JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

Comparative Study of the Main Top-down Approaches for the Estimation of Measurement Uncertainty in Multiresidue Analysis of Pesticides in Fruits and Vegetables

Paula Medina-Pastor,[§] Antonio Valverde,[#] Tuija Pihlström,[⊥] Sonja Masselter,[⊗] Miguel Gamon,[§] Milagros Mezcua,[§] Carmelo Rodríguez-Torreblanca,[△] and Amadeo R. Fernández-Alba*[,§]

[§]European Union Reference Laboratory for Pesticide Residues in Fruits and Vegetables (EURL-FV) (DG SANCO), Pesticide Residues Research Group at the Universidad de Almeria and the Laboratorio Agroalimentario de la Generalitat Valenciana, 04120 Almeria, Spain

[#]Pesticide Residues Research Group, Universidad de Almeria, 04120 Almeria, Spain

[⊥]National Food Administration, Box 622, 75126 Uppsala, Sweden

[⊗]AGES − Competence Center for Residues of Plant Protection Products, 6020 Innsbruck, Austria

[△]Statistics and Applied Mathematics Department, Universidad de Almeria, 04120 Almeria, Spain

ABSTRACT: Practical "top-down" approaches appear to be the most suitable for the evaluation of measurement uncertainty in pesticide residue testing laboratories, where analytical procedures are routinely applied to a large number of pesticide/food combinations. The opposite approach, "bottom-up" evaluation of measurement uncertainty, leads to great difficulties in evaluating all of the pesticides in a consistent way. Among the top-down approaches, there are two main ways in which measurement uncertainty can be estimated: One is based on default values, which are based on previous extensive interlaboratory experience and the proven accuracy of the laboratory; these include the Horwitz equation or the fit-for-purpose relative standard deviation (FFP-RSD). The other is based on experimental data from the quality control work of the laboratory: within-laboratory reproducibility, interlaboratory validation, or a combination of results obtained in proficiency tests. The principal existing guidelines from various bodies (Eurachem, Nordtest, and Eurolab) all propose different approaches for calculating measurement uncertainty. In this paper, the main top-down approaches are evaluated and compared using the data from the European Proficiency Test Database for Fruits and Vegetables and the Multiresidue Method validation databases obtained from the National Reference and Official Laboratories in Europe. The main conclusion of the comparative study is that a default expanded measurement uncertainty value of 50% could satisfy all of the requirements for facilitating and harmonizing, worldwide, the intercomparability of the pesticide residue confidence results between laboratories.

KEYWORDS: uncertainty estimation, top-down approach, multiresidue method, pesticide residue, harmonization

INTRODUCTION

Uncertainty is a parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand.[1]

To evaluate the pesticide analysis dispersion in the most adequate and cost-effective way, it is important to establish a consensus on the possible ways of calculating measurement uncertainty in multiresidue methods to facilitate their comparability.

The harmonization of these methods of evaluating measurement uncertainty would lead to a set practice dealing with MRL exceedance worldwide. Detection of violative samples is linked to measurement uncertainty consideration as presented in Figure 1. A well-accepted fixed measurement uncertainty applied in laboratories will facilitate trade worldwide.

There are different guidelines worldwide summarizing how measurement uncertainty should be calculated in different situations. In all of them, the way in which combined relative standard measurement uncertainty is obtained for different methodologies is explained. Almost all use a coverage factor (k) of 2 to obtain the expanded measurement uncertainty.[1] A coverage factor of 2 expresses a 95% confidence interval. The bigger this coverage factor, the greater the confidence given to the uncertainty measurement.

The different ways of calculating measurement uncertainty mentioned in the guidelines are for generic situations, allowing them to be applied to the most diverse examples. The GUM guideline[1] describes extensively the bottom-up approach or strict mathematical calculation based on the estimation of all separately measured uncertainty components and combining them by applying the propagation law. However, separate evaluation of all the measurement uncertainty components in a multiresidue method (MRM) for pesticide results is a very tedious process, with many calculations, leading to laborious approaches for the routine laboratories. Many authors have already presented their calculation procedures as bottom-up simplified approaches.[2−5]

Special Issue: Florida Pesticide Residue Workshop 2010

Received: October 17, 2010
Revised: November 21, 2010
Accepted: November 23, 2010
Published: December 14, 2010

dx.doi.org/10.1021/jf104060h | *J. Agric. Food Chem.* 2011, 59, 7609–7619

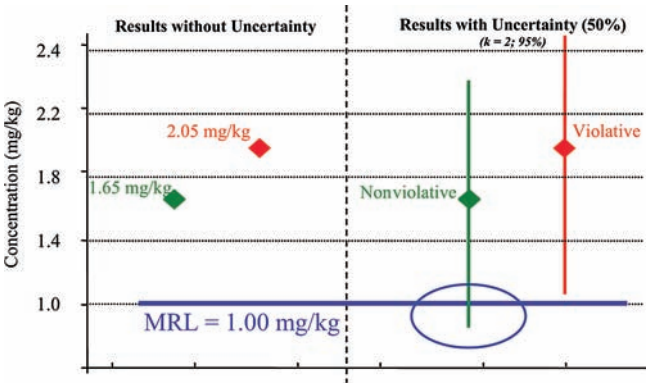

Figure 1. Example of violative and nonviolative samples before and after applying uncertainty intervals.

Top-down approaches are evaluated either by scientific judgments based on assumptions that simplify the process or by empirical data. There is a clearly accepted trend for such approaches that corresponds with the recent trend MRMs have followed over recent years allowing simultaneous analysis of hundreds of compounds.

Chronological guidelines have introduced different top-down approaches as the way to estimate measurement uncertainty:

EURACHEM Guideline.[6] Although in the first edition (from 1995) the measurement uncertainty calculus was based on the error propagation equation, in the second editon (from 2000), it incorporated the idea that chemistry laboratories should base the measurement uncertainty estimation on practical experiences as well as introduce formal quality assurance procedures in the calculus.

The NORDTEST[7] (from 2004) calculated measurement uncertainty on the basis of two components. The first component is the measurement uncertainty from the within-laboratory reproducibility (R_w) taken either from stable control samples covering the whole analytical process at both a low and a high concentration level or by control samples not covering the whole analytical process at varied concentrations or by unstable control samples. The second component is the measurement uncertainty from method and laboratory bias, $u'(\text{bias})$, which is estimated either from certified reference material, interlaboratory comparisons, or recoveries. The two components are then used following eq 1.

$$u' = \sqrt{u'(R_w)^2 + u'(\text{bias})^2} \quad (1)$$

Within-laboratory reproducibility (R_w) is the measure between repeatability and reproducibility, where operator and/or equipment and/or time and/or calibration can be varied, but within the same laboratory, in other words, intermediate precision, whereas (bias) is the difference between the mean measured value from a large series of test results and an accepted reference value (a certified or nominal value).

EUROLAB[8] (from 2007) based the measurement uncertainty calculus on the dispersion of the relative difference of the results given by a laboratory on different PT schemes. In this guideline it is assumed that the standard deviation (SD) of the relative differences of the results given on different proficiency tests for a specific class of pesticides is the relative standard measurement uncertainty (u').

The NATA guideline[9] concluded in 2009 that although GUM covers only physical measurements, the introduction of internal and external data from validation and quality control interlaboratory studies on chemical measurements maximizes the probability of including all potential contributions to the measurement uncertainty estimation. This guideline uses the Horwitz equation (eq 2) for calculating interlaboratory study relative standard deviation as the relative standard measurement uncertainty (u').

$$\sigma = 0.22c^{0.8495} \quad (2)$$

σ is the relative standard deviation, and c is the concentration.

Laboratories being accredited under ISO 17025[10] must estimate the measurement uncertainty of the results they submit, so laboratories base their calculus on the above-described guidelines to estimate this measurement uncertainty. In the case of pesticide residue determination, the laboratories must base their calculus on multiresidue methods most of the time.

Measurement uncertainty (MU) in pesticide residue analysis has been a discussion topic in the Codex Alimentarius Commission meetings for some years now. The Codex approved a Guideline on the Estimation of Uncertainty[11] in 2006. There, three major phases of analytical chemistry MU determination were established: sampling, preparation of test portions, and analysis. It was suggested that due to the high number of possible methods, analytes, and matrix combinations in MRM, laboratories could use a properly selected range of analytes and sample matrices which represent the residues and commodities to be analyzed in terms of their physicochemical properties and composition from validated data and long-term precision, thus facilitating uncertainty measurement. Interlaboratory tests are considered to be a useful tool in estimating the between-laboratory variability of the data, providing a reliable estimate of the method performance and the measurement uncertainty associated with their application. In this guideline, measurement uncertainty values for pesticide concentration ranging from 1 μg/kg to 1 mg/kg are expressed either for the within-laboratory reproducibility in the range of 16−53%[12] or for the average between-laboratory reproducibility set at 25%.[13]

The establishment of the standard deviation of a series of tests run by a single laboratory, as a measure of standard uncertainty, requires the results from a large data set that is not always available. The true standard deviation (σ) can be estimated from the sample standard deviation (S). To settle what numbers of results are required for MU estimation, taken either from validation results or how many batches of analysis from recovery data or from PT results, Table 1[11] shows the expected ranges of standard deviation for values of N at 95% probability. Thus, for an accepted error of σ^2 of 25%, that is, error limits of $0.75\sigma^2$ and $1.25\sigma^2$, the value of N should be 31. This will mean that the minimum number of results that should be taken into account for the measurement uncertainty estimation should not be lower than 31.

Later, in mid-2009, the Codex Alimentarius Commission[14] agreed that rigorous bottom-up approaches based on the calculation of individual values for countless commodity/pesticide combinations were, in general, impractical for MRM MU calculation. It was decided then that estimation based on method validation, quality control, and proficiency testing results was more appropriate and realistic for MRM purposes. Estimation of measurement uncertainty should consider the complexity of pesticide residue analysis; therefore, a straightforward guidance

Table 1. Expected Ranges of Standard Deviation for Different Numbers

N	$S_{min} = f_1\sigma$ f_1	$S_{max} = f_2\sigma$ f_2
5	0.35	1.67
7	0.45	1.55
15	0.63	1.37
31	0.75	1.25
61	0.82	1.18
121	0.87	1.13

based on an empirical top-down approach and ISO 17025 compliance needed to be prepared, not taking MU derived from sampling into account. The EU Codex Pesticide Residue Working Group came up with a proposal for calculating MU based on the European Proficiency Test (EUPT), introducing the idea of accepting 25% fit-for-purpose variability as a standard deviation leading to the assumption of ±50% expanded MU as the default value. Although some restrictions or different approaches were mentioned for laboratories applying this fixed value, it was an important step in simplifying and harmonizing measurement uncertainty.

The presented work discusses both the advantages and disadvantages of applying the main top-down approaches for calculating MU in pesticide residue analysis. These approaches take datafrom a default value, either based on interlaboratory data gained over the years, alongside the proven accuracy of the laboratory such as in the Horwitz equations or in the fit-for-purpose standard deviation (FFP-RSD) or, alternatively, based on experimental data focused mainly on the quality control of the laboratory from repeated analyses of spiked samples, within-laboratory reproducibility, or interlaboratory validation data.

Important focus is put on the harmonizing of the MU calculus in multiresidue methods because this will ensure compliance with residue legislation at an international level. This is a relevant point in the case of enforcement but not equally relevant to the case of risk assessment and its objectives.

RESULTS AND DISCUSSION

Influence of Data Dispersion on the Different Pesticides, Commodities, and Techniques Used. Conceptually, the measurement uncertainty value is associated with a single pesticide/commodity/concentration level result. As a consequence of the changes in the analytical method applied or its efficacy under different experimental conditions, such as with a change in the matrix, for example, this could lead to different MU values. It might lead to serious conflicts in pesticide residue laboratories where the number of commodity/residue combinations being evaluated may easily reach values in the thousands. Therefore, calculating MU values for every combination is, in general, impracticable.

For that reason, it is important to know to what degree those variations can differ according to personal skills and knowledge. To do this, the effects of varying the extraction solvent and the method techniques on the commodity/residue combination are evaluated in detail. This evaluation is applied to the EUPT database[15] with more than 12000 values. These values come from a variety of laboratory performances, not being selected upon their quality.

The evaluation of the EUPT database shows that personal skills can overcome commodity/residue combination problems by appropriately applying the correction factor. For that reason, it could be considered that if the knowledge and skills of the laboratory staff are good, then the estimation of MU calculus will not solely be dependent on the extraction method or the technique used.

Figure 2 shows how pesticides that are typically determined by gas chromatography, as is endosulfan, for example, showed no significant differences in the z score values with different extraction solvents used over two different years. Table 2 shows that for all the methods used, most of the absolute z score results achieved are ≤ 2.

There are four main extraction methods implemented for MRM, based on the EUPT-FV results database. Given that methods may suffer slight variations according to the laboratory's internal validation, work experience, or implementation of updated official methods, for clarity's sake, the methods commented on here will be simplified to the extraction solvent used instead of the method's name. These extraction solvents are acetone, based mainly on mini-Luke's method,[16] although Spetch[17] modification substituting dichloromethane with ethyl acetate and cyclohexane is also considered as part of this solvent. Acetonitrile solvent mainly refers to the QuEChERS[18] method including the main variations.[19] Ethyl acetate mainly refers to the developed National Food Administration (Sweden) method[20] and, finally, methanol extraction solvent.[21]

With regard to pesticides amenable to liquid chromatography and taking methomyl as an example, Figure 3 shows the different solvents used for its determination over 2 years. From this figure, no solvent clearly improves the performance of the laboratory in any of the tests (see also Table 3).

It can be confirmed that the solvent used does not result in significant differences in the z score results (achieved by conducting a simple ANOVA test). This study of the variance has been performed on all of the pesticides present over the past four EUPTs (EUPT-FV 8−11). None show significant differences. Another example of a specific nondependent pesticide can be seen in Figure 4.

From Figure 4, it can be seen that in the z score representation, whatever the solvent used represented by a color, the z score bar gives good results. The same is deduced from the confidence percentage interval graph that comes from the ANOVA test. As none of the bars overlap, this demonstrates that there are no significant differences using one solvent or another.

Moreover, taking the past four EUPT-FV evaluations mentioned into consideration, the number of unacceptable z scores (that is, results above an absolute z score value of 3.0) is constant, independent of the solvent used. A summary of these results is given in Table 4.

Another possible variation in the MRM results, caused by the commodity/residue interaction, can be the determination technique. For an experienced laboratory, the choice of the determination technique for a pesticide does not influence the results. It is well-known that some pesticides are amenable to only one specific technique, such as endosulfan or methiocarb sulfone. Others, which are sensitive to both techniques, do not influence the results achieved by the laboratories, as can be seen in Figure 5 and Table 5. Figure 5 shows the 103 laboratories' z score calculations for methamidophos in EUPT-FV10 for carrot matrix. Each of the z score results is plotted and colored depending on the determination technique used by each

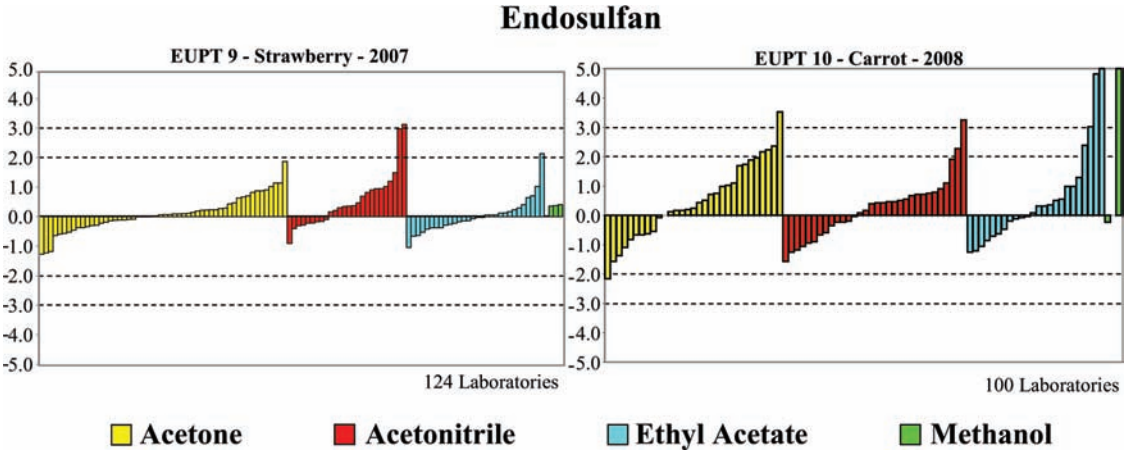

Figure 2. Different solvents used for the determination of endosulfan over a 2-year period.

Table 2. Number of *z* Score Results According to the Extraction Method Used in EUPT 9 and 10, Corresponding to Figure 2

	no. of *z* scores classified by extraction method used for EUPT 9				no. of *z* scores classified by extraction method used for EUPT 10			
classification	acetone	acetonitrile	ethyl acetate	methanol	acetone	acetonitrile	ethyl acetate	methanol
$\|z\ \text{score}\| \leq 2$	64	24	30	4	26	30	30	2
$2 < \|z\ \text{score}\| \leq 3$	0	1	1	0	4	1	2	0
$3 < \|z\ \text{score}\|$	0	1	0	0	1	1	2	1

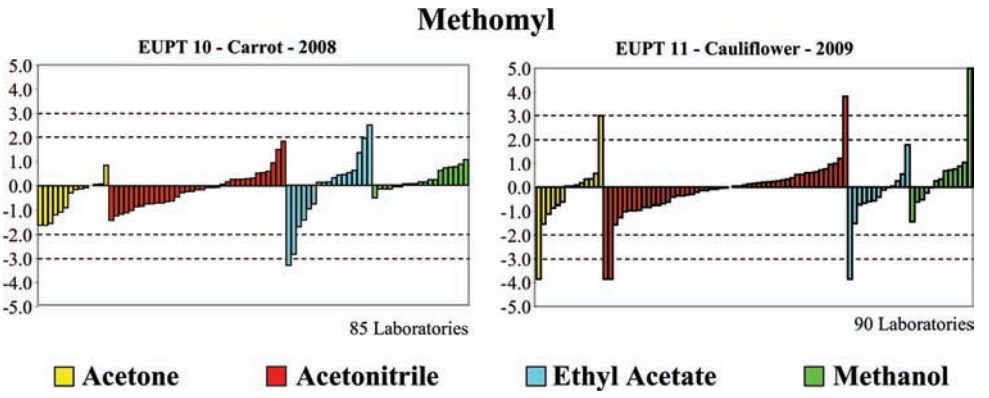

Figure 3. Different solvents used for the determination of methomyl in EUPT-FV over a 2-year period.

Table 3. Number of *z* Score Results According to the Extraction Method Used in EUPT 10 and 11, Corresponding to Figure 3

	no. of *z* scores classified by extraction method used for EUPT 10				no. of *z* scores classified by extraction method used for EUPT 11			
classification	acetone	acetonitrile	ethyl acetate	methanol	acetone	acetonitrile	ethyl acetate	methanol
$\|z\ \text{score}\| \leq 2$	14	35	14	19	12	49	11	12
$2 < \|z\ \text{score}\| \leq 3$	0	0	2	0	1	0	0	0
$3 < \|z\text{-score}\|$	0	0	1	0	1	3	1	1

laboratory. An ANOVA test was also carried out to verify that there were no significant differences between the techniques used.

Table 5 shows how the percentages of both techniques used in Figure 5 are practically the same. According to the *z* score results, there is a slightly higher percentage of laboratories conducting LC determination and obtaining higher accuracy results than those using GC.

It has been verified that, in experienced laboratories, there is not a strong dependency on the results as a consequence of the extraction method used or on the determination technique of the commodity/residue interaction, when the laboratory has proven experience and skilled personnel.

Examples of "Target Default Value" Measurement Uncertainty Approaches Based on Experience. For these types of approaches, a targeted value is taken. This is the result of

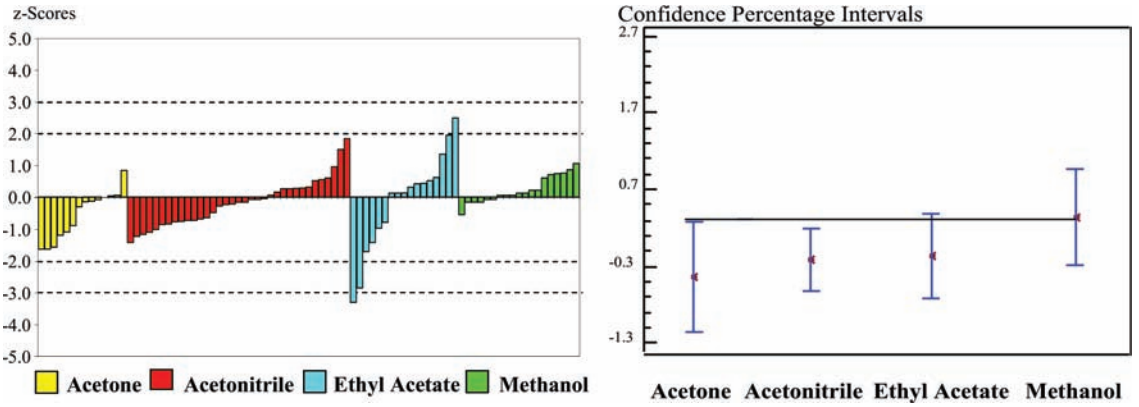

Figure 4. ANOVA test performed on z score results for kresoxim-methyl pesticide upon solvent extraction.

Table 4. Number of Unacceptable z Score Results Dependent on the Solvent

extraction solvent	no. of $\|z$ score$\|$ results > 3.0	% from the total no. of results
acetone	2564	5
acetonitrile	2546	4
ethyl acetate	1766	5
methanol	566	4

interlaboratory work based on experience. For a laboratory using a multiresidue method, in order to choose these approaches, sufficient experience and a minimum standard must be proven beforehand.

First Approach: Estimating Measurement Uncertainty Based on the FFP-RSD of PT Standards. To be able to apply this approach, based on proficiency test participation, laboratories are required to demonstrate a minimum analytical performance and quality criteria such as that considered in SANCO/10684/2009.[22]

The choice of 25% FFP-RSD for pesticide residue is an average estimation of the resultant dispersion of data. This target result, set before the participants submit their own results, is compared with the robust dispersion (Qn RSD),[23] which is the dispersion after the laboratories have submitted their results. In Figure 6, the different Qn values can be seen for the pesticides present in the three most recent EUPTs. Each bar represents the Qn value, and the assigned value achieved by each pesticide is given. The Qn values do not vary according to the pesticide concentration present in the sample. The means of these robust values are 24, 25, and 24, respectively, for the three EUPTs. These means are close to the fixed 25% FFP-RSD. The coefficients of variation of the Qn values (CV_{Qn}) are 13, 21, and 20, respectively, meaning a low dispersion of the data. Some of the coefficients are higher, especially for EUPT 10 and 11, because of some polar pesticides, such as metamidophos, that raised the dispersion of the laboratory results.

The EUPT database has applied the FFP-RSD for all of their EUPT-FVs. From the first to the tenth, 86% of the results fall into a range corresponding to a z score ≤ 2 in absolute value, meaning that the selected fixed value has been appropriate for many years.[24]

This approach fixes the combined relative standard measurement uncertainty (u') at 25%. Multiplying u' by a coverage factor of 2 for a 95% confidence interval gives $\pm$50% expanded relative measurement uncertainty (U').

Second Approach: Estimating Measurement Uncertainty Based on the Horwitz Equation or Further Modifications of PT Standards. The Horwitz equation (eq 2) was built from a collection of many different performed collaborative studies of many different analytes including pesticides and many different commodities including fruits/vegetables to determine how much variability should be allowed among laboratories, depending on the concentration, in order to make this dispersion interchangeable. It worked out how to interpret the values produced by different laboratories so as to know how much can be allowed in between-laboratory variability dependent on the concentration.[25,26] The Horwitz equation was based on interlaboratory studies. Later, Thompson introduced a slight variation on the Horwitz equation,[27] dividing it into three different concentration-dependent equations, known as the Thompson equations (eq 3).[28] The relative standard deviation (σ) becomes constant for a concentration (c) above 1.2×10^{-7} g/kg. For a concentration range from 1.2×10^{-7} to 0.138 g/kg, the Thompson equation is the same as the Horwitz equation. It is only for concentrations above 0.138 g/kg that the Thompson equation produces lower coefficients of variation.

$$\sigma = \begin{cases} 0.22c & \text{if } c < 1.2 \times 10^{-7} \\ 0.22c^{0.8495} & \text{if } 1.2 \times 10^{-7} \leq c \leq 0.138 \\ 0.01c^{0.5} & \text{if } c > 0.138 \end{cases} \quad (3)$$

Figure 7 represents a comparison of the three approaches: FFP-RSD, Horwitz, and Thompson. For most of the concentration intervals in which pesticide residues are usually present:, that is, from 1 mg/kg (1 ppm) to 1 μg/kg (0.001 mg/kg or 1 ppb), the three functions each take a different representation. Thompson and FFP-RSD are very similar: both have constant relative standard deviation values of 22 and 25%, respectively. In fact, the EUPTs conducted up to now have been held in this concentration range: this is why the line remains in this interval. The Thompson equation has a constant value up to 1.2×10^{-7} mg/kg. For higher concentrations, it practically becomes the Horwitz function.

On the basis of the Horwitz equation and the Thompson modification, the approach is to calculate the relative standard deviation as being the combined relative standard measurement uncertainty. This is then multiplied by a coverage factor of 2 to obtain the expanded relative measurement uncertainty.

It is of interest to compare the three relative uncertainties based on these targeted values for the typical pesticide concentration range of 0.01−1.0 mg/kg. This is done in Table 6.

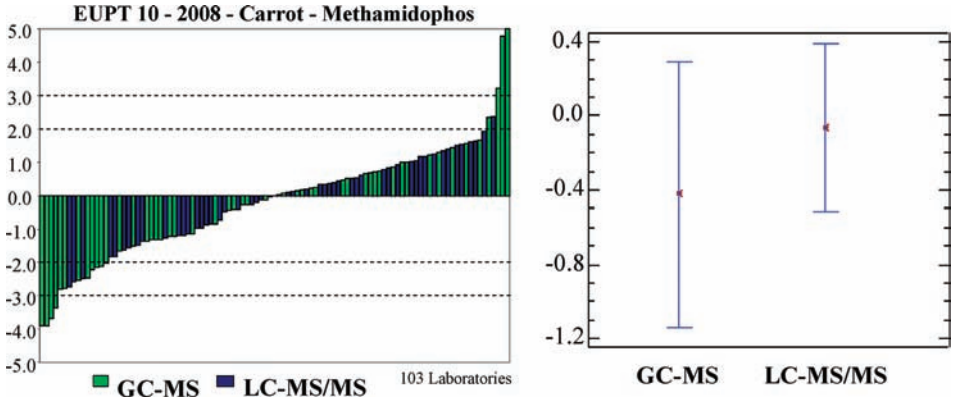

Figure 5. *z* score and ANOVA test graphs for methamidophos results for EUPT-FV10, conducted on carrots in 2008, corresponding to the determination techniques.

Table 5. Percent of *z* Scores According to the Determination Technique

	% of *z* scores classified by determination technique used for EUPT 10	
classification	gas chromatography	liquid chromatography
total	49	52
$\|z\ \text{score}\| \leq 2$	72	92
$2 < \|z\ \text{score}\| \leq 3$	16	8
$3 < \|z\ \text{score}\|$	12	0

From Table 6, it can be seen that for the two concentrations, the variation on the three uncertainties are great. For 0.001 mg/kg, the Horwitz function has the highest of all with 45% compared to the other two that have 25 and 22%, whereas for a 1.0 mg/kg concentration, it is the FFP-RSD that differs most from the other two, from 25 to 16%.

Examples Based on Experimental Data Approaches. These approaches are empirical. They are based on data taken from validation and quality control results from the same laboratory, but results from laboratory proficiency test participation may be included.

Third Approach: Estimating Measurement Uncertainty Based Only on Data from Proficiency Test Participation. This way of estimating MU is based on the bias the laboratory achieved when participating in proficiency tests over time. To take part in this test, the laboratory will previously have to have validated their routine multiresidue method and have quality control measurements taken routinely. This information is not used in this approach, but it will give the laboratory the necessary quality status to be able to take part in proficiency tests, in which each laboratory will run out their validated method, achieving accurate *z* score results.

To use this approach, it is recommended that the laboratory participates in a sufficient number of tests where the possible pesticides used to treat the sample cover those in the laboratory scope. Moreover, the number of results achieved in PT participation should be sufficient to give the laboratory a minimum of 31 results (see Table 2). For example, in EUPT-FV, the test sample is normally treated with 18−20 pesticides. This means that participation in at least two EUPTs would be sufficient to apply for this approach.

The example presented in Table 7 is based on the participation of the National Reference Laboratory of Sweden over the two most recent EUPTs for fruits and vegetables. The proficiency test results taken are independent of the type of matrix: high water content or acidic[22] and indistinctly of GC or LC determination, as long as they were in the scope of the laboratory. The EUPT-submitted results taken from the database are as they were reported: adjusted to three significant figures for concentrations above 0.010 mg/kg.[22]

The EUPT database, from which all of the needed parameters are taken, is located in the European Union Reference Laboratory in Fruits and Vegetable (EURL-FV) web page.[15] Once all of the results are collected, the bias is calculated, that is, the differences of the laboratory results from the median or assigned value achieved for each pesticide in each test divided by the median. Then the standard deviation of all the bias is calculated.

The standard deviation of the relative differences in percentage is 24%, which is equivalent to the relative standard measurement uncertainty (u') and the expanded measurement uncertainty (U') being 48%.

Fourth Approach: Estimating Measurement Uncertainty Based on Intralaboratory Validation/QC and Laboratory Bias Based on PT Data. This approach follows eq 1 commented on previously in the Introduction.

$$u' = \sqrt{u'(R_w)^2 + u'(\text{bias})^2} \quad (1)$$

The equation combines two components. The first one is the relative measurement uncertainty $u'(R_w)$, which is the within-laboratory reproducibility standard deviation involving quality control data and validation data (if needed to complement it, should the data number be below 31, see Table 2). The second component is the relative standard measurement uncertainty of the bias, $u'(\text{bias})$, which uses the method and the laboratory bias obtained from the proficiency test data. Here, other sources of bias can be used as mentioned in the Nordtest guideline.[7] The assigned value and the dispersion of interlaboratory relative standard deviations (Qn RSD) are used in this second component.

The following approach example uses data from the EURL-FV in Valencia. The data used are a combination of proficiency test and internal data, so, to choose this last one, the quality control recoveries must come from the same methods as the ones used in the PTs. The Valencia laboratory uses two methods, one for LC with 93 pesticides and one for GC with 66 pesticides. The pesticides used in each one are as follows: the 93 pesticides used in the LC method are acetamiprid, aldicarb, aldicarb sulfone,

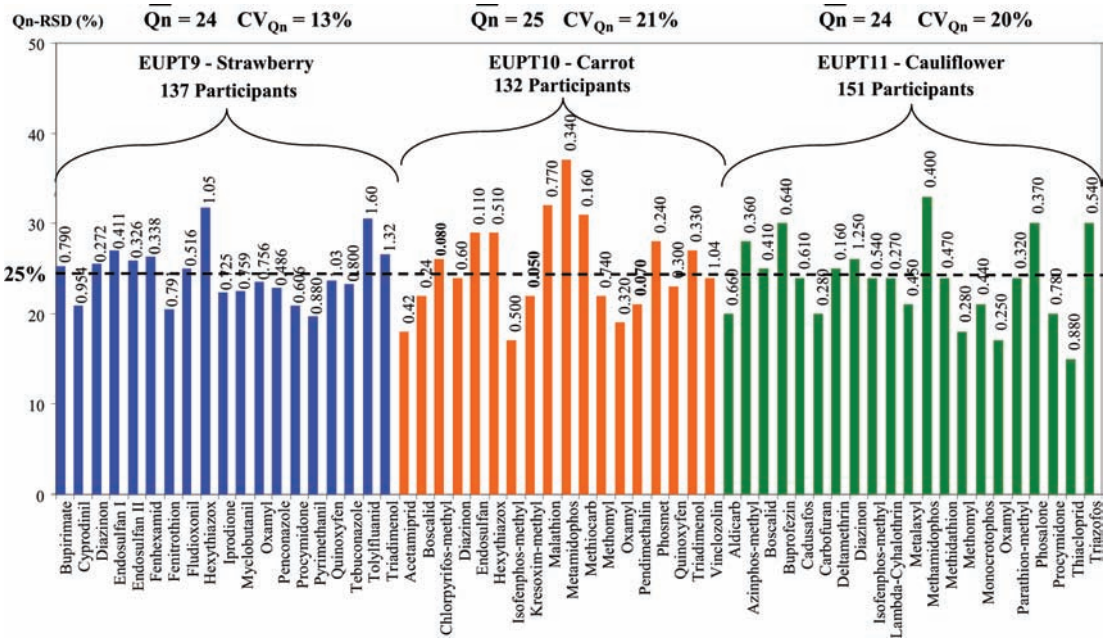

Figure 6. Average Qn-RSD for different concentration pesticides in EUPT 9−11 corresponding to the years 2007−2009.[24]

azinphos-methyl, azoxystrobin, boscalid, carbaryl, carbendazim, carbofuran, 3-OH carbofuran, chlothianidin, cyproconazole, demetonsulfone, demetonsulfoxide, dichlorvos, diflubenzuron, dimethoate, dimethomorph, fenbuconazole, fenhexamid, fluquinconazole, flutriafol, hexaconazole, imazalil, imidacloprid, iprovalicarb, linuron, mecarbam, metconazole, methiocarb, methiocarb sulfone, methiocarb sulfoxide, methomyl, tiophanate-methyl, monocrotophos, omethoate, oxamyl, paclobutrazole, paraoxon-methyl, penconazole, phosmet, prochloraz, propoxur, spinosyn A, spinosyn D, tebufenozide, thiabendazole, thiacloprid, thiamethoxam, triadimefon, triadimenol, triflumizole, triticonazole, amitraz, bitertanol, carbetamide, carboxin, carfentrazone, chlorotoluron, cymoxanil, dicrotophos, diethofencarb, dimefuron, dimethenamid, dimoxystrobin, diniconazole, DMST, epoxiconazole, etaconazole, ethiofencarb, ethiofencarb sulfone, ethiofencarb sulfoxide, ethirimol, fenamidone, fenobucarb, flufenacet, fluoxastrobin, flurtamone, flutolanil, forchlorfenuron, formetanate, fosthiazate, isoprocarb, isoxaflutole, mepronil, methacrifos, methoxyfenozide, mevinphos, nitenpyram, pethoxamid, prohexadione, propamocarb, and quinoclamine; the 66 pesticides used in the GC method are orthophenylphenol, 4,4-DDE, acrinathrin, benalaxyl, bifenthrin, bromopropylate, bupirimate, lambda-cyhalothrin, cypermethrin, cyprodinil, chlorfenapyr, chlorfenvinphos, chlorpyriphos, chlorpyriphos-methyl, chlorpropham, chlorthal-dimethyl, chlorthiophos, deltamethrin, diazinon, dichlofluanid, diphenylamine, diflufenican, endosulfan I, endosulfan II, endosulfan sulfate, esfenvalerate, ethion, etofenprox, ethoprophos, fenpropathrin, fenthion, fipronil, fluvalinate, phosalone, iprodione, isofenphos-methyl, lindane, metalaxyl, methidathion, myclobutanil, nuarimol, parathion-methyl, permethrin, pyridaben, pyrimethanil, pirimicarb, pirimiphos-methyl, pyriproxyfen, procymidone, profenofos, propiconazole, propyzamide, prothiofos, pyrifenox I, pyrifenox II, quinalphos, quinoxyfen, tebuconazole, tebufenpyrad, terbufos, tetraconazole, tetradifon, tolclofos-methyl, tolylfluanid, trifluralin, vinclozolin.

In Figure 8, the quality control charts are represented.

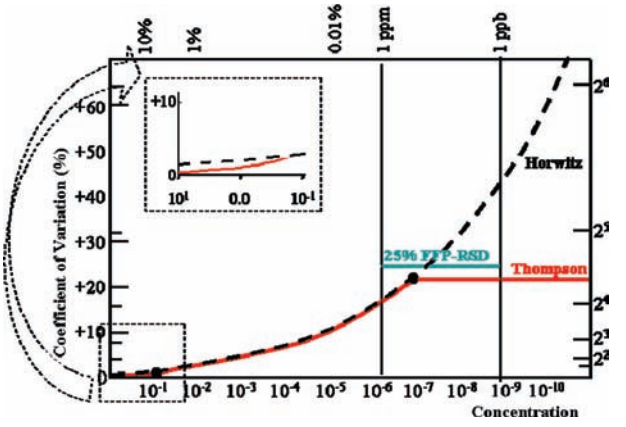

Figure 7. Horwitz, Thompson, and FFP-RSD 25% comparison functions graph.

Table 6. Comparison of the Relative Uncertainty for the Three Targeted Default Value Approaches

concentration (mg/kg)	FFP-RSD u' value (%)	Thompson u' value (%)	Horwitz u' value (%)
0.001	25	22	45
1.0	25	16	16

Method recovery data must be combined and a standard deviation of all the recoveries of all the pesticides calculated. The standard deviation gives 0.15. The $u'(R_w)$ applied is therefore 0.15.

To calculate the $u'(\text{bias})$, first, two components have to be calculated as indicated in eq 4.

$$u'(\text{bias}) = \sqrt{\text{RMS}'_{\text{bias}}{}^2 + u'(C_{\text{ref}})^2} \quad (4)$$

dx.doi.org/10.1021/jf104060h |J. Agric. Food Chem. 2011, 59, 7609–7619

Table 7. Standard Deviation of the Relative Differences from Two European Proficiency Test Results

EUPT-FV	pesticide	lab results (x_i)	assigned value (X)	bias′$_i$ $(x_i - X)/X$
EUPT-FV10, carrot	acetamiprid	0.500	0.419	0.193
	boscalid	0.287	0.238	0.208
	chlorpyrifos-methyl	0.096	0.078	0.231
	diazinon	0.762	0.603	0.264
	endosulfan sulfate	0.116	0.107	0.089
	hexythiazox	0.656	0.509	0.289
	isofenphos-methyl	0.679	0.499	0.361
	kresoxim:methyl	0.056	0.050	0.120
	malathion	1.08	0.771	0.402
	methamidophos	0.234	0.342	−0.315
	methiocarb	0.237	0.157	0.510
	methomyl	0.762	0.739	0.031
	oxamyl	0.337	0.322	0.047
	pendimethalin	0.095	0.074	0.284
	phosmet	0.317	0.236	0.343
	quinoxyfen	0.411	0.298	0.379
	triadimenol	0.398	0.330	0.202
	vinclozolin	1.30	1.04	0.250
EUPT-FV11, cauliflower	aldicarb	0.570	0.658	−0.134
	azinphos-methyl	0.275	0.355	−0.225
	boscalid	0.335	0.414	−0.191
	buprofezin	0.632	0.638	−0.009
	cadusafos	0.525	0.611	−0.141
	carbofuran	0.286	0.283	0.011
	deltamethrin	0.163	0.157	0.038
	diazinon	1.56	1.25	0.248
	isofenphos-methyl	0.467	0.540	−0.135
	λ-cyhalothrin	0.287	0.266	0.079
	metalaxyl sum	0.522	0.450	0.160
	methamidophos	0.260	0.405	−0.357
	methidathion	0.836	0.472	0.771
	methomyl	0.269	0.277	−0.029
	monocrotophos	0.391	0.438	−0.106
	oxamyl	0.237	0.249	−0.046
	parathion-methyl	0.321	0.320	0.003
	phosalone	0.280	0.368	−0.239
	procymidone	0.777	0.780	−0.004
	thiacloprid	1.00	0.880	0.138
	triazofos	0.449	0.538	−0.165
	standard deviation			0.239

The first componenet (RMS'_{bias}) is the square root of the sum of the squared bias divided by the number of results (m) taken from the PTs.

The second component is $u'(C_{ref})$, expressed in eq 5. It is the sum of the robust relative standard deviation (Qn) divided by the square root of the number of results reported by the laboratories for each of the pesticides in the scope $(No.)^{1/2}$. Then, this sum is divided by the number of results (m) taken from the PTs.

$$\frac{\sum_i \frac{Qn}{\sqrt{No.}}}{m} \times 1.253 \qquad (5)$$

Equation 5 is multiplied by a factor of 1.253 according to ISO 13528.[29] This ISO states that this factor must be multiplied by $u'(C_{ref})$ whenever the assigned value in PTs is the median.

Table 8 shows how the relative standard measurement uncertainty from method and laboratory bias based on PT data, $u'(bias)$, is calculated.

Substituting in eq 4 the results calculated in Table 8

$$u'(bias) = \sqrt{RMS'^{\,2}_{bias} + u'(C_{ref})^2} = \sqrt{0.2263^2 + 0.0239^2}$$
$$= 0.2283$$

dx.doi.org/10.1021/jf104060h |J. Agric. Food Chem. 2011, 59, 7609–7619

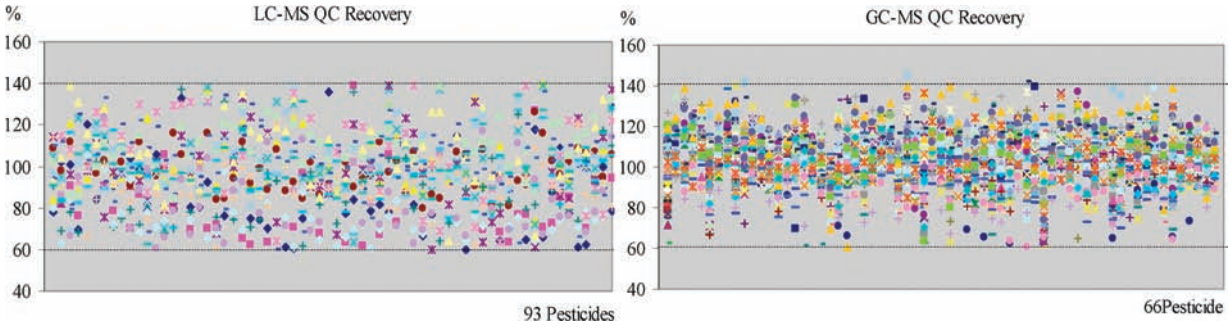

Figure 8. Recovery quality control charts for the two methods used by the laboratory.

Table 8. Calculation of u'(bias) for the EURL-FV Valencia Laboratory

EUPT-FV	pesticide	lab results	PT assigned values	$(\text{bias}'_i)^2$	Qn	no. of results	$(\text{No.})^{1/2}$	$(Qn)/(No.)^{1/2}$
EUPT-FV10, carrot	acetamiprid	0.337	0.419	0.0383	0.18	85	9.220	0.020
	boscalid	0.139	0.238	0.1720	0.22	74	8.602	0.026
	chlorpyrifos-methyl	0.056	0.078	0.0796	0.26	126	11.225	0.023
	diazinon	0.412	0.603	0.1003	0.24	125	11.180	0.021
	endosulfan sulfate	0.062	0.102	0.1538	0.29	110	10.488	0.028
	hexythiazox	0.396	0.509	0.0493	0.29	80	8.944	0.032
	isofenphos-methyl	0.436	0.499	0.0159	0.17	69	8.307	0.020
	kresoxim-methyl	0.028	0.050	0.1936	0.22	113	10.630	0.021
	malathion	0.697	0.771	0.0091	0.32	124	11.136	0.029
	methamidophos	0.245	0.342	0.0798	0.37	103	10.149	0.036
	methiocarb	0.096	0.157	0.1510	0.31	65	8.062	0.038
	methomyl	0.538	0.739	0.0740	0.22	88	9.381	0.023
	oxamyl	0.274	0.322	0.0222	0.19	84	9.165	0.021
	pendimethalin	0.056	0.074	0.0592	0.21	96	9.798	0.021
	phosmet	0.139	0.236	0.1689	0.28	95	9.747	0.029
	quinoxyfen	0.244	0.298	0.0328	0.23	95	9.747	0.024
	triadimenol	0.265	0.331	0.0398	0.27	103	10.149	0.027
	vinclozolin	0.90	1.04	0.0181	0.24	124	11.136	0.022
EUPT-FV11, cauliflower	aldicarb	0.679	0.658	0.0010	0.20	91	9.539	0.021
	azinphos-methyl	0.349	0.355	0.0003	0.28	128	11.314	0.025
	boscalid	0.373	0.414	0.0098	0.25	102	10.100	0.025
	buprofezin	0.453	0.638	0.0841	0.30	118	10.863	0.028
	cadusafos	0.810	0.611	0.1061	0.24	76	8.718	0.028
	carbofuran	0.245	0.283	0.0180	0.20	107	10.344	0.019
	deltamethrin	0.138	0.157	0.0146	0.25	130	11.402	0.022
	diazinon	1.140	1.25	0.0077	0.26	144	12.000	0.022
	isofenphos-methyl	0.498	0.54	0.0060	0.24	86	9.274	0.026
	λ-cyhalothrin	0.211	0.266	0.0428	0.24	138	11.747	0.020
	metalaxyl	0.445	0.45	0.0001	0.21	122	11.045	0.019
	methamidophos	0.341	0.4045	0.0246	0.33	109	10.440	0.032
	methidathion	0.453	0.472	0.0016	0.24	136	11.662	0.021
	methomyl	0.190	0.277	0.0986	0.18	84	9.165	0.020
	monocrotophos	0.322	0.4375	0.0697	0.21	95	9.747	0.022
	oxamyl	0.230	0.2485	0.0055	0.17	89	9.434	0.018
	parathion-methyl	0.277	0.32	0.0181	0.24	129	11.358	0.021
	phosalone	0.383	0.368	0.0017	0.30	136	11.662	0.026
	procymidone	0.750	0.78	0.0015	0.20	136	11.662	0.017
	thiacloprid	0.961	0.879	0.0087	0.15	82	9.055	0.017
	triazophos	0.612	0.538	0.0189	0.30	132	11.489	0.026
	$\Sigma(\text{bias}'_i)^2$			1.09973		$\Sigma_i(\text{Qn})/(\text{No.})^{1/2}$		0.9326
	no. of results (m) 39					no. of results (m) 39		
	$\text{RMS}'_{\text{bias}} = \sqrt{\frac{\Sigma(\text{bias}'_i)^2}{m}}$			0.2263		$u'(C_{\text{ref}}) = \Sigma_i(\text{Qn})/(\text{No.})^{1/2}/m \times 1.253$		0.0239

dx.doi.org/10.1021/jf104060h |*J. Agric. Food Chem.* 2011, 59, 7609–7619

Table 9. Comparison of the Five Approaches

approach	combined relative standard uncertainty (%), u'	expanded relative uncertainty (%), U' (k = 2, 95%)
1 (1.0−0.001) mg/kg	25	50
2 Thompson (1.0−0.001) mg/kg	16−22	32−44
2 Horwitz (1.0−0.001) mg/kg	16−45	32-90
3	24	48
4	27	54
5 (worst case)	24	48

Going back to eq 1 and substituting the within-laboratory reproducibility ($u'(R_w)$ = 0.15), the relative standard measurement uncertainty is

$$u' = \sqrt{u'(R_w)^2 + u'(\text{bias})^2} = \sqrt{0.15^2 + 0.2283^2} = 0.2732 = 27\%$$

So u' = 27% and the expanded measurement uncertainty therefore is $U = u' \times 2$ = 54%.

Fifth Approach: Estimating Measurement Uncertainty Based on Intralaboratory Data (Validation and Quality Control). This approach uses validation data carried out by the laboratory for a specific multiresidue method at a particular time, taking into account its repetition and reproducibility. The reproducibility over time has a greater influence than the repetition. From Table 1, for an accepting error of 25%, a minimum of 31 results should be taken when using this approach. These data can be taken either from the quality control recovery data or from validation data to gather enough data to perform a standard deviation of all of the results. This will be the relative measurement uncertainty.

Of course, this way of calculating MU should be revised as the quality control amasses data over time. The longer the data are taken, the more realistic the result will be.

When using this approach, the laboratory should take precautions and watch for possible sources of error that would not be highlighted without the proficiency test. This approach is the least consistent of all, as it only takes internal data into account. It is very difficult for a laboratory to internally detect sources of error from operators, from lack of stability in the standards, or from any other sources of error that could affect a result.

Some examples of this approach are calculated from recent published validation work carried out by our collaborators on this paper.[30−32] The three papers have validated their method and have a relative standard deviation for every pesticide in their scope. Calculating the standard deviation of the relative standard deviations, you will get the combined relative measurement uncertainty. For each of the three works mentioned, the uncertainties (u') are 15, 24, and 20%, respectively. As can be seen, these types of approaches can be useful for laboratories performing pesticide analysis with special characteristics, either in the pesticide or in the matrix (again, this works out to be the same).

Comparison of the Five Approaches. Considering the five approaches evaluated (see Table 9) in the majority of the cases, the values of both the combined relative standard and the expanded relative uncertainties are very close to each other, except for the Horwitz and Thompson approaches, with which, with a common pesticide concentration of 1 mg/kg, the combined relative standard decreases to 16%.

The obtained values are considered default values, achieved by the laboratories over a particular time, depending on the results used and able to change within the time or within other data taken from the same laboratory.

Conclusions. In this work, five major approaches have been described to calculate the expanded measurement uncertainty for laboratories analyzing pesticide residues using multiresidue methods.

From these approaches, both advantages and disadavantages can be raised for each.

The approaches that center on target default values based on experience have the advantage of making the harmonization of measurement uncertainty calculus easier, and at the same time they fix the threshold values for the laboratories. In contrast, in the case of the FFP-RSD (approach 1), the laboratory is forced to take part in proficiency tests on a routine basis to set acceptable results. In the case of the Horwitz and Thompson equations, the measurement uncertainty calculus is dependent on concentration, even though, nowadays, concentration exerts less and less influence on chromatography techniques in the pesticide field.

On the other hand, those approaches based on experimental data have the advantages of considering the internal quality data performed routinely to calculate the bias of the laboratory. However, with this limited internal information, possible sources of error might be missed, as systematic errors may not be detected. The approaches that combine the internal quality data with the laboratory's performance in proficiency tests will overcome these difficulties (so as will comparison with other laboratories highlight the deficiencies).

Independent of the approach selected, it is important to consider the benefits that the harmoniztation of measurement uncertainty can result in and that the value of U' = 50% can fit all of the approaches.

AUTHOR INFORMATION

Corresponding Author

*E-mail: amadeo@ual.es. Fax: +34 950015483.

DISCLOSURE

Mention of brand or firm name does not constitute an endorsement by the EURL over other similar products not mentioned.

REFERENCES

(1) Joint Committee for Guides in Metrology (JCGM). *Evaluation of Measurement Data − Guide to the Expression of Uncertainty in Measurements (GUM)*, 1st ed.; 2008.

(2) Konieczka, P.; Namiesnik, J. Estimating uncertainty in analytical procedures based on chromatographic techniques. *J. Chromatogr., A* **2010**, *1217*, 882–891.

(3) Ratola, N.; Santos, L.; Herbert, P.; Alves, A. Uncertainty associated to the analysis of organochlorine pesticides in water by solid-phase microextraction/gas chromatography-electron capture detection − evaluation using two different approaches. *Anal. Chim. Acta* **2006**, *573−574*, 202–208.

(4) Bettencourt da Silva, R. J. N.; Santos, J. R.; Filomena, M. Worst case uncertainty estimates for routine instrumental analysis. *Analyst* **2002**, *127*, 957–963.

(5) Bettencourt da Silva, R. J. N.; Lino, M. J.; Ribero, A.; Santos, J. R. Expression of results with uncertainty for the determination of pesticides in melon − experience in a proficiency test. *Analyst* **2001**, *126*, 743–746.

(6) EURACHEM Guide. *Quantifying Uncertainty in Analytical Measurements*, 2nd ed.; Laboratory of the Government Chemist: London, U.K., 2000.

(7) NORDTEST Report TR 537. *Handbook for Calculation of Measurement Uncertainty in Environmental Laboratories*, 2nd ed.; Espoo, Finland, 2004; http://www.nordicinnovation.net/nordtestfiler/tec537.pdf.

(8) EUROLAB Technical Report 1/2007. *Measurement Uncertainty Revisted: Alternative Approaches to Uncertainty Evaluation*; European Federation of National Associations of Measurement, Testing and Analytical Laboratories: Paris, France, 2007; www.eurolab.org.

(9) Technical Note 33. *Guideline for Estimating and Reporting Measurement Uncertainty of Chemical Test Results*; National Association of Testing Authorities of Australia (NATA): Australia, 2009; www.nata.com.au.

(10) ISO/IEC 17025. *General Requirements for the Competence of Testing and Calibration Laboratories*; International Organisation for Standarisation (ISO): Geneva, Switzerland, 2005.

(11) CAC/GL 59-2006. *Guidelines on Estimation of Uncertainty of Results*; Codex Alimentarius Commission; Rome, Italy, 2006; www.codexalimentarius.net/download/report/701/al31_24e.pdf.

(12) CAC/GL 40-1993. *Guidelines on Good Laboratory Practice in Residue Analysis*; Codex Alimentarius Commission, rev. 1, Rome, Italy, 2003; www.codexalimentarius.net/download/standards/378/cxg_040e.pdf.

(13) Alder, L.; Korth, W.; Patey, A. L.; Van Der Schee, H. A.; Schoeneweiss, S. Estimation of measurement uncertainty in pesticide residue analysis. *J. AOAC Int.* **2001**, *84* (5), 1569–1578.

(14) ALINORM 09/32/24. *Proposed Draft Revision of the Guidelines on the Estimation of Uncertainty of Results for the Determination of Pesticide Residues*; (CAC/RCP 59−2006) at Step 3, Codex Alimentarius Commission, Rome, Italy, 2009; 94−101; www.codexalimentarius.net/download/report/724/al32_24e.pdf.

(15) European Union Proficiency Test in Fruits and Vegetables (EURL-FV) Database. http://www.eurl-pesticides.eu/docs/public/tmplt_article.asp?CntID=693&LabID=500&Lang=EN.

(16) De Kok, A.; Hiemstra, M. Optimization, automation, and validation of the solid-phase extraction cleanup and on-line liquid chromatographic determination of *N*-methylcarbamates in fruits and vegetables. *J. AOAC Int.* **1992**, *75*, 1063–1072.

(17) Spetch, W.; Pelz, S.; Gilsbach, W. Gas-chromatographic determination of pesticide residues after clean-up by gel-permeation chromatography and mini-silica gel-column chromatograph. *Fresenius' J. Anal. Chem.* **1995**, *353* (2), 183–190.

(18) Anastassiade, M.; Lehotay, S. J.; Štajnbaher, D.; Schenck, F. J. Fast and easy multiresidue method employing acetonitrile extraction/partitioning and "dispersive solid-phase extraction" for the determination of pesticide residues in produce. *J. AOAC Int.* **2003**, *86* (2), 412–431.

(19) Lehotay, S. J.; Son, K. A.; Kwon, H.; Koesukwiwat, U.; Fud, W.; Mastovska, K.; Hoha, E.; Leepipatpiboon, N. Comparison of QuEChERS sample preparation methods for the analysis of pesticide residues in fruits and vegetables. *J. Chromatogr., A* **2010**, *1217* (16), 2548–2560.

(20) Andersson, A.; Palsheden, H. Comparison of the efficiency of different GLC multi-residue methods on crops containing pesticide residues. *Fresenius' J. Anal. Chem.* **1991**, *339*, 365–367.

(21) Klein, J.; Alder, L. Applicability of gradient liquid chromatography with tandem mass spectrometry to the simultaneous screening for about 100 pesticides in crops. *J. AOAC Int.* **2003**, *86* (5), 1015–1037.

(22) SANCO Document. *Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed*, http://www.eurl-pesticides.eu/docs/public/tmplt_article.asp?CntID=727&LabID=100&Lang=EN, No. SANCO/10684/2009, Almeria, Spain, 2009.

(23) Müller, C. H.; Uhlig, S. Estimation of variance components with high breakdown point and high efficiency. *Biometrika* **2001**, *88* (2), 353–366.

(24) Medina-Pastor, P.; Rodriguez-Torreblanca, C.; Andersson, A.; Fernandez-Alba, A. R. European Commission proficiency tests for pesticide residues in fruits and vegetables. *Trends Anal. Chem.* **2010**, *29*, 70–83.

(25) Horwitz, W. Evaluation of analytical methods used for regulation of food and drugs. *Anal. Chem.* **1982**, *54* (1), 67–76.

(26) Horwitz, W.; Kamps, L. R.; Boyer, K. W. Quality assurance in the analysis of foods and trace constituents. *J. Assoc. Off. Anal. Chem.* **1980**, *63*, 1344–1354.

(27) Thomposon, M.; Lowthian, P. J. A Horwitz-like function describes precision in a proficiency test. *Analyst* **1995**, *120*, 271–272.

(28) Thompson, M. Recent trends in inter-laboratory precision at ppb and sub-ppb concentrations in relation to fitness for purpose criteria in proficiency testing. *Analyst* **2000**, *125*, 385–386.

(29) ISO 13528. *Statistical Methods for Use in Proficiency Testing by Interlaboratory Comparisons*; International Organisation for Standarisation (ISO): Geneva, Switzerland, 2005.

(30) Pihlström, T.; Blomkvist, G.; Friman, P.; Pagard, U.; Österdahl, B. G. Analysis of pesticide residues in fruit and vegetables with ethyl acetate extraction using gas and liquid chromatography with tandem mass spectrometric detection. *Anal. Bioanal. Chem.* **2007**, *389*, 1773–1789.

(31) Ferrer Amate, C.; Unterluggauer, H.; Fischer, R. J.; Fernández-Alba, A. R.; Masselter, S. Development and validation of a LC-MS/MS method for the simultaneous determination of aflatoxins, dyes and pesticides in spices. *Anal. Bioanal. Chem.* **2010**, *397* (1), 93–107.

(32) Kmellár, B.; Fodor, P.; Pareja, L.; Ferrer, C.; Martínez-Uroz, M. A.; Valverde, A.; Fernandez-Alba, A. R. Validation and uncertainty study of a comprehensive list of 160 pesticide residues in multi-class vegetables by liquid chromatography−tandem mass spectrometry. *J. Chromatogr., A* **2008**, *1215*, 37–50.