Linear calibrations in chromatography: the incorrect use of ordinary least squares for determinations at low levels, and the need to redefine the limit of quantification with this regression model

Juan M. Sanchez

Chemistry Dept., University of Girona, Aurèlia Capmany 69, 17003-Girona (Spain)

Running Title: Incorrect use of ordinary least squares at low concentration levels

Correspondence to:

Juan M. Sanchez, PhD

Chemistry Department, Science Faculty

University of Girona

Aurèlia Capmany, 69

17003-Girona (Spain)

Email: juanma.sanchez@udg.edu

Phone: (+34) 636569984

Orcid ID: 0000-0002-4139-7273

Non-standard abbreviations:

CV	coefficient of variation
LOF	lack-of-fit
OLS	ordinary least squares
RSD	relative standard deviation
RSD_r	relative standard deviation under repeatability conditions
$PRSD_R$	predicted relative standard deviation of reproducibility
s_{bl}	standard deviation of the blank
TI	tolerance interval
WLS	weighted least squares

Keywords

Received: 29/01/2020; Revised: 27/03/2020; Accepted: 30/03/2020
This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/jssc.202000094.

Accuracy profiles, Analytical calibrations, Quantification limit, Least squares.

Abstract

Ordinary least squares is widely applied as the standard regression method for analytical calibrations, and it is usually accepted that this regression method can be used for quantification starting at the limit of quantification. However, it requires calibration being homoscedastic and this is not common. Different calibrations have been evaluated to assess whether ordinary least squares is adequate to quantify estimates at low levels. All calibrations evaluated were linear and heteroscedastic. Despite acceptable values for precision at limit of quantification levels were obtained, ordinary least squares fitting resulted in significant and unacceptable bias at low levels. When weighted least squares regression was applied, bias at low levels were solved and accurate estimates were obtained. With heteroscedastic calibrations, limit values determined by conventional methods are only appropriate if weighted least squares is used. A “practical limit of quantification” can be determined with ordinary least squares in heteroscedastic calibrations, which should be fixed at a minimum of 20 times the value calculated with conventional methods. Biases obtained above this “practical limit” were acceptable applying ordinary least squares and no significant differences were obtained between the estimates measured using weighted and ordinary least squares when analyzing real-world samples.

1. Introduction

Analytical methods used in laboratories must be evaluated and tested to ensure they produce valid results that are suitable for their intended purpose. It should be taken into account that the main objective in quantitative analysis is to provide an estimate of the content of an analyte with acceptable accuracy (i.e., trueness and precision [1]). Therefore, the most important factor should be the quality of the inverse predictions (i.e., back-calculated values) rather than the quality of fit [2,3].

Analytical calibrations are required to find a function that can describe the relationship between the instrumental response and the concentration of the target analyte. It is common to assume that the simplest model adequately describing this concentration-response relationship should be used [4]. In chemical and biological analysis many instruments show linear detector responses over several orders of magnitude; therefore, linear regression models, mainly ordinary least squares (OLS), are extensively used in practical applications; which are more intuitive and easier to fit than non-linear ones, and estimators are simpler to determine [5,6]. Despite mathematical simplicity being desirable, this is not a significant limitation as today's software programs can fit complex models without specialized knowledge, eliminating the need for a mathematical background to calculate parameters.

In any case, it must be understood that regression models are based on the fulfillment of certain preliminary conditions, which were adopted in formulating the model, and the failure to meet some of these conditions can lead to significant biases and imprecisions in the concentration estimates [7,8]. The most extensive fitting model used in laboratory calibrations is linear OLS, which requires variance of the dependent variable to be constant

at all values of the independent variable (homoscedasticity). However, in practice analytical and biological methods yield non-constant variance over the working range [7,9-19], unless this range is particularly narrow (usually up to one order of magnitude). Previous studies [16,19-29] have already demonstrated that the use of OLS with heteroscedastic data may result in significant bias and underestimation of the precision at concentrations that are close to the limit of detection (LOD), due to the overestimation of the high concentration standards. Despite this, it is still very common to see researchers applying OLS regression in calibrations without evaluating whether the calibration presents homoscedasticity or heteroscedasticity. Unfortunately, "a great number of people using least squares have just enough training to be able to apply it, but not enough training to see why it often shouldn't be applied" [30].

From a quantitative point of view, a basic parameter to be determined in the validation of an analytical method is the limit of quantification (LOQ), which has been defined as the lowest amount or concentration of an analyte at which performance is acceptable for a typical application [31], or that can be estimated with acceptable reliability [32] or precision [33]. Once this parameter has been determined, it is always assumed that the regression model used can be applied starting from the LOQ to obtain accurate estimates. Unfortunately, the common methodologies used for the determination of LOQ values are only based on analyte response at a single concentration, making it inconsistent in situations of non-constant variance [11]. Usually the determination of the LOQ is simply based on an extrapolation of the IUPAC limit of detection or determination (LOD), which is only based on instrumental repeatability [4,33,34]. Sometimes the precision chosen at the LOQ level is defined as 10% relative standard deviation (RSD), as was suggested by Currie [35]. In other cases the LOQ is taken as being a fixed multiple of the LOD or the

concentration that produces a signal of *k* times (usually *k*=10) the standard deviation of the blank (s_{bl}) [31,36]. The most significant limitation of these approaches is that they are only based on the characterization of the target parameter without an assessment of uncertainty and there is a lack of bias accountability for this limit [4]. Therefore, the risk of accepting an unsuitable assay and rejecting a suitable assay is unknown and uncontrolled [14,37]. Moreover, when some statistical test is applied, it only computes type I error (i.e., probability of false positives), which involves that methods giving imprecise results can be more easily validated than more precise ones [37].

Some recent validation guidelines require the evaluation of the trueness in the determination of the LOQ, and indicate that it is the lowest amount of an analyte that can be determined with acceptable precision and trueness [4,38]. Moreover, different studies [14,23,24] have found that the selection of the regression model may have a significant effect on the estimation of LOQ values when trueness and precision are taken into account. Scientists should be aware that a linear relationship between the dependent variable and the concentration does not guarantee method trueness when a bias is present [39], and this may easily occur at low concentration levels with heteroscedastic calibrations [21].

In 1997 the Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) published a guide on the validation of chromatographic bio-analytical methods proposing a new validation approach using “accuracy profiles” [39]. This approach considers that by taking into account tolerance intervals (i.e., intervals where the expected proportion of future results will fall with a predetermined β-expectation [40]), which contain both trueness and precision, better determination of LOQ values is achieved [2,3,14,16,28,33,37,39-45].

The aim of this study is to compare the two most common linear regression fittings, OLS and weighted least squares (WLS), and to assess whether the LOQ values determined by conventional methods based on precision at blank level genuinely give a limit where the precision and the trueness of estimates are adequate for quantification purposes with both regression models. The accuracy profile methodology has been used for assessing the accuracy of the results obtained.

2. Materials and Methods

Fifteen experimental analytical calibrations using different chromatographic methods (GC-FID, HPLC-UV and GC-MS) were evaluated. In all cases, a minimum of six calibration standards evenly distributed along the study range were used. The response at each calibration level was calculated as the mean obtained from at least seven series of independent standard replicates. Each series was prepared from new stock solutions by different analysts in different days. This procedure allowed (i) the assessment of the existence of heteroscedasticity in the calibrations, (ii) the evaluation of the linearity of the regression models, and (iii) the calculation of inter-series precision.

The standard deviation at blank level (s_{bl}) was determined from the analysis of a minimum of seven independent replicates of spiked blanks prepared at concentrations up to five times the LOD of each method. LOD and LOQ values based on the precision at blank level were calculated as $3.3s_{bl}$ and $10s_{bl}$, respectively, as proposed by different validation guidelines [31,35,36].

SPSS 15.0 for Windows (SPSS Inc., Chicago, IL, USA) was used for the statistical and regression calculations both with OLS and WLS. A difference was considered as significant

when $p<0.05$. The weighting factor applied in WLS regressions was determined as the inverse of the experimental variance ($w_i=1/s_i^2$). However, it was also evaluated the use of other factors, such as $w_i=1/x_i^2$ and $w_i=1/y_i^2$ (where x_i corresponds to the concentration of the i-standard and y_i to its analytical response).

2.1. Accuracy profile approach

As indicated by Feinberg [37], the basic idea behind this concept is to translate the fitness-for-purpose objective into the acceptability criterion (λ):

$$|Z - X| < \lambda \qquad \text{(eq. 1)}$$

where X is the analytical result, Z the true value and the limit λ is not arbitrary and depends on the goal of the analytical procedure.

This methodology is based on the calculation of a tolerance interval at each concentration level (TI_j). Briefly, it requires the determination of the bias for the back-calculated concentrations at each concentration level and the calculation of a β expectation tolerance interval (eq. 2). Full equations and detailed description of the methodology are described in the literature [2,16,39].

$$TI_j = Bias_j(\%) \pm Q_t\left(\vartheta, \frac{1+\beta}{2}\right)\sqrt{1+\frac{1}{p\cdot n\cdot B_j^2}}\cdot CV_{IP,j} \qquad \text{(eq. 2)}$$

To draw the accuracy profile plot, the relative error (%) for the back-calculated concentration at each concentration level is plotted against the corresponding standard concentration, together with the corresponding upper and lower tolerance limits. Two lines are drawn, one connecting the lower tolerance limits obtained at each concentration level and another connecting the upper tolerance limits, which allow showing specified acceptance limits in the graph.

When the whole accuracy profile, including the tolerance intervals, is within in the acceptance limits, the analytical method is expected to provide accurate results for its intended purpose. If some point of the profile steps outside the limits, the method should not be considered for that concentration level. In the present study, the acceptance limits at LOQ level have been set at ±20%, which is considered acceptable by the US-FDA for bioanalytical methods [4].

3. Results and Discussion

3.1. Evaluation of linearity and heteroscedasticity of calibrations

The linearity of all calibrations in the working ranges evaluated was assessed graphically by checking the residual plots, and statistically by applying the lack-of-fit (LOF) [46] and Mandel's [47] tests. The distribution of residuals and the *p*-values obtained ($p>0.10$) confirmed that the use of linear functions was satisfactory in all the calibrations evaluated.

Levene's test for homogeneity of variances [48,49] was applied to assess the distribution of variances in the calibration ranges. The results obtained showed that all the calibrations were heteroscedastic ($p<0.01$), including those where the working range only covered up to

one order of magnitude. Given this, the most adequate linear regression model should be WLS rather than OLS.

3.2. Conventional determination of LOD and LOQ

Table 1 shows the LOD and LOQ values obtained for the methods evaluated by measuring the standard deviation of spiked blanks and by applying the $3.3s_{bl}$ and $10s_{bl}$ criteria [31,35,36]. The calibrations evaluated were divided into three groups: a first group with an initial standard concentration above the calculated LOQ, a second group starting at a level equivalent to the LOQ, and a third group where the first standard was between the LOD and the LOQ. Table 1 also includes the bias for the back-calculated concentration (relative error, %) and the inter-series precision (RSD, %) for the first standard in each calibration (this standard yielded the largest RSD value in all calibrations). As all methods present heteroscedasticity, bias values were determined for both OLS and WLS.

3.3. Evaluation of precision and bias at low levels

The US-FDA suggests that a 15% precision between runs for all calibrators should be achieved except at LOQ level, which it sets at 20% [4]. The AOAC [50] recommends that for those methods where no collaborative studies have set the limits for reproducibility, acceptable values should be determined from the HorRat ratio:

$$HorRat = {RSD_r}/{PRSD_R} \qquad \text{(eq. 3)}$$

where RSD_r is the RSD obtained under repeatability conditions and $PRSD_R$ is the predicted relative standard deviation of reproducibility ($PSRD_R = 2C^{-0.15}$). The recommended limits of acceptance are set between 0.3 and 1.3.

The inter-series RSD obtained in the present study for the first standard (close to the LOQ level) were <15%, and <8% for the other standards, and HorRat ratios at LOQ level were between 0.3 and 0.9 (Table 1). These results indicate that all calibrations gave acceptable values for precision at all levels, which means that all assessed methods would pass acceptance criteria defined in many conventional validation guidelines, where only precision at LOQ level is taken into account.

When the percentages of relative error for back-calculated concentrations (i.e., bias) were calculated, it was observed that OLS regression only yielded acceptable bias (±20% at LOQ level [4]) in the first group of calibrations, when the first standard concentration was above the LOQ (Table 1). For the other calibrations, the bias obtained with OLS was always excessive at the LOQ level, reaching values ~~of~~ >100% in some cases. When WLS regression functions were applied the bias was always ±15%. This confirms that OLS regression fails to obtain accurate estimates (taking into account trueness and precision) close to the LOQ, whereas WLS regression does yield good estimates. Similar results have been obtained in other studies [21,23,24].

3.4. Accuracy profile plots

The accuracy profile plots obtained for the calibrations evaluated can be seen in Figure 1 and Supplementary Materials. In all calibrations where the first standard was clearly above

the LOQ, the accuracy profiles for OLS and WLS regression models gave results that were equivalent (Figures 1a and 1b) and the tolerance intervals were inside the acceptance limits for the whole calibration range. This indicates that OLS can be considered a correct regression method from a practical point of view when measuring at levels well above the LOQ. It is already known that, in linear heteroscedastic calibrations, OLS yield poorly and inaccurate assessment of the intercept but does not introduce significant changes in the slope [19,21], which means that discrepancies in estimates obtained between OLS and WLS can be found at the lowest levels of concentrations and can be very important.

When the first standard was at a level equivalent to the LOQ (Figures 1c and 1d), the accuracy profile for the OLS model clearly showed that this regression gave excessive bias and incorrect results at LOQ level, whereas the WLS model gave adequate tolerance limits in the whole range. Finally, when the first standard was set at a concentration between LOD and LOQ (Figures 1e and 1f), despite WLS regressions always giving a bias inside a ±20% limit, tolerance intervals were beyond the acceptance limits.

3.5. Weighted least squares

It is clear that WLS should be the golden regression method for linear analytical and biological calibrations in laboratories due to the existence of heteroscedasticity. However, despite the results presented in this study do not really introduce a novelty, and different studies [16,19-29] have already confirmed this fact, OLS is still the routine linear regression model applied in the majority of laboratories, also for trace analysis and quantification at low levels, near LOQ, where OLS always fails to obtain accurate results. The revision of some recent volumes published in three high impact journals (first quartile, with impact factors >3.8) have shown that in a total of 31 articles the analytical performance of the proposed

calibration method was evaluated and was found to be linear. From these, 29 (93.5%) applied OLS, without evaluation of heteroscedasticity, and only 2 (6.5%) WLS. In many cases, the authors used the proposed OLS regression for quantification at trace levels, with some samples yielding results at levels close to the reported LOQs, which were determined by conventional ways. When different routine laboratories were asked about the type of calibration they use, >90% answered that they always apply OLS for linear methods. Here, a question arises: why are scientists so reticent to change their linear calibration calculations despite many studies have demonstrated the significant limitations of OLS at low level? A possible answer is that in many university sciences degrees OLS is the only linear calibration model that is taught. A second explanation is that many researchers indicate that one of the main drawbacks for applying WLS is the need to perform a large number of replicate standards at each level to obtain the weighting factors ($w_i=1/s_i^2$), which is not practical in the daily routine of many laboratories. However, it has been demonstrated that standard deviation is proportional to the concentration [18,22,25,26,51]. For this reason, different simplified experimental approaches have been proposed to avoid the requirement of replicate measurements at each calibration level [18,22,28,51,52], using empirical weighting factors, such as $1/x_i^{1/2}$, $1/x_i$, $1/x_i^2$, $1/y_i^{1/2}$, $1/y_i$, and $1/y_i^2$; from which $1/x_i^2$ and $1/y_i^2$ usually yield the best results. In this study, the use of WLS regression with $1/x_i^2$ and $1/y_i^2$ as weighting factors was evaluated. Twenty-five independent replicates of a sample lot were measured in five different days, with a new calibration each day measuring three independent replicates for each standard. Each sample was analyzed directly and with a dilution yielding a concentration of around 5·LOQ. Concentrations of samples from diluted and undiluted solutions were calculated by OLS and WLS (using $1/s_i^2$, $1/x_i^2$, and $1/y_i^2$ as weighting factors). As can be seen in Figure 2, a statistical significant difference was

obtained between the results (ANOVA test, $p=4\cdot10^{-23}$). The HSD Tukey post-hoc test confirmed that only OLS regression from diluted samples yielded a significant bias (-17% bias), with all other calibration methods giving equivalent results ($p=0.701$). This confirms that, in routine analysis, WLS can be applied without the need for replicate measurements at each calibration level using empirical factors such as $1/x_i^2$ and $1/y_i^2$ to obtain non-biased results at low levels.

3.6. Estimation of a "practical LOQ" using OLS

The evaluation of the accuracy profile plots shows that the bias and tolerance limits obtained at levels ≥5 times the LOQ were always acceptable for both OLS and WLS models (Figure 1 and Table 1). This suggests that the use of OLS regression does not introduce any significant error, from a practical point of view, with heteroscedastic calibrations provided that the first standard is set at least 5 times above the LOQ and no quantification measurements are done below this point. From these plots, a "practical LOQ" could be defined as at least 5·LOQ to obtain non-biased results with OLS regression.

It was decided to assess whether the use of a minimum concentration of 5·LOQ, determined from accuracy plots, can really be applied in the analysis of real-world samples using OLS. Therefore, a set of experiments were performed to study the bias obtained with OLS when measuring at low levels (between LOQ and 20·LOQ). First, three different batches of different commercial teas were evaluated (one green tea and two pu-erh teas), and independent replicate samples of each batch were analyzed for theobromine content. The solution obtained after the extraction of the target analyte was measured directly and after being diluted at a level close to the LOQ. Each replicate sample was analyzed on different

days over an eight-week period, with new analytical calibrations being made each day. The content of the target analyte in tea leaves (in mg g^{-1}) was calculated from each replicate (both from the diluted and undiluted solutions analyzed) applying OLS and WLS regressions (Figure 3). In a first group of samples (Figure 3a, n=13, pu-erh tea), the solution was diluted until a level that was around 10·LOQ (a level where accuracy profiles plots did not yield differences between OLS and WLS); a second group (Figure 3b, n=12, green tea) was diluted until 5·LOQ; and a last group (n=28, pu-erh tea) was diluted to 0.5·LOQ. In all cases, the signals measured for the diluted samples agreed with the dilution factor applied, as expected for a linear response function. Despite accuracy profile plots suggested that no significant differences should be expected for the first group of samples (10·LOQ), the evaluation of the results yielded statistical significant differences (ANOVA test, $p<0.02$) for all the contents determined with OLS regression from the diluted solutions. Only for those samples in which the dilution was performed at 10·LOQ can the results be considered equivalents at a 99% significance ($p=0.012$, Figure 3a). On the other hand, the results obtained from the undiluted samples, both with OLS and WLS, and the diluted sample applying WLS yielded equivalent contents in all cases (Tukey HSD post-hoc test, $p=0.982$ with dilution to 10·LOQ, Figure 2a; $p=0.994$ with dilution to 5·LOQ, Figure 3b; and $p=0.409$ with dilution to 0.5·LOQ, Figure 3c). These results also confirmed that the percentage of bias obtained applying OLS regression increased when the content of the solution measured decreased (-6.2% bias at 10·LOQ, -12.1% at 5·LOQ and -32.5% at 0.5·LOQ). These results agree with those obtained with the accuracy profile plots as the percentage of bias obtained at concentration levels between 5-10 times the LOQ were inside the prefixed acceptance limits at this level (±20%); however, the analysis of a large number of replicate samples showed that a systematic bias was still found applying OLS.

In a second set of experiments, 24 different tea samples were analyzed with the help of final year Biotechnology degree students. Each student analyzed one sample with his/her own calibration. For each sample, the final solution was measured directly and after being diluted 10 times because preliminary tests of this type of samples showed that this dilution factor yielded solutions between LOQ and 20·LOQ. As in previous experiments, the analyte content in tea leaves was determined from both the diluted and the undiluted solutions using the two regression methods. The paired t-test was applied to assess whether the results obtained from the diluted and undiluted samples yielded equivalent results. No significant differences were obtained between the contents determined by OLS and WLS from the undiluted solutions (mean difference=-0.002, s=0.006, p=0.104), as well as between the amounts found between the diluted and undiluted solutions with WLS (mean difference=0.006, s=0.020, p=0.158). However, a significant difference was obtained between the diluted and undiluted solutions with OLS (mean difference=-0.090, s=0.009, p<0.001). The bias obtained from the diluted solutions compared to the undiluted solutions was calculated using the two regression methods (Figure 4). It was found that the use of OLS regression yielded negative bias for all samples when the amount determined was <20·LOQ, which increased exponentially when the concentration detected from the diluted solutions was <10·LOQ (LOQ=0.06 mg L^{-1}, Figure 4a). In the case of WLS regression (Figure 4b), no systematic errors were observed (the results were randomly distributed around 0% bias) and although the percentage of bias tended to increase as the concentration of the diluted solution decreased, it was always inside ±10% until the LOQ level.

These results indicate that, despite the accuracy profile plots yielded acceptable accuracy using OLS regression for heteroscedastic calibrations when the first standard was set at a level of 5·LOQ, systematic errors were still found in estimations determined with OLS at

levels between 5-10 times the LOQ. Therefore, a more rigorous criteria should be implemented if OLS wants to be used and a minimum quantification value of around 20 times the LOQ should be taken into account as "practical LOQ" if acceptable accuracy, without appreciable systematic errors (<10%), wants to be obtained. This clearly limits the applicability of OLS for quantifications at low levels.

4. Concluding remarks

OLS is still nowadays the regression function most commonly used in linear analytical calibrations despite they are usually heteroscedastic, and it is known that OLS always fail to determine estimates at low levels in this conditions. In order to be able to use OLS with heteroscedastic calibrations it is required to change the way how LOQ is calculated. Conventional methods to determine LOQ only takes into account precision at blank level and always lead to limits that overestimate the real possibilities of OLS. The results obtained in the present study confirm that WLS regression should be the golden regression method applied with heteroscedastic calibrations for trace analysis because at low levels, close to the LOQ, is where OLS yields higher and systematic bias. If scientists want to apply OLS for linear calibrations, they have to take account that the percentage of bias in their estimations increases exponentially when the content measured approaches the LOQ level. In this sense, to obtain valid results without systematic errors (or <10%) using OLS, the conventional estimation of the LOQ needs to be increased by a minimum factor of 20. This results in a "practical LOQ" that ensures the validity of the estimates. However, it clearly limits the applicability of OLS regression for determinations at trace levels with

heteroscedastic calibrations. At this level, only WLS can be guaranteed to give true and precise results.

Acknowledgements

The author would like to thank A. Hughes for his help in the writing and proof reading of the manuscript. The help of the Biotechnology degree students of the Science Faculty at the University of Girona is also acknowledged.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

References

1. ISO 3534-1. Statistics – vocabulary and symbols. Part 1: Probability and general statistical terms. ISO, Geneva 1993
2. Hubert Ph., Nguyen-Huu J.J., Boulanger B., Chapuzet E., Cohen N., Compagnon P.A., Dewé W., Feinberg M., Lallier M., Laurentie M., Mercier N., Muzard G., Valat L., Rozet E., Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal-part II. *J. Pharm. Biomed. Anal.* 2007, *45*, 82-96
3. Bouabidi A., Rozet E., Fillet M., Ziemons E., Chapuzet E., Mertens B., Klinkenberg R., Ceccato A., Talbi M., Streel B., Bouklouze B., Hubert Ph., Critical analysis of several

analytical method validation strategies in the framework of the fit for purpose concept. *J. Chromatogr. A* 2010, *1217*, 3180-3192

4. US Food and Drug Administration (US-FDA). Bioanalytical Method Validation: Guidance for Industry. Silver Spring, MD 2018. https://www.fda.gov/downloads/drugs/guidances/ucm070107.pdf. Accessed 04 Apr 2019.
5. Yan X., Su X.G., Linear Regression Analysis: Theory and Computing. World Scientific, Singapore 2009. ISBN: 0789812834119
6. Burger S.V., Introduction to Machine Learning with R. O'Reilly Media Inc, Sebastopol 2018. ISBN: 9781491976432
7. de Souza S.V.C., Junqueira R.G., A procedure to assess linearity by ordinary least squares. *Anal. Chim. Acta* 2005, *552*, 25-35
8. Rozet E., Ceccato A., Hubert C., Ziemons E., Oprean R., Rudaz S., Boulanger B., Hubert P., Analysis of recent pharmaceutical regulatory documents on analytical method validation. *J. Chromatogr. A* 2007, *1158*, 111-125
9. Oppenheimer L., Capizzi T.P., Weppelman R.M., Mehta H., Determining the lowest limit of reliable assay measurement. *Anal. Chem.* 1983, *55*, 638-643
10. Vial J., Jardy A., Experimental comparison of the different approaches to estimate LOD and LOQ of an HPLC method. *Anal. Chem.* 1991, *71*, 2672-2677
11. Zorn M.E., Gibbons R.D., Sonzogni W.C., Weigthed least-squares approach to calculating limits of detection and quantification by modeling variability as a function of concentration. *Anal. Chem.* 1997, *69*, 3069-3075

12. Zorn M.E., Gibbons R.D., Sonzogni W.C., Evaluation of approximate methods for calculating the limit of detection and quantification. *Environ. Sci. Technol.* 1999, *33*, 2291-2295
13. Van der Berg R.A., Hoefsloot H.C.J., Westerhuis J.A., Smilde A.K., Van der Werf M.J., Centering, scaling, and transformation: improving the biological information content of metabolomics data. *BMC Genomics* 2006, *7*, 142
14. Hubert Ph., Nguyen-Huu J.J., Boulanger B., Chapuzet E., Chiap P., Cohen N., Compagnon P.A., Dewé W., Feinberg M., Lallier M., Laurentie M., Mercier N., Muzard G., Nivet C., Valat L., Rozet E., Harmonization of strategies for the validation of quantitative analytical procedures.: a SFSTP proposal-part II. *J. Pharm. Biomed. Anal.* 2007, *45*, 70-81
15. Desimoni E., Brunetti B., About estimating the limit of detection of heteroscedastic analytical systems. *Anal. Chim. Acta* 2009, *655*, 30-37
16. González O., Blanco M.E., Iriarte G., Bartolomé L., Maguregui M.I., Alonso R.M., Bioanalytical chromatographic method validation according to current regulations, with a special focus on the non-well defined parameters limit of quantification, robustness and matrix effects. *J. Chromatogr. A* 2014, *1353*, 10-27
17. Uhrovcik J., Strategy for determination of LOD and LOQ values – Some basic aspects. *Talanta* 2014, *119*, 178-180
18. Raposo F., Evaluation of analytical calibration based on least-square linear regression for instrumental techniques: a tutorial review. *TRAC Trends Anal. Chem.* 2016, *77*, 167-185
19. Sanchez J.M., Estimating detection limits in chromatography from calibration data: Ordinary least squares regression vs. weighted least squares. *Separations* 2018, *5*, 49

20. Meites L., Smith H.C., Kateman G., The effects of errors in measuring the independent variable in least-squares regression analysis. *Anal. Chim. Acta* 1984, *164*, 287-291
21. Ketkar S.N., Bzik T.J., Calibration of analytical instruments: impact of nonconstant variance in calibration data. *Anal. Chem.* 2000, *72*, 4762-4765
22. Almeida A.M., Castel-Branco M.M., Falcao A.C. Linear regression for calibration revisited: weighting schemes for bioanalytical methods. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci*. 2002, *774*, 215-222
23. Ceccato A., Klinkenberg R., Hubert Ph., Streel B., Sensitive determination of buprenorphine and its *N*-dealkylated metabolite norbuprenorphine in human plasma by liquid chromatography coupled to tandem mass spectrometry. *J. Pharm. Biomed. Anal.* 2003, *32*, 619-631
24. Streel B., Ceccato A., Klinkenberg R., Hubert Ph., Validation of a liquid chromatographic-tandem mass spectrometric method for the determination of loperamide in human plasma. *J. Chromatogr. B* 2005, *814,* 263-273
25. Tellinghuisen J., Weighted least-squares in calibration: what difference does it make? *Analyst* 2007, *132*, 536-543
26. Zeng Q.C., Zhang E., Dong H., Tellinghuisen J., Weighted least squares in calibration: estimating data variance functions in high-performance liquid chromatography. *J. Chromatogr. A* 2008, *1206*, 147-152
27. McDowall R.D., Are you validating methods for bioanalytical samples? *LCGC Eur.* 2009, *22*, 204-210
28. Dossou K.S.S., Edorh P.A., Chiap P., Chankvetadze P., Servais A.C., Fillet M., Crommen J., Determination of enantiomeric purity of S-amlopidine by chiral LC with emphasis on reversal of enantiomer elution order. *J. Sep. Sci.* 2011, *34*, 1772-1780

29. Zabell A.P.R., Lytle F.E., Julina R.K., A proposal to improve calibration and outlier detection in high-throughput mass spectrometry. *Clin. Mass Spectrom.* 2016, *2*, 25-33
30. ClockBackward Essays, Ordinary least squares linear regression: flaws, problems and pitfalls, 2009. https://www.clockbackward.com/2009/06/18/ordinary-least-squares-linear-regression-flaws-problems-and-pitfalls/. Accessed 14 Oct 2019
31. Eurachem Guide. The fitness for purpose of analytical methods – a laboratory guide to method validation and related topics. Magnusson B. and Örnemark U. (Eds.), 2nd ed., 2014. ISBN: 978-91-87461-59-0
32. AOAC International. AOAC guidelines for single laboratory validation of chemical methods for dietary supplements and botanicals. Association of Official Analytical Chemists, 2002
33. Thompson M., Ellison S.L.R., Wood R., Harmonized guidelines for single-laboratory validation of methods of analysis (IUPAC technical report). *Pure Appl. Chem.* 2002, *74*, 835-855
34. Mermet J.M., Limit of quantitation in atomic spectrometry: an unambiguous concept? *Spectrochim. Acta Part B*. 2008, *63*, 166-182
35. Currie L.A., Limit for qualitative detection and quantitative determination: application to radiochemistry. *Anal. Chem.* 1968, *40*, 586-593
36. Curie L.A., Nomenclature in evaluation of analytical methods including detection and quantification capabilities. *Pure Appl. Chem.* 1995, *67*, 1699-1723
37. Feinberg M., Validation of analytical methods based on accuracy profiles. *J. Chromatogr. A* 2007, *1158*, 174-183
38. International Conference on Harmonisation (ICH). ICH harmonised tripartite guideline - validation of analytical procedures: text and methodology Q2(R1), step 4 version, 2005.

https://www.ich.org/products/guidelines/quality/article/quality-guidelines.html. Accessed 04 Apr 2019

39. Hubert Ph., Chiap P., Crommen J., Boulanger B., Chapuzet E., Mercier N., Bervoas-Martin S., Chevalier P., Grandjean D., Lagorce P., Lallier M., Laparra M.C., Laurentie M., Nivet J.C., The SFSTP guide on the validation of chromatographic methods for drug bioanalysis: from the Washington Conference to the laboratory. *Anal. Chim. Acta* 1999, *391*, 135-148
40. Hubert Ph., Nguyen-Huu J.J., Boulanger B., Chapuzet E., Chiap P., Cohen N., Compagnon P.A., Dewé W., Feinberg M., Lallier M., Laurentie M., Mercier N., Muzard G., Nivet C., Valat L., Harmonization of strategies for the validation of quantitative analytical procedures: a SFSTP proposal-part I. *J. Pharm. Biomed. Anal.* 2004, *36*, 579-586
41. Feinberg M., Boulanger B., Dewé W., Hubert P., New advances in method validation and measurement uncertainty aimed at improving the quality of chemical data. *Anal. Bioanal. Chem.* 2004, *380*, 502-514
42. Gustavo González A., Ángeles Herrador M., Accuracy profiles from uncertainty measurements. *Talanta* 2006, *70*, 896-901
43. Ciccheti E., Chaintreau A., Quantitation of the main constituents of vanilla by reverse phase HPLC and ultra-high-pressure-liquid-chromatography with UV detection: method validation and performance comparison. *J. Sep. Sci.* 2009, *32*, 3043-3052
44. Dharuman J.G., Vasudevan M., Reliability-targeted HPLC-UV method validation – A protocol enrichment perspective. *J. Sep. Sci.* 2014, *37*, 228-236
45. Sghaier L., Cordella C.B.Y., Rutledge D.N., Watiez M., Breton S., Sassiat P., Thiebaut D., Vial J., Validation of a headspace trap gas chromatography and mass spectrometry

method for the quantitative analysis of volatile compounds from degraded rapeseed oil. *J. Sep. Sci.* 2016, *39*, 1675-1683

46. Massart D.L., Vandeginste B.G.M., Buydens L.M.C., De Jong S., Lewi P.H., Smeyers-Verbeke J., Handbook of chemometrics and qualimetrics: part A. Elsevier, Amsterdam 1997
47. Mandel J., The statistical analysis of experimental data. Dover Pub. Inc., New York 1964
48. Levene H., Robust test for equality of variances, in: Olkin I., Ghurye S.G., Hoedffding W., Madow W.G., Mann H.B. (Eds.), Contributions to probability and statistics. Stanford University Press, Palo Alto 1960, pp. 278-292
49. Brown M.B., Forsythe A.B., Robust test for the equality of variances. *J. Am. Stat. Assoc.* 1974, *69*, 364-367
50. AOAC International. Appendix F: Guidelines for standard method performance requirements, in: Official Methods of Analysis (OMA) 2016. http://www.eoma.aoac.org/app_f.pdf Accessed 04 Apr 2019
51. Gu H., Liu G., Wang J., Aubry A.F., Arnold M.E., Selecting the correct weighting factor for linear and quadratic calibration curves with least-square regression algorithm in bioanalytical LC-MS/MS assays and impacts of using incorrect weighting factors on curve stability, data quality, and assay performance. *Anal. Chem.* 2014, *86*, 8959-8966
52. Kiser M.M., Dolan J.W., Selecting the best curve fit. *LCGC Eur.* 2004, *17*, 138-143

Figure Captions

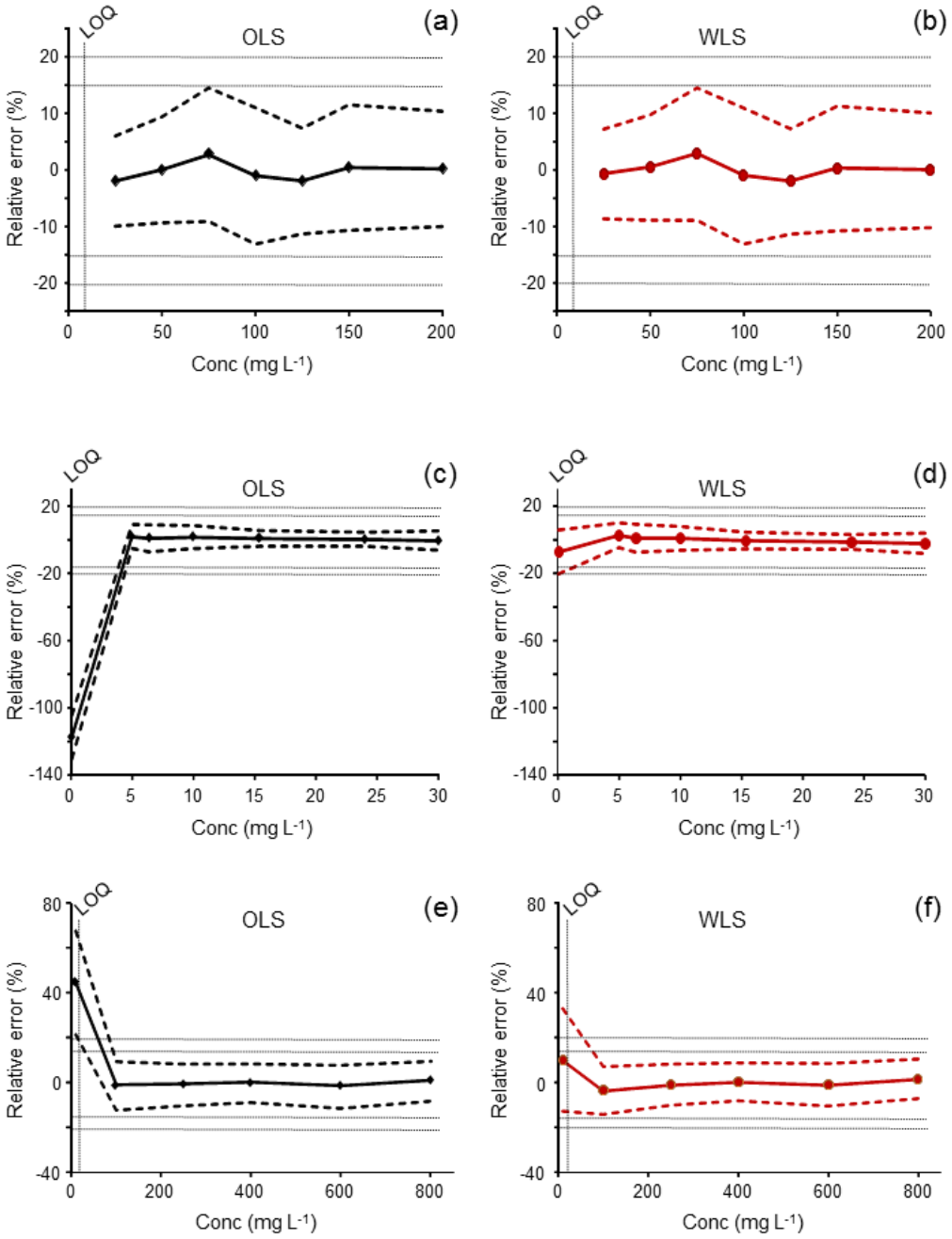

Figure 1. Accuracy profile plots obtained using OLS regression (diamonds, left side plots) and WLS regression (circles, right side plots, $w_i=1/s_i^2$) for (a,b) a method with a working range starting at a concentration above the LOQ determined from the standard deviation at blank

level, (c,d) a method where the first standard was set at the LOQ level, and (e,f) a method that used a first standard between the LOD and the LOQ. Dashed lines in each graph show the corresponding upper and lower tolerance intervals. Vertical dotted lines show the LOQ for each method. Horizontal dotted lines correspond to acceptance limits of ±20% (for LOQ level) and ±15% (>LOQ level) [3].

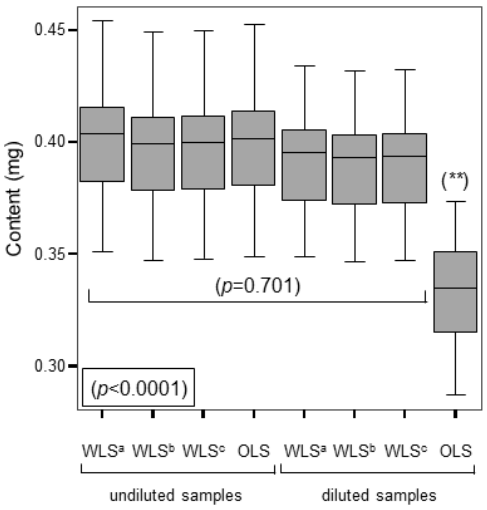

Figure 2. Box-plots obtained in the content determination of a sample lot (analysis of 25 independent replicates). The amount of the target analyte for each replicate was calculated using OLS and WLS regression models, from the starting solution and the same solution diluted until 5 times the LOQ of the method.

(a) WLS regression applied with a weighting factor $w_i=1/s_i^2$, (b) $w_i=1/x_i^2$, and (c) $w_i=1/y_i^2$.

(**) p<0.01

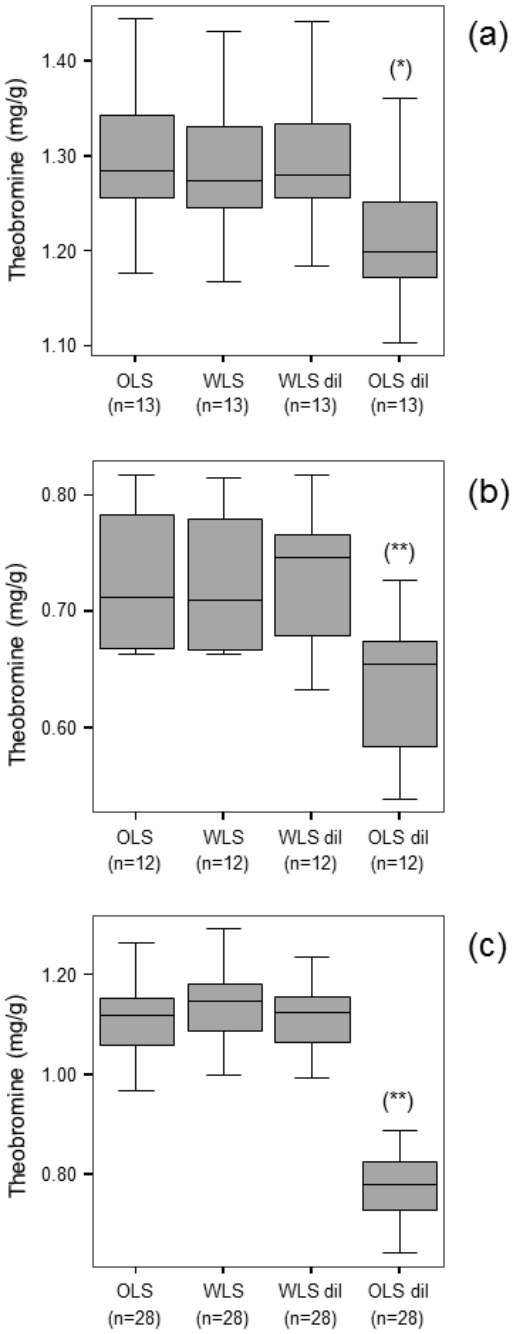

Figure 3. Results obtained in the analysis of independent replicates for three batches of tea samples. The amount of the target analyte was calculated using OLS and WLS regression

models from the final solution and the same solution diluted until (a) 10 times the LOQ of the method applied, (b) 5 times the LOQ, and (c) 0.5 times the LOQ.

(*) $p<0.05$, (**) $p<0.01$

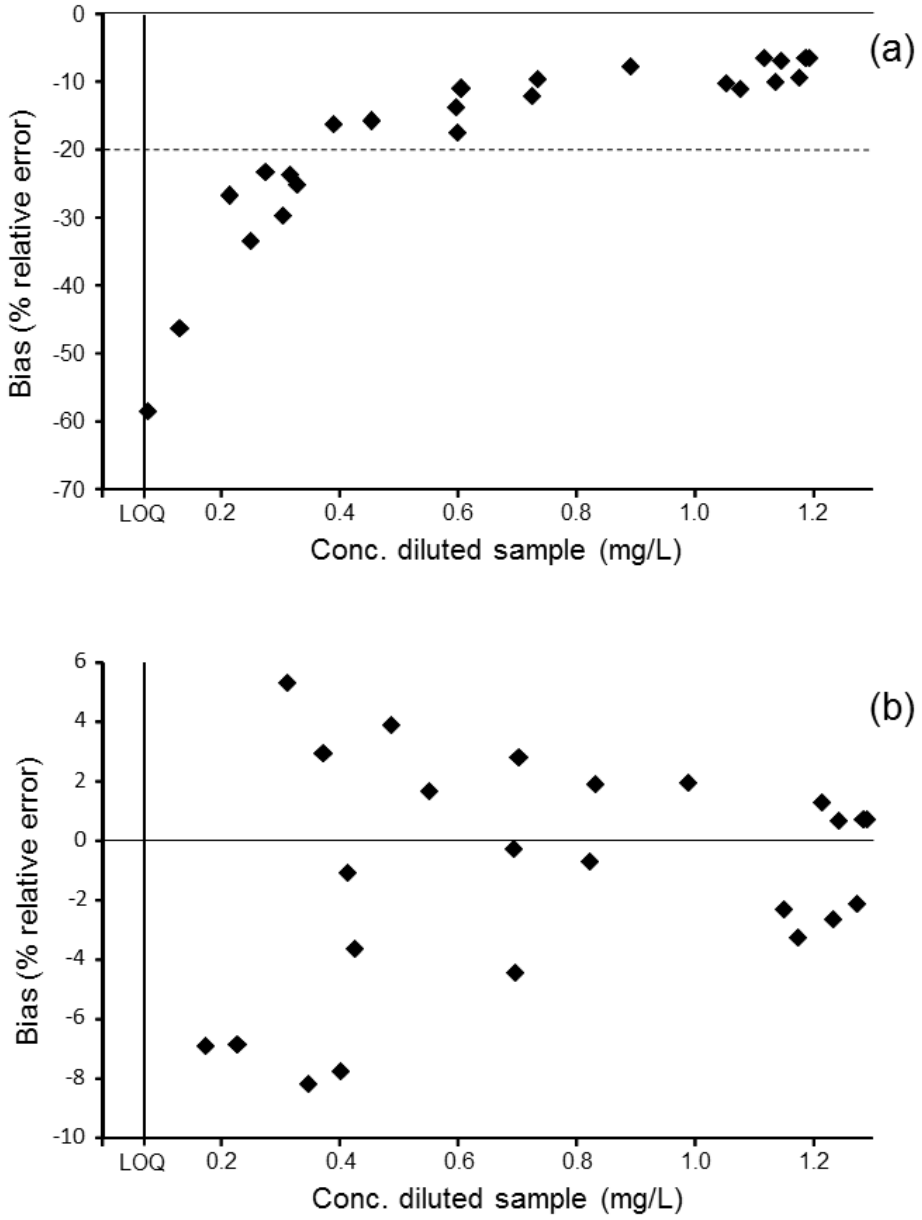

Figure 4. Bias (%) obtained using diluted solutions at levels close to the LOQ respect to the results determined from the corresponding undiluted solutions. Results obtained using (a) OLS regression and (b) WLS regression. Vertical lines correspond to the LOQ of the method determined from the standard deviation at blank level. Dashed line in (a) correspond to a ±20% bias acceptance limit at LOQ level [3].

Table 1. Some validation parameters obtained with the two linear regression methods evaluated (OLS and WLS). LOD and LOQ were determined from the standard deviation of spiked blanks as 3.3 s_{bl} and 10 s_{bl} respectively. Back-calculated bias corresponds to the relative error (%) obtained for the first standard used in each calibration. RSD is the relative standard deviation (%) determined for the first standard. HorRat values were calculated as indicated in eq. 5.

Method	LOD	LOQ	Working range	Back-calculated bias (%)		RSD (%)	HorRat value
				OLS	WLS		
First standard above LOQ							
GC-FID	3 mg L^{-1}	9 mg L^{-1}	25-200 mg L^{-1}	-1.9	-0.8	3.2	0.4
HPLC-UV	2 µM	6 µM	50-500 µM	-5.3	-1.1	4.9	0.5
HPLC-UV	4 µM	12 µM	50-500 µM	4.6	0.4	7.2	0.8
HPLC-UV	10 µM	30 µM	50-500 µM	7.1	1.4	6.3	0.7
HPLC-UV	8 µM	24 µM	50-500 µM	-9.0	-2.2	5.3	0.6
First standard at a level equivalent to the LOQ							
HPLC-UV	0.02 mg L^{-1}	0.06 mg L^{-1}	0.1-30 mg L^{-1}	-117.6	-7.1	5.2	0.3

HPLC-UV	0.02 mg L^{-1}	0.06 mg L^{-1}	0.1-20 mg L^{-1}	101.7	1.5	10.0	0.5
HPLC-UV	0.4 mg L^{-1}	1.2 mg L^{-1}	1-50 mg L^{-1}	-17.6	-1.6	11.0	0.7
GC-MS	0.2 ppbv	0.6 ppbv	0.6-30 ppbv	18.2	-4.0	14.7	0.3
First standard between LOD and LOQ							
GC-FID	6 mg L^{-1}	18 mg L^{-1}	10-800 mg L^{-1}	44.6	10.2	9.4	0.9
GC-FID	9 mg L^{-1}	27 mg L^{-1}	10-800 mg L^{-1}	74.8	6.4	7.4	0.7
GC-FID	9 mg L^{-1}	27 mg L^{-1}	10-800 mg L^{-1}	35.8	2.8	9.2	0.8
GC-FID	8 mg L^{-1}	24 mg L^{-1}	10-800 mg L^{-1}	24.2	14.8	9.1	0.8
GC-FID	5 mg L^{-1}	15 mg L^{-1}	10-400 mg L^{-1}	16.2	2.9	7.9	0.7
HPLC-UV	0.8 mg L^{-1}	2.4 mg L^{-1}	1-50 mg L^{-1}	-81.4	-7.7	8.7	0.5