

Keywords:
Limit of detection
Estimation approaches
Arsenic
ICP-MS

ABSTRACT

The analytical properties of an analytical method must be evaluated through validation protocols. Beside specificity and/or selectivity, linearity of calibration, repeatability and accuracy, the most important parameters are: *LOD* (limit of detection) and *LOQ* (limit of quantification). Through these limits, it is possible to define the smallest concentration of analyte that can be reliably detected and quantified. To establish these limits, an analyst should apply several estimation methods and test a large number of sample replicates. It is difficult to make a compromise between complex statistical programs and the simple analytical demand to have reliable analytical parameters. The differences and equivalency of estimation methods and approaches for analytical limits could be overcome by an experimental comparison. In this paper, the focus is the *LOD* of inductively coupled plasma-mass spectrometry (ICP-MS) measurements employed for the determination of arsenic. The current approaches for the calculation of the *LOD* are summarized and critically discussed.

1. Introduction

In last 10 years, our group was involved in the separation and ICP-MS measurements of arsenic species in natural waters, and it was of interest to define the limits of the procedure and compare them with those of other arsenic speciation analyses [1–5]. Even when the separation and preparation methods were excluded and only ICP-MS results were compared, discrepancies in the analytical properties are notable. Our attention was focused on the *LOD*, which describes the smallest concentration of an analyte that can be reliably measured by an analytical procedure. The *LOD* is the lowest quantity of a substance that can be distinguished from the absence of that substance (a blank value) within a stated confidence limit. A low *LOD* is not sufficient, but it is a necessary condition for a method to be suitable for trace analysis of certain samples. Some recent articles dealing with ICP-MS measurements of total arsenic and arsenic speciation are analyzed concerning the *LOD* determined by one of the standard protocols. The *LOD* was mainly calculated as three times the standard deviation of the background signal or replicate analysis of deionized spiked water samples. It is interesting that the *LOD* values were very similar for all arsenic compounds [6,7] with even the matrix (pure water or seawater) having no notable influence [8]. Noticeable is the rule that a very low value of *LOD* accompanies a low *LOQ* and high precision. According to the reviewed papers, the *LOQ* was assumed as the value corresponding to triple the limit of detection. Only a few authors gave details. Precision was determined as the value of the relative standard deviation (*RSD*).

2. The approaches for estimation of limit of detection for trace analysis

The different values of the *LOD* are a consequence not only of the different approaches for *LOD* estimation, but also because they are random variables. Analysts usually use the simplest approach, although there are several appropriate approaches for the estimation of the *LOD*. According to the ICH guideline [9], the *LOD* should be determined always when the method is to be applied to limit tests of analytes, and optionally for the quantitative tests for impurities. The *LOD* and the method used for determining the *LOD* should be presented. Additionally, it is stated that in cases where an estimated value for the *LOD* was obtained by calculation, this estimate may subsequently be validated by the independent analysis of a suitable number of samples known to be near or prepared at the *LOD*. Usually, it is not reported whether this check was performed and since it is not obligatory, one can never be sure whether it was done or not.

* Corresponding author.
E-mail address: ljubinka@tmf.bg.ac.rs (L.V. Rajaković).

0039-9140/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2012.08.016

Nomenclature

A	regression coefficient of the response/concentration functional relationship (intercept, by ordinary, linear least square regression)
B	regression coefficient of the response/concentration functional relationship (slope, by ordinary, linear least square regression)
a_w	regression coefficient of the response/concentration functional relationship (intercept, by weighted, linear least square regression)
b_w	regression coefficient of the response/concentration functional relationship (slope, by weighted, linear least square regression)
k	coefficient
n	number of data points of the calibration
N	number of measurement of the blank
r	determination coefficient
S	estimate of standard deviation of sample having concentration equal to *LOD* in units of concentration
s	estimate of standard deviation of the blank in units of concentration
s_a	standard deviation of intercept
s_0	estimate of the standard deviation of the response variable of the blank
s_{low}	estimate of standard deviation of the response variable of the sample having concentration equal to the *LOD*
$s_{y/x}$	residual standard deviation (in ordinary, linear least square regression)
$t_{1-p,v}$	quantile of the one tail t-distribution at the level of confidence 1–p and v degrees of freedom
$t_{1-q,v}$	quantile of the one tail t-distribution at the level of confidence 1–q and v degrees of freedom
x	concentration
X	true concentration
$\overline{X}$	the mean value of the concentrations of calibration standards
x_C	critical value, that is limit of decision in the concentration domain (lowest concentration that can be distinguished from the blank)
x_D	limit of detection in the concentration domain (concentration under which, a priori, a sample may erroneously taken for the blank)
y_C	critical value of the response variable, that is limit of decision in the net response domain (lowest response that can be distinguished from the background)
y_D	limit of detection in the response domain (response under which, a priori, a sample may erroneously taken for the background)
z_{1-p}	one tail standardized normal variable
z_{1-q}	one tail standardized normal variable
p	significance level, probability of a false positive error
q	significance level, probability of a false negative error

Greek symbols

α	true intercept of the calibration line
β	true slope of the calibration line
$\delta_{\alpha,\ \beta,\ n-2}$	non-centrality parameter of the non-central t distribution
v	degrees of freedom
σ_0	population standard deviation of the response of a blank
σ_{low}	population standard deviation of the response of the sample having concentration equal to the *LOD*
μ_B	mean response of a blank

There are various methods (approaches) and their variations for estimation of LOD. They can be divided to the ones based on the statistical or mathematical parameters and those based on the authorized institutions or authors who promoted the protocols for the estimation of the LOD. The position of the analyst is not easy. The best test for LOD would be experiments with estimated LOD concentration repeated many, many times (an example: for $N=100$ only 5 measurements should be inadequate). Since there is no time, this is not a usual protocol, one has to choose the most applicable and adequate method for estimation. Which method should be chosen?

2.1. Basic definitions

There are two distinct questions connected with the chemical measurement process, that raise in front of an analyst: How high the instrument response has to be in order to be ascribed to the presence of an analyte? and What is the smallest concentration of an analyte that will be reliably detected if actually present in a probe? It was Currie [10–12] who first formulated these two questions and offered the solution based on statistical theory of hypothesis testing. When writing the hypothesis in his original work, Currie used generic symbol L to represent a few values at the same time: a net (blank corrected) response or amount or concentration, but for the sake of clarity, it is wise to separate definitions in raw signal (insturmental response), y, domain and in the content (concentration), x, domain—two domains connected by the calibration curve. The following definitions are basicaly as by Currie [12], with slightly changed notation.

The critical level or as is often called the limit of decision, denoted with subscript C, y_C, is the net assay signal above which a blank corrected response is reliably attributed to the presence of the analyte. The statistical definition of the critical level is based on the rejection of the null hypothesis, H_0: concentration equal to zero, at the significant level p (type I-error rate, false positive): $P(y > y_C|X=0) \le p$, where P is the probability, y is the net signal and X is the true concentration of analyte. The *LOD* in signal domain denoted with subscript D, y_D, is the net signal corresponding to an analyte concentration level which may a priori be expected to be recognized. It is established by invoking the type II-error, false negative error rate, q: $P(y < y_C|X=x_D)=q$. Actually, at the concentration x_D, y_D is the mean of the net responses but a single net response lies under y_C with a probability q. Different values of p and q can be adopted, however they are usually set at 0.05 or 0.01.

Whenever statistics is on the stage it is of vital importance to make distinction between population parameters and sample test statistics. Population parameters are theoretical values almost never known in real life, while the sample test statistics as estimates of true parameters are computed from the real experimental measurements. According to Voightman [13] values in the theoretical domain should also have their own notation. He suggested that decision and detection limit calculated from population parameters should be in upper case notation, while the decision and detection limits calculated from experimental

data should be in lower case. In the light of this constatation it can be noticed that the definitions given above are connected to the experimental values. Finally, limit of decision and limit of detection are both acctually collections of four values in different domains, as shown in Table 1.

Relationships between limits of decision, detection and quantification and false positive and false negative rate, in the experimental concentration domain are illustrated and presented in Fig. 1. Before going into the details of the calculation process for different kinds of measurement systems and different kinds of approaches, it is important to emphasize the necessity of proper interpretation of the final results of an analysis concerning the critical value, *LOD* and *LOQ*.

If the measured concentration value is

- $x < x_C$ result: analyte not detected, concentration is less than the *LOD*,
- $x_C < x <$ *LOQ* result: analyte detected, concentration is less than the *LOQ*,
- $x >$ *LOQ* result: analyte detected, concentration can be quantified with reliable accuracy and precision.

Definitions introduced above are of general validity, while methods for *LOD* calculation have to be separately given for different types of chemical measurement systems. Our focus is on three groups of univariate chemical measurement systems (CMS), depending on the characteristics of noise and calibration procedure

I. CMS with homoscedastic, Gaussian measurement noise and ordinary least squares (OLS) processing of the calibration curve data.

Table 1
Definition of limit of decision and limit of detection in dependence of type of parameters and type of domain.

Type of parameters/ type of domain	Limit of decision and limit of detection in the net signal domain	Limit of decision and limit of detection in the concentration domain
Population parameters	Y_C, Y_D	X_C, X_D
Sample test statistics	y_C, y_D	x_C, x_D

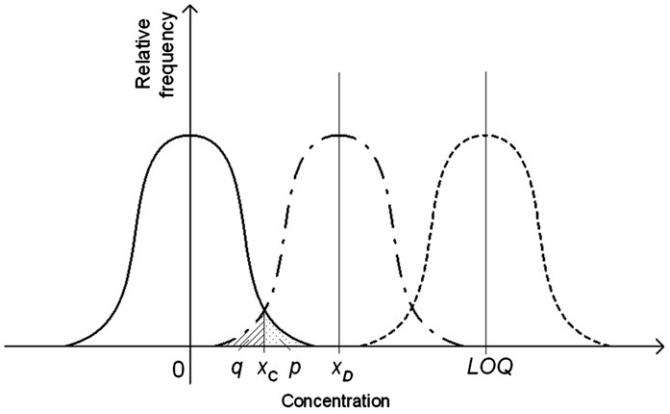

Fig. 1. Principal relationships between the critical value, *LOD* and *LOQ*.x_C—the critical concentration, also called the limit of decision, y_D—limit of detection, *LOQ*—limit of quantification, *p*—probability of false positive error, *q*—probability of false negative error.

II. CMS with heteroscedastic, Gaussian measurement noise and weighted least squares (WLS) processing of the calibration curve data.

III. CMS with nongaussian distributions of measurement error and nonlinear calibration curve.

I. CMS with homoscedastic, Gaussian measurement noise and ordinary least squares (OLS) processing of the calibration curve data

Assuming that there are no systematic errors, an ideal linear system model can be postulated in the following equation form:

$$y = \alpha + \beta X + \varepsilon \tag{1}$$

where X is an independent variable (true analyte concentration), y is a dependent variable (recorded instrumental response), α is the true intercept, β is the true slope, and ε is additive, zero mean, white, Gaussian noise with population standard deviation σ_0, written in short notation: $\varepsilon \sim N(0, \sigma_0)$. When σ_0 is independent of the analyte concentration, the system is called homoscedastic and measurements at any level of concentration are distributed as $y_i \sim N(\alpha+\beta X, \sigma_0)$. The average result of M repeated measurements in signal domain, denoted by $\bar{y}$, is distributed as $\bar{y} \sim N(\alpha+\beta X, \sigma_0/M^{1/2})$. Obtainment of concentration is why the measurements are performed at first place, so simple calculation is performed as transfer from signal domain

$$x = \frac{\bar{y}-\alpha}{\beta} \tag{2}$$

since α and β are errorless values, x is distributed as $x \sim N(X, \sigma_0/(M^{1/2}\beta))$, which is simplified to $x \sim N(X, \sigma_0/\beta)$ if the measurement on a specimen under test is performed only once.

If M measurements are performed on a true blank, $X=0$, $\bar{y}_B \sim N(\alpha, \sigma_0/M^{1/2})$, and in concentration domain $x_B \sim N(0, \sigma_0/(M^{1/2}\beta))$. Specially, for $M=1$ $y_B \sim N(\alpha,\sigma_0)$ and $x_B \sim N(0, \sigma_0/\beta)$. Now, directly from Currie's definitions we get the formulas for critical level and *LOD*. The following are pairs of equations accounting for general value of M and $M=1$, respectively:

$$X_C = z_p\sigma_0/(M^{1/2}\beta) \;\&\; X_C = z_p\sigma_0/\beta \tag{3}$$

$$X_D = (z_p+z_q)\sigma_0/(M^{1/2}\beta) \text{ and } X_D = (z_p+z_q)\sigma_0/\beta \tag{4}$$

It is clear that Currie did not base the definitions in signal domain on gross, but on net signal in order to match the ones in concentration domain.

$$Y_C = z_p\sigma_0/M^{1/2} \text{ and } Y_C = z_p\sigma_0 \tag{5}$$

$$Y_D = (z_p+z_q)\sigma_0/M^{1/2} \text{ and } Y_D = (z_p+z_q)\sigma_0 \tag{6}$$

True parameter α in Eq. (2) is not known, so it has to be estimated somehow. One possible method for obtaining unbiased experimental estimate, denoted by $\hat{\alpha}$, is by performing N replicate measurements of a true blank and calculating $\hat{\alpha}$ as their average value. Variance in x (Eq. (2)) is then the sum of variances of $\bar{y}$and of $\hat{\alpha}$and thus it may be written $x \sim N(X, \sigma_0(1/M + 1/N)^{1/2})$. In literature [12,13], the use of so called composite blank subtraction factor denoted by $\eta^{1/2}$ is suggested. Using η notation $x \sim N(X, \sigma_0\eta^{1/2})$, the following form of true limits of decision and detection is obtained:

$$X_C = z_p\sigma_0\eta^{1/2}/\beta \text{ and } Y_C = z_p\sigma_0\eta^{1/2} \tag{7}$$

$$X_D = (z_p+z_q)\sigma_0\eta^{1/2}/\beta \;\&\; Y_D = (z_p+z_q)\sigma_0\eta^{1/2} \tag{8}$$

In practice, instead of Eq. (1), calibration experiment, including n data pairs, followed by ordinary least squares regression procedure

leads to equation

$$y = a + bx \tag{9}$$

where a is the intercept and b is the slope.

Now, apart from replicate measurements of blank, another way of blank subtraction is possible while performing transformation from signal to concentration domain. Actually, a is an unbiased point estimate test statistics for α, and is normally distributed with population standard deviation σ_a: $a \sim N(\alpha, \sigma_a)$. Thus, the result of M repeated measurements on a specimen under test, using a as $\hat{\alpha}$, is $x \sim N(X, (\sigma_0^2/M + \sigma_a^2)^{1/2})$.

Population standard deviation of the intercept can be calculated from the population standard deviation of the regression, or residual standard deviation or residual standard error, denoted by $\sigma_{y/x}$ using the following equation [14]:

$$\sigma_a = \sigma_{y/x}\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{1/2} \tag{10}$$

where $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ is the mean of n standards. From experimentally obtained OLS only s_a, sample standard deviation of the intercept, can be calculated by the equation

$$s_a = s_{y/x}\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{1/2} \tag{11}$$

where $s_{y/x}$ is sample standard error about the regression defined as follows:

$$s_{y/x} = \left(\frac{\sum_{i=1}^{n}(y_i - a - bx_i)^2}{n-2}\right)^{1/2} \tag{12}$$

For homoscedastic noise OLS processing of calibration curve and constant number of replicates per standard deviation of noise is equal to sample standard error about the regression, $s_0 = s_{y/x}$.

When formulating experimental domain limit of decision and limit of detection expressions, there are two notations: because of the use of negatively biased point estimate test statistics s_0 instead of population parameter σ_0 the critical z values must be replaced with critical central t distribution values for a proper number of degrees of freedom. Blank correction factor has to be modified in the following way:

$$\eta^{1/2} = \left(\frac{1}{M} + \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)^{1/2} \tag{13}$$

Finally, limit of decision and limit of detection in experimental domain, should be calculated using the Eqs. (14) and (15)

$$x_C = t_p s_0 \eta^{1/2}/b \text{ and } y_C = t_p s_0 \eta^{1/2} \tag{14}$$

$$x_D = (t_p + t_q) s_0 \eta^{1/2}/b \text{ and } y_D = (t_p + t_q) s_0 \eta^{1/2} \tag{15}$$

The rates of false positives and false negatives are as expected to be: this is very important conclusion Voigtman derived performing rigorous testing of this method by Monte Carlo simulations [15].

II. CMS with heteroscedastic, Gaussian measurement noise and weighted least squares (WLS) processing of the calibration curve data

In the case when the variance of measured response is not constant at any of the concentration values the system is called heteroscedastic. Processing of calibration experiment data is modified introducing weights. Data that are more reliable (having smaller variability) are given greater emphasis, or weight. Weights are often the inverse standard deviation or the inverse variance.

Ideal linear calibration model can be presented in the form of Eq. (16), following the notation introduced by Voigtman [16]:

$$f(X) = \alpha + \beta X + \varepsilon(\beta X) \tag{16}$$

where X is errorless independent variable in the concentration domain, α is the true intercept, β is the true slope and $\varepsilon(\beta X)$ is additive, zero mean white Gaussian (normal) noise ε $(\beta X) \sim N(0, \sigma(\beta X))$. The result of one measurement of specimen of unknown true concentration is distributed as $f(X) \sim N(\alpha + \beta X, \sigma(\beta X))$, or if M repeated measurements are performed as $\bar{f}(X) \sim N(\alpha + \beta X, \sigma(\beta X)/M^{1/2})$.

In practice, N data pairs of the form $X_i, \bar{f}(X_i)$ are obtained by the calibration experiment and processed by WLS regression. Each $\bar{f}(X_i)$ is the sample mean of M_i independent replicate measurements per X_i standard, which is assumed to have zero error. The result is WLS calibration curve being simply

$$f(x) = a_w + b_w x \tag{17}$$

where a_w is WLS intercept, b_w is WLS slope. As with the OLS both of them are Gaussian distributed

$a_w \sim N(\alpha, \sigma_{(a)w})$ and $b_w \sim N(\beta, \sigma_{(b)w})$. x is the result of experimental measurement, actually a random variate point estimate of X. Voigtman [16] emphasized the importance of recognizing the equality

$$f(x) = (X)\bar{f} \tag{18}$$

Then, if the a_w is taken as $\hat{\alpha}$ in the blank subtraction procedure, Eq. (19) is valid

$$f(x) - a_w \sim N(\beta X, (\sigma_{(a)w}^2 + \sigma^2(\beta X)/M))^{1/2} \tag{19}$$

For this model it is of great importance that true noise precision model is known, so that it can be incorporated in calculation procedure. All the details are explained in the paper of Voigtman [16–20], but it should be stressed that blank correction factor should be calculated according to the equation

$$\eta^{1/2} = \left(\frac{s_0^2}{NN_S} + \frac{1}{n} + \frac{\bar{x}_w^2}{\sum_{i=1}^{n} w_i (x_i - \bar{x}_w)^2}\right)^{1/2} \tag{20}$$

where N_S is weight normalization factor.

III. CMS with nongaussian distributions of measurement error and nonlinear calibration curve

In the case when the basic assumption of Currie detection concept that the noise is Gaussian distributed is not fulfilled the alternative approaches are suggested in the literature. According to nonparametric approach suggested by Linnet and Kondratovich [21] limit of decision and limit of detection are calculated directly in experimental concentration domain. Limit of decision is calculated as the value of the $(N(95/100)+0.5)$th ordered observation of the repetitions of the blank measurements for the usual value of $p = 0.05$. Further on, *LOD* is estimated by the following equation:

$$x_D = x_C + c_q s_{low} \tag{21}$$

where c_q is correction factor for v degrees of freedom: $c_q = z_q/(1 - 1/4v)$.

Recently, Lavagnini et al. [22] suggested the use of nonparametric Theil-Sen regression technique combined with the statistical approach of Lancaster and Quade. They used one-sided simultaneous tolerance interval to calculate the *LOD* which is similar to the Hubaux–Vos approach which is discussed below.

Traditional method of LOD estimation [23,24]: It can be noted that in the every day analysis the so called 3 sigma rule is the one most often applied. Only one numerical value is calculated and used at the same time both as decision and detection limit. The

calculation formula is as simple as

$$x_D = ks_0/b \quad (22)$$

where k is a constant.

Many simplifications are accepted: the number of future replicates of a specimen under test is restricted to one, the correction of standard deviation because of blank subtraction is neglected (net signal and concentration are noisier than the gross instrumental signal). As a result Type II errors are neglected, it is implicitly specified $q=50\%$ and accepting a fixed value of k (very often $k=3$), Type error I rate is out of control, usually unknown to the analyst.

The same equation is used in two cases, when the value of blank is calculated by independent measurements of the blank (10–20 times) [23] and when the signal and standard deviation of the blank are estimated from the characteristics of the calibration line [9,14]. In the second case the authors [9,14] suggested that the value of the intercept should be used to represent the signal of the blank and the residual standard deviation of the regression or the standard deviation of intercept should be used to represent the standard deviation of the blank.

US EPA (United States Environmental Protection Agency), Method 200.8 [24]: EPA introduced Method Detection Limit (*MDL*). The *MDL* is the minimum concentration of an analyte that can be identified, measured, and reported with 99% confidence that the analyte concentration is greater than zero. To determine *MDL* values, seven replicate aliquots of fortified reagent water are subjected to the entire analytical method. It is suggested that *MDL* should be calculated as follows:

$$MDL = t_{99\%,6}S \quad (23)$$

where: t is Student's value for a 99% confidence level and with 6 degrees of freedom, $t=3.14$ for seven replicates, and S is an estimate of standard deviation of the replicate analysis of spiked samples.

MDL values should be determined annually. When a new operator begins to work or when the analyst changes analytical performance (caused by a change in instrument hardware or operating conditions) then *MDL* values need to be determined again.

Method based on prediction intervals, (*Hubaux–Vos*) [25]: The Hubaux–Vos method [25] is based on the prediction interval of the calibration line. The authors tended to incorporate false positive and false negative errors into upper and respectively lower prediction limits of the calibration line. This method can be represented by the following equations:

$$x_C = \frac{t_{(p,n-2)}s_{y/x}}{b}\left(1+\frac{1}{n}+\frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{1/2} \quad (24)$$

and

$$x_D = x_C + \frac{t_{(q,n-2)}s_{y/x}}{b}\left(1+\frac{1}{n}+\frac{(x_D-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{1/2} \quad (25)$$

where: b is the slope of calibration line, n is the number of data points of the calibration, $\bar{x}$is the mean value of the concentrations of calibration standards, x_i is the concentration of standard used for calibration. The x_D value can be obtained through iterative calculations or graphically (x_D can be calculated as the abscissa of the intersection of the parallel line to the x axis passing through y_C with the lower one-sided $(1-q)$ 100% prediction function). It can be seen that in the equations giving x_C and x_D, the terms inside the square root represent the two contributors to the variance, arising from the variability of the measurement and from the uncertainty of the calibration curve. The second contributor depends on the chosen experimental design, thus indicating that the values of x_C and x_D depend on the adopted experimental design.

Method based on non-centrality parameter of the non-central t-distribution [26]: ISO [26] is an alternative approach to calculate x_D. The critical level in the concentration domain is calculated, as above, *via* a central t-distribution. The limit of detection, x_D, is calculated by a non-central t-distribution using the chosen protection against a false negative error. The value of x_D in the concentration domain is immediately obtained from the calibration function as

$$x_D = \frac{\delta_{(p,q,n-2)}s_{y/x}}{b}\left(1+\frac{1}{n}+\frac{\bar{x}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right)^{1/2} \quad (26)$$

$\delta_{p,q,n-2}$ is the non-centrality parameter of the non-central t-distribution.

The equation for x_D was derived using the experimental domain calculating expressions and at the same time making the assumption that true σ_0 value is known. They matched two things that do not go together. As a result this method is biased as shown by Voigtman [13] who performed Monte Carlo simulations.

Signal to noise ratio [27,28]: This method is mainly used for chromatographic techniques. The noise and the signal are measured manually on the chromatogram printout. The height of noise magnitude is taken as an estimate of standard deviation of the blank and the height of the peak as an estimate of analyte signal. The *LOD* corresponds to the amount of analyte for which the signal-to-noise ratio is equal to the value 3. The advantage of this approach is that it is very simple to implement, but it is much dependent on the subjective assessment of the operator.

A review of the estimation approaches for the calculation of the *LOD* with critical comments is presented in Table 2.

3. The approaches for estimation of limit of detection applied for ICP-MS measurements of arsenic

To present the results in the most appropriate manner, the different approaches were tested on our ICP-MS data. The determination of arsenic was realized by ICP-MS using an Agilent 7500ce ICP-MS system (Waldbronn, Germany) equipped with an octopole collision/reaction cell, a MicroMist nebulizer and a Peltier cooled (2.0 °C) quartz Scott-type double pass spray chamber, employing Agilent 7500 ICP-MS ChemStation software. The optimal operating conditions of the Agilent 7500ce are presented in Table 3.

Experimental data and the procedure is described in details in our previous work [1–5]. The concentrations of the standards were: 0.0, 0.8, 2.0, 4.0, 10.0 and 20.0 μg L^{-1}. The analysis were performed in a one week period under conditions of intermediate precision (same analyst, same equipment, same reagents but different days). The measurements were repeated in eight cycles, each cycle containing one replicate of each standard. The order of measurements in each cycle was from the lowest to the highest concentration. Additionally, eighteen blank samples and seven samples spiked at level of 0.5 μg L^{-1} As, were analyzed. For each measurement ICP-MS gives the mean value of three automatic scans. This mean value was used for the further calculations.

3.1. Traditional model

The *LOD* values for the ICP-MS measurements accomplished in our laboratory calculated using the most common traditional method with its eight variations are presented in Table 4.

Table 2
A review of the estimation approaches for the *LOD* calculation with critical comments.

Author/s	Lit.	Approach		Comments
		Description	Equations	
I. Traditional model				
Currie	[10]	Calculation based on net signal value, blank estimated as standard deviation of blank or intercept of calibration.	$x_C=t_p s_0\eta^{1/2}/b$ $y_C=t_p s_0\eta^{1/2}$ $x_D=(t_p+t_q)s_0\eta^{1/2}/b$ $y_D=(t_p+t_q)s_0\eta^{1/2}$	The method which is proven to be correct. Blank subtraction, number of repetitions of measurements, and usage of estimated standard deviation are taken into account. Equations are generally valid, only η factor is changed to match the particular case.
IUPAC	[23]	Measured signal, y, is based on standard deviation of blank (theoretically σ_0, practically s_0).	$y=y_0+k\sigma_0$ $y=y_0+ks_0$	σ_0 is a theoretical value which can be true value if there are indefinite numbers of replicates. Since this is not performed, instead of σ_0 is used s_0, which presents its point estimate. k is not strictly defined, but it should be ascribed to the number of blank replicates. The usual value of k is 3. The mean of replicates of a blank and the intercept of calibration line are interchangeably used although they are not numerically the same. They are both unbiased estimates of the mean blank signal.
US EPA	[24]	Seven replicate aliquots of fortified reagent water are subjected to entire analytical method.	$MDL=t_{99\%,6}\,S$	If the *MDL* is taken as the *LOD* the probability of false negative conclusion is as high as 50%.
II. Methods based on prediction intervals and non-centrality parameter, (Hubaux–Vos and ISO)				
Hubaux–Vos	[25]	Based on the prediction intervals of calibration line.	Eqs. (24) and (25)	According to Voigtman the approach is false [13] due to systematic error of prediction. The calculated LOD is biased. Inconvenient to the analyst and should not be recommended.
ISO	[MEP_L_bib2626]	Based on the non-centrality parameter of the non-central t distribution.	Eq. (26)	The same as previous.
III. Method based on signal to noise ratio				
USP, EP	[27,28]	Based on the signal of measurement when *signal* to *noise*, ratio is equal to value 3.	$Signal/Noise=3$	Convenient for chromatographic methods. Negative aspect: a subjective interpretation of parameters.

Table 3
Optimal instrumental (Agilent 7500ce) operating conditions.

Operation parameters	
RF frequency (MHz)	27
RF power (W)	1500
Plasma gas flow (l/min)	15
Nebulizer gas flow (l/min)	0.9
Sample uptake rate (rps)	0.3
Data acquisition	
Acquisition mode	Peak hopping
Dwell time (ms)	100
Integration time (s)	0.1–0.3/point
Repetition	3 (FullQ)

Eighteen replicates of the blank probes were measured and the calibration curve was obtained eight times. In order to be concise, only one calibration curve is presented in Fig. 2.

The original data as provided by the instrument software correspond with the data presented as an example in Table 4. All calibration curves were simple ordinary linear least square regressions, *OLS*, (correlation coefficient $r \geq 0.9999$ each time). From each calibration curve, one set of results was obtained.

All sets of results are presented in Fig. 3 as their mean value and standard deviation. For each of eight variations, the critical value of the signal domain and the *LOD* value were calculated and transformed to the concentration domain through the calibrations. It can be noticed, that by variations 1 and 3, the obtained *LOD* values in the concentration domain are negative. This is a consequence of intercept of the calibration line being greater than the mean value of the blank signal, even greater than the calculated y_D. For the same reason, the calculated critical concentrations can also be negative values.

The calculated critical and *LOD* values are unique for each variation and for each calibration and they are not directly comparable. The most appropriate values were obtained according to variation 1. The critical value was estimated as the mean value of eight calibration curves (not presented) and its value was 0.011 $\mu g\,L^{-1}$. The *LOD* was estimated as the mean value of eight calibration curves (presented in Fig. 3) and its value was 0.022 $\mu g\,L^{-1}$.

Advantages of Currie's variation are the following: the strict student t parameter is employed and this is the best way to control the false positive and false negative rate (p and q) and the calculation is simplified by the introduction of blank subtraction factor that can take into account both cases when mean of blank replicates or intercept of calibration line is used as the value of blank. The critical value of the concentration domain and the *LOD* may be directly obtained using Eqs. (14) and (15), respectively. This approach includes the lowest standard deviation of the *LOD* values, which is 0.0026 $\mu g\,L^{-1}$.

Attention should be paid to the way the standard deviation of noise is estimated. The residual standard deviation as a characteristic of the calibration curve or standard deviation of the blank repetitions may be used as an estimate. Pooling of the two mentioned standard deviations is also possible. If the residual standard deviation is used it depends on the design of the calibration experiment: number of standards used, type of regression analysis (ordinary or weighted), and the distribution of the calibration points.

The seven replicates were tested by US EPA 200.8 method. The calculated *MDL* is 0.04 $\mu g\,L^{-1}$. It is very close to the value obtained by variation 2. According to the given definition the

Table 4
LOD values calculated according to the traditional approaches. The data of ICP-MS measurements accomplished in our lab was the base of calculation ($N=26$).

#	Ref.	Approach		Signal (counts/s)		Concentration ($\mu g\,L^{-1}$)	
				Critical values y_C	*LOD* y_D	Critical values x_C	*LOD* x_D
1.	[12,13]	Blank correction factor is implemented as most adequate.	$x_C=t_p s_0\eta^{1/2}/b$ $y_C=t_p s_0\eta^{1/2}$ $x_D=(t_p+t_q)s_0\eta^{1/2}/b$ $y_D=(t_p+t_q)s_0\eta^{1/2}$	95	190	0.011	0.022
2.	[23]	Respecting Student *t* parameter	$y_C=y_B+ts_0$ $y_D=y_C+ts_0$	247	366	0.025	0.039
3.	[10,14]	Intercept is used as value of blank.	$y_C=a+ts_0$ $y_D=y_C+ts_0$	151	269	0.014	0.028
4.	[6]	*k* is equal 3	$y_D=y_B+3s_0$	–	337	–	0.035
5.	[6]	Intercept is used as value of blank and *k* is equal 3.	$y_D=a+3s_0$	–	241	–	0.025
6.	[9,14]	Instead of standard deviation of blank, s_0, residual standard deviation, $s_{y/x}$, is used.	$y_C=a+ts_{y/x}$ $y_D=a+2ts_{y/x}$	525	1017	0.058	0.116
7.	[9,14]	nstead of standard deviation of blank, s_0, residual standard deviation, $s_{y/x}$, is used and *k* is equal 3.	$y_D=a+3s_{y/x}$	–	765	–	0.086
8.	[9,14]	Instead of standard deviation of blank, s_0, standard deviation of intercept s_a is used and *k* is equal 3.	$y_D=a+3s_a$	–	430	–	0.047

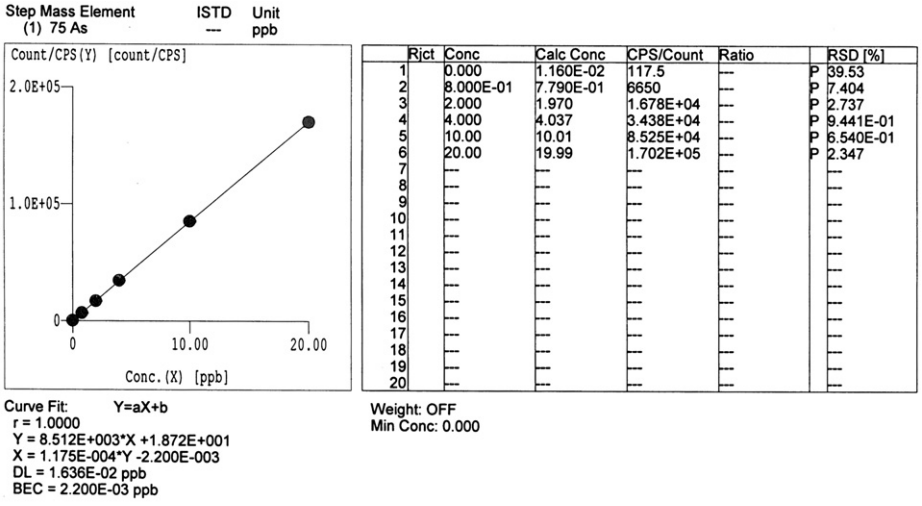

Fig. 2. One representative calibration curve obtained directly from ICP-MS. The original data provided by the instrument software correspond with the data presented as an example in Table 4.

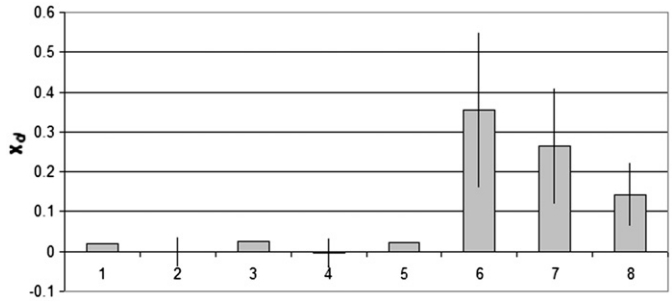

Fig. 3. Mean *LOD* values based on the eight calibrations and their standard deviation for each 7 variants of the traditional method.

MDL is supposed to be the analog of the *LOD*. Analyzing the Eq. (12) it can be concluded that actually the *MDL* corresponds to critical value, and that reporting concentration less than *MDL* in the absence of detection is not appropriate. If the *MDL* is taken as the *LOD* the probability of false negative conclusion is as high as 50%.

3.2. Method based on prediction intervals, (Hubaux–Vos) [25]

The data obtained from eight times repeated calibration standard measurements were subjected to Hubaux–Vos and the ISO calculation methodology. In Fig. 4 estimation of *LOD*

according to Hubaux–Vos model is shown, using ordinary linear least square regression calibration line.

The *LOD* values calculated according to the Hubaux–Vos and ISO approaches are presented in Table 5.

The method was applied to calibration lines obtained in two ways: ordinary least square (OLS) regression and weighted least square (WLS) regression. After fitting the data to the calibration curve, the limits were calculated, according to Eqs. (12)–(14). Since the obtained values calculated by this method (OLS type) were extremely high, the *LOD* was more than 1 μg L^{-1}, which is not realistic. It is considered that this approach is not applicable and adequate for this type of measurements. Regardless of this fact, the same data were tested by the WLS approach. Totally different values for the *LOD* were obtained, lower and more realistic, in the vicinity of 0.1 μg L^{-1}. However, the drawback of the method is obvious because which approach should be chosen is not known in advance. The differences between the results obtained by the two approaches can be ascribed to the heteroscedasticity of the data (the standard deviation is not the same for all concentration levels). A detailed analysis of heterosedasticity lies outside the scope of this paper. However, we tested the data according to the Bartllet test for homogeneity of variances [29]. Null hypothesis that variances are the same at all concentrations was rejected at the significance level of 0.01. When the same data was tested with ISO (OLS and WLS regression methods), the results were expectable and in harmony with the Hubaux–Vos method. Recently, Voigtman [13,15–20] performed mathematical modeling of analytical properties, among them the *LOD*, and he proved that this prediction method was not correct from the beginning because it considers predictions that are biased. ISO method was also explained to be intrinsically wrong [14], since the improper combination of theoretical parameters and experimental equations was used through its derivation.

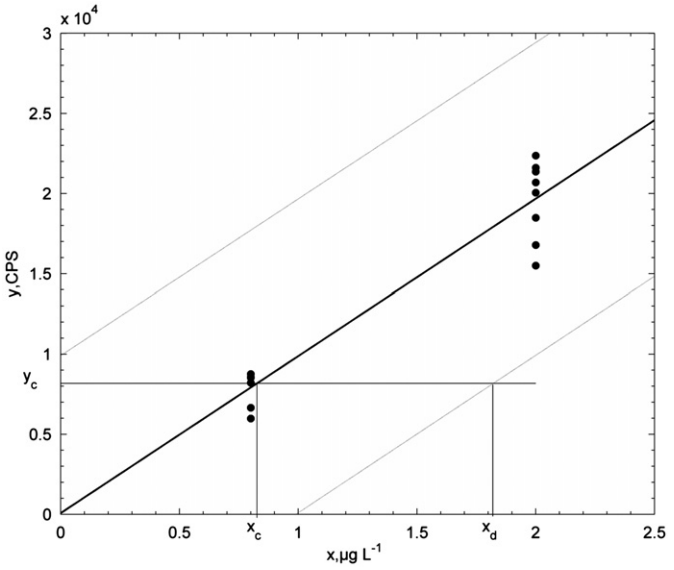

Fig. 4. Graphical estimation of *LOD* according to the Hubaux–Vos method (real data of the ICP-MS measurements were applied).The middle line is the calibration line which include all eight replicates of the calibration standard measurements when an ordinary least square regression was performed. Dashed lines present the prediction function.

Intuitively, prediction intervals are inconsistent with the assumption of homoscedasticity because their width varies with concentration [13]. Given that the noise is homoscedastic the Currie's detection question can be formulated this way: "what value of y_D is such that the distribution of blank, shifted up so that is centered at $y=y_D$, will have no more than 100q % of its lower tail area below the line $y=y_C$". This is shown in Fig. 1. The very hearth of the problem is that Hubaux–Vos gave the answer to another question: "what value of y_D is such that the OLS prediction interval at $x=x_D$ will have no more than 100q% of its lower tail area below the line at $y=y_C$". As it was mentioned, Voigtman [13,15–20] performed extensive Monte Carlo simulations. Analyzing tens of millions independent data sets for various types of chemical measurement systems they have shown that experimental *LOD*s (both in signal and concentration domain) obtained by Hubaux–Vos and ISO methods are significantly negatively biased. As a result, rates of false negative errors are higher than specified, and out of control. Through experimental comparison, bias of these two methods (in case of proper treatment of calibration data) cannot be properly tested, due to limited, small number of measurements obtainable in practice. Voigtman's conclusion derived from numerical experiments should be respected in the first place because there is no need to use complicated methodologies that are wrong in their essence. In our future work we will follow his way of modeling the *LOD* for ICP-MS and IC measurements.

The third approach, the signal-to-noise ratio, was not experimentally tested because this review considers ICP-MS measurements.

4. Conclusion

In this paper, approaches for the estimation of analytical limits from the analyst's viewpoint were addressed and reviewed. The scope of the experimentalists is to provide reliable analytical results. By validation protocols, the analytical properties of an analytical method should be specified. The focus was on the *LOD* of ICP-MS determination of arsenic. The current approaches for the calculation of the *LOD* are summarized and discussed. In reviewed papers dealing with trace analysis and *LOD*, two points of interest were distinguished, one related to the analysts who are the final users of commercial models or standard protocols for *LOD* estimation and the other relates to the theoreticians who develop new models and new approaches. A reasonable balance and compromise between these two groups is necessary to create a new protocol for *LOD* determination, a protocol which will harmonize complex and respectable mathematical models with the simple demands of analysts to have clear and reliable procedures.

Table 5
LOD values calculated according to the prediction interval approaches. The data of ICP-MS measurements accomplished in our lab was the base of calculation (eight times repeated) calibrations.

Author/s	Ref	Approach	Equation	*OLS*		*WLS*	
				Critical values x_C	*LOD* x_D	Critical values x_C	*LOD* x_D
Hubaux–Vos	[25]	Prediction intervals	Eqs. (24) and (25)	0.826	1.81	0.018	0.069
ISO	[26]	Non-centrality parameter of the non-central t distribution	Eq. (26)	–	1.66	–	0.127

The true *LOD* can never be experimentally obtained, only estimated. A unique value for the *LOD* calculated by a certain model cannot be directly compared to those calculated by other models. *LOD*s calculated using various estimation models are random varieties characterized by probability density functions, the shape of which depends on the number of standards used for the calibration, the specific values of the standards, the number of replicates per standard and the number of blank replicates. Each method for *LOD* estimation has its advantages and demands experimental trial. The *LOD* is not a permanent and constant value. It needs to be rechecked and adopted for each specific case. The experimental assessment accomplished in this work beside the theoretical analysis indicates inconsistency in all approaches. In all models some assumptions are introduced that are not always adequate and realistic.

The original traditional IUPAC model for *LOD* estimation was adopted and simplified in a large number of studies to the rule, $LOD = 3s$, without consideration of the type of blank subtraction and number of degrees of freedom. The calculated *LOD* is at the same time critical value. However, the original concept of the Currie's method provides more intrinsic elements of analysis, which means that the parameter $\eta^{1/2}$ is introduced in order to account for correction due to blank subtraction and the number of repetitions of measurements of specimen under test. For our experimental results, the most appropriate values were obtained according to Currie's variation of the traditional method; the critical value was 0.011 $\mu g\, L^{-1}$ and *LOD* was 0.022 $\mu g\, L^{-1}$.

Acknowledgments

The authors are grateful for the financial support from the Research Fund of Serbia, University of Belgrade, and Projects nos. III43009, ON172007 and TR 37010.

References

[1] N.B. Issa, V.N. Rajaković-Ognjanović, B.M. Jovanović, Lj.V. Rajaković, Anal. Chim. Acta 673 (2010) 185–193.
[2] N.B. Issa, V.N. Rajaković-Ognjanović, A. Marinković, Lj.V. Rajaković, Anal. Chim. Acta 706 (2011) 191–198.
[3] N.B. Issa, A.D. Marinković, Lj.V. Rajaković, J. Serbian Chem. Soc. 77 (6) (2012) 775–788.
[4] B. Jovanović, V.L. Vukašinović-Pešić, Lj.V. Rajaković, Water Environ. Res. 83 (2011) 498–506.
[5] B.M. Jovanović, V.L. Vukašinović-Pešić, Đ.N Veljović, Lj.V Rajaković, J. Serb. Chem. Soc. 76 (2011) 1–17.
[6] I. Komorowicz, D. Baralkiewicz, Talanta 84 (2011) 247–261.
[7] S.N. Ronkart, V. Laurent, P. Carbonnelle, N. Mabon, A. Copin, J.P. Barthélemy, Chemosphere 66 (2007) 738–745.
[8] T. Nakazato, H. Tao, T. Taniguchi, K. Isshiki, Talanta 58 (2002) 121–132.
[9] ICH Q2(R1), Validation of Analytical Procedures: Text and Methodology, International Conference on Harmonisation of Technical Reguirements for Registration of Pharmaceuticals for Human Uses, 2005.
[10] L.A. Currie, Pure Appl. Chem. 67 (1995) 1699–1723.
[11] L.A. Currie, Anal. Chem. 40 (1968) 586–593.
[12] L.A. Currie, Chemom. Intell. Lab. Syst. 37 (1997) 151–181.
[13] E. Voigtman, Spectrochim. Acta B 63 (2008) 115–128.
[14] J.C. Miller, J.N. Miller, Statistics and Chemometrics for Analytical Chemistry, 5th ed., Pearson Prentice Hall, Edinburgh, 2005.
[15] E. Voigtman, Spectrochim. Acta B 63 (2008) 129–141.
[16] E. Voigtman, Spectrochim. Acta B 63 (2008) 142–153.
[17] E. Voigtman, Spectrochim. Acta B 63 (2008) 154–165.
[18] E. Voigtman, K.T. Abraham, Spectrochim. Acta B 66 (2011) 822–827.
[19] E. Voigtman, K.T. Abraham, Spectrochim. Acta B 66 (2011) 828–833.
[20] E. Voigtman, K.T. Abraham, Spectrochim. Acta B 66 (2011) 105–113.
[21] M.Kondratovich Linnet, Clin. Chem. (2004) 732–740.
[22] I. Lavagnini, D. Badocco, P. Pastore, F. Magno, Talanta 87 (2011) 180–188.
[23] IUPAC, Compendium of Chemical Terminology—The Gold book, 1997, 〈http://goldbook.iupac.org/〉.
[24] US EPA (United States Environmental Protection Agency), Method 200.8., USA, 1994.
[25] A. Hubaux, G. Vos, Anal. Chem. 42 (1970) 849–855.
[26] International Organization for Standardization (ISO), ISO 11843-2 Capability of Detection—Part 2: Methodology in the Linear Calibration Case, 2000.
[27] United States Pharmacopeia 30, National Formulary 25, General Chapter 1225, Validation of compendial methods, The United States Pharmacopeial Convention, Rockville, Maryland, USA, 2007.
[28] European Pharmacopeia 6.2, Council of Europe, Strasbourg, France, 2008.
[29] D.L. Massart, B.G.M. Vandeginste, S.N. Deming, Y Michotte, L Kaufman, Chemometrics: A Textbook, Elsevier, Amsterdam, 1988.