LINEARITY ASSESSMENT ACCORDING TO IUPAC GUIDELINES FOR THE DETERMINATION OF PLASTICIZERS IN PLASTIC FOOD PACKAGING BY GAS CHROMATOGRAPHY

Mohd Marsin Sanagi,*[1,2] Zalilah Nasir,[1] Susie Lu Ling,[1] Wan Aini Wan Ibrahim,[1] Ahmedy Abu Naim[1]

[1]Ibnu Sina Institute, [2]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

***Corresponding author: E-mail: marsin@kimia.fs.utm.my**

ABSTRACT

Linearity assessment as a required performance in method validation has always been subject to different interpretations and definitions by various guidelines and protocols. However, there are very limited applicable implementation procedures that can be followed by laboratory chemist in assessing linearity. Thus, this work proposes a simple method for linearity assessment in method validation by a regression analysis that covers experimental design, estimation of the parameters, outlier treatment, and evaluation of the assumptions according to IUPAC Guidelines. The suitability of this procedure was demonstrated by its application to an in-house validation for the determination of plasticizers in plastic food packaging by gas chromatography.

Keywords: *Linearity Assessment, Plasticizers, Plastic Food Packaging, Gas Chromatography*

Introduction

Method validation is an important requirement in the practice of chemical analysis. It is indeed, one of the measures required by a laboratory to be recognized and compliance with national and international regulations in all areas of analysis. Guide (1) interpreted the International Standard Organization definition of method validation as being the process of defining an analytical requirement, and confirming that the method under consideration has performance capabilities consistent with that the application requires. In other words, method needs to be validated to provide evidence that the method used is fit for purpose.

General requirements in method validation for performance characteristics shall include but not limited to applicability, selectivity, calibration and linearity, trueness, precision, recovery, range, detection limit, limit of determination or limit of quantification, sensitivity, ruggedness, fitness for purpose, matrix variation as well as measurement uncertainty. Method validation is therefore an essential component of the measures that a laboratory should implement to allow it to produce reliable analytical data (2).

Calibration is a procedure that determines the systematic difference that may exist between a measurement system and a reference system represented by the reference materials and their accepted values (3). Considering that the majority of the analytical methods use linear relationship in one way or another, examination of a calibration function for linearity is

important in validating an analytical method, as well as an everyday task in routine analytical operations (4).

There are several definitions concerning linearity in the literatures (1,5-7). However, the linearity definition can be summarized as the ability of the method to elicit test results that are directly proportional to analyte concentration in a given range. Range is the interval between the upper and lower levels of analyte (inclusive) that have been demonstrated to be determined with precision, accuracy and linearity using the defined method. In practice, the linearity study should be designed to be appropriate for the intended analytical method.

Different guidelines, protocols and papers provide recommendations for linearity assessment in chemical analysis (1-3,6,8). Among the recommended statistical methods to be used for the assessment are ordinary least squares regression (OLS), weighted least squares regression (WLS), or least median of squares regression (LMS). Unfortunately, the recommendations are sometimes complicated or controversial and do not detail the experimental designs, the statistical calculation and the respective assumptions that need to be checked.

Thompson et al. (2) suggested that linearity can be tested informally by examination of a plot of residuals produced by linear regression of the responses on the concentrations in an appropriate calibration set. Any curved pattern suggests lack of fit due to a nonlinear calibration function. A test of significance can be taken by comparing the lack-of-fit variance with that due to pure error. However, there are other causes of lack of fit other than nonlinearity that can arise at certain type of analytical calibration, so the significance test must be used in conjunction with a residual plot. Despite its current widespread of indication of quality of fit, the correlation coefficient (R^2) is misleading and inappropriate as a test for linearity and should not be used (2,8,9-11).

Considering the need for a simpler practical approach to evaluate linearity, this paper presents an application procedure based on the IUPAC Harmonized Guidelines for Single-Laboratory Validation of Methods of Analysis (2). We selected this guideline because it was a collaboration effort by the IUPAC, ISO and AOAC International and later adopted by CODEX Committee of Method of Analysis and Sampling in Joint FAO/WHO Food Standard Program by reference for CODEX purposes in CAC/GL 49-2003 document.

The proposed application procedure

The objective of this work is to establish a practical approach to evaluate linearity range applied to an in-house validated method for the determination of plasticizers in plastic food packaging by gas chromatography based on IUPAC Guidelines (2).

Generally, there are very limited standard methods available for chemical analysis of plastics. Analysis of plastics is rather complex because there is a great variety of possible techniques for both the extraction and subsequent quantitative analysis (13). There is no common procedure for sample preparation because different solvents (acetone, dichloromethane, hexane, diethyl ether, tetrahydrofuran, methanol) can be used for various type of extraction such as by Soxhlet apparatus or solvent-solvent extraction. Besides, certified reference materials are lacking, making it difficult to compare one's own results with a standard value. However, a common technique for plasticizer quantification is gas chromatography with flame ionization (GC-FID) or mass spectrometry detection (GC-MSD).

In this work, an in-house method for the determination of plasticizers in plastic food packaging has been validated to quantify as many as 21 types of plasticizer simultaneously within a short analysis time (35 minutes with GC-FID detection). Only ten types of plasticizers are presented and discussed for the purpose of linearity testing. The plasticizer

compounds studied are listed in Table 1. The abbreviation for each plasticizer is used for the following explanation and discussion.

Experimental design

Experimental design on linearity was based on guidelines by Thompson et al. (2). It involved a study whether the calibration function (a) is linear, (b) passes through the origin, and (c) is unaffected by the matrix of the test material. Schematic diagram for assessing the linear range is shown in Figure 1 whereas matrix effect study is shown in Figure 2.

For an experi Codex Alimentarius Commission, Rome mental design to characterize the linearity domain, let *n* as a total number of calibration levels, each level was carried out in *p* replicates. So, the linear model can be expressed as follows:

$$y_{ij} = a + b \cdot x_i + \varepsilon_{ij}$$

where, y_{ij} is the j^{th} replica peak area measurement of the i^{th} calibration level, x_i is the calibration standard value, b is the slope of the regression line, a is the y-intercept point of the regression line, $a + b \cdot x_i$ represents the predicted peak area measurement of the i^{th} calibration level, and ε_{ij} is the difference between y_{ij} and the predicted of the peak area measurement of the i^{th} calibration level (experimental instrumental error). This model is very easy to use, because model coefficients can be estimated simply by least squares regression technique when the error ε_{ij} is normally distributed.

Estimation of linear parameters

Parameters of the regression line would be obtained using the following equations:

(a) Mean of p peak area measurement of the i^{th} calibration level,

$$\bar{y}_i = \frac{1}{p}\sum_{j=1}^{p} y_{ij}$$

(b) Mean of all the calibration standard value of n calibration levels,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

(c) Mean of all the peak area measurement is,

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} \bar{y}_i$$

(d) Estimated slope,

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\bar{y}_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

(e) Estimated y-intercept point,

$$a = \bar{y} - b \cdot \bar{x}$$

(f) Regression or predicted value associated with the i^{th} calibration level,

$$\hat{y}_i = a + b \cdot x_i$$

(g) Residual,

$$\varepsilon_{ij} = y_{ij} - \hat{y}_i$$

(h) Residual standard deviation,

$$S_{res} = \sqrt{\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}(y_{ij} - \hat{y}_i)^2}{np-2}}$$

(i) Standard deviation at y-intercept,

$$S_a = S_{res} \bullet \sqrt{\left(\frac{1}{np} + \frac{\bar{x}^2}{\sum_{i=1}^{n} p(x_i - \bar{x})^2}\right)}$$

(j) Standard deviation of the slope,

$$S_b = \frac{S_{res}}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$$

Charts construction

The residuals ε_{ij} are plotted against each respective concentration level (*y*-residual plot). Two horizontal dotted dashed lines corresponding to $\pm\, t_{(0.95,\, np\text{-}2)}.S_{res}$ are used to indicate the accepted variation of each single point in the residual plot.

Visual inspection of y-residual plot

The *y*-residual plot is a good indicator of the deviation in relation to the linearity assumption: the linear dynamic range is valid if the residual values are fairly distributed between positive and negative values. However, rejection of outlier should be performed for those data which exceeds $\pm t.S_{res}$ where ***t*** is critical value from Student ***t*** table for *np*-2 degree of freedom. It can also be done by an outlier test as discussed by Weisberg (14) or Jacknife residuals test presented by Horwitz (15).

Test of the linearity assumption

Several error or variation values linked to the calibration should be defined and estimated using the data collected during the experiment after rejection of outlier data. A lack-of-fit test is then performed on the basis of these results, making it possible to test the assumption of non-validity of the linear dynamic range. To perform the lack-of-fit test, the total variability of the responses is decomposed into the sum of squares due to regression (SS_{reg}) and the residual (about regression) sum of squares (SS_{res}). However, the residual sum of squares is separated into lack-of-fit (deviation from linearity) (SS_{lof}) and pure error (from repeated points) sum of squares (SS_{pe}). Finally, sums of squares produced by lack-of-fit is obtained by difference:

$$SS_{lof} = SS_{res} - SS_{pe}$$

This technique has been extensively described (11,12) and was selected because the test is simple and can be easily implemented on much spreadsheet software, using the internal functions. The lack-of-fit test is derived from analysis of variance (ANOVA) applied to regression, computed and summarized in a table comparable to Table 2.

The significance test interpretation is performed in two steps.

(a) First hypothesis: Is the linear regression model acceptable?

- If the ratio F_{reg} is higher than the critical value $F_{(1-\alpha,1,np-n)}$ the hypothesis is that the variation of y are explained by a regression model can be accepted.
- $F_{(1-\alpha,1,np-n)}$ is the value of Fisher distribution for 1 and np-n degrees of freedom at risk level α.
- If the first hypothesis is acceptable, then the second hypothesis would be tested. If not, the regression model is not valid.

(b) Second hypothesis: Is the nonlinear/lack-of-fit model rejected?

- If the ratio F_{lof} is lower or equal to the critical value $F_{(1-\alpha,\, \boldsymbol{n}-2,\boldsymbol{np}-\boldsymbol{n})}$, the hypothesis that the regression model is a linear model can be accepted. The linearity domain is validated.
- If it does not, the limits of the explored domain must be restricted and another lack-of-fit test is performed.

Another statistical test for an intercept significantly different from zero should be carried out after the linear range has been determined. A Student ***t*** test is used to determine whether an intercept, a is significantly different from zero.

$$t_{cal} = \left| \frac{a}{S_a} \right|$$

Where, t_{cal} is calculated t value, a is an intercept, and S_a is an intercept standard deviation. If t_{cal} is lower or equal to the critical value t distribution for n-2 degree of freedom at the risk level of α, the hypothesis of an intercept not significantly different from zero is accepted.

Test for general matrix effect

It simplifies calibration enormously if the calibration standards can be prepared as simple solutions of the analyte. The effects of a possible general matrix mismatch must be assessed in validation if this strategy is adopted. A test for general matrix effect can be made by applying the method of analyte additions (also called "standard additions") to a test solution derived from a typical test material. The test should be done in a way that provides the same final dilution as the normal procedure produces, and the range of additions should encompass the same range as the procedure-defined calibration validation. Once the calibration is linear, the slopes of the usual calibration function and the analyte additions plot can be compared for significant difference.

At first, parameters for ordinary linear regression such as intercept, slope, standard deviation of residual, intercept and slope for both calibration lines (aqueous and matrix) are calculated. An F-test should be used to determine the differences between both (aqueous and matrix solution) residual variances, S^2_{res}. If both residual variances are equal, the following formula is used to calculate t value for further t-test:

$$t_{cal-1} = \frac{|b1 - b2|}{\left[\frac{(n1-2)S^2{}_{res1} + (n2-2)S^2{}_{res2}}{n1+n2-4} \left(\frac{1}{\sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2} + \frac{1}{\sum_{i=1}^{n}(x_{i2} - \bar{x}_2)^2} \right) \right]^{1/2}}$$

where, $b1$ is slope of regression line for aqueous solution and $b2$ is slope of regression line for matrix solution, $n1$ is total calibration levels for aqueous solution and $n2$ is total calibration levels for matrix solution, S^2_{res1} is residual variance for aqueous solution and S^2_{res2}

is residual variance for matrix solution, x_i is calibration standard value of the i^{th} calibration level, x_{i1} for aqueous solution and x_{i2} for matrix solution, $\bar{x}_1$ is mean of all the calibration standard value for aqueous solution and $\bar{x}_2$ is mean of all the calibration standard value for matrix solution.

However, when both residual variances are not equal, the following formula is used to calculate *t* value for further Student *t*-test:

$$t_{cal-2} = \frac{|b1 - b2|}{\left(S^2_{b1} + S^2_{b2}\right)^{1/2}}$$

where, *b*1 is slope of regression line for aqueous solution and *b*2 is slope of regression line for matrix solution, S^2_{b1} is standard deviation of regression line slope for aqueous solution and S^2_{b2} is standard deviation of regression line slope for matrix solution

Following that, the hypothesis testing is performed for comparing the slope of both lines using Student *t*-test:

H_o: the slopes are equal

H_1: the slopes are not equal

If t_{cal-1} is used to calculate the *t* value, it is then compared with n1+n2-4 degree of freedom at the chosen significance level. For t_{cal-2}, the calculated *t* value is compared with a Student's *t*-distribution:

$$t' = \frac{\left(t_1 S^2_{b1} + t_2 S^2_{b2}\right)}{\left(S^2_{b1} + S^2_{b2}\right)}$$

where t_1 and t_2 are the theoretical *t* values at the chosen level of significance with *n1*-2 and n2-2 degree of freedom, respectively.

When the t_{cal} is less or equal with the tabulated *t*, it can be concluded that the method is selective and the calibration for routine use can be carried out in aqueous solution. Otherwise, the calibration line constructed for routine used should be prepared in sample matrix solution.

Method

Equipment and Apparatus

(a) *Gas chromatograph.*- Shimadzu Model 2010 (Kyoto, Japan) equipped with 30 m x 0.25 mm ID, 0.25 µm 100% dimethylpolysiloxane DB-1 capillary column (J&W Scientific, Folsom, CA, USA) with flame ionization detector; autosampler Shimadzu AOC-20i.

(b) *Data collection.*- Shimadzu *GCSolution* software, Version 2.1 (Kyoto, Japan) was used for data acquisition, analysis and instrument control.

(c) *Magnetic stirrer plate.*- Fisher Scientific model SM6 (UK), or equivalent.

(d) *Rotary evaporator.*- Heidolph model Laborota 4002 (Germany), or equivalent.

(e) Syringes.- 1 mL, slip tip, nonsterile clean

(f) Syringe filters.- Disposable 0.45 µm pore size, PTFE membrane, 4mm filter size, Alltech (USA), or equivalent.

(g) Microfiber filter papers.- Disposable, 125 mm diameter size, CHMLAB Group (Barcelona, Spain), or equivalent.

(h) Conical tubes.- 15 mL disposable polypropylene tube, Corning (USA) or equivalent.

(i) Evaporation flasks.- 250 mL borosilicate round bottom flask

(j) Glasswares.- Class A, volumetric pipets and flasks were used to prepare all calibration standards and spiking solutions.

Reagents

(a) *Solvents*.-Acetone gas chromatographic grade (Merck, Germany); methanol LC grade (Fisher Scientific, UK), tetrahydrofuran ACS grade (Merck, Germany), and chloroform LC grade (May & Baker, England).

(b) *Plasticizer standards*.- DMP, DnBP, DPep, ATBC, BBP, DEHA, TOP, DEHP, DnOP and TOTM analytical reference materials were supplied by Kanto Chemicals (Japan) with purity of more than 99%.

Standard solution preparation

Stock solutions of ten plasticizers at a concentration of 10000 mg/L were separately prepared in acetone. The intermediate mixture of ten plasticizers solution at 500 mg/L was prepared by diluting appropriately from respective stock standard solution with acetone solvent:

(a) Experiment for linear range.- The calibration solutions were prepared between 0.8 to 50 mg/L at equally spaced (every 5 mg/L) by diluting in acetone the intermediate mixture of 500 mg/L. The solutions were prepared and determined by the instrument in two independent replicates. The levels of concentration studied are equivalent to the range from 8 to 500 mg/kg of each plasticizer in plastic food packaging.

(b) Experiment for matrix effect.- Two calibration solutions at the predetermined linear range were prepared in aqueous acetone and matrix blank solution in at least six levels.

Matrix blank extraction

In a clean 500 mL Erlenmeyer flask, add 20 mL tetrahydrofuran, followed by 100 mL methanol. Stir with a magnetic stirrer bar on a magnetic stirrer plate for 5 minutes. Then, filter the extract with microfiber filter paper into a 250 mL evaporating flask. Rinse twice with 20 mL methanol each. Concentrate until almost dryness at 50°C water bath using rotary evaporator. Quantitatively transfer this concentrated extract to 10 mL volumetric flask with acetone and mark up to the volume. Filter the extract with PTFE disposable syringe filter into GC vials.

Analysis

The standard solution series prepared for the first experiments were determined for GC-FID peak area response in a random order (different day) while the standard solutions prepared for the second experiment were determined by the GC-FID on the same day.

Results and discussion

Method development began with evaluation of extraction and instrumental set-up from methods reviewed. Emphasis was placed on steps that had potential to be extended to multiple classes of synthetic plasticizers in plastic food packaging. Finally, the in-house established method was able to determine simultaneous 21 plasticizer types using one sample preparation and determination with the previously stated GC-FID temperature program. But , for the purpose of proposed linearity study application procedure, only ten types of plasticizer are used and discussed. Figure 3 illustrates a typical chromatogram of mixture of ten plasticizers studied. The elution order is *(1)* DMP (retention time, 6.064 min), *(2)* DnBP (retention time, 11.808 min), *(3)* DPeP (retention time, 13.703 min), *(4)* ATBC (retention time, 14.746 min), *(5)* BBP (retention time, 15.398 min), *(6)* DEHA (retention time, 16.094

min), *(7)* TOP (retention time, 16.739 min), *(8)* DEHP (retention time, 17.241 min), *(9)* DnOP (retention time, 18.681 min) and *(10)* TOTM (retention time, 24.379 min). It clearly shows that the analytical method proposed in this work completely separates all the analytes.

For this linearity study, eleven concentration levels were selected, ranging from 0.8 to 50 mg/L of plasticizer compound. This chosen interval corresponds to the usual content that is expected in plastic food packaging. Each level was measured twice. Draper & Smith (10) proposed an experimental design with three levels (two extremes and a central) with a larger number of replicates in the lower or upper levels. Nevertheless, references related to method validation suggested a minimum of five or six concentrations levels (1,2,5) equally spaced across the concentration range, at least in duplicate (2). Replicates of each calibration point give information about the intrinsic variability of the response measurements (pure error). In order to respect measurement independence, each replicate must be performed on a newly prepared standard solution and measured in a random order to avoid the problem of confusing non-linearity with temporal effects, such as calibration drift (2). If the replicates are just repetitions of the same reading or obtained by successive dilutions, the residual variance S^2_{res} will tend to underestimate the variance σ^2 and the lack-of-fit test will tend to wrongly detect non-existent lack-of-fit (10).

Data collected for linearity study with ten plasticizers is further statistically estimated using a simple ordinary least squares regression method and transformed into *y*-residual plots as shown in Figure 4.

Visual examination of *y*-residual plots in Fig. 4 indicated possible outliers and revealed no other obvious deficiency. The points that were outside the accepted confidence interval ($\pm\, t_{(0.95,\, p-2)}.S_{res}$) were regarded as outliers. There are five plasticizers were shown as having only one outlier point each which comes from the first replicate for DnBP, ATBC, BBP and TOTM at 40 mg/L while for DEHA at 35 mg/L. Another two plasticizers exhibited outlier data from the second replicate only that is DPeP at 25 & 30 mg/L whereas DnOP at 30, 45 and 50 mg/L. For DMP, TOP and DEHP, one outlier point was detected in each replicate. This may be due to the instrument fluctuation during analysis. With this assumption, no further statistical test for outliers need to be carried out.

The residual plots in Fig. 4 could also be used to determine for any heteroscedastic data formation. Graphically, homogeneous scatter across is seen in the linearity domain studied (between 0.8 to 50 mg/L) for all ten plasticizers. However, further homoscedasticity test can be conducted as mentioned in the proposed application procedure. The distribution data in this experiment is fairly homoscedastic due to the narrow selected study range (factor of 50). If the study range is selected at a wider scope (example factor of 200 and above), the analytical data would normally tend to become heteroscedastic where the deviation between replicates becomes bigger at higher concentration tested (8,11,17).

As for data distribution, it clearly shows that the points pattern for the ten plasticizers are randomly distributed about the straight line to assume linearity. However, Weisberg (14) stated that the assumption of normal errors plays only a minor role in regression analysis and needed for inference with small samples. Furthermore, abnormality of the unobservable errors is very difficult to diagnose in small samples by examination of residuals. After rejection of outlier data was done, parameters of ordinary least squares regression method were estimated again.

Further significant tests on regression model and linearity were carried out for those ten plasticizers and the final findings are shown in a form of Analysis of Variance (ANOVA) in Table 4. At the beginning of regression test, it has been validated that all the plasticizers are significantly correlated to the analyte concentration studied between 0.8 to 50 mg/L, where the *F* calculated or observed values are much higher than the corresponding *F* critical value of $F_{(1-\alpha,1,\boldsymbol{p}(\boldsymbol{n}-1))}$. However, when the same data is used to test the error of model, only

nine plasticizers namely DnBP, DPeP, BBP, DEHA, TOP, DEHP, DnOP, and TOTM are actually not significant at the risk of 1 % to accept the proposed linearity range (0.8-50 mg/L). The linearity range for ATBC is between 0.8 to 45 mg/L where the Fisher variable associated to the test of error of model is smaller than the critical value of $F_{(1-\alpha,\ p-2,p(n-1))}$.

The understanding of these findings should be based on the properties of the *F* lack-of-fit test. If a calibration line has a significant curvature, the null hypothesis of linearity will be rejected and attempts must be made to find a more appropriate model (12). An obvious alternative would be a polynomial fitting, but the question of how complex a model would need to be is difficult and fundamental (7,8). On the other hand, if null hypothesis is not rejected, it does not mean that the linear model is correct, only that the model is not contradicted by the data (10) or that insufficient data exist to detect the inadequacies of the model (17). In addition, there are causes of lack-of-fit other than non-linearity that can arise in calibration curves (2), so the significant test must be used in conjunction with a residual plot.

Finally, the linear calibration curves of peak area response versus plasticizer concentrations were constructed (Figure 5). Since the linearity range for the ten plasticizers has been determined, other tests were carried out to determine whether the linear calibration curve passes through the origin. Results of the test are shown in Table 4. Based on the significant test, it shows that all plasticizer calibration function are linear and pass through the origin except for DMP and TOTM. Despite the fact that a much simpler of two-levels calibration standard can be applied for eight plasticizers namely DnBP, DPeP, ATBC, BBP, DEHA, TOP, DEHP, and DnOP, it actually would not be practical for routine use in this particular method. This is due to the nature that this method was established to determine simultaneous ten plasticizers in a single run. As a result, those eight plasticizers would be treated the same as DMP and TOTM i.e. at least three-levels calibrations standard are required for daily use.

For matrix effect study, data from the second experiment was calculated, tested and tabulated in Table 5. A lack of significance in this test will often mean that the matrix variation effect will also be absent. Results show that there are no matrix effect for all the ten plasticizers namely DMP, DnBP, DPeP, ATBC, BBP, DEHA, TOP, DEHP, DnOP and TOTM. Thus, a simple aqueous calibration standard solution can be used for routine analysis of plasticizers using GC-FID.

At the end of the work, the calibration function for daily routine use and the linear range for ten plasticizers studied are summarized and presented in Table 6. It demonstrates that the linear range of the in-house validated method for determination of plasticizers in plastic food packaging by GC-FID has been verified to lay between 8 to 500 mg/kg for nine plasticizers namely DMP, DnBP, DPeP, BBP, DEHA, TOP, DEHP, DnOP, and TOTM. However, ATBC has shown to be linear between 8 to 450 mg/kg only.

Conclusions

The proposed procedure to assess linearity range in this paper was straightforward, highly practical and sufficient to be applied for an in-house validated method. This work confirms that the practical procedures are able to fulfill the minimum requirements in Section A3-Calibration and linearity of IUPAC Guidelines.

Acknowledgements

The authors would like to thank Universiti Teknologi Malaysia and the Public Health Laboratory, Health Department of Johor for technical support, research facilities and studentship to Zalilah Nasir and Susie Lu Ling.

References

(1) EURACHEM Guide (1998). *The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics*. United Kingdom: LGC (Teddington) Ltd.
(2) Thompson, M., Ellison, S.L R., & Wood, R. (2002) *Pure Appl. Chem*., **74**, 835-855
(3) International Standard ISO 11095 (1996) *Linear calibration using reference materials*, International Organization for Standardization, Geneva
(4) de Souza, S.V.C. & Junqueira, R.G. (2005) *Anal. Chim. Acta*, **552**, 25-35
(5) Huber, L. (1998*) LC/GC*, February, 96-103
(6) Association Official of Analytical Chemist (AOAC) (1998). *Peer-Verified Methods Program Manual on Policies and Procedure*, AOAC International: Arlington
(7) Mark, H. J. (2003) *Pharm. Biomed. Anal.*, **33**, 7-20
(8) Miller, J.N., & Miller, J.C. (2000) *Statistics and Chemometrics for Analytical Chemistry*. Fourth Edition. England: Prentice Hall
(9) Kiser, M.M., & Dolan, J.W. (2004) *LC.GC Europe*, March, 138-143
(10) Draper, N.R., & Smith, H. (1998) *Applied Regression Analysis*, New York: Wiley, p. 706.
(11) Montgomery, D.C., & Runger, G.C. (2003) *Applied Statistics and Probability for Engineers*, New York: John Wiley and Sons Inc.,.
(12) Codex Alimentarius Commission (2003*) FAO/WHO Food Standards, CAC/GL 49,* Codex Alimentarius Commission, Rome
(13) Ritter, A., & Meyer, V.R. (2005) *Polymer Testing*, **24**, 988-993
(14) Weisberg, S. (2005) *Applied Linear Regression*, 3rd Ed., New York: John Wiley & Sons, p. 194.
(15) Horwitz, W. (1995) *Pure Appl. Chem*., **67**, 331-343
(16) Garcia, I., Ortiz, Sarabia, L., Vilches, C., & Gredilla, E. (2003) *J. Chromatogr. A*, **992**, 11-27
(17) Meyer, P.C., & Zund, R.R. (1993) *Statistical Methods in Analytical Chemistry*, New York: John Wiley & Sons, p. 81.

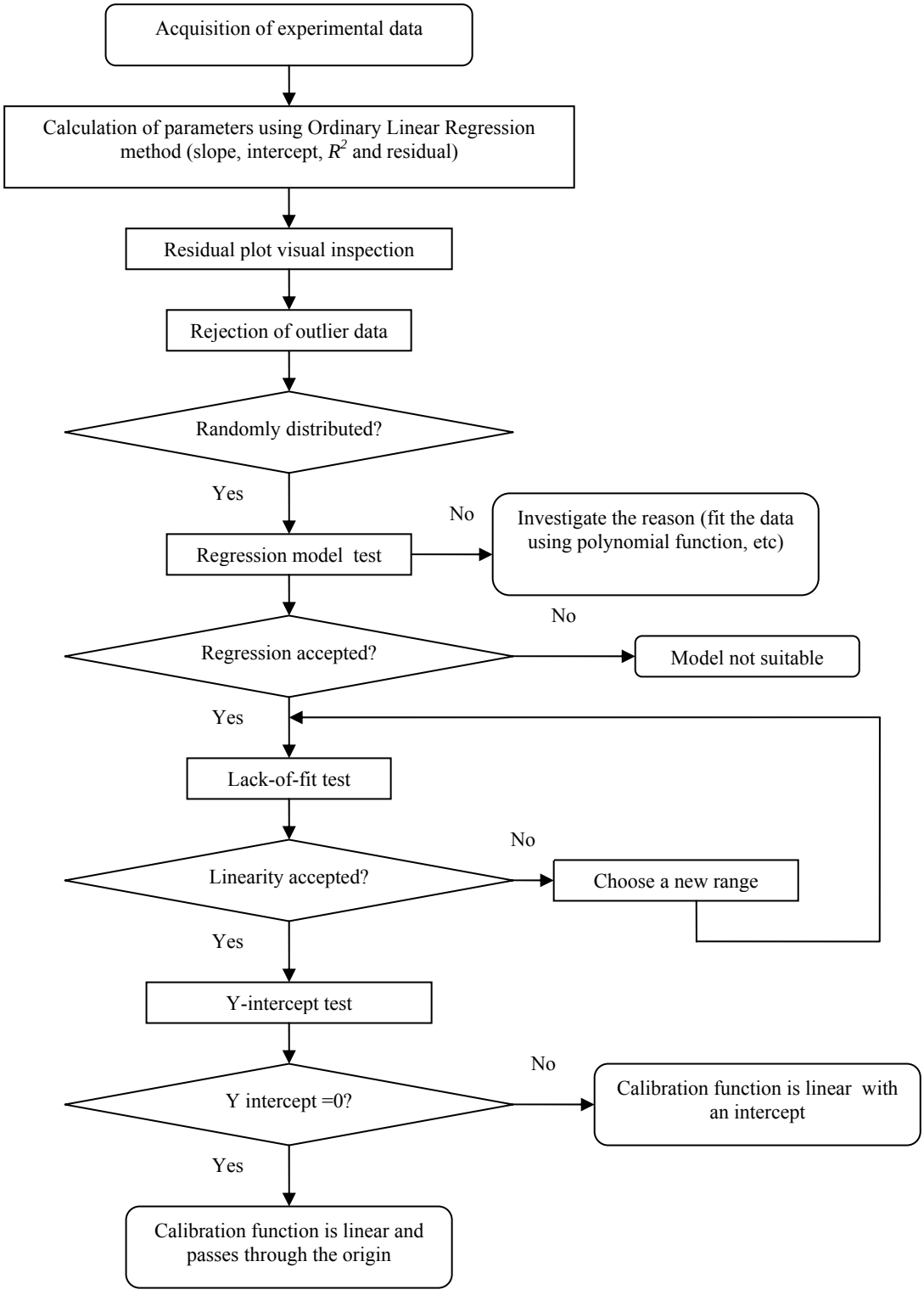

Figure 1. Schematic diagram for linear range testing

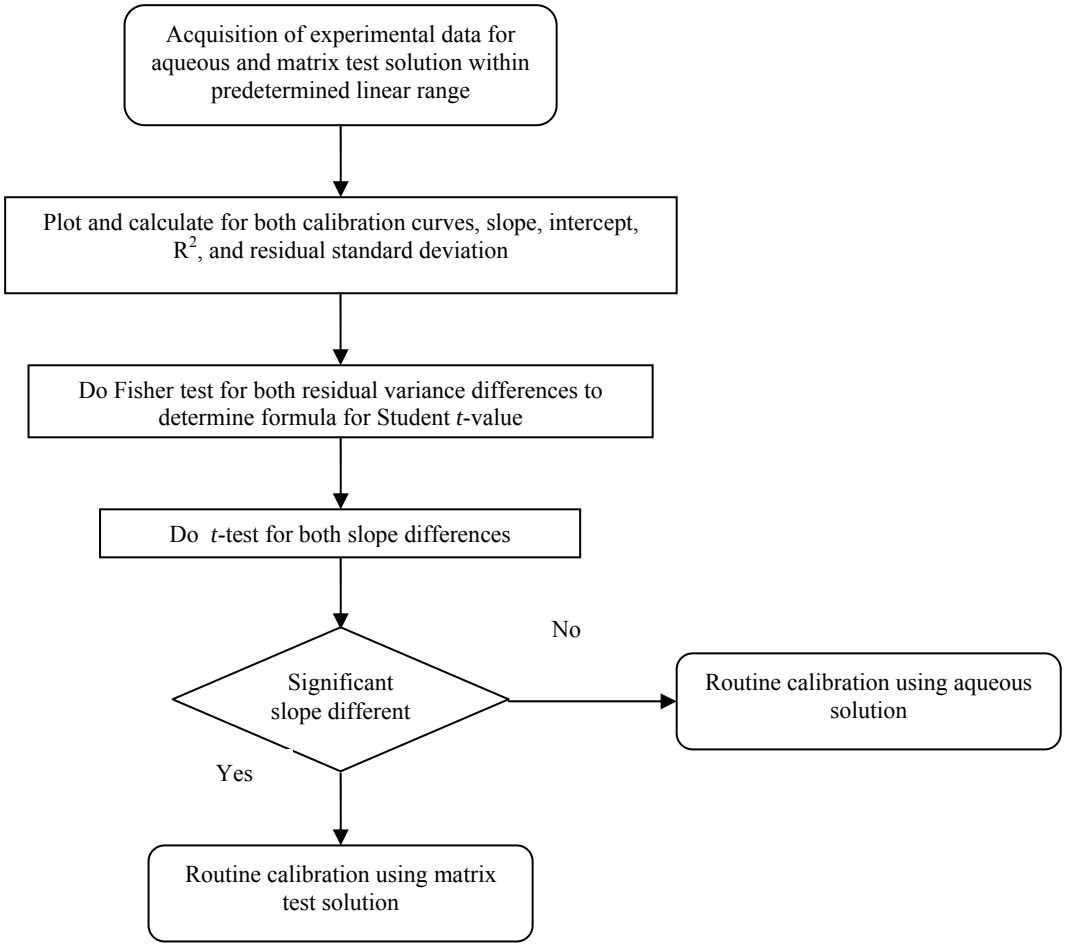

Figure 2. Schematic diagram for general matrix effect study

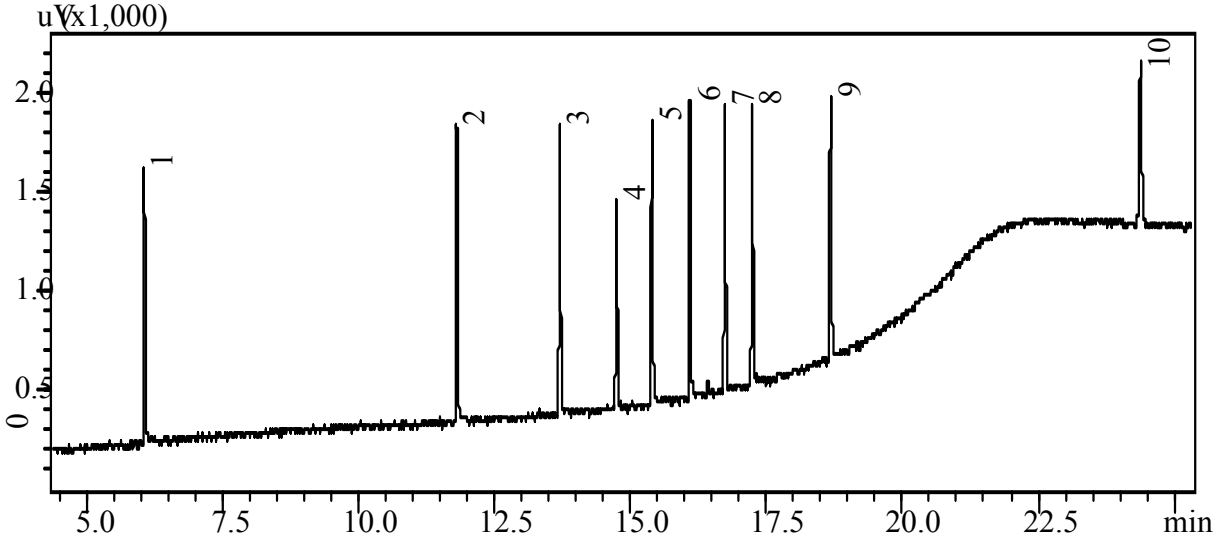

Figure 3. GC-FID chromatogram of a mixture of plasticizer standard which contains 10 mg/L each. Peaks: 1, DMP; 2, DnBP; 3, DPeP; 4, ATBC; 5, BBP; 6, DEHA; 7, TOP; 8, DEHP; 9, DnOP; 10, TOTM.

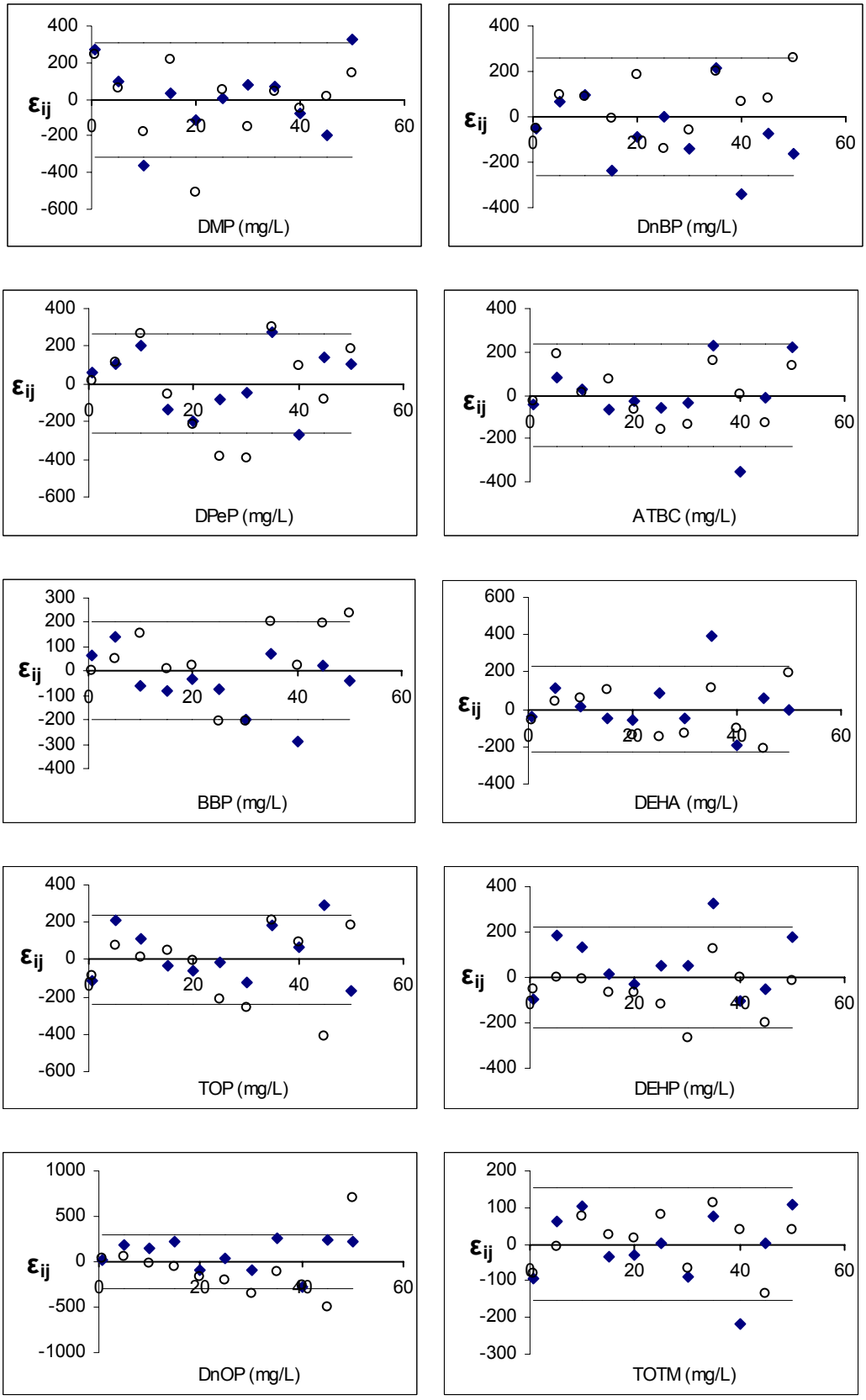

Figure 4. ***y***-residual plots of ε_{ij} versus concentration for ten plasticizers, ♦ indicate the first replicate data, ○ indicate the second replicate data and dashed lines are $\pm\, \boldsymbol{t}_{(9.95,p-2)}.S_{res}$

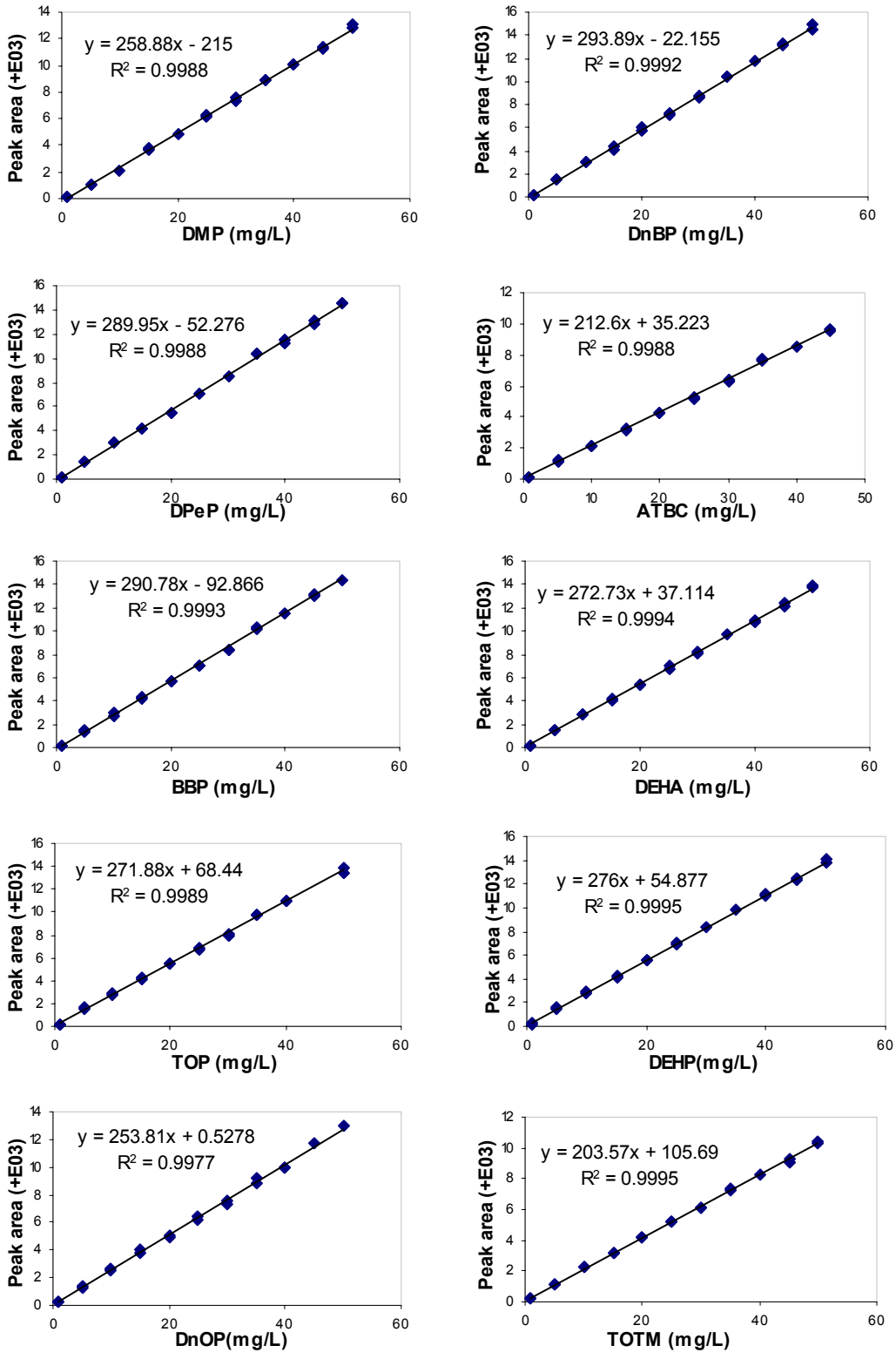

Fig. 5. Calibration curves of GC-FID peak area response versus concentration for ten plasticizers after discarding outliers with respective Ordinary Least Squares method statistics.

Table 1. List of plasticizer compounds selected for linearity testing with respective abbreviation, CASRN and molecular mass

No.	Plasticizer name	Abbreviation	CASRN	Molecular mass (g/mol)
1.	Dimethyl phthalate	DMP	131-11-3	194.2
2.	Di-n-butyl phthalate	DnBP	84-74-2	278.4
3.	Dipentyl phthalate	DPeP	131-18-0	306.4
4.	Acetyl-tri-n-butyl citrate	ATBC	77-90-7	402.0
5.	Benzyl butyl phthalate	BBP	85-68-7	312.4
6.	Di-2-ethylhexyl adipate	DEHA	103-23-1	370.6
7.	Tris(2-ethylhexyl) phosphate	TOP	78-42-2	434.6
8.	Di-2-ethylhexyl phthalate	DEHP	117-81-7	390.6
9.	Di-n-octyl phthalate	DnOP	117-84-0	390.6
10.	Tri-2-ethylhexyl trimellitate	TOTM	3319-31-1	546.8

Table 2. ANOVA test used to determine the linearity domain

Sources of variation	Sum of squares, SS	Degree of freedom, d.f	Variance, S^2	Fisher ratio, F
Residual	$SS_{res} = \sum_{i=1}^{n}\sum_{j=1}^{p}\left(y_{ij} - \hat{y}_i\right)^2$	np-2	$S^2_{res} = \frac{SS_{res}}{np-2}$	
Regression	$SS_{reg} = \frac{\left[\sum_{i=1}^{n}\sum_{j=1}^{p}(x_i - \bar{x})(y_{ij} - \bar{y})\right]^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$	1	$S^2_{reg} = \frac{SS_{reg}}{1}$	$F_{reg} = \frac{S^2_{reg}}{S^2_{pe}}$
Lack-of-fit	$SS_{lof} = SS_{res} - SS_{pe}$	n-2	$S^2_{lof} = \frac{SS_{lof}}{n-2}$	$F_{lof} = \frac{S^2_{lof}}{S^2_{pe}}$
Pure error	$SS_{pe} = \sum_{i=1}^{n}\sum_{j=1}^{p}\left(y_{ij} - \bar{y}_i\right)^2$	np-n	$S^2_{pe} = \frac{SS_{pe}}{np-n}$	

n, total calibration level; p, total replicate for each calibration level

Table 3. Results of ANOVA statistics on significance test for regression model for ten plasticizers

Source	SS	d.f	MS	Fcal	Fcrit	Conclusion
1. DMP (0.8-50 mg/l)						
Regression	345000208	1	345000208	35982.50	10.56	Regression model accepted
Lack-of-fit	312625	9	34736	3.62	5.35	Linearity accepted
Pure Error	86292	9	9588			
Total	345399125	19	345044532			
2. DnBP (0.8-50 mg/l)						
Regression	448071830.982	1	448071830.982	24530.106	10.044	Regression model accepted
Lack-of-fit	185118.161	9	20568.685	1.126	4.942	Linearity accepted
Pure Error	182662.000	10	18266.200			
Total	448439611.143	20	22421980.557			
3. DPeP (0.8-50 mg/l)						
Regression	453611722	1	453611722	40902.98	10.56	Regression model accepted
Lack-of-fit	453588	9	50399	4.54	5.35	Linearity accepted
Pure Error	99810	9	11090			
Total	454165120	19	453673211			
4. ATBC (0.8-45 mg/l)						
Regression	168811094	1	168811094	37576.20	11.26	Regression model accepted
Lack-of-fit	165090	8	20636	4.59	6.03	Linearity accepted
Pure Error	35940	8	4493			
Total	169012124	17	168836223			
5. BBP (0.8-50 mg/l)						
Regression	380277452	1	380277452	50443.97	10.56	Regression model accepted
Lack-of-fit	213205	9	23689	3.14	5.35	Linearity accepted
Pure Error	67848	9	7539			
Total	380558504	19	380308680			
6. DEHA (0.8-50 mg/l)						
Regression	395562922	1	395562922	32746.93	10.56	Regression model accepted
Lack-of-fit	135894	9	15099	1.25	5.35	Linearity accepted
Pure Error	108715	9	12079			
Total	395807531	19	395590101			
7. TOP (0.8-50 mg/l)						
Regression	336140698	1	336140698	30193.86	10.04	Regression model accepted
Lack-of-fit	255573	8	31947	2.87	5.06	Linearity accepted
Pure Error	111328	10	11133			
Total	336507599	19	336183778			
8. DEHP (0.8-50 mg/l)						
Regression	402770732	1	402770732	43773.61	10.56	Regression model accepted
Lack-of-fit	110645	9	12294	1.34	5.35	Linearity accepted
Pure Error	82811	9	9201			
Total	402964188	19	402792228			
9. DnOP (0.8-50 mg/l)						
Regression	277138360	1	277138360	12409.33	10.56	Regression model accepted
Lack-of-fit	432644	9	48072	2.15	5.35	Linearity accepted
Pure Error	200998	9	22333			
Total	277772002	19	277208765			
10. TOTM (0.8-50 mg/l)						
Regression	214979590	1	214979590	99573.69	10.04	Regression model accepted
Lack-of-fit	82134	9	9126	4.23	4.94	Linearity accepted
Pure Error	21590	10	2159			
Total	215083314	20	214990875			

SS, sum squares; d.f., degrees of freedom; MS, mean squares; Fcal, Fisher ratio; Fcrit, Critical value of F-distribution for a one-tailed test at α=0.01

Table 4. Results for significant test on y-intercept for ten plasticizers compound calibration curves

Plasticizer type	Intercept, *a*	Standard deviation, *Sa*	Calibration level, *n*	$\lvert t_{cal} \rvert$	t_{crit}	Zero y-intercept
DMP	-215.0	63.53	11	3.38	3.25	No
DnBP	-22.15	56.00	11	0.40	3.25	Yes
DPeP	-52.28	71.06	11	0.74	3.25	Yes
ATBC	35.22	45.92	10	0.77	3.36	Yes
BBP	-92.87	51.28	11	1.81	3.25	Yes
DEHA	37.11	45.58	11	0.81	3.25	Yes
TOP	68.44	58.37	10	1.17	3.36	Yes
DEHP	54.88	41.72	11	1.32	3.25	Yes
DnOP	0.5278	77.62	11	0.01	3.25	Yes
TOTM	105.69	29.74	11	3.55	3.25	No

$\lvert t_{cal} \rvert = a/S_a$; t_{crit}, critical value of $\lvert t \rvert$ for a two-tailed test for n-2 degree of freedom at α=0.01

Table 5. Results for significance test on matrix effect for calibration standard solution

Plasticizer Type	*	Aqueous	Matrix	*F*- test for variance (S^2_{res})	*t*-test for slopes	Conclusion
DMP	*b*	248.34	232.05	Not same	Same	No matrix effect
	a	-126.00	-177.75			
	S_b	7.83	4.52			
	S_a	215.98	124.78			
	S_{res}	367.36	212.24			
	n	10	10			
DnBP	*b*	368.68	363.07	Not same	Same	No matrix effect
	a	-360.99	-333.08			
	S_b	14.98	11.46			
	S_a	295.52	225.89			
	S_{res}	696.29	532.24			
	n	10	10			
DPeP	*b*	296.53	294.57	Not same	Same	No matrix effect
	a	-59.512	-5.66			
	S_b	3.81	1.82			
	S_a	42.09	20.14			
	S_{res}	69.37	33.20			
	n	6	6			
ATBC	*b*	221.27	199.00	Not same	Same	No matrix effect
	a	230.06	262.51			
	S_b	17.91	17.18			
	S_a	361.72	346.97			
	S_{res}	820.52	787.05			
	n	8	8			
BBP	*b*	311.16	319.19	Not same	Same	No matrix effect
	a	-180.83	-186.50			
	S_b	4.33	3.59			
	S_a	85.31	70.74			
	S_{res}	201.02	166.67			
	n	10	10			
DEHA	*b*	288.19	298.43	Not same	Same	No matrix effect
	a	-119.62	-133.54			
	S_b	3.33	3.41			
	S_a	65.88	67.17			
	S_{res}	154.77	158.27			
	n	10	10			

**b*, slope; *a*, intercept; S_b, standard deviation of slope; S_a, standard deviation of intercept; S_{res}, residual standard deviation; *n*, number of data to build the calibration curve

Table 5. *(continued)*

Plasticizer Type	*	Aqueous	Matrix	*F*- test for variance (S^2_{yx})	*t*-test for slopes	Conclusion
TOP	*b*	285.81	286.99	Not same	Same	No matrix effect
	a	-129.57	-65.07			
	S_b	3.33	3.41			
	S_a	65.88	67.17			
	S_{res}	154.77	158.27			
	n	10	10			
DEHP	*b*	289.57	299.56	Not same	Same	No matrix effect
	a	-136.97	-2.75			
	S_b	4.42	2.62			
	S_a	87.23	51.58			
	S_{res}	205.52	151.53			
	n	10	10			
DnOP	*b*	265.73	271.24	Not same	Same	No matrix effect
	a	-147.08	-112.49			
	S_b	3.60	2.85			
	S_a	70.89	56.16			
	S_{res}	167.03	132.31			
	n	10	10			
TOTM	*b*	215.58	221.36	Not same	Same	No matrix effect
	a	-37.04	-41.62			
	S_b	2.13	1.68			
	S_a	41.94	33.12			
	S_{res}	98.83	78.04			
	n	10	10			

Table 6. Summary of findings for linearity testing and calibration function based on the IUPAC Guidelines (2) for simultaneous determination of ten plasticizers in plastic food packaging using GC-FID

Plasticizer type	Calibration function for routine use		Linear range (mg/kg)
	Level (at least)	Medium	
DMP	Three	Aqueous	8-500
DnBP	Three	Aqueous	8-500
DPeP	Three	Aqueous	8-500
ATBC	Three	Aqueous	8-450
BBP	Three	Aqueous	8-500
DEHA	Three	Aqueous	8-500
TOP	Three	Aqueous	8-500
DEHP	Three	Aqueous	8-500
DnOP	Three	Aqueous	8-500
TOTM	Three	Aqueous	8-500