Journal of Chromatography B, 877 (2009) 2380–2387

Contents lists available at ScienceDirect

Journal of Chromatography B

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

Advances in quality control for dioxins monitoring and evaluation of measurement uncertainty from quality control data☆

Gauthier Eppe*, Edwin De Pauw

Mass Spectrometry Laboratory, Center of Analysis of Residues in Traces (CART), Chemistry Department, University of Liège, 3 Allée de la chimie, B6C, Sart-Tilman, B-4000 Liège, Belgium

ARTICLE INFO

Article history:
Received 7 November 2008
Accepted 10 May 2009
Available online 15 May 2009

Keywords:
Multivariate quality control
Hotelling's T^2
Exponentially weighted moving average (EWMA)
Measurement uncertainty
Polychlorinated dibenzo-*p*-dioxin (PCDDs)
Polychlorinated dibenzofurans (PCDFs)
Polychlorinated biphenyls (PCBs)

ABSTRACT

This paper describes an application of multivariate and multilevel quality control charts with the aim of improving the internal quality control (IQC) procedures for the monitoring of dioxins and dioxin-like PCBs analysis in food. Dioxin analysts have to use the toxic equivalent concept (TEQ) to assess the toxicity potential of a mixture of dioxin-like compounds. The TEQ approach requires quantifying individually 29 dioxin-like compounds. Monitoring the congeners separately on univariate QC charts is misleading owing to the increase of false alarm rate. We propose to subdivide the TEQ value into 3 sub-groups and to control simultaneously the 3 variables in a T^2 chart. When a T^2 exceeds the upper control limit, it acts as a warning to trigger additional investigations on individual congeners. We discuss the minimum number of runs required to reliably estimate the QC chart parameters and we suggest using data from multilevel QC charts to properly characterize the standard deviations and the correlation coefficients. Moreover, the univariate QC chart can be sensitised to detect systematic errors by using exponentially weighted moving average (EWMA) technique. The EWMA chart provides an additional guidance on setting appropriate criteria to control the method bias and to support trend analysis. Finally, we present an estimate of measurement uncertainty by computing the accuracy profile in a retrospective way with the QC data generated and we discuss assessment of compliance with regulatory maximum levels.

1. Introduction

Laboratories must be able to produce reliable data when performing analytical tests for a customer or for official control purposes regardless of the methodology used. Consequently, laboratories should implement an analytical quality assurance (AQA) management program to ensure that high-quality data are achieved. As defined by Tavernier et al. [1], AQA is a complete set of measures a laboratory must undertake. It encompasses method validation, estimation of measurement uncertainty, effective internal quality control (IQC) procedures, participation at relevant proficiency testing (PT) schemes and accreditation to an international standard, e.g. ISO/IEC 17025. Method validation forms the first level of an AQA system since an analytical method must first be validated before it can be implemented for routine use. Once validated, the method can be used in routine work within the pre-established framework. However, a series of measures has to be taken at regular intervals to ensure the constancy of the results. Specific quality control samples are usually employed to monitor the performance of the routine method. As defined in the Harmonized Guidelines for Internal Quality Control in Analytical Chemistry Laboratories: 'IQC is a set of procedures undertaken by laboratory staff for the continuous monitoring of operations and the results of measurements in order to decide whether results are reliable enough to be released.' [2]. Interpreting reliability can be achieved using any of several statistical methods. One such well-known statistical tool, the Shewhart control chart [3], provides a simple plot of the concentration measured on a QC material on the *y* axis against the run number on the *x* axis (i.e. time). This decision tool found widespread application in improving both the quality of manufacturing processes and the quality of data released by analytical laboratories. The univariate Shewhart control chart requires that the IQC results tend to follow the normal distribution, characterized by a mean (m) and a standard deviation (s). Different limits are directly calculated from s. These limits are warning limits (±2s) and control limits (±3s). The theory, construction and interpretation of the Shewhart chart are detailed in many papers, books and ISO standards [4–8].

Although the Shewhart chart performs well with respect to long-term effects from random error monitoring, it lacks the sensitivity to detect systematic errors. In 1959, Roberts first introduced the exponentially weighted moving average (EWMA) control scheme [9]. Using simulation to evaluate its properties, he showed that the

☆ This paper is part of a special issue entitled "Method Validation, Comparison and Transfer", guest edited by Serge Rudaz and Philippe Hubert.

* Corresponding author. Tel.: +32 4 3663422; fax: +32 4 3664387.
E-mail address: g.eppe@ulg.ac.be (G. Eppe).

1570-0232/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2009.05.009

EWMA is useful for detecting small shifts in the mean of a process. EWMA chart can be easily superposed on Shewhart chart. A combined Shewhart-EWMA chart provides protection against both large and small shifts in a control process [10].

Modern hyphenated techniques such as gas chromatography (GC)–mass spectrometry (MS) are used to perform multi-components analyses (e.g. pesticides, veterinary drug residues, environmental contaminants). To control multi-residues analysis, usually multiple Shewhart charts will be used equivalent to the number of quality control variables. In the specific case of dioxins, the monitoring of polychlorinated dibenzo-*p*-dioxin (PCDDs), polychlorinated dibenzofurans (PCDFs) and dioxin-like polychlorinated biphenyls (DL-PCBs) in food requires the individual measurement of 29 analytes when using a congener-specific GC-high resolution mass spectrometry (HRMS) method. For IQC purposes, it is inappropriate to separately plot them on 29 individual Shewhart charts and examining them one at a time. Interpretation based on multiple rules, e.g. Westgard rules [11], only works in case of univariate independent control chart. These rules cannot provide any overall interpretation regarding the decision making as a whole. Then, the final decision would amount to a global rejection of the QC if one of the individual control charts exhibits an out-of control situation. Dioxin laboratories usually overcome the problem by applying decision rules only on the toxic equivalent (TEQ) value reported in a univariate QC chart, e.g. the sum of PCDD/Fs or the sum of PCDD/Fs and DL-PCBs. This simplified approach is considered as necessary to release results in the context of official food control [12]; but probably not sufficient owing to the TEQ value that can overlook several minors or relevant analytical problems for congeners with low toxic equivalency factors (TEFs). To improve the effectiveness of IQC procedures, multi-analytes method can be controlled by multivariate quality control charts. Hotelling introduced multivariate statistical process control and T^2 chart in his 1947 pioneer paper [13]. These statistics have been widely used in the industry for manufacturing and process control purposes. Surprisingly, few applications have emerged in chemical analytical chemistry laboratories. In 1995, the harmonized guidelines for IQC stated that multivariate approach was still a subject of research and cannot be regarded as sufficiently established for inclusion in their document [2]. Later, Nijhuis et al. applied multivariate control charts to a GC analysis of fatty acids in oil [14].

In this paper, we investigate the use of multivariate-EWMA QC charts to strengthen the IQC of PCDD/Fs and DL-PCBs monitoring in foodstuffs. We propose warning tools using the T^2 chart and EWMA control limits in addition to the TEQ univariate QC chart. In particular, we consider the issue of the number of runs necessary to characterize univariate and T^2 charts. Further, we suggest using the large quantity of QC data gathered to assess measurement uncertainty (MU). For this purpose, the MU is computed by using the accuracy profile in a retrospective way.

2. Multivariate quality control charts

2.1. Correlation between variables

The existence of correlations between the compounds measured simultaneously with a single multi-analyte GC-HRMS method is often evident by simple visual inspection of the univariate Shewhart charts. To facilitate the graphic representation of data, the following example is limited to two variables. In Fig. 1, different situations (A and B) are given for a 2-dimensional (2D) TEQ data set composed of PCDD/Fs and *non-ortho* (N-O) PCBs results. The univariate control charts of PCDD/Fs and N-O PCBs are depicted next to the vertical and horizontal-axis in order to compare the univariate process control and the multivariate approach. In both univariate

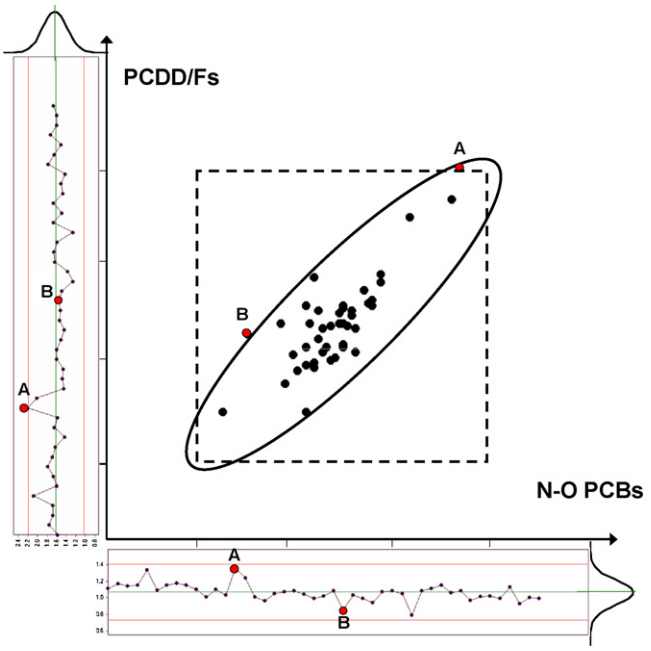

Fig. 1. An example of 2-dimensional plot observed for a pair of variables (PCDD/Fs and N-O PCBs) along with the ellipse determined from the characterization correlation estimates for these two variables.

charts, the red solid lines show the upper and lower control limits (±3s). These control limits are represented by the dashed square in 2D. In terms of process control, it would mean that the square area indicates an in-control situation when the two control charts are considered individually. The correlation structure between the two variables modifies the distribution of the data points in the square. They mainly move along the diagonal axis. The ellipse in the scatterplot represents the multivariate control limits and the area inside the ellipse is the true in-control situation. The larger the correlation between the two variables the more the ellipse is stretched and deviates from the square. In situation A, the process is out-of control in a univariate sense for PCDD/Fs fraction but in-control in the multivariate sense. Such situations can occur because the univariate approach does not account for the correlation structure in the data set. The correlation structure is not changed, the N-O PCBs increase in the same direction as the PCDD/Fs. However, point A is quite close to the edge of the control ellipse. In situation B, the reversed is observed. The process is in-control in the univariate charts while the point B falls outside of the control ellipse in the multivariate approach. It could be an assignable cause of variation or a false alarm. In this case, it is caused by an abrupt break in the correlation between the levels. The mutlivariate control chart takes into account the correlation between the levels. It is more sensitive than the univariate approach to data points that depart from the original correlation structure. A measure that takes into account multivariate covariance structure was proposed by Hotelling and is called Hotelling's T^2 [13].

2.2. Hotelling's T^2 chart

The Hotelling's T^2 chart is the most popular tool used to monitor multivariate control process. The T^2 distance is a measure that accounts for the covariance structure of a multivariate normal distribution. In most application it is assumed that the observations follow a multivariate normal distribution. When the process is in control and the in-control true parameter values (mean and covariance) are known, the chi-square distribution can be used. The

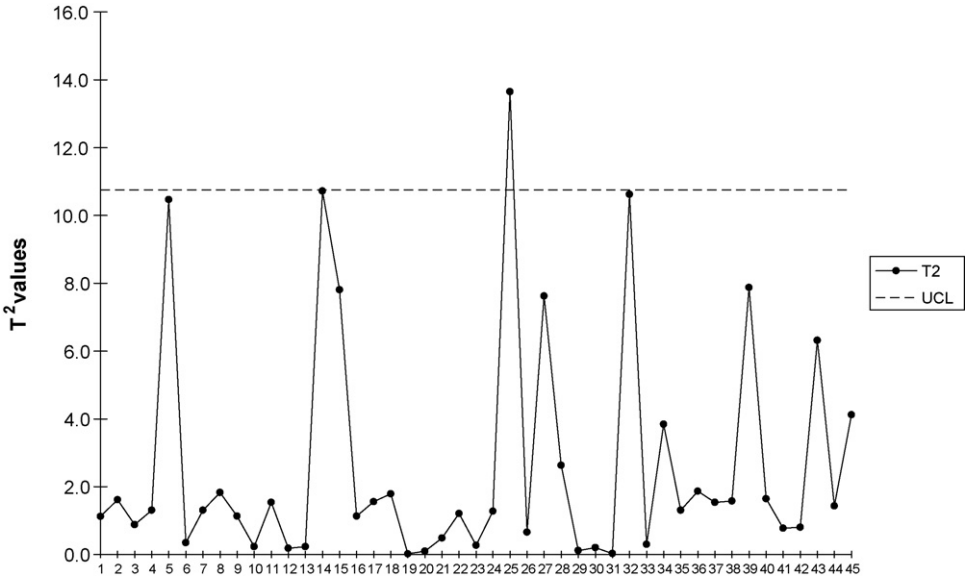

Fig. 2. Hotelling T^2 chart computed from 2 variables: the 17 PCDD/Fs and the 4 N-O PCBs.

statistic plotted on the $\chi^2_{(p)}$ control chart is given by:

$$\chi^2_{(p)} = (X - \mu_0)' \Sigma_0^{-1} (X - \mu_0) \quad (1)$$

where X is a $(p \times 1)$ vector at the nth QC sample, μ_0 is a $(p \times 1)$ vector of in-control true means of the process and Σ_0 is the $(p \times p)$ variance-covariance matrix (p = the number of variables). The upper control limit (UCL) on the chart is $\chi^2_{\alpha,p}$, where $\chi^2_{\alpha,p}$ is the (1-α)th percentile point of chi-square distribution with p degrees of freedom and α is the probability of false alarm. In practice, the process parameter values are not known, the covariance, mean, and control limit are estimated from a limited pool of data or from a preliminary phase when the process is believed to be in control. The unbiased estimates of the mean $\bar{X}$ $(p \times 1)$ vector and the $(p \times p)$ empirical covariance matrix S are given by

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad S = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X})' \quad (2)$$

respectively; n the number of replicates from the pool. For an individual observation vector X, the T^2 control chart is then constructed by using the following equation:

$$T^2 = (X - \bar{X})' S^{-1} (X - \bar{X}) \quad (3)$$

The T^2 limit plotted on the T^2 chart is given by:

$$UCL_{T^2} = \frac{p(n^2-1)}{n(n-p)} F_{\alpha,(p,n-p)} \quad (4)$$

With $F_{(p,n-p)}$ representing the F distribution with p and $(n-p)$ degrees of freedom for a suitability chosen α.

The data from Section 2.1 can be used to construct the T^2 chart by computing equation (3) and the UCL using equation (4), $\alpha = 0.01$. Fig. 2 shows the T^2 chart, the T^2 values are reported on the vertical axis against the observation number on the horizontal axis. Because the T^2 distance is always a positive number, the T^2 chart only contains an UCL. The higher the T^2, the more distant is the observation from the mean. As long as the points plotted on the T^2 control chart fall below the UCL, the process is assumed to operate under control. When one or more points exceed the UCL, the process is deemed out-of-control due to one or special causes and an investigation is carried out to detect these special causes. The main advantage is therefore its simplicity to interpret data. However, it is not a panacea. The scale of the value displayed on the chart is not related to the scale of any monitored variables. It merely provides an informative value. When T^2 exceeds the UCL, the analyst does not know which particular variable caused the break in the correlation structure between variables. It simply acts as a warning. In a second step, it requires analysis of individual curves that should clarify the origin and the nature of the interference.

Neither the univariate Shewhart QC chart nor the multivariate charts are sensitive to small shifts of the mean of a process (e.g. 1–0.5s). More advanced charts like the cumulative sum (CUSUM) and the EWMA charts attempt to detect these small shifts. One of the advantages of the EWMA chart is that it can be easily superposed on Shewhart QC chart or multivariate QC charts.

2.3. Exponentially weighted moving average (EWMA)

Briefly, EWMA averages the data in a way that gives less and less weight to data as they are further removed in time from the current measurement [15].

It is based on the recurrence formula:

$$\text{EWMA}_t = \lambda x_t + (1-\lambda)\text{EWMA}_{t-1} \quad \text{for } t = 1, 2, \ldots, n. \quad (5)$$

where

- EWMA_0 is the mean of historical data (target)
- x_t is the observation at time t, in a univariate chart
- n is the number of observations to be monitored including EWMA_0
- $0 < \lambda \leq 1$ is a constant that determines the depth of memory of the EWMA.

The parameter λ, called the smoothing factor, determines the rate at which 'older' data enters into the calculation of the EWMA statistic. A value of $\lambda = 1$ implies that only the most recent measurement influences the EWMA (degrades to Shewhart chart). Thus, a large value of λ gives more weight to recent data and less weight to older data; a small value of λ gives more weight to older data. The value of λ is usually set between 0.2 and 0.3 although this choice is somewhat arbitrary.

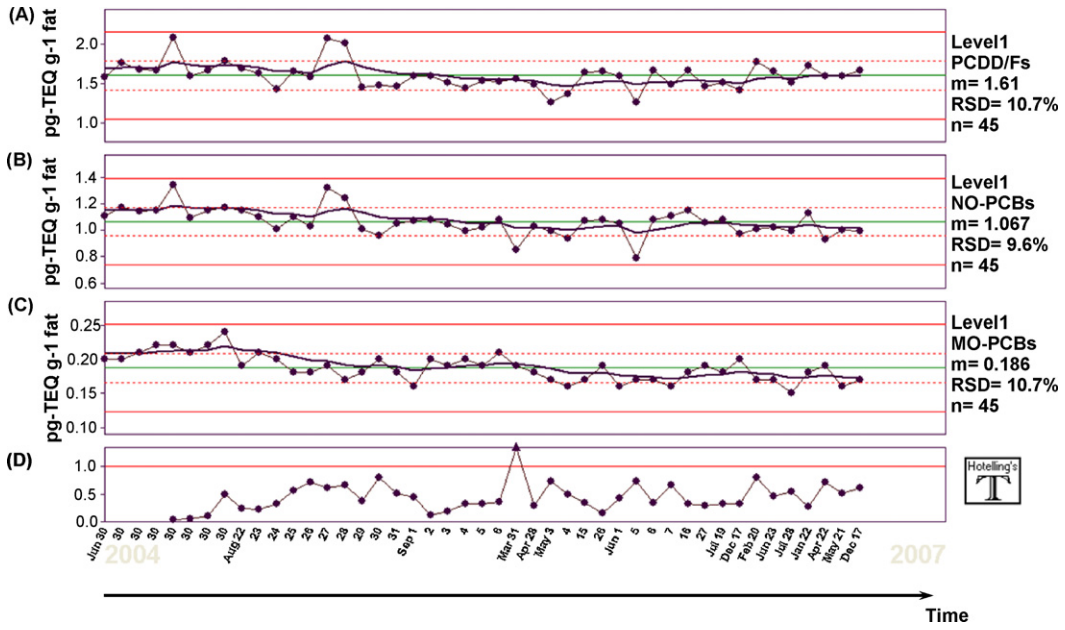

Fig. 3. Internal quality control charts for PCDD/Fs (A), NO-PCBs (B) and M-O PCBs present in QC beef fat (level 1). (D) The Hotelling *T* chart.

3. Experimental

3.1. Quality control sample

During the course of routine analysis, a procedural blank and a QC sample are included in each series of real samples. For foodstuffs, QC sample consisted of purified beef fat fortified with the 17 PCDD/Fs (EPA 1613 stock solution, Wellington Laboratories, BCP instruments, Lyon, France). The 12 DL-PCBs are from Dr Ehrenstorfer (Augsburg, Germany, concentration of 100 μg/ml). Each QC batch is prepared as follow: approximately 1 kg of beef fat was transferred into a beaker and thawed at 60 °C on a magnetic hot plate stirrer. The fat is thoroughly homogenised by stirring at 600 rpm. The 29 target compounds are spiked at the following concentrations, expressed as total TEQ: approximately 0.5× (level 1), 1× (level 2), 1.6× (level 3) and 3× (level 4) of the maximum level of 3 pg-TEQ g^{-1} fat for the sum of PCDD/Fs and 4.5 pg-TEQ g^{-1} fat for the sum of PCDD/Fs and DL-PCBs) [16]. Only pool 1 and 2 were spiked with all the DL-PCBs at the maximum level. Pool 3 and 4 are older pools and they contain only PCDD/Fs and N-O PCBs. They have been prepared when the EU legislation on DL-PCBs was not yet enforced. The QC material is homogenized for 24 h. Portions between 4 g and 7 g are bottled in umber glass vial and stored at −20 °C until analysis.

3.2. Analytical procedure

The whole congener-specific analytical procedure including extraction by pressurized liquid extraction (PLE), Power-Prep system (Fluid Management Systems, Inc. Waltham, MA) clean-up and detection by GC-ID-HRMS is exhaustively described elsewhere [17].

3.3. Statistical software

Mutlivariate quality control-EWMA charts were implemented by using the software MultiQC4.0 (http://www.multiQC.com, quality control software, Metz, France). The e.noval software was used to get the accuracy profiles as well as to compute the MU associated with the analytical method (http://www.arlenda.com, Arlenda, Liège, Belgium).

4. Results and discussion

4.1. Multivariate quality control approach

When using a congener-specific GC-HRMS method, the TEQ approach requires quantifying individually 29 dioxin-like compounds. Analysts have to accurately quantify the individual congeners in order to release a reliable TEQ content of a sample. The IQC approach proposed here consists to pool the 29 toxics congeners in 3 groups: the sum of the 17 PCDD/Fs, the sum of the 4 NO-PCBs and the sum of the 8 *mono-ortho* (MO)-PCBs, all expressed in TEQ units. There are several reasons for this choice. Firstly, multivariate approach requires that all data be observed for all variables in every run; the proposed sub-groups fulfil this requirement. Secondly, potential sources of contamination in a laboratory (samples, reagents, solvents, disposables, glassware, etc.) can be completely different for PCDD/Fs and PCBs. Thirdly, the M-O PCBs fraction is collected separately from the PCDD/F and N-O PCB fraction during the clean-up procedure [17]. Finally, the European Commission has for a transitional period set separate maximum levels for the sum of PCDD/F TEQs and for the sum of PCDD/F and DL-PCB TEQs [16]. Foodstuffs must comply with both maximum levels.

The sample preparation method considers two separate fractions requiring two separate injections on the GC-HRMS system. The TEQ content is calculated using the 1998-World Health Organisation (WHO-TEFs). Fig. 3A–C represents the univariate control charts for PCDD/Fs, N-O PCBs and M-O PCBs respectively (level 1). When starting a new QC batch, the QC chart parameters (m, s and the correlation matrix) are not known and they cannot be reliably estimated until a reference pool of data has been collected during a preliminary period. The QC chart can however start in provisional floating parameters, i.e. each new QC data recorded are used to recalculate the means (m_p), the standard deviations (s_p) and the correlation matrix coefficients, *p* the number of variables. The central green lines define the m_p values with the upper and lower $\pm 3s_p$ control limits drawn in plain (red). There is no observation outside the $\pm 3s_p$ control limits in univariate control charts. Fig. 3D depicts the corresponding Hotelling's *T* chart. One should note that the software used here reports a relative *T* value (ratio of T/T_{limit}) instead of a T^2. Interpretation of data is based on the same rules as defined in

Table 1
Relative bias, intermediate precision, combined standard uncertainty, expanded standard uncertainty and relative expanded standard uncertainty computed for PCDD/Fs from 4 levels of QC beef fat test material.

Level	Spiked concentration pgTEQ g^{-1} fat	Measured mean level pgTEQ g^{-1} fat	Relative bias (%)	Intermediate precision SD pg-TEQ g^{-1} fat	Method bias uncertainty pg-TEQ g^{-1} fat	Combined uncertainty pg-TEQ g^{-1} fat	Expanded uncertainty (k = 2) pg-TEQ g^{-1} fat	Relative expanded uncertainty (k = 2) (%)
1	1.69	1.61	−4.9	0.163	0.019	0.164	0.327	19.4
2	3.04	3.06	0.5	0.225	0.028	0.227	0.453	14.9
3	5.00	4.87	−6.6	0.431	0.051	0.434	0.868	17.4
4	10.75	10.81	0.6	0.910	0.107	0.916	1.832	17.0

Section 2.2 but it provides here a unique UCL value of 1. In floating mode, the first T value requires at least $p + 1$ data point to be computed. During the first 24 measurements, one can observe the high correlation coefficient between PCDD/Fs and N-O PCBs ($r_{12} = 0.958$). The r_{13} and r_{23} values are lower than 0.5 between PCDD/Fs and M-O PCBs and between N-O PCBs and M-O PCBs, respectively. The main reason of low correlation coefficients is related to the separated GC-HRMS injections. The tuning and the calibration are performed separately on each instrument. An out-of-control signal is observed at the 25th data point in Fig. 3D. In such a case, the interpretation of the T chart is straightforward. It is caused by an abrupt break in the correlation structure between PCDD/Fs and N-O PCBs. However, identifying an out-of-control variable from a T^2 chart is not always as simple, specifically when a large number of variables are monitored ($p > 5$). Whenever new data are validated in floating mode, new averages, standard deviation and correlations between variables are re-calculated with an additional degree of freedom. Thus, it gives a r_{12} value of 0.881 between PCDD/Fs and NO-PCBs by including the 25th point in the data set. The slight drop in the correlation coefficient decreases the sensitivity of the T chart to detect breaks in the correlation structure between variables. This explains why, on several occasions later, some additional points are not in out-of-control situations (e.g. the 32th, 34th and 39th points). Once the 45 data points are validated r_{12} equals to 0.8 (r_{13} and $r_{23} < 0.5$). It could be hazardous to keep floating parameters too long. One can easily imagine how a slow drift could affect the mean and the control limits of univariate charts. It is necessary to lock m_p, s_p and the correlation coefficients as soon as data from a stable reference period have been collected. One of the key issues in constructing multivariate charts are therefore to determine the minimum number of runs needed to correctly assign the working values of m_p, s_p, and the correlation coefficients. For example, if the first 20 runs were selected to calculate the working values, five points (25th, 32th, 34th, 39th and 45th) would have been considered as out-of-control situations on the T chart (Fig. 3). However, most of these warnings could be seen as false alarms by visual inspection of univariate charts. Marquis studied the probability of false alarm outside the range [m ±3s] by numerical simulation [18]. He concluded that a preliminary period corresponding to 100 runs should be sufficient to guarantee an acceptable range of false alarm [0.06–1%], theoretically 0.27%. He also demonstrated that s is the predominant factor compared to m that affects the probability of false alarm in univariate charts. In such a case, it would mean in Fig. 3 that only half of the necessary runs would have been recorded to adequately characterize our method. It is clear that increasing the preliminary period to hundred data points is practically and economically unrealistic to apply in contaminants and residues analytical laboratories. In specific cases, it is however possible to solve the problem without any additional costs. Indeed, it is quite common in many analytical disciplines, including dioxin analysis, to observe that s_p is proportional to level as shown in Table 1 where the intermediate precision is calculated for the sum of PCDD/Fs. It follows that relative standard deviations (RSDs) are rather constant across the entire concentration range (level 1–4). The RSDs can be pooled and an estimate of the pooled RSD is obtained using the following equation:

$$\mathrm{RSD}_{p,\mathrm{pool}} = \sqrt{\left(\frac{(n_{p,1}-1)\mathrm{RSD}_{p,1}^2 + (n_2-1)\mathrm{RSD}_{p,2}^2 + \dots}{(n_{p,1}-1) + (n_{p,2}-1) + \dots.}\right)} \quad (6)$$

where $\mathrm{RSD}_{p,\mathrm{pool}}$ is the RSD pooled for the pth variable, $\mathrm{RSD}_{p,1}$ is the relative standard deviation of pth variable calculated from the QC sample at level 1, $n_{p,1}$ is the number of replicates for that QC sample, etc.

In this study, 4 levels of QC test materials were prepared. By selecting the number of runs at 25 per level, it provides an estimated $\mathrm{RSD}_{p,\mathrm{pool}}$ with 96 degrees of freedom. The control limits of the QC charts are then constructed with:

$$[\mathrm{m}_{p,i} \pm 3 \times \mathrm{RSD}_{p,\mathrm{pool}} \times \mathrm{m}_{p,i}] \quad (7)$$

where $m_{p,i}$ is the average of the pth variable at the ith QC level. Fig. 4 illustrates the 4 levels of QC charts for the PCDD/Fs where the $\pm 3s_{\mathrm{pool}}$ control limits are constructed from a $\mathrm{RSD}_{\mathrm{pool}}$, 25 runs per level. The correlation matrix must also be reliably estimated. Each level of monitoring provides an estimate of the correlation coefficients. For example, r_{12} between PCDD/Fs and N-O PCBs equals to 0.88; 0.67; 0.77; 0.74 at levels 1; 2; 3; 4, respectively (25 runs per level). The higher value at level 1 (Fig. 3) cannot be explained. It has been observed during 25 runs and this is not a coincidence but it seems totally artificial. On the light of other r values, a r_{12} value of 0.75 reflects a better approximation of the correlation between PCDD/Fs and N-O PCBs when taking into account the 4 levels of monitoring.

At this stage, a T chart can be constructed for each level of monitoring. It requires $m_{p,i}$, $\mathrm{RSD}_{p,\mathrm{pool}}$ and the estimated correlation matrix. The QC material is inserted at a frequency of one per ten routine samples. The level of the QC material is selected on a random basis. However, it should match as much as possible the concentration of routine samples. The T value is calculated and used as a warning. Once a T value is above its UCL, the cause of a possible problem has to be investigated promptly. The first action consists to analyse the univariate QC charts of the 3 pooled TEQ variables to determine which sub-group of the TEQ is incriminated. The $\pm 3s_p$ decision rule is directly applied to decide whether the series of routine samples can be released or not. One can note that a T value above its UCL is not only caused by a break in the correlation structure between the variables, it could also happen when the 3 variables deviates from the mean in the same direction. In that specific case, the T chart is less sensitive compared to the univariate QC charts.

Even if the results can be released, the origin of an out-of-control T value must be found. When a sub-group of the TEQ has been identified, the analyst can go further by analysing individual congeners. One or several congeners may be the cause of the out-of-control situation. The following features should be sought: high procedural blank level, internal standard recovery rates, GC peak shape and integration, retention time, isotope ratio, relative response factors. Based on our experience, most answers to an out-

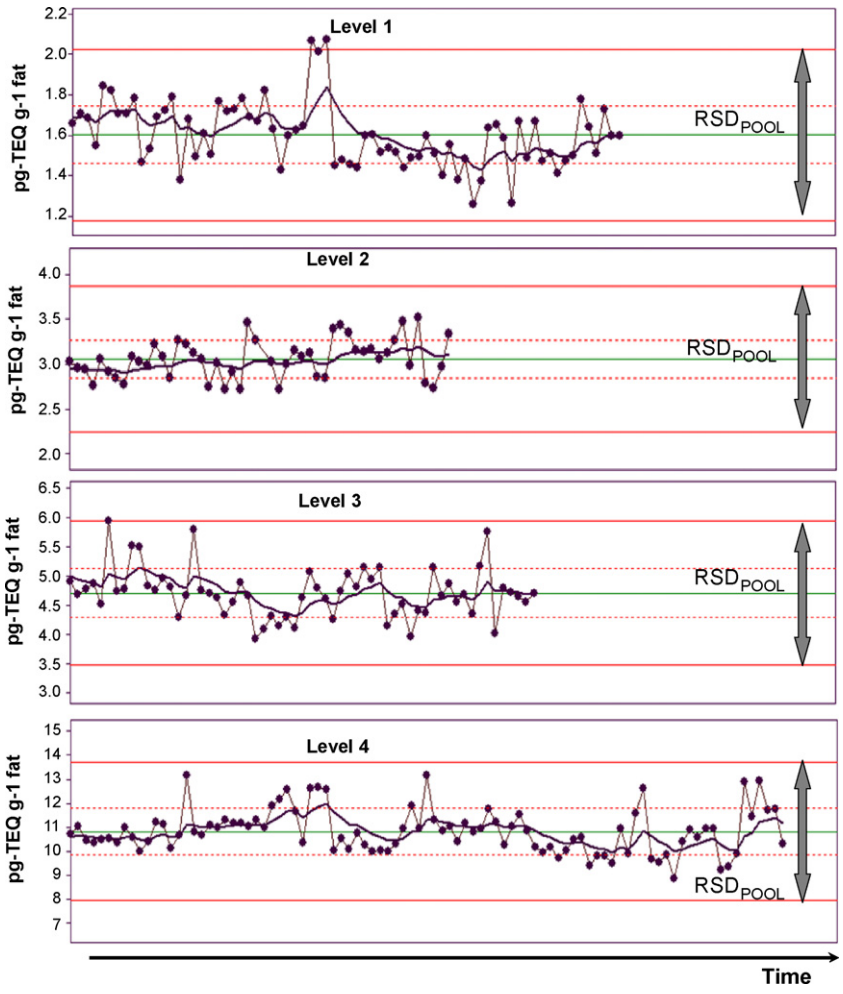

Fig. 4. Control charts at 4 different levels (1.69 pg-TEQg^{-1} fat; 3.04 pg-TEQg^{-1} fat; 5.00 pg-TEQg^{-1} fat; 10.00 pg-TEQg^{-1} fat, for the sum of the 17 PCDD/Fs). RSD$_{p,pool}$ calculated with $n = 25$ per level, 4 levels of QC.

of-control situation are found in these listed features. However, if these actions yield no insight, then further QC material measurements are needed.

4.2. Trend analysis

To detect small shifts in QC process and to support trend analysis, EWMA charts can be superposed on univariate QC charts. Figs. 3 and 4 illustrate the EWMA chart by the bold curve and its control limits (dashed lines, m $\pm 3s_{EMWA}$). The smoothing factor λ is set at 0.2 and the relationship between s and s_{EMWA} is expressed by the following equation:

$$s_{EWMA} = s\sqrt{\frac{\lambda}{2-\lambda}} \quad (8)$$

By setting $\lambda = 0.2$, Equation (8) gives a s_{EMWA} equals to one third of s. Thus, the control limits of the EWMA curve (dashed lines) represent a range between ± 1s. The EWMA monitoring chart provides an additional guidance on setting an appropriate criterion to support trend analysis: the method bias should lie between the control limits of ± 1s with a λ value of 0.2. A significant method bias is observed once the bold curve is beyond the dashed control limits of ± 1s. Moreover, the EWMA curves should be randomly distributed on both sides of the mean indicating that no systematic method bias was observed during the whole period. For instance, in Fig. 4, the EWMA curves fall within the dashes lines demonstrating that the method bias does not exceed the RSD$_{pool}$ of 8.8%, except for a short period of time in level 1. One can observe 2 out of 3 QC data beyond the control limits of ± 3s, leading to an EWMA curve exceeding the ± 1s limits. An efficient and appropriate means to evaluate the relevance of the proposed criterion is first to participate in proficiency tests (PTs) and evaluate the performance your laboratory obtains. It provides relevant information in terms of trueness assessment of the analytical method (laboratory and method bias) and it gives you the opportunity to adapt the degree of stringency of your internal QC criteria.

4.3. Quality control data and measurement uncertainty

It has already been reported that the use of QC data are an appropriate way to evaluate MU [19]. This 'top-down' approach takes into account most of the relevant contributions to MU by performing long term precision and trueness studies. In the context of contaminants in food, strict regulations are enforced in Europe [16]. MU clearly has implications for interpreting analytical results for compliance with maximum limits. In the specific case of dioxin-like compounds, measurement is given in TEQs for comparison with maximum levels, i.e., the uncertainty associated with a measurement must also be stated in TEQs.

The large quantity of data produced by the implementation of multivariate and multilevel IQC procedures is shrewdly used to estimate the MU. The study case presented here deals with data collected for the sum of PCDD/Fs; it can be easily extended to

DL-PCBs. The approach only considers the uncertainty associated with the analytical procedure. In particular, it takes into account the repeatability effects, the intermediate precision effects and the method bias (spiking study). The laboratory bias is not assessed as it requires a trueness study carried out with certified reference materials (CRMs). The uncertainties related to sampling, sample homogeneity or stability are not discussed here.

There are different ways and manners to estimate MU. A practical and direct way of using the data collected during in-house validation step to estimate MU can be deduced from the accuracy profile and the concept of total error [20]. Feinberg et al. demonstrated that the standard deviation used to compute the β-expectation tolerance interval is nothing more than the standard combined uncertainty as defined in the ISO/TS 21748 guide [21]. The tolerance interval calculated for each concentration level gives an estimate of the expanded uncertainty. Considering the multilevels QC data gathered for the sum of PCCD/Fs (Fig. 4), the accuracy profile is computed in a retrospective way for levels 1 to 4. The interval constructed here is a confidence interval instead of a tolerance interval, characterized with a large number of j series (between-runs) of one replicate. Table 1 gives an overview of the statistical parameters computed to estimate MU. The relative method bias varies between −6.6% and 0.6% within the working range and meets the Commission Regulation criteria, i.e. ±20% [12]. The contribution of the method bias to the overall uncertainty can be neglected compared to the intermediate precision (s_R). Indeed, the uncertainty associated with the estimated method bias is proportional to $s_R \times j^{-1/2}$ ($n = 1$) and tends to be negligible when the number of between-runs series increase [21]. The expanded uncertainty (U) is calculated by using a coverage factor (k) of 2 at 95% level of confidence [12]. The relative U for the sum of PCDD/Fs expressed in TEQ varies between 15% and 19% within the working range. When performing precision and trueness studies from validation data, we have already reported relative U between 15% and 24% associated with the TEQ value for the sum of the PCDD/Fs [22,23]. As already mentioned, the estimated MU reported here does not take into account the laboratory bias. Previous internal trueness studies with matrix match CRMs showed a contribution of the bias uncertainty (laboratory and method bias) between 23% and 65% to the overall uncertainty budget, giving a contribution between 4% and 16% to the combined standard uncertainty [23].

The expanded uncertainties at the levels 1, 2 and 3 around the maximum limit of 3 pg-TEQ g^{-1} are graphically illustrated in Fig. 5. The limits of ±U at 95% level of confidence are represented by the 2 sides of the narrow white path and the associated curves indicate the inferred probability density function for the value of the measurand. The upper and lower limits of U are simply connected by straight lines in order to interpolate the behaviour of the limits between the 3 measured levels and to support the graphical representation.

Decision rules are needed in the view of acceptance or rejection of a sample or a lot. In the specific framework of dioxin control in foodstuffs, the decision rule implies to subtract the expanded uncertainty (U) from the measurement result (x-U) and compare this with the maximum limit [16]. If (x-U) exceeds the maximum limit, it will result in a decision of non-compliance. On the basis of this decision rule, an acceptance and a rejection zone are determined as shown in Fig. 5. If the measurement result lies in the acceptance zone the product is declared compliant and if in the rejection zone it is declared non-compliant [24]. Thus, the decision rule for the PCDD/Fs implies to extend the acceptance zone beyond the maximum limit, i.e. maximum limit plus U at 95% level of confidence. It corresponds to the dashed B line in Fig. 5. The region located between A and B is usually more complicated for decision-making. Decision rules may include additional measurements or sharing the risk between the laboratory and the end-user of the data. The issue

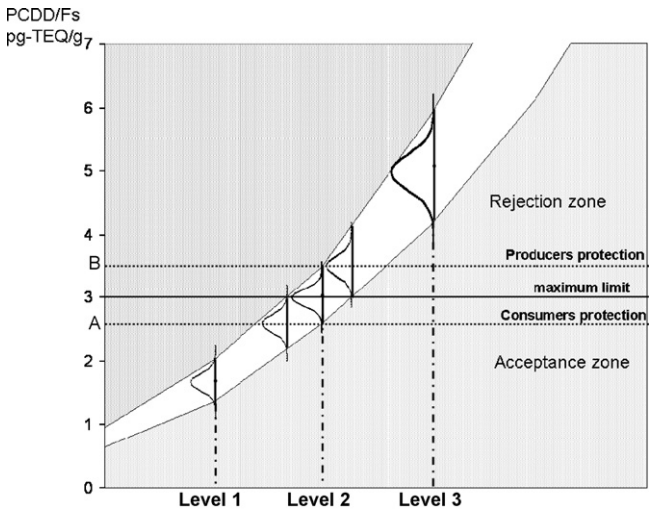

Fig. 5. Graphical display of expanded measurement uncertainty calculated from QC data at 1.69 pg-TEQ g^{-1} fat (level 1); 3.04 pg-TEQ g^{-1} fat (level 2) and 5.00 pg-TEQ g^{-1} fat (level 3) for the sum of the 17 PCDD/Fs across the maximum limit of 3 pg-TEQ g^{-1} fat.

of compliance assessment within that zone lends itself to debate. To avoid any discussions and misinterpretation, European dioxin regulation considers the whole region between A and B as an acceptance zone. This decision rule is clearly in favour of food producer's interest. It should however be noted that concentrations between 1.5 pg-TEQ g^{-1} fat (so called action level) and B correspond to a subregion within the acceptance zone where actions have to be taken to identify the source(s) of contamination [25].

5. Conclusions

In this paper, multivariate and multilevel quality control charts are studied for the analysis of 29 dioxin-like compounds in foodstuffs by GC-HRMS. We proposed to pool the 29 congeners in 3 groups and to monitor their toxic equivalent values in univariate charts and multivariate T^2 chart. The T^2 chart acts as a warning to trigger further investigations in the data set when a T^2 value exceeds the upper control limit. The ±3s decision rule remains applicable in univariate charts to release results. One of the criticisms that can be addressed to the proposed approach is its lack of sensitivity for the congeners that contribute to a lesser extent to the TEQ value. If analytical problems occur for those congeners, it is left to the analyst opinion to decide whether extensive investigations and works are needed. This decision, however, has to take into account the method's fitness-for-purpose requirements.

The EWMA chart provides an additional guidance on setting appropriate criterion to control the method bias and for trend analysis. It is a useful tool to support trend analysis according to ISO/IEC 17025 requirements. The number of trends analysis to carry out during a year depends on the frequency of QC materials used. With the approach described in this paper, twice a year allows recording relevant information.

Implementing IQC procedures in a laboratory has a non-negligible cost in the overall price of analysis. The multilevel QC data generated for monitoring can be used to estimate measurement uncertainty across the maximum limits for compliance assessment. The sources of uncertainty covered by the study must be specified.

Acknowledgements

The authors acknowledge the University of Liège, the Walloon Region, the European Regional Development Fund (ERDF) and the National Fund for Scientific Research (FNRS) for financial support.

References

[1] I. Tavernier, M. De Loose, E. Van Bockstaele, Trends Anal. Chem. 23 (2004) 535.
[2] M. Thompson, R. Wood, Pure Appl. Chem. 67 (1995) 649.
[3] W. Shewhart, The Economical Control of Quality of Manufactured Products, D. Van Nostrand, New York, NY, 1931.
[4] ISO 8258, Shewhart control charts, Geneva, 1991.
[5] ISO 7873, Control charts for arithmetic average with warning limits, Geneva, 1993.
[6] ISO 7870-1, Control Charts part 1: General Guidelines, Geneva, 2007.
[7] ISO 7966, Acceptance control charts, Geneva, 1993.
[8] S. Levey, E.R. Jennings, Am. J. Clin. Pathol. 20 (1950) 1059.
[9] S.W. Roberts, Technometrics 1 (1959) 239.
[10] J.M. Lucas, M.S. Saccucci, Technometrics 32 (1990) 1.
[11] J.O. Westgard, P.L. Barry, M.R. Hunt, Clin. Chem. 27 (1981) 493.
[12] Commission Regulation (EC) No 1883/2006 of 19 December 2006 laying down the sampling methods and the methods of analysis for the official control of dioxins and the determination of dioxin-like pcbs in foodstuffs, Official Journal of the European Union L364/32 (2006).
[13] H. Hotelling, in: C. Eisenhart, M.W. Hastay, W.A. Wallis (Eds.), Multivariate Quality Control—Techniques of Statistical Analysis, MacGraw-Hill, New York, 1947, p. 111.
[14] A. Nijhuis, S. de Jong, B.G.M. Vandeginste, Chemometr. Intell. Lab. Syst. 47 (1999) 107.
[15] A.S. Neubauer, Clin. Chem. 43 (1997) 594.
[16] Commission Regulation (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs, Official Journal of the European Union L364/5 (2006).
[17] J.-F. Focant, G. Eppe, C. Pirard, E. De Pauw, J. Chromatogr. A 925 (2001) 207.
[18] P. Marquis, Ann. Biol. Clin. 59 (2001) 214.
[19] EURACHEM/CITAC Guide, Quantifying uncertainty in analytical measurement, 2nd ed., 2000, http://www.eurachem.org.
[20] M. Feinberg, B. Boulanger, W. Dewé, Ph. Hubert, Anal. Bioanal. Chem. 380 (2004) 502.
[21] ISO 21748, Guidance for the Use of Repeatibility, Reproducibility and Trueness Estimates in Measurement Uncertainty Estimation, Geneva, 2004.
[22] G. Eppe, E. De Pauw, Organohalogen compounds 59 (2002) 403.
[23] G. Eppe, Thesis, ISSN 0075-9333, http://bictel.ulg.ac.be/ETD-db/collection/available/ULgetd-04022008-071117/, 2007.
[24] EURACHEM/CITAC guide, Use of uncertainty information in compliance assessment, 1st ed., 2007, http://www.eurachem.org.
[25] Commission Recommendation No 2006/88 of 6 February 2006 on the reduction of the presence of dioxins, furans and PCBs in feedingstuffs and foodstuffs, Official Journal of the European Union L42/26 (2006).