**SZENT ISTVÁN UNIVERSITY**
**FACULTY OF FOOD SCIENCE**

# OPTIMIZATION OF SAMPLING PROCEDURES FOR VERIFYING COMPLIANCE OF COMMODITIES WITH MAXIMUM RESIDUE LIMITS OF PESTICIDES

ZSUZSA FARKAS

GÖDÖLLŐ
2017

**PhD School/Program**

| | |
|---|---|
| **Name:** | PhD School of Food Sciences |
| **Field:** | Food Sciences |
| **Head of School:** | Gyula Vatai, DSc,<br>Szent István University, Faculty of Food Science, Department of Food Engineering |
| **Supervisor:** | Dr. habil Árpád Ambrus, CSc,<br>honorary professor, retired chief scientific advisor |

**Signature of acceptance of the Head of School and the Supervisor:**

The applicant met all requirements of the PhD regulation of Szent István University and the thesis is accepted for the defence process.

| | |
|---|---|
| …… | ……... |
| Signature of the Head of PhD School | Signature of the Supervisor |

## 1. INTRODUCTION

Control of pesticides is particularly important because of the large-scale use and potential toxicity of active ingredients both from food safety and environmental protection points of view. There are many factors that may influence the distribution and the concentration of pesticide residues present on treated crops. These factors contribute to the inevitable variability of residues in the sampled lot, and the uncertainty of sampling. In order to be able to correctly assess the residue values obtained from laboratory analysis, information on the combined uncertainty of measured residues should be available. Most commonly, the uncertainty of sampling is the main contributor to the combined uncertainty of the measured residues, nonetheless it is often ignored in practice when the results are evaluated.

The maximum permitted level of pesticide residues in products is regulated by statutory limits. International organizations require the use of standard sampling procedures for different measurements, but ignore the uncertainty of sampling, as the maximum residue limits refer to the average residue content of a sample that is taken according to the requirements. Verifying the compliance of a product is the responsibility of the producer. For the product not to be rejected by the competent authorities, the product should be tested by the owner before placing it on the market. During the self-control of commodities, the combined uncertainty of the measured residues, including also sampling uncertainty, has to be considered. Considering the high cost of analyses, optimal use of available resources is essential.

## 2. OBJECTIVES

- Enumeration of the factors contributing to the uncertainty of the results of pesticide residue analysis, in particular, the uncertainty of sampling.
- Determination of sampling uncertainty for analysis of pesticide residues in different crops based on previously published results of analysis of primary samples and samples obtained from supervised residue trials.
- Determination of the 95% relative confidence interval of the estimated sampling uncertainty as the function of the number of sampled lots and replicate samples.
- Elaboration and presentation of a method for the optimization of the process of pesticide residue analysis that considers the factors affecting the uncertainty of the results.
- Determination of the action limit, based on the combined uncertainty of the results, for the self-control of crop commodities before placing them on the market.
- Elaboration of an MS Excel-based method for the automatic calculation of the action limit.

- *It was not my intention* to check the findings of previous peer reviewed publications regarding the distribution of pesticide residues. I have accepted them based on the publications.

## 3. MATERIALS AND METHODS

### 3.1 Estimation of sampling uncertainty of pesticide residues based on primary samples

For the modelling of sampling uncertainty, databases of residue values of 100-300 primary samples from 182 pesticide-crop combinations available from previous studies were used. The database comprised of samples from different origin:

(a) Residue values from carrot and parsley primary samples taken from fields treated according to normal agricultural practice. One piece of carrot root and one handful of parsley leaves made up a primary sample. A dataset comprised of 120 primary samples taken from 1 treated field. Altogether, 18 datasets were available.

(b) Normalized datasets; all values above the limit of quantification (LOQ) of each dataset were divided by the average value of the corresponding dataset. The given normalized residue values were combined into two datasets (carrot: number of residue values (A) = 1183, $\mu = 1$, CV = 0.60; parsley A = 959, $\mu = 1$, CV = 0.79).

(c) 120-300 residues measured in primary samples taken from different fruit and vegetable fields with known pesticide treatment history. 95 datasets were available from 13 countries.

(d) 100-100 residue data measured in primary samples of fruits taken from a lot with unknown pesticide treatment history. The samples were analysed with multi-residue method in 69 pesticide-crop combinations.

(e) Synthetically generated datasets with lognormal distribution; according to previous results, the distribution of pesticide residues can be best described by lognormal distribution, therefore some of the modelling were conducted with datasets of lognormal distribution containing generated samples of size 10.

Datasets listed in points (a), (c) and (d) contained altogether 182 pesticide-crop combinations. Only residue values above the LOQ were considered. All of the 182 datasets were used for the determination of the confidence interval of sampling uncertainty values estimated from supervised residue trial database (chapter 4.5.1).

#### *3.1.1 Modelling the factors that affect the uncertainty of sampling*

*In order to estimate the uncertainty of sampling,* samples of sizes (*n*) of 5, 10 and 25 were taken 1000-10,000 times (*N*), respectively, from datasets (a) and (b) with validated MS Excel macro.

*In order to examine the effect of sample size (n), replicate samples (p) and number of sampled lots (L)* on the uncertainty of sampling, modelling were conducted with residue values in samples taken with random sampling with replacement from datasets of (a) (b) and (e).

*3.1.2 Calculation methods*

$CV_R$ values (combined uncertainty of the measured residue, $R_i$) were calculated by two different ways.

1) With the conventional method above 10 replicates (p > 10):

$$SD_R = \sqrt{\frac{\sum(R_i - \bar{R})^2}{A-1}} \qquad (1)$$

$$CV_R = \frac{SD_R}{\bar{R}} \qquad (2)$$

2) In case of 10 and less replicates (p ≤ 10): the method of range statistics was used, that approximates the $SD_R$ for small number of samples. In this case, the relative difference of residue values was divided by the corresponding factor (*d*):

$$SD_R = \frac{R_{max} - R_{min}}{d}; \quad CV_R = \frac{SD_R}{\bar{R}} = \frac{R_{max} - R_{min}}{\bar{R} \times d} \qquad (3)$$

The value of the factor *d* is 1.128; 2.059; 2.534; 2.847 and 3.078 for 2, 4, 6 and 8 replicates, respectively.

As the measurement results are obtained with the multiplication of the contribution of the consecutive steps of the determination process, the uncertainty of the measured value is calculated based on the rule of error propagation:

$$CV_R = \sqrt{CV_S^2 + CV_{SS}^2 + CV_{Sp}^2 + CV_A^2} \qquad (4)$$

$$CV_L = \sqrt{CV_{SS}^2 + CV_{Sp}^2 + CV_A^2} \qquad (5)$$

Where $CV_S$, $CV_{SS}$, $CV_{Sp}$, and $CV_A$ are the uncertainty of sampling, sample size reduction, sample processing and analysis, respectively.

Calculation of $CV_L$ (uncertainty of the laboratory phase) can be done from residues of test portions analysed on different days. It incorporates the uncertainty of sample size reduction (in case of large samples), sample processing and analysis:

$$CV_L = \left(\frac{\sum \Delta}{n}\right) / 1.128 \qquad (6)$$

$$\Delta = \frac{|R_1 - R_2|}{\bar{R}} \qquad (7)$$

The uncertainty of sampling ($CV_S$) cannot be calculated directly. It can be determined from the uncertainty of the measured residue ($CV_R$) and the uncertainty of the laboratory phase ($CV_L$):

$$CV_S = \sqrt{CV_R^2 - CV_L^2} \qquad (8)$$

As $CV_L$ values were below 30% of the $CV_R$ values in all cases, they practically did not affect the value of the estimated $CV_S$, therefore the main contributor to the uncertainty was the variability of residues measured in primary samples. Because of that, $CV_R$ values calculated with equations 2 and 3 were considered as the uncertainty of sampling.

The relative 95% confidence interval of sampling uncertainty values ($CI_{r0.95}$) were calculated with the following equation:

$$CI_{r0.95} = \frac{CV_{P0.975} - CV_{P0.025}}{\overline{CV}} \qquad (9)$$

Where $CV_{P0.975}$ and $CV_{P0.025}$ are the 97.5$^{th}$ and 2.5$^{th}$ percentiles of the $CV_R$ values of replicate samples taken 1000-10,000 times (*N*), and $\overline{CV}$ is the average of the 1000 or 10,000 $CV_R$ values.

### 3.2 Estimation of sampling uncertainty based on supervised residue trial data

The uncertainty of sampling was estimated based on supervised residue trial data that were used by FAO/WHO Joint Meeting on Pesticide Residues (JMPR) for recommending Maximum Residue Limits (MRL). The dataset, containing 25,876 single residue values, comprised of 12,087 replicate samples (in most cases, >99.95%, duplicate samples) from the trial plots.

The uncertainty of sampling was evaluated for 106 crops and 24 crop groups based on the food and feed classification of Codex Alimentarius. The confidence intervals of the estimated sampling uncertainty values were determined based on the results of experimental data. The advantage of the method is that it does not have any assumptions about the distribution of CV values.

#### *3.2.1 Methods for the calculation of sampling uncertainty and the effect of other relevant factors*

The $CV_R$ values of replicate samples taken simultaneously from one trial plot (called 'sample pairs' as they comprised of duplicate samples in most of the cases; indicated with $CV_{R1}$) was calculated with range statistics.

For a given crop, more than one sample pairs were available. The $CV_R$ value of a crop was the average value of the $CV_{R1}$ values of the sample pairs and is indicated with $CV_{Ri}$.

$$CV_{Ri} = \frac{\sum_i CV_{R1}}{N_i} \qquad (10)$$

As previous studies showed that the distribution of pesticide residues is independent from the type of the active ingredient and formulation, therefore the best estimate for the characteristic sampling uncertainty values of a given crop is the average of $CV_{R1}$ values calculated from trials conducted with different pesticides.

$CV_L$ values are usually not indicated in JMPR reports, therefore approximately 6500 recovery studies were collected. The average $CV_A$ value calculated from the recovery studies was 8.56%.

Taking into account the strict circumstances of sample preparation in case of supervised residue trials, assuming that the uncertainty of sample size reduction and the sample processing is approximately 5%, the calculated rounded $CV_L$ value is 10%. Sampling uncertainty value of the $i^{th}$ crop ($CV_{Si}$) can be calculated from $CV_{Ri}$ and $CV_L$ with equation 8.

Sample sizes applied in case of official control and supervised residue trials may differ. Assuming that OECD Guideline was followed when conducting supervised residue trials, sampling uncertainty values were corrected with a factor. It was assumed that in supervised residue trials 24, and 12 primary samples made up a composite sample in case of small and medium sized, and large sized crops, respectively, while in case of official control, sample sizes are 10 and 5, the values of the factor was f=1.55 ($\sqrt{24}/\sqrt{10} = \sqrt{12}/\sqrt{5} = 1.55$). Corrected sampling uncertainty value of crops ($CV_{SiKorr}$) can be calculated as:

$$CV_{SiKorr} = CV_{Si} \times f \qquad (11)$$

Sampling uncertainty values of crops belonging to the same crop group ($CV_{SiKorr}$) was estimated from different number of sample pairs, therefore the sampling uncertainty values of crop groups ($CV_{Scs}$) was calculated as the weighted average of the sampling uncertainty values of different crops ($CV_{SiKorr}$).

$$CV_{Scs} = \sqrt{\frac{\sum(df_i \times CV_{SiKorr}^2)}{\sum df_i}} \qquad (12)$$

## 4. RESULTS

### 4.1 Effect of sample size on the variability of pesticide residues

The composite samples taken from both the normalized and the field trial primary sample datasets show that the average residue values of primary samples and theoretically expectable average residue values in composite samples are close to each other. The $CV_R$ values calculated form the results of random sampling with replacement are similar to the theoretically expectable $CV_R$ values. According to the results, modelling done with random sampling with replacement gives an unbiased estimate.

The range of residue values decreases with the increasing number primary samples making up a composite sample (sample size). In case of sample size $\geq$ 25, the distribution of residue values in composite samples is close to a normal distribution (Figure 1.). It is not advisable to apply sample size larger than 25 in case of medium and large size crops, as the sampling uncertainty does not decrease considerably, while for such large sample sizes, sample processing requires special equipment and can result in larger $CV_{SS}$ and $CV_{Sp}$ values.

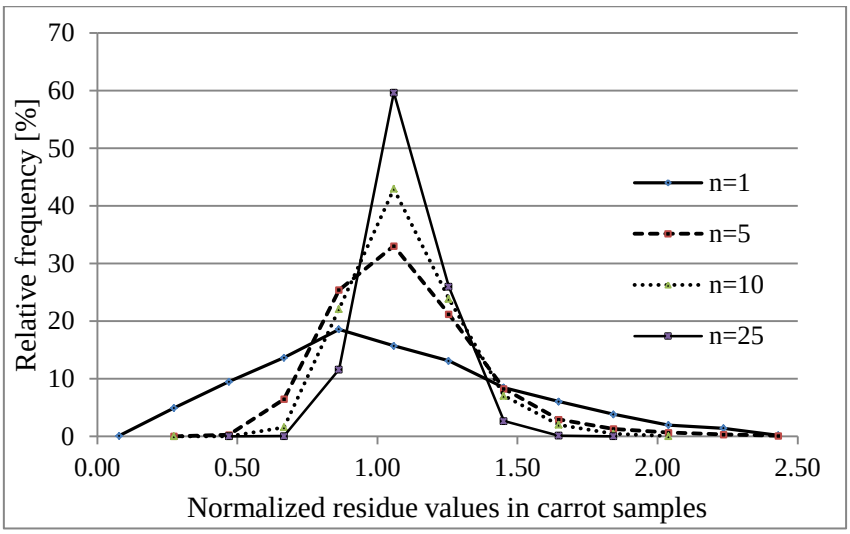


Figure 1. Distribution of normalized residues in carrot composite samples of different sample sizes.

**4.2 Effect of the number of replicate samples on the variability of pesticide residues**

The relative range of the estimated $CV_S$ values decreases with the increasing number of replicate samples and the most likely value (the maximum value of the relative frequency distribution curves) is close to the real value as it is shown in Figure 2. The systematic error, (deviation between the average $CV_S$ value of composite samples taken repeatedly from one parent population, and the $CV_n$ value calculated from the $CV_1$ value of the parent population and the size samples ($n$) as follows):

$$CV_n = \frac{CV_1}{\sqrt{n}} \qquad (13)$$

decreases with the increasing number of replicate samples and is less than 10% above 12 replicates.

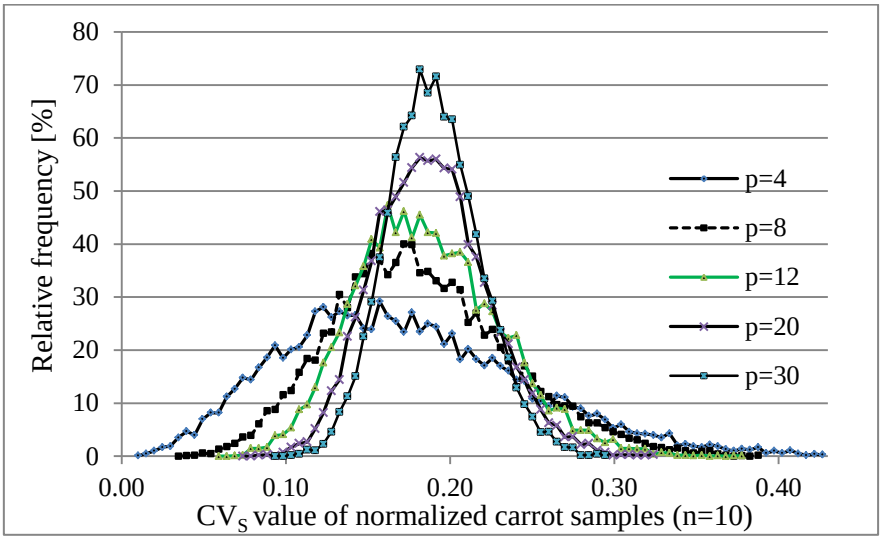


Figure 2. Distribution of $CV_S$ values of composite carrot samples (n=10) taken from normalized dataset with $p$ replicates. Expectable $CV_S$ value is 0.19.

### 4.3 Effect of the variability of pesticide residues of the parent population

The $CV_1$ value measured in primary samples varied between 0.11 and 1.44. Therefore, in order to examine the effect of the residue variability of the parent population on the $CV_S$, composite samples of size 10 (*n*) were taken with random sampling with replacement 10,000 times (*N*) from generated lognormal parent populations of different $CV_1$ values (0.11; 0.18; 0.25; 0.26; 0.36; 0.40; 0.55; 1.06 and 1.44), and $\mu = 1$. The parameters of the generated datasets of parent populations represented the true variability of field trial residues. The random sampling was repeated 4, 8, 12, 20 and 30 times (*p*).

The calculated relative confidence interval values ($CI_{r0.95}$) are illustrated in the function of number of replicate samples in Figure 3.

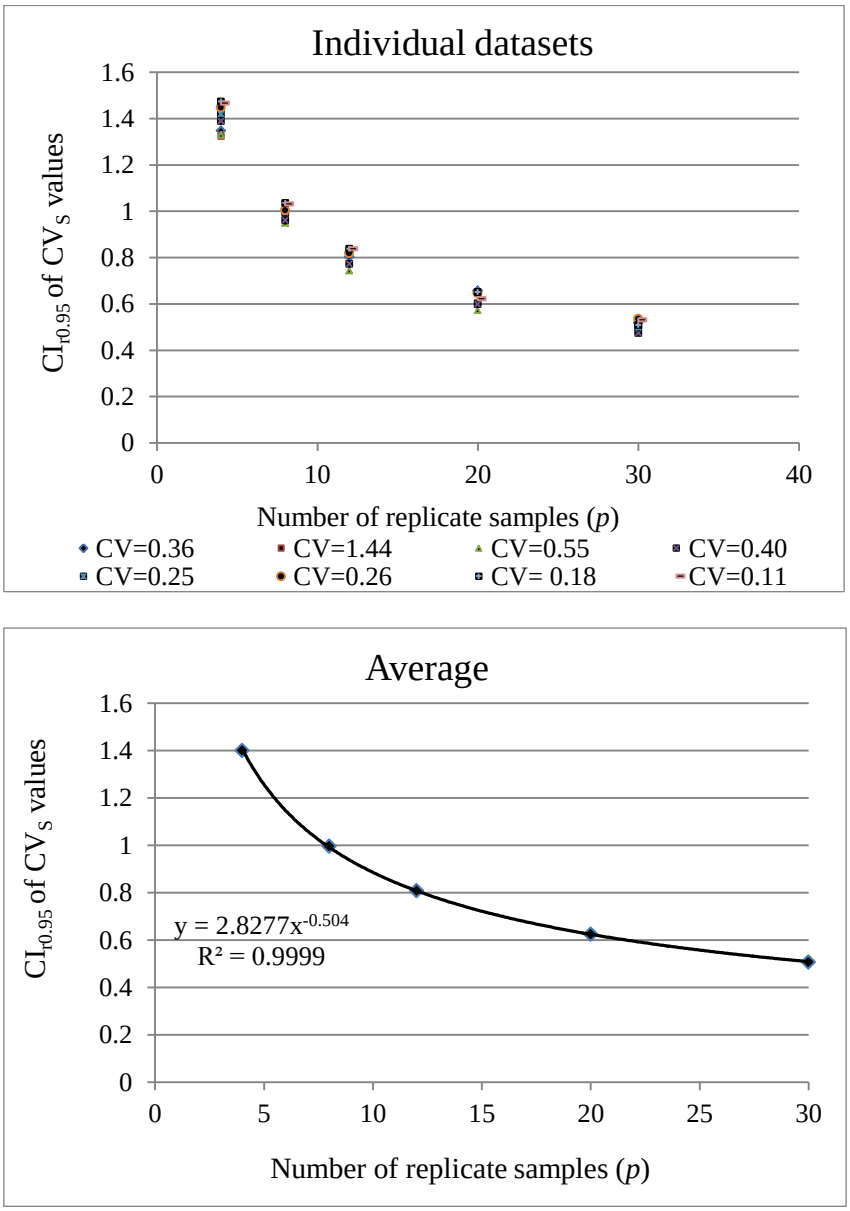


Figure 3. Upper: Relative 95% range of $CV_S$ values of samples taken with *p* replicates of samples of size 10 from datasets with $CV_1$=0.11; 0.18; 0.25; 0.26; 0.36; 0.40; 0.55; 1.06 and 1.44. Lower: Average of relative ranges.

Figure 3. clearly shows that the relative range of estimated sampling uncertainty values is independent from the variability of the parent population, therefore the results of the modelling are generally applicable. The equation fitted on the calculated average values indicates a close relationship ($R^2$ = 0.9999), proving that the relative variances are constant.

### 4.4 Effect of the number of sampled lots on the variability of pesticide residues in composite samples

Samples of size 10 were drawn 10,000 times (*N*) with 2, 4, 6 and 8 replicates with replacement from synthetic lognormal populations (μ = 1, CV = 0.8; that represents the distribution of residues measured in primary samples). The random sampling was repeated 20 times from 20 different generated populations, therefore it represented 20 different sampling units (lots).

Relative range of $CV_S$ values decreased with the increasing number of sampled lots. According to the results, above 20 tested lots, the gain is getting marginal and between 8 and 20 lots the results must be handled with care (Figure 4.). There is no optimum number of samples and sampled lots alone, the cost, risks and other relevant circumstances always have to be taken into account.

The same modelling was conducted with the datasets of parsley and carrot; and market samples. The modelling gave similar results as the one conducted with synthetic datasets.

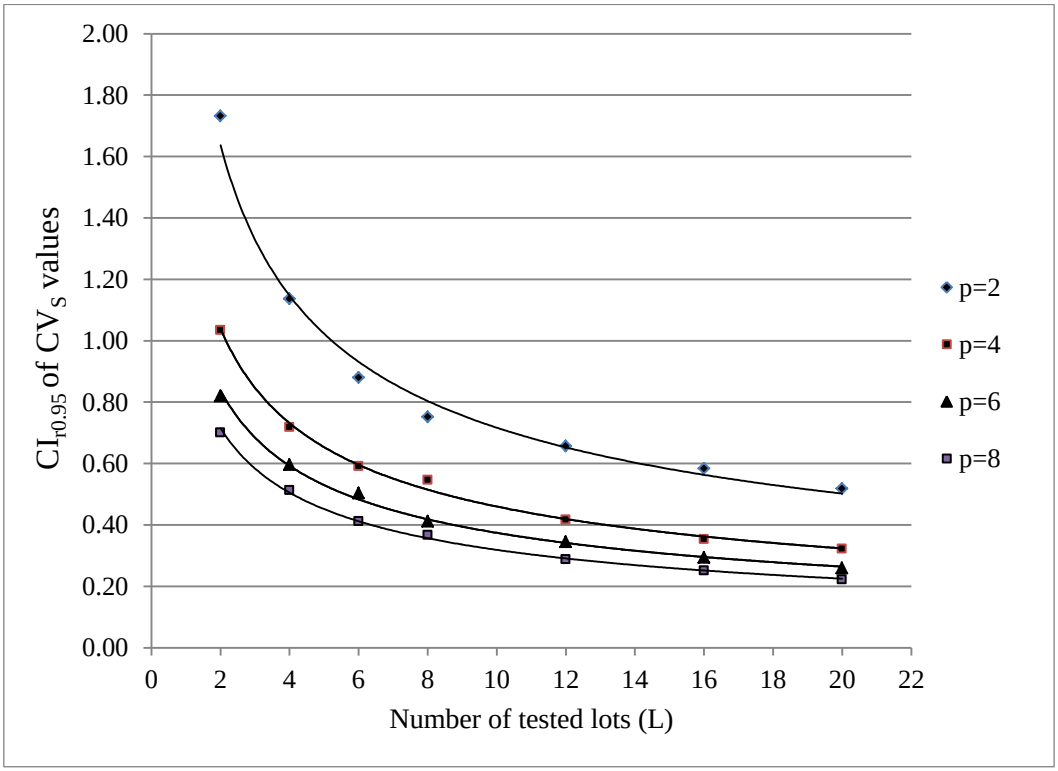


Figure 4. Relative 95% range of $CV_S$ values of samples of size 10 taken from *L* independent lots with *p* replicates. The parameters of the parent population: μ = 1; SD = 0.8.

### 4.5 Sampling uncertainty values estimated from supervised residue trial data

Supervised residue trials are conducted under strictly controlled circumstances (shortest permitted time between pesticide application and harvest and highest permitted dose and application frequency, as the goal is to gather information on the maximum level of residues present in/on crops). Therefore, higher variability may occur in case of residues measured in samples deriving from fields treated according to general agricultural practice.

Therefore residue data were collected for crops treated in supervised trials and according to normal field practice.

Based on the comparison of the variability of residues, a factor of 1.2 is recommended for adjusting the sampling uncertainty values estimated from supervised trials to reflect the variability expectable under field conditions.

Sampling uncertainty values were determined for 106 crops and 24 crop groups from the dataset of 12,087 replicate samples derived from independent supervised residue trials. The characteristic sampling uncertainty values of crop groups and the confidence intervals of sampling uncertainty values calculated with two different methods are indicated in Table 1. $CV_{Sprim}$ values were calculated with equation 13 from the average sampling uncertainty values of composite samples. The values are corrected with the factor of 1.2.

#### *4.5.1 Calculation of confidence intervals*

The confidence intervals of sampling uncertainty values were determined based on $Chi^2$ distribution and from the experimental data. The modelling of the relationship of confidence intervals of sampling uncertainty values and the number of tested lots was extended to 182 independent lots. Samples of size 10 were drawn 1000 times with 2 replicates in order to represent the circumstances of supervised residue trials.

The relative 95% range of $CV_S$ values were plotted in the function of the 2, 4, 8, 16, 40, 80, 130, 160 and 182 lots (Figure 5.). The equation of the fitted curve is ($R^2 = 0.981$):

$$CI_{r0.95} = 2.4679N^{-0.439} \qquad (14)$$

Relative confidence intervals were calculated from the equation (14) by replacing $N$ with the corresponding number of sample pairs. Absolute confidence intervals were calculated by multiplying relative confidence interval values with the corresponding sampling uncertainty value of crops ($CV_{SiKorr}$) and crop groups ($CV_{Scs}$).

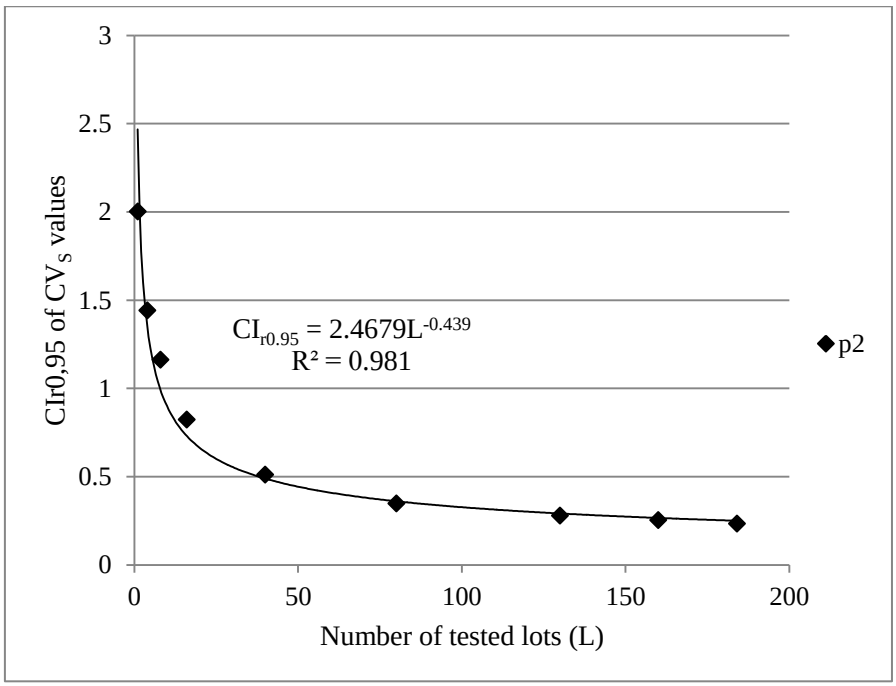


Figure 5. Relative 95% range of $CV_S$ values calculated from residue values of samples of size 10 taken from 182 lots with 2 replicates (*p*).

As the distribution of CV values is asymmetric, the upper and lower confidence limits were calculated according to the proportion of upper and lower confidence limits calculated from $Chi^2$ distribution. Confidence interval values calculated from results of experimental data were somewhat higher than those calculated with $Chi^2$ distribution assumption.

In case of fresh fruits and vegetables, the uncertainty of sampling cannot be determined based on a few measurements because of the large residue variability. Underestimating the sampling uncertainty may result in serious consequences when controlling the compliance of the products before placing them on the market. Therefore, besides the values of sampling uncertainty, the upper confidence limits of sampling uncertainty values ($UCL_{prim}$), calculated from the results of experimental data, are also indicated in Table 1.

It is the responsibility of the decision maker whether to choose the average sampling uncertainty values of primary samples ($CV_{Sprim}$) or their upper confidence limits ($UCL_{prim}$) for the calculation of action limits before placing the products on the market. It is easy to calculate the sampling uncertainty values for the appropriate sample size for the actual situation with equation 13, from the $CV_S$ values, that refer to primary samples indicated in Table 1, when preparing a sampling plan for pre-market self-control.

Table 1. Sampling uncertainty values and confidence intervals of crop groups.

| | | | | $Chi^2$ distribution assumption | | | From experimental data | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| M[1] | Crop groups | N | $CV_S$ | LCL | UCL | CI | LCL | UCL* | CI | $CV_{Sprim}$ | $UCL*_{prim}$ |
| 4 | Small sized fruits | 768 | 0.33 | 0.31 | 0.34 | 0.03 | 0.30 | 0.35 | 0.04 | 1.24 | 1.32 |
| 12 | Medium sized fruits | 2139 | 0.27 | 0.26 | 0.28 | 0.02 | 0.25 | 0.28 | 0.02 | 1.01 | 1.05 |
| 4 | Large sized fruits | 560 | 0.30 | 0.28 | 0.32 | 0.03 | 0.28 | 0.32 | 0.05 | 0.80 | 0.86 |
| 4 | Medium sized vegetables | 1211 | 0.36 | 0.35 | 0.38 | 0.03 | 0.34 | 0.38 | 0.04 | 1.37 | 1.45 |
| 4 | Bush berries | 171 | 0.18 | 0.16 | 0.20 | 0.04 | 0.16 | 0.20 | 0.05 | 0.67 | 0.76 |
| 7 | Legume vegetables | 211 | 0.33 | 0.31 | 0.37 | 0.06 | 0.30 | 0.38 | 0.08 | 1.27 | 1.43 |
| 4 | Brassica vegetables | 698 | 0.32 | 0.31 | 0.34 | 0.03 | 0.30 | 0.35 | 0.05 | 0.87 | 0.93 |
| 5 | Cucurbits | 337 | 0.37 | 0.34 | 0.40 | 0.06 | 0.33 | 0.40 | 0.07 | 0.98 | 1.08 |
| 11 | Leafy vegetables | 1872 | 0.29 | 0.28 | 0.31 | 0.02 | 0.28 | 0.31 | 0.03 | 1.12 | 1.17 |
| 6 | Root and tuber vegetables | 256 | 0.30 | 0.28 | 0.33 | 0.06 | 0.27 | 0.34 | 0.07 | 1.14 | 1.28 |
| 2 | Stalk and stem vegetables | 276 | 0.20 | 0.18 | 0.22 | 0.04 | 0.18 | 0.22 | 0.04 | 0.76 | 0.85 |
| 4 | Pulses | 346 | 0.40 | 0.38 | 0.44 | 0.06 | 0.37 | 0.45 | 0.08 | 1.53 | 1.69 |
| 6 | Cereal grains | 340 | 0.21 | 0.20 | 0.23 | 0.03 | 0.19 | 0.23 | 0.04 | 0.81 | 0.89 |
| 1 | Grasses for sugar or syrup production | 15 | 0.71 | 0.52 | 1.09 | 0.57 | 0.53 | 1.06 | 0.53 | 2.68 | 4.04 |
| 2 | Tree nuts | 101 | 0.19 | 0.17 | 0.23 | 0.05 | 0.17 | 0.23 | 0.06 | 0.74 | 0.87 |
| 5 | Oilseeds | 247 | 0.33 | 0.30 | 0.36 | 0.06 | 0.30 | 0.37 | 0.07 | 1.26 | 1.41 |
| 1 | Seeds for beverages and sweets | 22 | 0.55 | 0.42 | 0.78 | 0.35 | 0.43 | 0.77 | 0.35 | 2.08 | 2.93 |
| 4 | Legume forage and fodder | 288 | 0.28 | 0.26 | 0.31 | 0.05 | 0.26 | 0.31 | 0.06 | 1.07 | 1.19 |
| 6 | Straw, hay (of legume feeds) | 523 | 0.30 | 0.27 | 0.30 | 0.03 | 0.26 | 0.31 | 0.05 | 1.14 | 1.16 |
| 10 | Cereal forage, fodder and straw | 1176 | 0.29 | 0.28 | 0.30 | 0.02 | 0.27 | 0.32 | 0.05 | 1.10 | 1.20 |
| 2 | Grass forage | 19 | 0.22 | 0.17 | 0.32 | 0.15 | 0.17 | 0.32 | 0.15 | 0.84 | 1.21 |
| 1 | Grass hay | 18 | 0.15 | 0.12 | 0.23 | 0.11 | 0.12 | 0.23 | 0.11 | 0.59 | 0.86 |
| 2 | Dried herbs | 99 | 0.23 | 0.20 | 0.26 | 0.06 | 0.20 | 0.27 | 0.07 | 0.86 | 1.02 |
| 3 | By-products for animal feed | 391 | 0.23 | 0.21 | 0.24 | 0.03 | 0.21 | 0.25 | 0.04 | 0.86 | 0.94 |

[1]Number of crops belonging to a crop group.

### 4.6 Optimization of sampling plans

When testing the compliance of a product with the legal limits, there are two different situations:

1. After a product is placed on the market, official control only takes into account the uncertainty of the measurement, as the legal limit refers to the average residue content of the samples taken in accordance with the relevant standards. According to the EU regulations, a product is considered non-compliant only if the measured residue value minus the 50% extended sampling uncertainty exceeds the maximum residue limit (MRL).
2. Before placing a product on the market, the producer must ensure that average residue value of all composite samples taken from the lot in accordance with the standards meets the legal criteria. Therefore, the use of a so-called 'action limit' (AL) is recommended, that takes into account the combined uncertainty of the measured residues, including sampling uncertainty.

#### *4.6.1 Approximate calculation of action limit*

Action limit can be calculated based on the following equation:

$$MRL = AL + k \times CV_R \times AL; \qquad AL = \frac{MRL}{1+kCV_R} \qquad (15)$$

where the value of $k$ depends on the targeted compliance level. The approximate calculation considers the combined uncertainty of the measured residues as the only source of variation, and assumes that the standard deviation of the results is known. $CV_R$ can be calculated based on the uncertainty of the measurement steps and sampling, standard deviation can be determined from the measured residue value and the $CV_R$ value. The distribution of average residue value of samples of size $\geq 25$ is practically normal, in this case, $k$ can be replaced with standard normal variate, $Z$. In case of lognormal distribution, transformation is necessary.

#### *4.6.2 MS-Excel based template for the calculation of the action limit and optimization of sampling plans*

A practically applicable model was elaborated in MS Excel as a template, that takes into account the combined uncertainty of the measured residues, including the sampling uncertainty values determined for crops and crop groups. The model is recommended for use when checking the compliance of crops before placing them on the market, and for the optimization of practically applicable sampling procedures for producers. With the template, the producer is able to check if the product to be placed on the market is compliant with the relevant legal limits. The model was elaborated based on the characteristic distribution of pesticide residues, but in practice, it can be used for other contaminants as well, when sufficient information is available regarding the distribution of the given contaminant.

The template calculates the action limit with built-in equations. In order to reach the required compliance level, the action limit should not be exceeded. In case of replicate samples, the residue values of all samples have to be below the action limit. Figure 6. shows the operational characteristic curves drawn by the template. Under given test conditions and parameters, when testing one, two or four samples, none of the measured residue values can exceed the values of 0.41, 0.6 and 0.81 mg/kg, respectively, in order to reach the targeted 98% compliance.

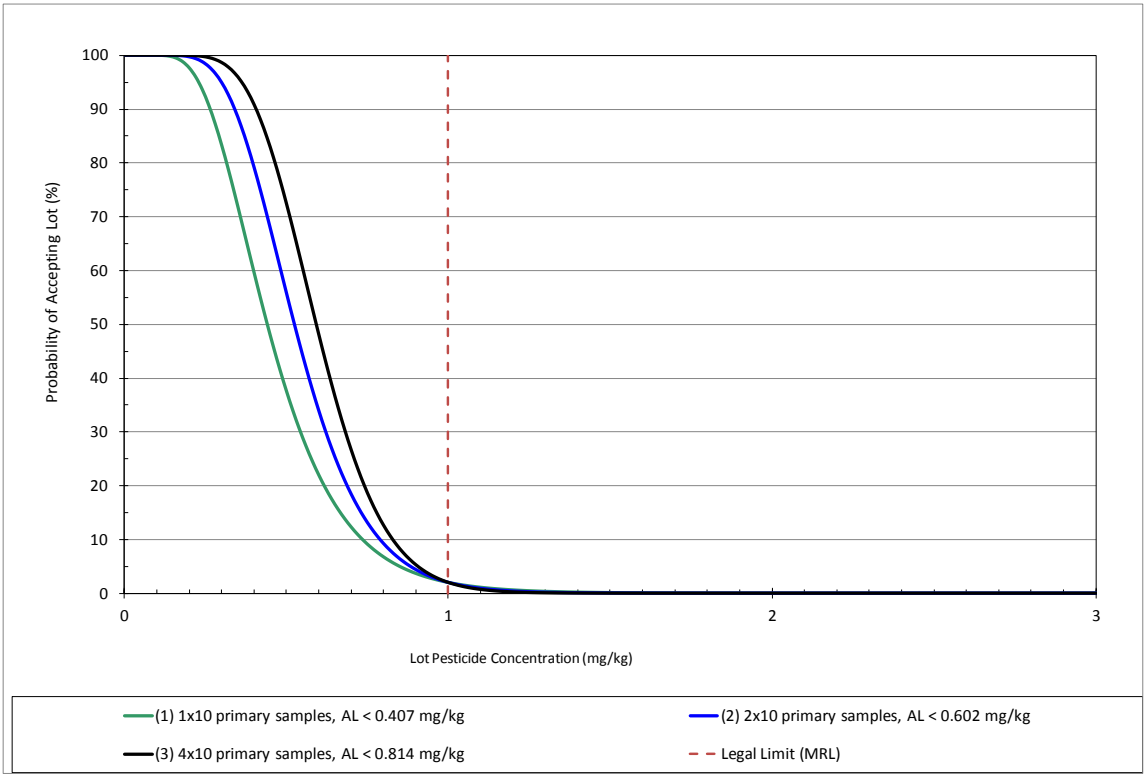


Figure 6. Operational characteristic curves drawn by the template with given parameters.

## 5. NEW SCIENTIFIC RESULTS

1. The uncertainty of sampling was determined for 106 crops and 24 crop groups from the available approximately 20,000 random primary samples and 12,087 replicate composite samples deriving from independent supervised residue trials.
2. Sampling uncertainty and the relative 95% confidence interval of sampling uncertainty values ($CI_{r0.95}$) was determined based on modelling with random sampling of residue values of primary samples. It was concluded that $CI_{r0.95}$ values depend on the number of replicate samples and the number of sampled lots, at the same time, independent from the variability of the $CV_{R1}$ value of the parent population (0.11-1.44).
3. Relationship of characteristic $CI_{r0.95}$ values calculated from different parent populations and the number of replicate samples ($p$) can be described by equation $2.8277p^{-0.504}$ ($R^2$=0.9999), that indicates that the relative variances are practically constant.
4. A method for the calculation of action limits to be applied in self-control of products was elaborated, that assumes one source of variation (the combined uncertainty of the measured residues) and known standard deviation. It calculates the action limit by taking into account the targeted MRL compliance level.
5. An MS Excel template was elaborated, which automatically calculates the action limit and draws the operational characteristic curve of the sampling by taking into account the uncertainty of sampling, the number of replicate samples and the factors affecting the laboratory testing of samples.

## 6. CONCLUSIONS AND RECOMMENDATIONS

It is the distributor's responsibility to decide on the minimum percentage of the product compliance with legal limits and other specific requirements, and by that, to decrease the risk taken regarding the product and lot to be considered non-compliant.

Optimized sampling plans can be developed, when the targeted compliance level and costs of sampling and analysis are known. In case of testing the compliance of vegetable crops before placing them on the market, the use of an action limit – that is lower than the legal limit and takes into account the combined uncertainty of the measured residues – is recommended. With the help of the elaborated MS Excel template, action limit can be easily determined on a given compliance level, and on the other hand, compliance level can be determined when action limit and a measured residue value is known. Sampling plans can be optimized for given circumstances and situations with the use of the sampling uncertainty values determined for 106 crops and 24 crop groups and other factors that contribute to the combined uncertainty of the measured residues, and relevant cost.

There is no optimum for sample size, number of replicate samples and lots to be tested according to the conclusions drawn from modelling field trial data. These parameters can only be determined together with the costs. The same way, there is no optimal solution for the preparation of sampling plans. In case of those crops and crop groups that are not covered by the estimated sampling uncertainty values, taking 6-6 replicate samples from a minimum of 8 lots is recommended for the calculation of sampling uncertainty values with relative 95% confidence interval within 50%.

## 7. PUBLICATIONS

### In the subject of the dissertation

Peer reviewed

1. **Farkas, Zs.**, Horváth, Zs., Kerekes, K., Ambrus, Á., Hámos A., Szeitzné Szabó, M. (2014): Estimation of sampling uncertainty for pesticide residues in root vegetable crops. *Journal of Environmental Science and Health, Part B* 49:01, 1-14.

   IF (2016): 1,362

2. **Farkas, Zs.**, Horváth, Zs., Szabó, IJ., Ambrus, Á. (2015a): Estimation of sampling uncertainty of pesticide residues based on supervised residue trial data. *Journal of Agricultural and Food Chemistry*. 63:18, 4409-4417.

   IF (2016): 3,154

3. **Farkas, Zs.**, Slate, A., Whitaker, TB., Suszter, G., Ambrus Á. (2015b): Use of Combined Uncertainty of Pesticide Residue Results for Testing Compliance with Maximum Residue Limits (MRLs); *Journal of Agricultural and Food Chemistry* 63:18, 4418-4428.

   IF (2016): 3,154

Book chapter:

1. **Zsuzsa Farkas**, Jo Marie Cook, Árpád Ambrus: Estimation of Uncertainty of Measured Residues and Testing Compliance with MRLs, in Ambrus, Á., Hamilton, D. (szerk.) Food Safety Assessment of Pesticide Residues, World Scientific, New Jersey, 2017. 404-466. ISBN 978-1786341686.

Not peer reviewed

1. Ambrus, Á., **Farkas, Zs**., Horváth Zs., Kötelesné Suszter G. (2014): Principles and practices of control of pesticide residues in food, *Journal of Food Investigation* LX, 2, 8-32.

### Other publications

Peer reviewed

1. Horváth, Zs., Sali, J., Zentai, A., Dorogházi, E., **Farkas, Zs.**, Kerekes, K., Ambrus, A. (2014): Limitations in the determination of maximum residue limits and highest residues of pesticides: Part I. *Journal of Environmental Science and Health, Part B* 49:03, 143-152.

   IF (2016): 1,362

2. Ambrus, Á., Horváth, Zs., **Farkas, Zs**., Szabó, IJ., Dorogházi E., Szeitzné-Szabó, M. (2014): Nature of the field-to-field distribution of pesticide residues. *Journal of Environmental Science and Health, Part B* 49:4, 229-244.

IF (2016): 1,362

3. Trevisani, M., **Farkas, Zs**., Serraino, A., Zambrini, AV., Pizzamiglio, V., Giacometti F., Ambrus Á. (2014): Analysis of industry generated data Part I. A baseline for the development of a tool to assist milk industry in designing sampling plans for controlling aflatoxin M1 in milk. *Food Additives & Contaminants: Part A* 31:7, 1246-1256.

   IF (2016): 2,047

4. **Farkas, Zs**., Trevisani, M., Horváth, Zs., Serraino, A., Szabó, IJ., Kerekes, K., Szeitzné-Szabó, M., Ambrus, Á. (2014): Analysis of industry generated data Part II. Risk based sampling plan for efficient self-control of aflatoxin M1 contamination in raw milk. *Food Additives & Contaminants: Part A* 31:7, 1257-1273.

   IF (2016): 2,047

5. Kerekes, K., Bonilauri, P., Serraino, A., Giacometti, F., Piva, S., Zambrini, V., Canever, A., **Farkas, Zs.**, Ambrus, Á. (2016): An effective self-control strategy for the reduction of aflatoxin M1 content in milk and to decrease the exposure of consumers. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess*. 12, 1840-1849.

   IF (2016): 2,047

Not peer reviewed

1. Ambrus, Á., Horváth, Zs., **Farkas, Zs**., Cseh, J., Petrova, S., Dimitrov, P., Duleva, V., Rangelova, L., Chikova-Iscener, E., Ovaskainen, ML., Pakkala, H., Heinemeyer, G., Lindtner, O., Schweter, A., Naska, A., Sekuła, W., Guiomar, S., Lopes, C., Torres, D. (2013): Pilot study in the view of a Pan-European dietary survey - adolescents, adults and elderly. *EFSA Journal*. Available online: www.efsa.europa.eu/publications.

2. **Farkas, Zs.**, Kerekes, K., Szabó, IJ., Ambrus, Á. (2014): MS Excel-based method for the preparation of target-oriented sampling plans, *Élelmiszervizsgálati Közlemények / Journal of Food Investigation* LXI, 2, 588-609.