CONVENTION ON LONG-RANGE TRANSBOUNDARY AIR POLLUTION
INTERNATIONAL CO-OPERATIVE PROGRAMME ON ASSESSMENT AND MONITORING
OF AIR POLLUTION EFFECTS ON FORESTS
AND
EUROPEAN UNION SCHEME
ON THE PROTECTION OF FORESTS AGAINST ATMOSPHERIC POLLUTION

Economic Commission
for Europe   European Commission   Flemish Community

Evaluation of the third FSCC Interlaboratory Comparison

and

Practical Recommendations towards the Future
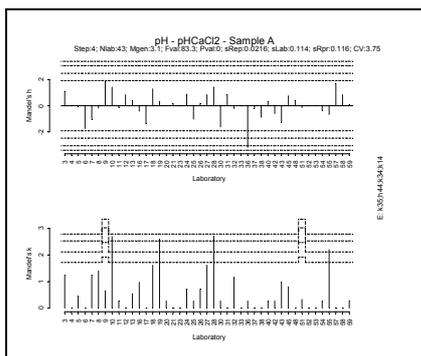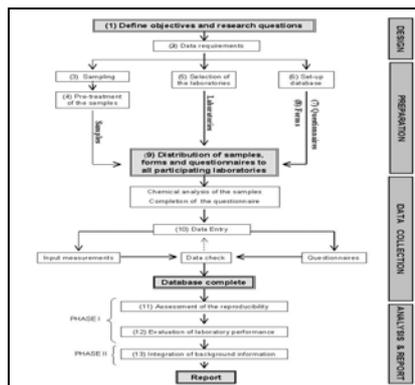
Document based on the third FSCC interlaboratory comparison
IBW.Bb.R.2004.005

Forest Soil Co-Ordinating Centre
Institute for Forestry and Game Management
April 2004
Elaborated by Delanote V., Cools N., De Vos B., Roskams P., Quataert P.

TABLE OF CONTENTS

Introduction

History

During the period 1985-2003, FSCC organised three ring tests among the laboratories of the ICP-Forests programme.

A <u>First Intercalibration Exercise</u>, with 22 participating laboratories from 22 different countries, used 4 standard soil samples and aimed at comparing different national analysis methods (Van der Velden and Van Orshoven, 1992). This comparison revealed a high variance between the results obtained by different methods and established the need for harmonisation of the methodologies.

Therefore a <u>Second Intercalibration Exercise</u> (Vanmechelen *et al.,* 1997), with 26 participating laboratories from 25 different countries, using 2 soil samples, was conducted in 1993, simultaneously with the analysis of the collected soil samples of the Level I plots. Laboratories using national methods were recommended to analyse the standard soil samples with both national and reference methods, in order to provide a basis for comparison. Once more the existing variance, especially between different methods, asked for the uniform use of reference methods.

In view of a second European wide soil survey, harmonisation and improvement of the analytical techniques is considered indispensable. In order to assure the quality of the data obtained by soil-analysis, the 10th Forest Soil Expert Panel (Warsaw, 2000) decided to proceed a <u>Third Intercalibration Exercise</u> (FSCC, 2003a).Fifty two laboratories from 26 different countries participated. FSCC finalised this third ring test in June 2003.

On the 11th Forest Soil Expert Panel meeting (24-26 March 2003, Ghent) it was stated that the results of the third interlaboratory comparison show that further improvement of the quality of the chemical analyses is needed. In order to ensure data quality, a new <u>Fourth Intercalibration Exercise</u> should be carried out prior to the next survey, or in 2005 at the latest.

This planned ring test was the impetus to evaluate the conduct of the previous interlaboratory comparisons and to develop practical recommendations towards the future. Based on the experience from the organisation of the third interlaboratory comparison, this document will set out guidelines for the organisation of ring tests on soil.

Scope of the document

First of all, it should be clear that this document is mainly based on the experience of the organisation of the third FSCC interlaboratory test. It is thus strongly recommended to read this document together with the ringtest report (FSCC, 2003a).

This document will discuss the four major phases of the organisation of the interlaboratory comparison:
> (I.)   Design,
> (II.)   Preparation,
> (III.)   Data Collection and
> (IV.)   Statistical Analysis and reporting.

In stead of making an abstract discussion, each paragraph describes three items: the **description** of the arrangements of the third FSCC interlaboratory comparison and the **evaluation** of the followed procedure, concluded by practical **recommendations** towards the organisation of future ring tests.

The global procedure of the organisation of an interlaboratory comparison is set out in Figure 1. This flow-chart shows all steps to guarantee a successful interlaboratory comparison. All steps of the procedure will be discussed in depth. The numbers in Figure 1 refer to the corresponding paragraphs.

The use of an uniform methodology for ring tests should stimulate the comparability of results. Moreover, experience from former ring tests might contribute to avoid practical problems. This document could be a first step towards a more uniform approach for interlaboratory comparisons.
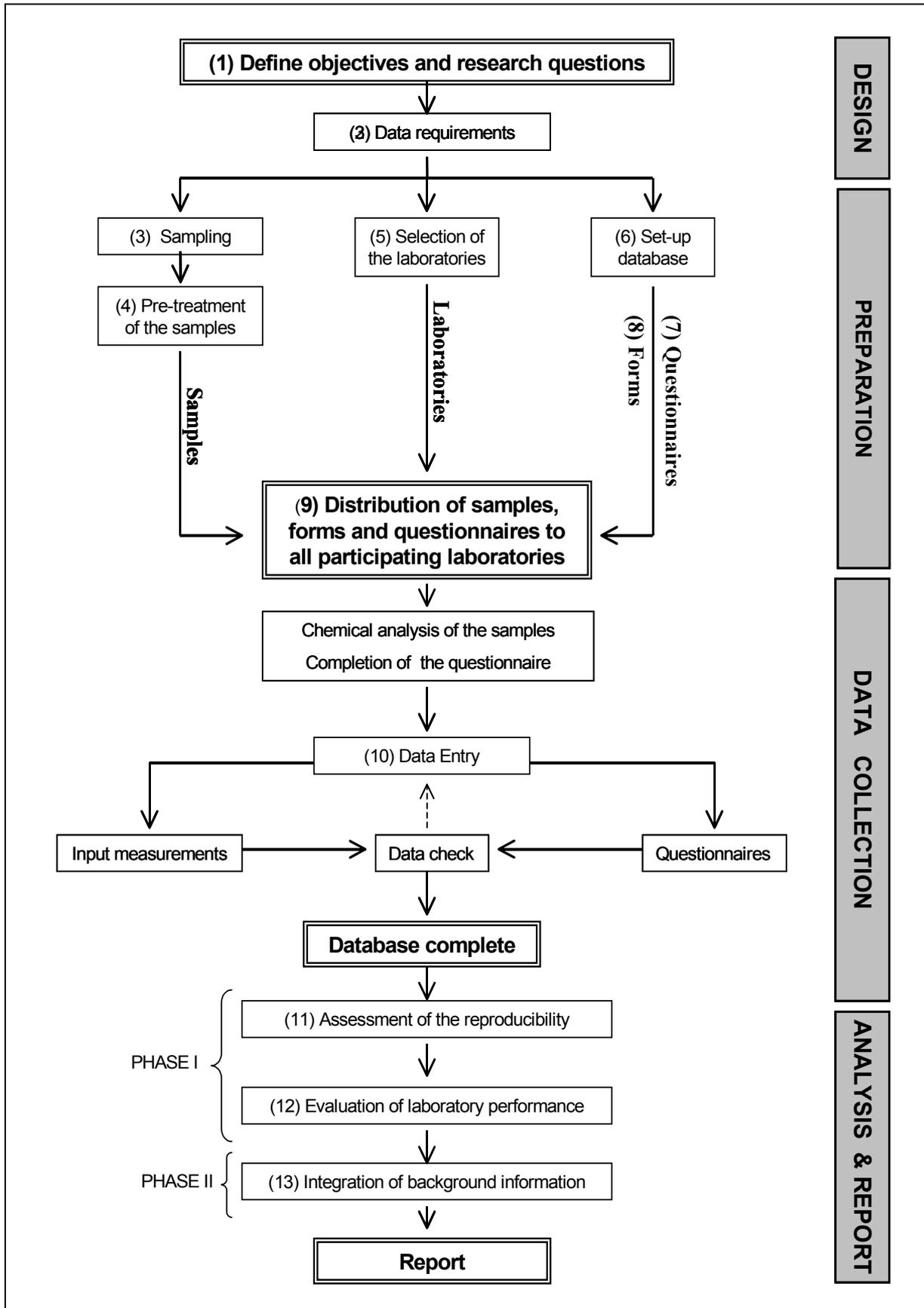
Figure 1: Flow-chart for the organisation of an interlaboratory comparison

---

I. DESIGN

---

1. Definition of the objectives and research questions

Description

The 3rd FSCC interlaboratory comparison tries to answer following research questions:

(1)   Obtain an overall view on the quality of the measurements;

(2)   Evaluate the performance of the participating laboratories;

(3)   The influence of the characteristics of the laboratories on their performance.

The third ring test aimed in the first place at answering the first two research aims. In a second phase, also the third research question was considered.

Evaluation

The chosen design depends strongly on the priority of the objectives. Objective one (and two) profit from a design based on voluntary participation (closer to the real situation of a survey). On the contrary, a more balanced design (see also paragraph 14) is necessary for the third objective.

Recommendations

- Be aware that the objectives of the interlaboratory comparison will define the design of the test. A careful evaluation of the research questions is indispensable to set out a suitable design.

- Realise that, based on one single intercalibration exercise, it is difficult, if not infeasible, to answer simultaneously the three research questions. Each research question demands a specific approach.

---

2. Data requirements

Description

The third FSCC ring test had the intention to answer the three research questions stated above. Therefore, two types of data were required; the results of the chemical analysis on one hand, and background information on the laboratories on the other hand.

Choice of the soil samples

In view of the future Pan-European survey, it is advisable to include different types of forest soils into the interlaboratory calibration exercise. As forest soils vary widely across Europe (Van Mechelen *et al.*, 1997), a ring test should try to include totally different samples representing as much as possible the variety of European forest soils. It is also important to include organic samples (L-layer, F+H layer) in the test. As this type of samples are typical for forest soils, laboratories not specialised in forest soils might have little experience analysing organic soil layers.

Although representative for large forested areas, some problems might occur when extremely poor samples are included in the test (cfr. sample A in the third FSCC ring test). As all participating laboratories should be able to handle this type of samples, it is not necessary to discard them from the exercise. As results may be below detection limit (DL), this type of samples requires strict rules concerning the reporting of the results (see paragraph 7).

Choice of the parameters

In view of the third FSCC Interlaboratory, all parameters (mandatory and optional) described in the manual (FSCC, 2003b) were included in the ring test. This means that in total 42 parameters had to be analysed, the analyses can be grouped as follows:

- **Group I: Particle size distribution (SA03)**
    Particle Size – Clay
    Particle Size – Silt
    Particle Size – Sand

- **Group II: pH (SA06)**
    $pH(CaCl_2)$
    $pH(H_2O)$

- **Group III: Carbonate content (SA07)**
    Carbonates

- **Group IV: Organic Carbon (SA08)**
    Organic Carbon

- **Group V: Total nitrogen content (SA09)**
    Total N

- **Group VI: Exchangeable cations (SA10)**
    Exchangeable Acidity
    Exchangeable Al, Ca, Fe, K, Mg, Mn, Na,
    Free H^+ Acidity

- **Group VII: Aqua Regia Extractable elements (SA11)**
    Extractable Al, Ca, Cd, Cr, Cu, Fe, Hg, K, Mg, Mn, Na, Ni, P, Pb, S, Zn

- **Group VIII: Total Elements (SA12)**
    Total Al, Ca, Fe, K, Mg, Mn, Na

- **Group IX: Acid oxalate extractable Fe and Al (SA13)**
    Reactive Al
    Reactive Fe

The choice to include the complete set of parameters into the interlaboratory comparison was made in view of the amendments made to the manual. The interlaboratory comparison was seen as an excellent occasion to evaluate the use of the revised methods (based on ISO-standards) of the manual.

Questionnaires

The questionnaire (see also paragraph 8) should provide the necessary data to obtain a view on the characteristics of the laboratories.
The questionnaire of the 3rd ringtest asked for general information on the laboratory and on the analysis.

Evaluation

Data requirements depend on the objectives of the interlaboratory comparison. The number and the type of samples (see paragraph 3) will be determined by the priority of the objectives.

If the ring test is in the first place organised to get a view on the overall data quality, it is interesting to include some "difficult" samples, as it is of great importance to know how well laboratories can handle this type of samples. For instance, including an extremely poor sample can provide a lot of information of the way laboratories deal with data close to detection limit.

The number of parameters included in an interlaboratory comparison can be discussed. A large amount of parameters increases the costs and efforts of participation. On the other hand, this is the only way of getting a complete view on the quality of the data.

If the main objective of the interlaboratory test is to search for the causes of differences between the laboratories, secondary information is needed. The level of detail of the questionnaire depends on the aimed objectives.

Recommendations

- Once the priority of the objectives is stated, the data requirements can be defined. These data requirements will completely depend on the objectives of the interlaboratory comparisons.

- Try to find an acceptable <u>number of samples</u>, in order to make the ring test both representative (the more samples, the better) and the financially feasible. If financially feasible, three to five soil samples are advised, including three samples in a 'normal range' and two 'extreme samples'.

- A <u>variety of samples</u> is preferred, so include organic samples and mineral samples of different soil types. In this way, laboratories get the chance to demonstrate their skills on different types of samples.

- Set out strict rules concerning reporting of results, especially when dealing with <u>extremely poor soil samples</u>. These samples are more difficult to analyse, nevertheless they are representative for large forest areas in Europe. It is important that laboratories learn to handle this type of samples.

- The number of parameters included in the test strongly depends on the research questions.

II.	PREPARATION

3. Sampling

Description

Sampling period

The sampling for the third FSCC ring test was carried out on two different days:

Sample A & C :   06/06/2002
Sample B: 12/06/2002

Sampling Location

Three samples were included in the 3ʳᵈ FSCC ring test, of which 2 mineral soil samples (A,B) and one organic sample (C). These samples were taken under forest conditions in different regions of Belgium.

Table 1 : Characterisation of the soil samples

	Sample A	**Sample B**	**Sample C**
Type	Mineral layer	Mineral layer	Organic layer
Location	'Withoefse heide'	'Mons trou des sarassins'	'Withoefse heide'
Depth	Ah – Horizon	0-10 cm depth	L+F layer
Soil texture	Loamy sand soil	Silty clay soil	Organic material
Chemical characterisation	Chemically poor soil	Very calcareous, High nutrient content	Organic material
Vegetation	Coniferous forest, Scots Pine dominant	Mixed deciduous forest Ash, pedunculate oak, maple	Coniferous forest, Scots Pine dominant

Sampling Method

The mineral soil samples (sample A & B) were taken using a regular push probe. As seen on Plate 1, several samples were taken in the sample area, in order to sample a sufficient amount of material.

15-20cm

Plate 1: Sampling of the mineral samples (sample A)

Based on the analyses of all parameters included in the 'Submanual on analysis and sampling of soil' (FSCC, 2003b), including three replications, the necessary amount of material is calculated at 500 g dry

matter (DM) per sample. This amount of material should allow the laboratories to carry out all analyses in triplicate, leaving some material in reserve.

Plate 2 : Sampling of the organic layer   **Plate 3: Sampling method organic layer**

As described in the 'Submanual on sampling and analysis of soils' (FSCC, 2003b), the F and H layer of the organic layer are sampled together. The sampling of the organic layer is done by use of a frame (50 cm by 50 cm) as shown in Plate 3.

Evaluation

There were no practical problems experienced when taking the samples in the field. Nevertheless, it could be useful to document the sampling more in the future.

Recommendations:

- Document the sampling in detail (exact sampling location, number of samples etc.).

- Before sending the samples to the participating laboratories, analyse the samples at least for the most important parameters in order to obtain a broad idea on the nature of the samples. Possible obstacles might pop up during the analyses (e.g. presence of $CaCO_3$ might hamper other analyses).

---

4. Pre-treatment of the samples

Description

FSCC conducted the pre-treatment of the samples, including three major steps; drying, the separation of the fine earth fraction (< 2mm) and homogenisation.

The first two steps (drying and sieving) are done according to the manual (FSCC, 2003b). Special care is taken concerning the homogenisation of the samples.

The soil must be thoroughly mixed before sub-samples are taken which will be sent to the laboratories. As the ring test focuses on small differences in measurements of different laboratories, it is of major importance that the difference between the constitution of the sub-samples is minimised. The soil samples were homogenised by an independent institute (VITO[1]) by means of riffling.

---

[1] VITO Vlaams Instituut voor Technologisch Onderzoek / Flemish Institute for Technological Research
Boertang 200, B-2400 Mol

Due to time and financial constraints, the riffling was repeated twice on sample B, but not repeated twice on sample A and C.

Evaluation

Note that the texture of sandy material might cause problems as this type of material is, due to its physical nature, very difficult to homogenise correctly. Since sample A, was a sandy sample, it was extremely difficult to homogenise.

Annex I (Homogeneity test) discusses the test of homogeneity in depth. The results of the homogeneity test are summarised in Table 2.

The tests of homogeneity have been carried out twice in the soil laboratory of the IBW[2]: The first time before the samples were sent out to the laboratories in July 2002 (texture and total organic carbon). Later on the test was repeated in March 2003 (Kjeldahl N and loss on ignition) as it was realised that no appropriate statistical design was applied in July 2002. Towards next ring tests, it is recommended to follow the second procedure illustrated in Annex I.

Table 2: Results of testing the hypothesis that there are no significant differences between the subsamples

Parameter	Sample A	Sample B	Sample C
Clay content	Not significant	Not significant	NA
Median texture class	Significant	Not significant	NA
Total Carbon	Significant	Not significant	NA
Loss on Ignition	Significant	Not significant	Significant
Total N	Not significant	Not significant	Not significant

Based on the homogeneity tests, for sample B none of the tests showed significant differences. There were some significant differences for sample A and C, however these results could be explained by the low values (close to the detection limit). Yet, there are some issues to be considered in the future.

Recommendations:

- Homogeneity of the samples is the <u>absolute condition</u> for a good ring test. For this reason absolute attention should be paid to the homogenisation of the samples. Before sending the samples to the participating laboratories, the homogeneity of the samples should be guaranteed. A homogeneity test before the shipment of the samples is indispensable.

- Be aware of the fact that homogenisation of the samples <u>costs time and money</u> (eg. € 2000 for homogenisation of the three samples at VITO)! Special equipment is needed for the homogenisation (eg. riffling). A certificate of homogeneity might be an unexpected cost (eg. € 5000 at VITO).

- Realise that the physical nature of sandy samples might hamper the homogenisation. It is very difficult to subsample these sandy soils correctly. As was discussed on the 11th FSEPM, sandy soils can only be homogenised thoroughly when grinded, but this might release more nutrients and give thus a distorted view on the results of the analysis.

---

[2] IBW – Instituut voor Bosbouw en Wildbeheer / Institute for Forestry and Game Management
  Gaverstraat 4
  B-9500 Geraardsbergen

5. Selection of the laboratories

Description

It was the task of the NFC's[3] to select the appropriate laboratories. FSCC advised to select more than one laboratory per country for ring test evaluation. In order to estimate the necessary amount of soil samples, all participating labs register for participation.

The identity of the participating laboratories is kept confidential. FSCC gave each laboratory a unique code. These codes are only known by FSCC and the laboratory itself. NFC's could become a list of the identification numbers of the participating laboratories of their country. The ring test report only mentions the code of the laboratories.

Evaluation

As the NFC's are free in their choice of laboratories, the design of the ring test might become unbalanced (e.g. the German NFC chose 13 laboratories, whereas other countries selected a single laboratory). Moreover, no specifications were given towards the type of laboratories, what implements an unbalanced design of the laboratories (see paragraph 13).

Recommendations

- The selection of the participating laboratories can best be made by an institution that has background information on the laboratories. NFC's are ideal contact points, as they are familiar with the structure of ICP-Forests and know most the laboratories of their country.

- Depending on the main objective of the ring test laboratories can be selected randomly, or carefully selected in view of a balanced design. See 'Integration of secondary data' (paragraph 13).

- The coding of the laboratories should be completely randomised. The use of the code is strictly limited to the statistical analysis. Make sure that the code is not used in address lists, e-mails, etc.

6. Set-up of the database

Description

As laboratories are asked to provide two types of information, results of the chemical analyses on one hand and background information on the other hand, FSCC prepared two documents for data submission: forms and questionnaires. The set-up and testing of these two documents will be discussed in paragraphs 7 and 8. This paragraph will focus on the set-up of the database where the results of both inquiries will be stored.
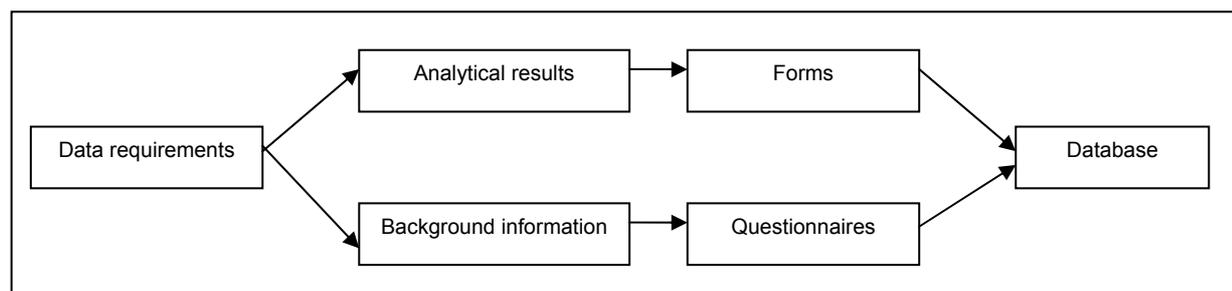
Figure 2: Components of the database

---

[3] NFC: National Focal Centre

All results were stored in an Microsoft Access database. This database has, as shown in Figure 2, a bipartite structure. Next to the reported results (forms), the answers of the questionnaires and the general data of the participating laboratories are linked to each other as shown in Figure 3.

The table 'Data' contains all analytical results of the laboratories. The table 'Questionnaire' is based on the extra information provided in the questionnaire and the third table 'Registered laboratories' contains the contact information of the participating laboratories. These three major tables are linked together by the key 'Labo id' (unique and secret code number for each of the participating laboratories).
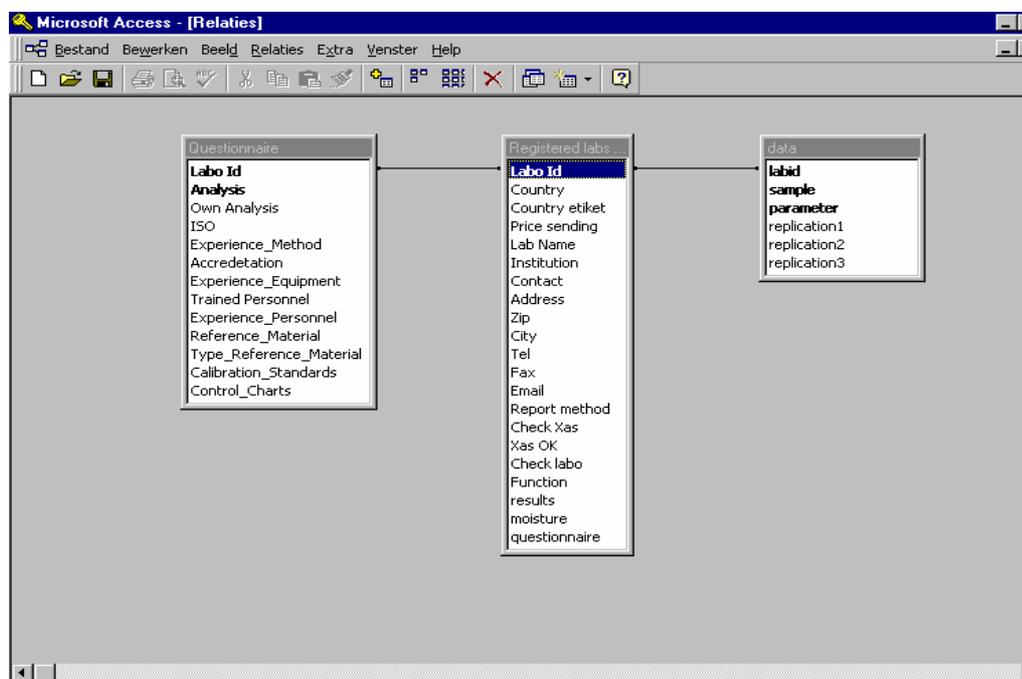
Figure 3: Set-up of the database

Evaluation

An input of the data through the FSCC-website could facilitate the data entry. Currently not all laboratories have easy access to the internet. However, it might be an option towards the future.

Recommendations

- The database design must be set-up in advance, before reporting of the results. Experience learns that Microsoft Access provides enough possibilities for the set-up of the database. If the number of laboratories is limited, an Microsoft Access-database is powerful enough to handle the data.

- The data of the questionnaire can easily be linked to the table with the results of the analysis. The code of the laboratory is used as an unique key.

---

7. Forms

Description

The standard forms (MS Excel) for the input of the results are shown in Annex II. Each from is provided in triplicate to the laboratories (one for each of the three samples). Each form consists of a table including all parameters, the correct unit and space to report the results for each of the three replicates.

Evaluation

When creating the forms, some problems that might hamper the reporting later on (see paragraph 10), must be anticipated. The problems that laboratories faced, reporting for the 3rd FSCC ring test are listed below. Most of these problems could have been avoided by creating forms of high quality, combined with clear guidelines for reporting.

Problems of rounding:

Laboratories have been asked to report a fixed number of digital numbers according to ISO standards and/or to the Forest Soil Condition database. Though, it has not been taken into account that sample A was extremely poor in nutrients. Concentrations were sometimes extremely low. Reporting a limited fixed number of digits, could cause a false reporting of zero values. After entering the results in the database, no distinction could be made between real zero's (reported as "0") and false zero's (truncated small values).

For this reason it is advised to allow laboratories to give their results in the highest precision they can achieve with the available equipment. If considered necessary it is always possible to convert the data into a lower precision afterwards.

Oven dry versus air dry:

Laboratories were asked to report their results on oven dry basis. Nevertheless a certain amount of laboratories reported on air dry soil. If the analysis were carried on air dry soil, a correction factor (moisture content of the soil) is necessary to convert the data.

The conversion from air dry soil to over dry soil can be done according the following formula:

Results on oven dry soil = Results on air dry soil * [1- moisture content(%)]

Note: The soil moisture content is a mandatory parameter for both organic and mineral layers, reporting in % with one decimal place.

Problems of detection limit:

It is of major importance that a clear distinction is made between values below detection limit and un-measured parameters.

For each parameter, the detection limit should be given (see further 'possibility of a DAR'). Values below the detection limit are considered too uncertain to report.

$$DL = 3s \qquad \text{(includes 99.7\% of the observations)}$$

Where s is the standard deviation of the measurement.

The reporting of values below the detection limit can lead to confusion and misunderstanding. For this reason, a clear and straight-forward handling of these values should be set out. It was agreed in the third FSCC ring test that values below detection limit should be reported as negative values (-1 in the ICP-forms).

In the third FSCC ringtest we met another problem related to the reporting of detection limits. Confusion arose whether the laboratories needed to report the detection limit or the quantification limit. In the ringtest report (Annex 10, p300) we recommended to use quantification limit. Both approaches are of equivalent value, as long as they are not mixed up.

Since in the other programmes of ICP-Forests, priority is given to the use of the detection limit, reported as –1 in the forms, we recommend a similar approach for the purpose of uniformity.

Possibility of a separate DAR

In view of the 3rd FSCC Interlaboratory Comparison, the detection limit was asked as a part of the questionnaire. Experience of the third (and former) intercalibration exercises learns that if this essential information is asked in the questionnaire, laboratories might forget to provide this essential information due to the large amount of information that is asked for.
For this reason it seems more appropriate to ask for the DL of the different analyses in a separate form, the so-called DAR (Data Accompanying Report).

This DAR should be seen as a table, containing all information necessary to interpret the reported results correctly (e.g. DL). The information of the DAR is essential and closely linked to the data itself.
When a DAR is used, a clear distinction is made between data-linked information and background information. This should stimulate the laboratories to pay more attention to the reporting of detection limits etc..

Recommendations

- Consider to develop a DAR (data accompanying report) in which all essential information to interpret the data correctly can be gathered.

- It is of major importance that clear arrangements are made before the reporting of the results. The handling of values below detection limit, rounding of results, etc. must be agreed on before data reporting.

- Care should be taken that the definition of DL is well understood. There is some discussion whether the quantification limit should be used in stead of the detection limit. One should clarify this issue before starting the interlaboratory comparison.

- One should ask (preferably in a DAR) for the detection limit or quantification limit for each procedure. Be careful, ask for <u>one</u> value: <u>or</u> detection limit, <u>or</u> quantification limit. Be consequent, so do not ask the QL for one procedure and the DL for another.

---

8.   Questionnaires

Description

The ring test was accompanied by a lab questionnaire (see Annex III) to collect data on the participating laboratories. Each laboratory was asked to provide:

1) General information on the laboratory:
   Address, Contact information, Type and Statute of the laboratory, Certificates, Number of personnel, Years of experience.

2) Analysis information. This information was collected under the format of multiple choice questions concerning analysis method, used equipment, experience of the personnel and procedures of quality assurance.

Evaluation

The objectives of the questionnaire should be clearly stated in advance. The questionnaire was meant to get a general view on the methods and procedures that were used for the analyses of this ring test in particular. This point was not made very clear to the laboratories since some of the participating laboratories answered the questions in view of their daily laboratory practices.

As sub-questions might cause confusion, it is advisable to avoid the use of sub-questions as much as possible.

e.g.   1. Is the lab working according to the reference method?
                   2. Has the lab experience with the reference method

Some laboratories answer no to the first question, but answer the second question with 'high experience'.

This type of confusion could be minimised by thoroughly checking the phrasing before distributing the questionnaires. A restricted test phase of the questionnaire, (sending the questionnaire to a limited number of non-participating laboratories, asking them to check the questionnaire thoroughly) could point out the questions that might cause problems. This will avoid misunderstandings during the operational phase of the questionnaire.

Preference should be given to multiple choice questions, this type of questions gives less confusion than questions with open answers. Moreover binary questions facilitate the input of the questionnaires in a database. Afterwards, binary contrasts are statistically easier interpretable (binary factor levels give single contrasts).

Recommendations

- Restrict the number of questions in the questionnaire to a minimum. Less data of good quality are preferred to an abundance of data of poor quality.

- The objective of the questionnaire must be clearly stated: the questionnaire tends to give a general view on the methods and procedures of the analyses of the ring test in particular. Daily laboratory practices are not under discussion.

- To avoid confusion and misunderstanding, questions must be formulated clearly. It is advised to test the questionnaire before sending it to the participating laboratories. Testing can be done for instance by asking some laboratories (that don't participate in the ring test) to fill in the questionnaire and report all possible pit-falls.

- Closed questions (multiple choice) are advised for several reasons: they minimise confusion, facilitate the input of the questionnaires in the database, moreover, binary contrasts are statistically easily manageable.

III. Data Collection

9.   Distribution of the samples

Description

Samples were sent to the participating laboratories during the first week of July. This timing caused for certain laboratories some difficulties due to summer holidays.

All NFC's, members of the Forest Soil Expert Panel an lab responsibles were informed on the shipment of the soil samples. All samples were sent by Taxipost. They were all identified by a unique code (possibility to trace the packages if necessary). The lab responsibles were asked to confirm the arrival of the samples.

Shipment of soil samples to Russia needed to be accompanied by a letter which declares that the samples are of no commercial value (see Annex IV).

Evaluation

The distribution of the samples for the third ring test passed smoothly, except for the Russian laboratory that faced some problems with customs.

The timing of the distribution of the samples was perhaps a bit inconvenient, as laboratories received the samples during summer holidays.

Recommendations

- Make sure that soil samples are well packed, humidity might change the constitution of the soils.

- Inform laboratories on sending date, and ask them to acknowledge the receipt. Identify the packages by an unique number (e.g. invoice number given by mailing service), in order to be able to track the packages if necessary.

- All parcels sent outside Europe must be accompanied by a letter which declares that the packages are of no commercial value.

---

10.   Data entry

Description

Reporting of results was done by a standard file format (Excel), provided by FSCC (see Annex II). As soon as laboratories reported to FSCC, results were added to the MS Access database.

After data have been entered in the database, each of the laboratories was sent a copy of their data in a PDF-file by the beginning of November 2002. The pdf-files were in exactly the same format (same number of digits, etc.) as the data were entered in the system. Laboratories could check, comment and correct the reported values if needed till the beginning of December 2002.

Evaluation

No problems arose importing the MS Excel files into the database. However, a web-application could facilitate the input of the data.

Recommendations

- Reporting directly into the database, if feasible, is not yet advised but would be interesting towards the future. Practically this means that laboratories enter their data in an MS Access-form that is directly linked to the database. This can possibly be done through the FSCC-website, where laboratories can log in by use of their laboratory number and a secured password. Although practical, this way of reporting has a few disadvantages; one must secure the reporting-procedure carefully, in order to avoid misuse of the system.

- All laboratories should get the opportunity to check their data and correct their results if necessary. After this check, participating laboratories confirm their results.

- Give the participating laboratories the opportunity to check (and if necessary correct) their results.

IV. STATISTICAL ANALYSIS AND REPORT

Introduction

The central *concept* of an interlaboratory study or ring test is **reproducibility**. Reproducibility reflects how confident we can be about the measurements of a certain item (and hence how useful they are).

The **reproducibility variance** expresses the (dis)agreement between the results obtained with the same method on identical test or reference material under different conditions (carried out by different persons, in different laboratories, with different equipment and at different times).

A ring test (*under repeatability conditions*; see further) allows decomposing the reproducibility variance into two (random) components: (1) the **repeatability variance** and (2) the **between-laboratory variance**:

$$\sigma_R{}^2 = \sigma_r{}^2 + \sigma_L{}^2$$

1. The **repeatability variance** ($\sigma_r{}^2$) is a measure of (dis)agreement between results obtained with the same method on identical test or reference material <u>under the same so-called **repeatability** conditions</u>: measurements repeated by one person, in the same laboratory, with the same equipment, within a short time interval. It reflects the within-laboratory variance (under repeatability conditions). The corresponding quality *concept* is **precision**.

2. The **between-laboratory variance** ($\sigma_L{}^2$) is a measure of (dis)agreement between results obtained with the same method on identical test or reference material <u>in different laboratories</u>. It reflects the laboratory effects or the systematic differences between laboratories. The corresponding quality *concept* is **accuracy** (or its opposite **bias**).

An alternative expression of the reproducibility is the **coefficient of variation**. It is a relative measure with respect to the mean value (and often expressed as a percentage). This gives a better impression how large the uncertainty on the measurements is. Very often the coefficient of variation is constant for increasing mean which implies that in absolute terms the reproducibility variance increases.

$$CV(\%) = 100 \frac{\sigma_R}{\mu}$$

To obtain a realistic estimate of the reproducibility, an important question is which laboratories to include in the study (see also data requirements / sampling). One point of view is that any sample of laboratories will do. Then we will get an impression of the general quality of all possible laboratories. This is not very informative if we are interested in **the highest quality achievable at current standards** (FSCC, 2003b). For this question a more restrictive sample is necessary. By accepting only laboratories working at - or at least approximating - current standards, we will get an estimate of the best we can achieve at the moment. It should be clear however that a restrictive policy risks giving a too optimistic and hence unrealistic view on the quality of the measurements if in practice we cannot realise a good selection of the laboratories.
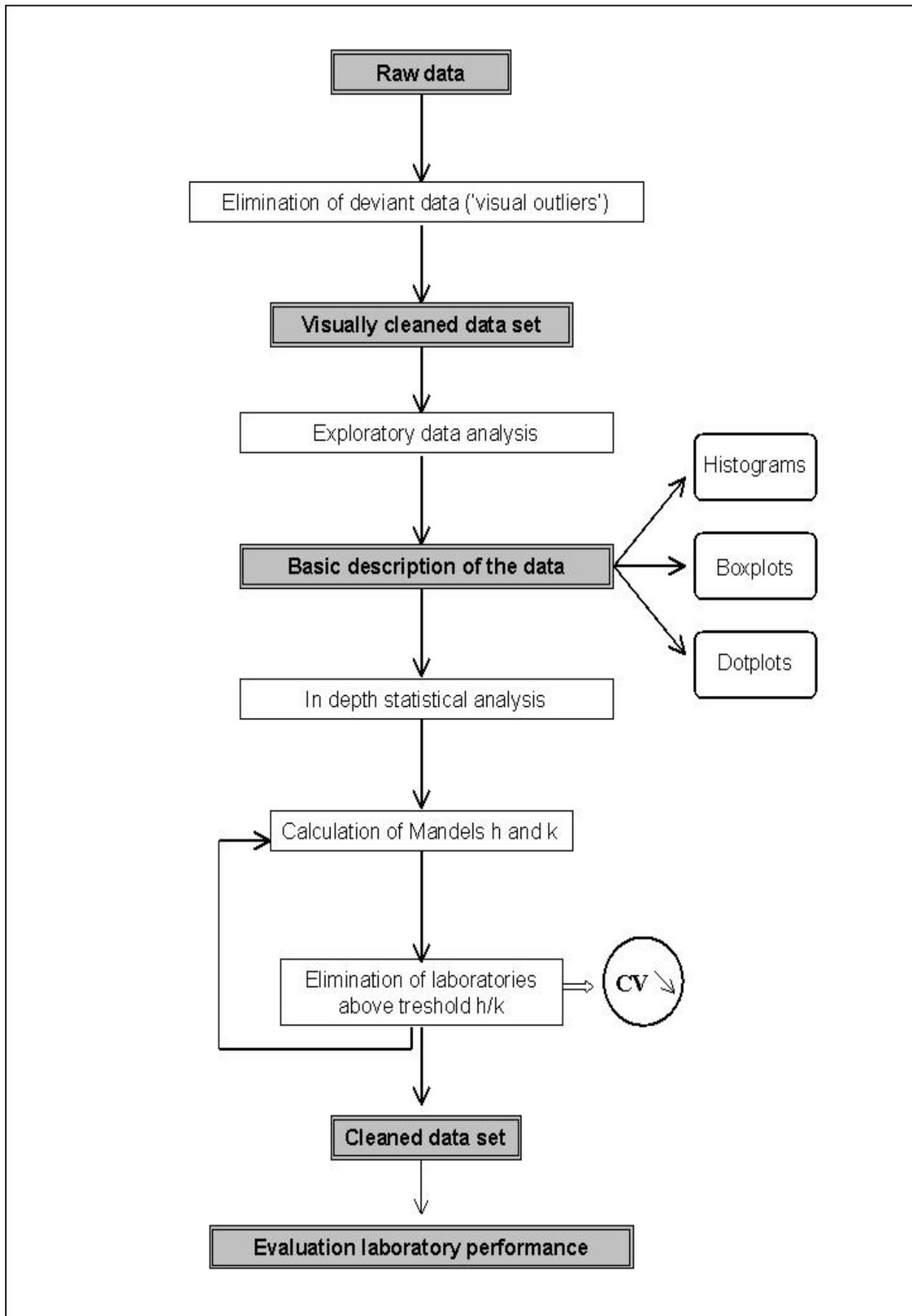
Figure 4: Flowchart of the first phase of the statistical analysis

11. Assessment of the reproducibility

See 'Quality Assurance and Quality Control in Forest Soil Analysis: 3rd FSCC Interlaboratory Comparison – Phase I (FSCC, 2003a)

Description

Step 1: exploratory data analysis.

As for any statistical analysis, the first step is (visual and tabular) data exploration to get a feeling with the data and to screen for major anomalies. For each element and each sample, the average by laboratory as well as the spread of the measurement within the laboratories is compared with respect to each other by *histograms*, *boxplots* and *ordered dotplots* (an ordered dotplot orders the laboratories according to their mean or median value to show the ranking). Deviant values which would distort the graphical representation are excluded and reported separately (without definitive exclusion, therefore a statistical approach will be used in the next step). In this way a first impression is obtained about the distribution of the measurements and the relative position of the laboratories.

Step 2: statistical evaluation of the consistency of the measurements and elimination of outliers.

The next step is a more formal statistical analysis of the consistency of the dataset. Two consistency statistics evaluating respectively **location** and **spread** of the measurements are used:

1. <u>Mandel's h</u> controls if the average location of a laboratory (m_j) is not too deviant from the global mean m ('the bulk of the data') by comparing it with the spread of the means (s_b). If so, the laboratory lacks **accuracy** (at least for the element and method considered).

$$h_j \triangleq \frac{e_j}{s_b} = \frac{m_j - m}{s_b} \quad j : 1 \rightarrow p$$

2. <u>Mandel's k</u> evaluates if the spread of the measurements within a laboratory (s_j) is not too deviant from the average spread (s_w). It is a measure for the **precision** of a certain laboratory with respect to the others. The assumption behind this test is that the measurements are a done under of repeatability conditions (see higher).

$$k_j \overset{\Delta}{=} \frac{s_j}{s_w} \quad j : 1 \rightarrow p$$

Mandel's h and k are related closely with the t- and F-distribution and hence we can use this link to assess their significance to judge whether a laboratory is deviant or not (*p* = number of laboratories)

$$h_j = \frac{p-1}{\sqrt{p}} \frac{t_j}{\sqrt{t_j^2 + p - 2}} \rightarrow t_j \propto t_{p-2}$$

$$k_j = \sqrt{\frac{p f_j}{(p-1) + f_j}} = \sqrt{p C_j} \rightarrow f_j \propto F_{n-1;(p-1)(n-1)}$$

This procedure tests p laboratories. To compensate for multiple testing, a correction of the significance level is necessary to protect the null hypothesis. Otherwise too many outliers will be found and eliminated such that the reproducibility will be overestimated. A sufficient (but sometimes unduly conservative) correction is dividing the classical α by p. This so-called Bonferroni correction guarantees that the global significance level at least as small as α. Further the standard chooses a small value for the significance level, namely α = 0.01. Only very deviant laboratories are declared as **outliers**. For α = 0.05 a laboratory is flagged as a **straggler**, i.e. an exceptional point to be further investigated, whereas the outliers are eliminated.

This approach keeps large values in the sample unless they are very deviant. The philosophy behind this rule is to get a realistic view on the real variability of the measurements. Figure 5 shows the difference between **tail values, stragglers and outliers**. A tail value is not exceptional; it is just a large value, which is not abnormal, given the number of laboratories involved.
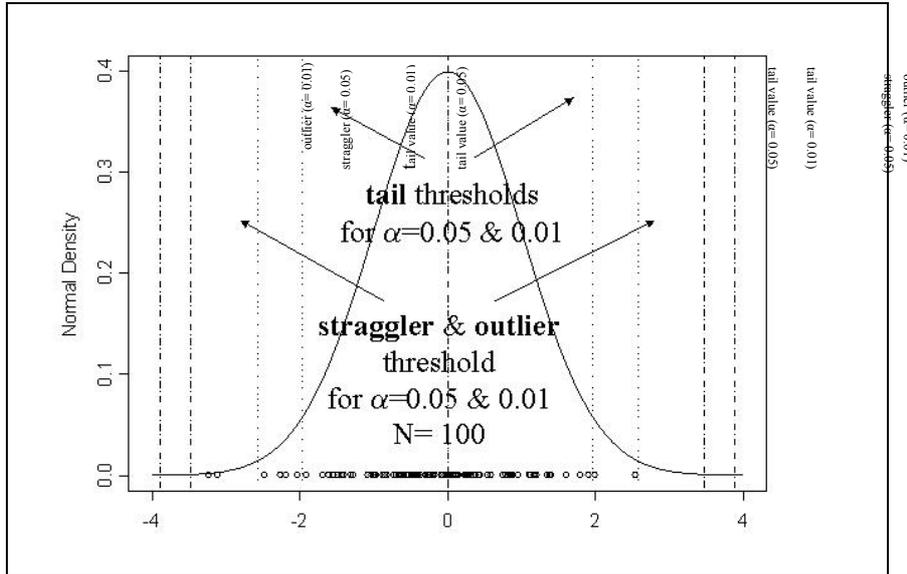
Figure 5: Different types of threshold values

The screening for outliers is an iterative procedure. By each step, outliers are eliminated, which leads to a shift in the distribution. As outliers can mask each other, new outliers may pop-up. The procedure is repeated until no outliers are detected, at that moment the dataset is 'cleaned' and we can proceed with the next step.

Step 3: estimation of the reproducibility.

The theoretical formula for the reproducibility is:

$$\sigma_R{}^2 = \sigma_r{}^2 + \sigma_L{}^2$$

The corresponding expression for estimation from a sample is:

$$\hat{\sigma}_R{}^2 = s_R^2 = s_r^2 + s_L^2 \quad with \; s_L^2 = \frac{s_b^2 - s_r^2}{n}$$

The first component of the reproducibility (the repeatability variance) can be estimated directly from the within-variances of the laboratories (s_r^2). For the second component, a somewhat more complicated approach is necessary. As the variance between the laboratory averages (s_b^2), contains also the within-variance (s_r^2), this term should be subtracted first to estimate the effect of the laboratory purely. That is what is shown in above formula.

Now, the coefficient of variation can be estimated by dividing s_R by the mean value (after elimination of the outliers).

$$CV\,(\%) = 100 * \frac{s_R}{m}$$

Step 4: Summary of the results into one single picture.

Before arriving at the final result, the dataset went through many, sometimes iterative manipulations. To get a concise overview (by element and sample measured) of what happened at the end one single picture is

generated with all essential information. Figure 6 is an example. The right margin specifies which inconsistent points are excluded and for which reason (by Mandel's h or k or both). The title gives the number of iterations (steps), the number of laboratories involved, the general mean, a significance test for the laboratory effects, the different variance components (repeatability, between-laboratory variance), the reproducibility variance and the coefficient of variation. The lines in the graph represent the consistency statistics compared with their thresholds (tail values, stragglers and outliers).
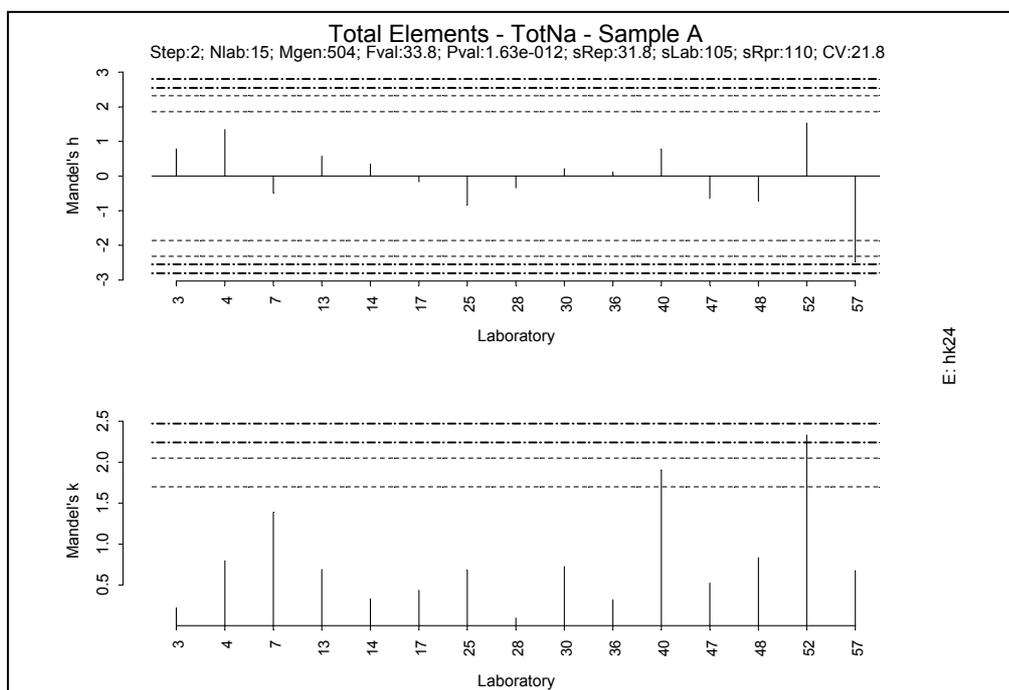
Figure 6: Example of the visualisation of the Mandel's h and k

Step 5: Summary of the coefficients of variation.

The final step is to make an overview of all the coefficients of variation (for each sample and each type of measurement). In this way we can judge the relative difficulty of the measurements and the samples and compare this with results in literature. Another use of the results is to compare it with a previous (or any other) ring tests.

Evaluation

(0) The development of an own procedure.

A strategic choice was to develop an own statistical procedure to master as much as possible the logic behind a ring test. This is because, in the (near) future, we will need this methodology on several occasions, not only for interlaboratory studies but also for quality control in general.

Instead a package as Ring 4.0 (Bartels, 2002) could have been used. A disadvantage of a commercial package is that not all steps are fully **transparent** ('black box'). Often the analysis is **closed**, i.e. one can analyse only what is provided and it is hard to work further on the results for further validation.

A general statistical package does not have these drawbacks. However, considerable time was necessary to develop and validate the routines.

(1) Relation to international standards

The statistical data analysis presented is **equivalent** with an international standard; the ISO 5725-2 "Accuracy (trueness and precision) of measurement methods and results – part 2: Basic method for determination of repeatability and reproducibility of a standard measurement method" (ISO, 1994c).

However, compared to the protocol, some items were adapted or added to facilitate analysis and interpretation:

1. All steps of the procedure have **a graphical representation.** This improves understanding of the analysis and allows to control the results visually (e.g. to look for unexpected or special cases). The procedure starts with an exploration of the original measurements to screen for **visual outliers**. Thereafter a rigorous statistical approach based on the ISO-standard is used to clean the data from outliers. The next step is the estimation of the reproducibility variance and the coefficient of variation. Finally the history and results are condensed into **one single picture** specifying the performance of the different laboratories with respect to each other.

2. It was possible **to reduce the number of tests for outliers and to eliminate the use of tables**. In fact, calculations show that only two consistency statistics (h & k) are necessary to judge the laboratories with respect to each other. The two statistics h & k are directly related to a t- and an F-distribution respectively, so no special statistical tables are necessary.

    ▪   The **h-statistic** judges if the average of the measurements of a certain laboratory are consistent with all other laboratories (evaluation of the location or *accuracy*), and,

    ▪   The **k-statistic** evaluates if the variation in the measurements of a certain laboratory do not exceed the average variation (evaluation of the spread or *precision*).

3. **The treatment of outliers is now based on clear unambiguous rules.** The procedure is iterative. If outliers are present, the statistical analysis is repeated without these outliers until no new outliers are found. The threshold for being an outlier is based on the Bonferroni approach to correct for multiple testing and to prevent that the dataset becomes too pure.

The same two principles are used in the ISO-standard 5725; but instead of the Mandel's h and k-statistics, two other tests are used. The ISO-standard evaluates the interlaboratory variance by means of the Grubb's test; the evaluation of the intra-laboratory variance is based on the Cochran's test. However it can be demonstrated that both approaches are equivalent.

(2) Repeatability conditions

Before starting the ring test, laboratories should be informed better about the way of evaluation. The evaluation of the internal variability (by Mandel's k) is uncommon in ring tests. This should be stated clearly and laboratories should be convinced to work under 'repeatability conditions'. This is important for the correct estimation of the reproducibility and for their performance as they can be excluded based on a too high internal variability.

The within laboratory variance is tested one-sided, only a too large laboratory variance is punished. The question remains if a too small within laboratory variability (e.g. exactly the same result for each of the three repetitions) could be seen as a problem.

Note that the calculation of the Mandel's k statistic is only possible if the laboratory reported at least the results of 2 **replicates**. For this reason, all laboratories reporting a single measurement for a certain analysis are rejected from the statistical analysis.

(3) The absence of a golden standard

The 'true' values of the soil parameters of each of the three samples are not known which means that the accurateness of the analyses can not be assessed. Instead the procedure uses "the bulk" of the data as the reference. The precision of the laboratory results is evaluated by comparing the measurements with the mean value, determined after exclusion of the outliers.

The implication is that laboratories using a different, more accurate technology might be penalised, because their results are too deviant from the general mean. An example is the determination of the particle size by laser diffraction (no reference method), which is considered to be a valid and possibly a more accurate alternative than by pipette method (Agrawal et al., 1991, Loizeau et al, 1994).

(4) Comparison of the Coefficients of Variation

Several elements hamper a direct comparison with the previous ring test and the results should be interpreted carefully.

A first problem is that outlying values have a large impact on the CV. So, different rules for treatment of the outliers can cause difference in CV. Hence, it was decided to reanalyse the previous ring test with the current method.

Also it is hard to judge if the degree of difficulty of both ring tests was equal. Especially the nature of the sample can have an important impact.

Recommendations

- Analysis of the data according to an accepted and **transparent standard** is an important prerequisite. A good example is ISO 5725-2 "Accuracy (trueness and precision) of measurement methods and results; part 2: Basic method for determination of repeatability and reproducibility of a standard measurement method" (ISO, 1994c). However, other good alternatives exist.

- The requirement to work under "repeatability conditions" should be **communicated** better to the laboratories. This is an important assumption behind the statistical calculations. Otherwise laboratories tend to repeat the measurements on different occasions to optimise accuracy. In the future one also should look at measurements which are systematically too close to each other (two-sided test of Mandel's k).

- Be aware there is **no 'golden standard'**. The procedure uses the bulk of the data as the reference. In some cases this can be totally wrong. More generally, there can be two (or more) populations of laboratories which are outliers with respect to each other. In this case the largest group can be considered falsely as the reference one.

- The **comparison of CVs** between two ring tests should be made carefully. A different approach in the elimination of outliers can have an important impact. Further, it is hard to judge if the samples of the two studies have the same level of difficulty. Especially, the nature of the sample can be quite different from study to study.

- Use a clear and straightforward statistical program that offers information on the whole of the statistical procedure. A disadvantage of commercially packages is their black box character and that **the analysis is closed**, i.e. one can analyse only what is provided and it is sometimes hard to work further on the results (for instance for validation). A general statistical package does not have these drawbacks, but in this case considerable skill and time can be necessary, certainly to develop the routines. The gain is much more insight.

---

12. Evaluation of the laboratory performance

See 'Quality Assurance and Quality Control in Forest Soil Analysis: 3rd FSCC Interlaboratory Comparison – Phase I (FSCC, 2003a)

Description

To classify the laboratories, the idea was that being an outlier for one element / sample does not need to be a problem. Only if systematically deviant results are reported, the performance of the laboratory is considered poor. For this reason it was counted how many times a laboratory was classified as straggler or outlier.

Evaluation

An important issue is how laboratories cope with values below the detection and quantification limit. However no clear rules were present in advance, so it was hard to make a comparison between the laboratories.

Another problem was that is was insufficiently clear to the laboratories they had to work under repeatability conditions. As a consequence repeatability variance and within laboratory variance is confounded in current study.


Recommendations

- Realise there is no golden standard for this type of ring tests. The bulk of the data is used as the reference and this can be wrong.

- Another important criterion is the handling of problems linked to low values (quantification and detection limit).

- Specify as clearly as possible that the laboratories should work under repeatability conditions.

---

13. Integration of background information

See 'Quality Assurance and Quality Control in Forest Soil Analysis: 3rd FSCC Interlaboratory Comparison – Phase II' (FSCC, 2003a)

Description

The next step was to find out whether the performance of the laboratories could be linked to the background information obtained from the questionnaire (third research question). First, a univariate analysis (ANOVA and Kruskall-Wallis) screened which factors were linked separately with the laboratory performance. Afterwards, regression models tested their combined influence (see Figure 7).
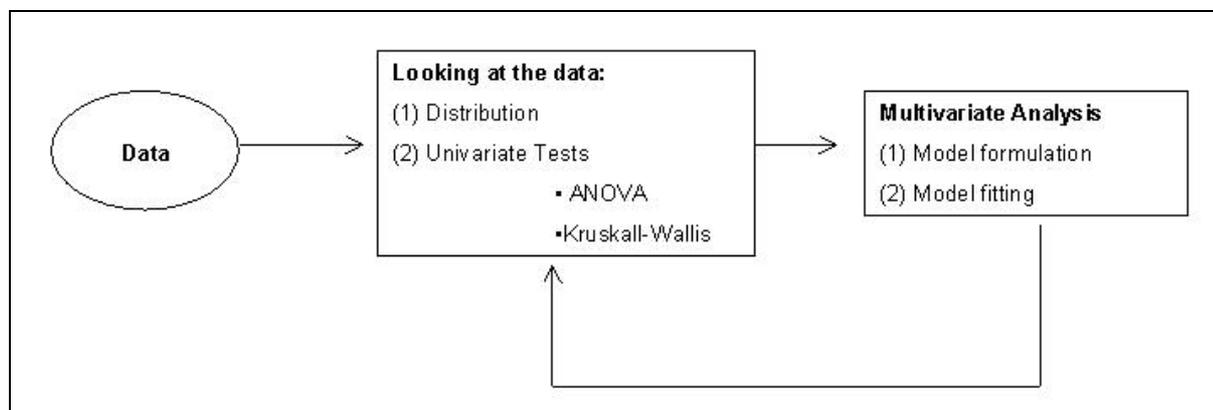
Figure 7: Flow-chart of the integration of secondary information

The following factors were investigated; the factor name has been put between brackets

(1)	Use of reference method (RefmC)
(2)	Experience level (ExpLevC)
(3)	Training of the personnel (Trained)
(4)	Accreditation (Accr)
(5)	Statute of laboratory (Statute)
(6)	Type of laboratory (Type)
(7)	Specialised in forestry (Forest)
(8)	Region (Region)

To meet normality requirements, QQ-plots were drawn to check the values of two <u>transformations</u>: the square root transformation and the logarithmic transformation. Since the values of Mandel's h, can either

be positive or negative, the absolute value is taken before transformation. This is not necessary for Mandel's k.

Table 3: Transformations used in the third ring test

Transformation	Mandel's h	Mandel's k
None	`abs(Hv)`	`Kv`
Square root	`sqrt(abs(Hv))`	`sqrt(Kv)`
Logaritmic	`log (abs(Hv))`	`log (Kv)`

As seen in the figures II.3 and II.4 of the ring test report (FSCC, 2003), for most cases the square root (sqrt) transformation gives the best fit.

Evaluation

In summary the analysis did not reveal consistent relations, nor for the inter-laboratory variance, nor for the intra-laboratory variance. The different factors only show significant influences for some groups, but none of the factors is significant for all groups of analyses. The reason for these differences is not that clear. Significance also depends on the type of statistical analysis giving a confuse picture.

Another problem is that the different variables do not have a clear meaning because many of the variables are interrelated. Also, if one tests enough, in the end, one always finds significant relations for which afterwards a good explanation can be found. Be aware of this overinterpretation of post-hoc analyses.

A lot of work was necessary to turn the questionnaire into meaningful variables. In advance, one should define more clearly which questions to answer and to make the hypotheses more explicit. In this way, the questionnaire will become sharper: fewer focused questions but more information (not at least because the answer to fewer questions will be done more careful).

Another problem was the imbalance in the design. This decreases the power of the statistical testing. This is very common in observational studies with rather a small number of samples. However if looking for important factors is an important study question, then the design should be adapted to prevent this, e.g. by stratifying for the most important issues. Another way to increase the power of the design is to change the factors within the laboratories. For instance by asking to analyse the data by two methods: their own preferred method and the method asked.

As mentioned earlier, the research questions are a bit contrary. Getting a representative picture of the quality asks for a different design than finding explicative variables for that quality. It is possible, however a random sample should be large to overcome possible imbalances and this is not the case here. Conversely a stratified sample is not representative anymore to get a good estimate of the quality unless correction factors are used.

Recommendations

- The results of this type of analysis should be interpreted carefully. It is a post-hoc analysis. Be aware of over-interpretation.

- An unbalanced design limits the scope and discriminate power of the data. In general, a random sample guarantees the characteristics to be balanced. For small samples, yet the design can be quite unbalanced. This is a general problem for observational studies.

- A way to arrive at a more balanced design is to stratify for the main questions or hypotheses to test. Remember however that this stratification is not an easy task and can prevent to have a representative sample to assess the reproducibility.

CONCLUSIONS

For the first time FSCC organised a ring test on this scale using an evaluation procedure based on the international standard ISO 5725-2 'Accuracy (trueness and precision) of measurement methods and results – part 2: Basic method for determination of repeatability and reproducibility of a standard measurement method' (ISO, 1994c).

The strategic choice was to analyse in depth the mathematical background and to develop an own tool to get a transparent and easy to understand procedure to overcome the 'black box character' of formerly used programmes. Generally, the developed statistical procedure satisfied the need for a solid and transparent way of evaluating the ring test results. In spite of some problems concerning confused rules for reporting, the new procedure gave good results.

Clearer instructions to the laboratories concerning the way of reporting are of major importance. Before sending the samples, laboratories should be informed in detail on the reporting rules (eg. values below detection limit, minimum number of repetitions, oven or air dry).

The questionnaire included in the third ring test was confusing for some laboratories. It is advised to develop shorter and straightforward questionnaire, better than asking for a large amount of information. In view of the statistical procedure, the use of closed questions is advisable. A testing phase of the questionnaire is useful to guarantee the efficient questioning of the laboratories.

It is important to choose an appropriate approach from the beginning; directed to a representative estimate of the reproducibility comparison or to a study trying to find the factors influencing the performance of the laboratories. Combining both strategies in one single ring test with a small number of laboratories is not an option. The sampling design depends on the chosen strategy, the first strategy asks for a representative sample, the second strategy requires a more specific sampling design (stratification to balance for the most important questions).

Hopefully, this document will stimulate the standardisation of methodology concerning the organisation of interlaboratory comparisons within the ICP-Forests programme.

REFERENCES

Agrawal, Y.C., McCave, I.N. and Riley, J.B., 1991. Laser diffraction size analysis. In: J.P.M. Syvitski (Editor), Principles, Methods, and Application of Particle Size analysis. Cambridge University Press, Cambridge, pp. 119-128.

Bartels, U., Busch, U., Fermer, H.J., 2002. ICP-Forests 5th Needle/Leaf Interlaboratory Test 2001/2002. International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests and European Union Scheme on the Protection of Forests against Atmospheric Pollution.

Forest Soil Co-Ordinating Centre, 2003a. Quality Assurance and Quality Control in Forest Soil Analyis: 3rd FSCC Interlaboratory Comparison. UN/ECE Convention on Long-Range Transboundary Air Pollution, International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests.

Forest Soil Co-Ordinating Centre, 2003b. Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Part IIIa. Sampling and Analysis of Soil. Version 4.0. UN/ECE Convention on Long-Range Transboundary Air Pollution, International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests.

ISO, 1994c. ISO 5725-2. Accuracy (trueness and precision) of measurement methods and results – part 2: Basic method for determination of repeatability and reproducibility of a standard measurement method'.

Loizeau, J.-L., Arbouille, D., Santiago, S. and Vernet, J.-P., 1994. Evaluation of a wide range laser diffraction grain-size analyser for use with sediments. Sedimentology, 41: 353-361.

Van der Velden, M. and Van Orshoven J. in cooperation with the Flemish Soil expert group. November 1992. Interne publicatie N°. 3. Bestuur Natuurbehoud en Ontwikkeling, Dienst Waters en Bossen, Administratie Milieu en Landinrichting, Network Integrated Monitoring of Forest Ecosystems, International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests in the ECE-region.

Vanmechelen, L., Groenemans R., Van Ranst E. 1997. Forest Soil Condition in Europe. Results of a Large-Scale Soil Survey. Prepared by Forest Soil Co-Ordinating Centre in Co-Operation with the Ministry of the Flemish Community. EC-UN/ECE, Brussels, Geneva.

Van Reeuwijk, L.P., Houba, V.J.G., 1998. Guidelines for quality management in soil and plant laboratories. FAO Soils Bulletin-74, Rome.

ANNEXES

Annex I: Test of homogeneity

Homogeneity test based on 'Texture' ('clay content' and 'medium texture class')

Texture ('clay' and 'medium texture class') was measured on five subsamples of the samples A (names of the subsamples: I5, B5, D3, G2 and G6) and B (names of the subsamples: C2, C8, E5, F3 and H7). The median texture class and the clay content measurements were repeated 5 times on one of the subsamples and once on the remaining 4 subsamples. The results of the analysis are presented in Figures 1 to 4.

This resulted in an extremely unbalance design, which did not allow for an ANOVA. Though, it was possible to calculate coefficients of variation. The within-laboratory coefficient (indicated as a full line in the figure) of variation is based on the 5 replicates on one subsample, which were for sample A, subsample I5 and for sample B, subsample H7). The between-laboratory coefficient of variation (dotted line in the figure) is based on the results of the four other subsamples. It is not allowed to use one measurement of the first subsample because then the question rises on which sample to choose.

Figure I.1 : Set-up of the homogeneity test
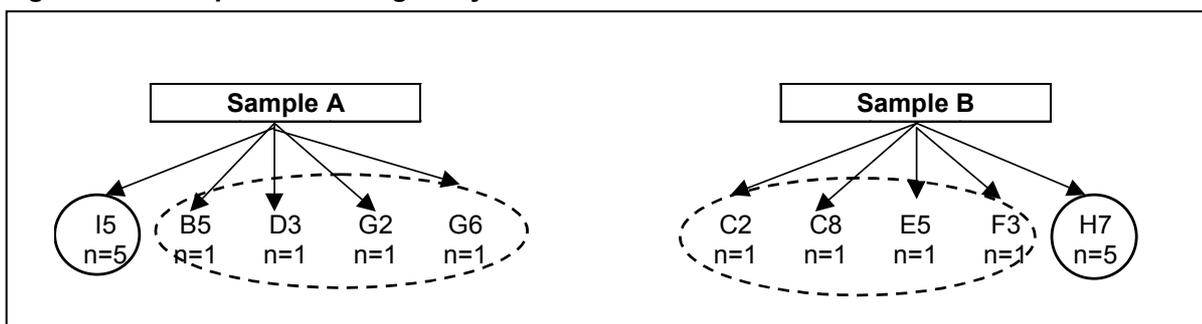
For sample A, the CV's within sample I5 were 6 % for 'medium texture class' and 7.6 % for the 'clay content', and between the subsamples, 13 % for the 'medium texture class' and 6.22 % for the 'clay content'. For sample B, the CV's between the subsamples were smaller than the CV's within the sample. For the 'medium texture class', the within-subsample CV was 9.6% compared to 5.2% between the subsamples. For the 'clay content', the within-subsample CV was 2.3 % versus 1.3 % between the subsamples.

So only the variability between the subsamples of sample A for the analysis of the 'medium texture' was larger than within one subsample. To be homogeneous, the variance should have been of similar magnitude.
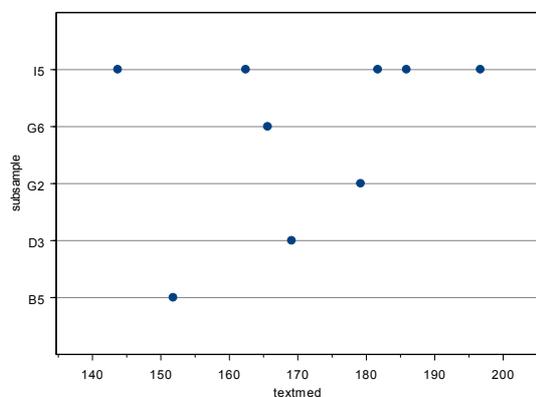
Figure I.2: Homogeneity test for sample A: Median texture
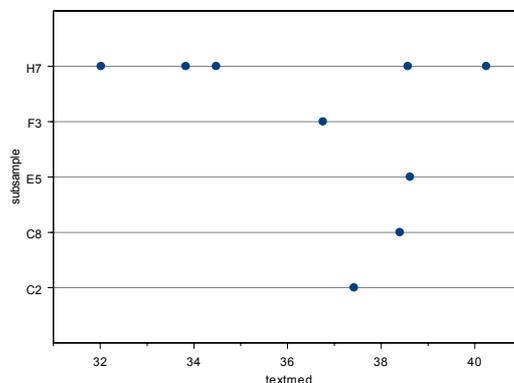
Figure I.3: Homogeneity test for sample B: Median texture

Figure I.4 : Homogeneity test for sample A: Clay content

Figure I.6: Homogeneity test for sample B: Clay content

Homogeneity test based on 'Total Carbon'

In this test, the same 5 subsamples of sample A and B were used. On one subsample, the TC measurement was repeated 15 times (I5), on the other 4 subsamples the TC was measured in triplicate. The results of the laboratory analyses are presented in Figure 8 and 9.

Figure I.6 : Homogeneity test of sample A: Total carbon

Figure I.7 : Homogeneity test of sample B: Total carbon

The results of sample A were not completely normally distributed but transformations did not offer a lot of help. The statistical evaluation revealed that sample A was heterogeneous: subsample G6 was significantly bigger than sample B5, D3 and I5. This could have been due to one very high measurement in one of the subsamples, increasing the within-sample variance and as well the between-sample variance enormously. After the elimination of one possible outlier, there was still a significant difference between the groups. Even by removing the whole subsample, the dataset did not become homogenous.

The results of the two ANOVA models are given below.

```
** Analysis of Variance Model Sample A***
Short Output:
Call:
   aov(formula = logtc ~ Subsample, data = homogeniteitsdata.TC.A, na.action =
na.exclude)
Terms:
             Subsample Residuals
 Sum of Squares   0.0807    0.0409
Deg. of Freedom       4        22
Residual standard error: 0.0431
Estimated effects may be unbalanced
         Df Sum of Sq Mean Sq F Value    Pr(F)
Subsample  4    0.0807  0.0202    10.9 0.0000515
Residuals 22    0.0409  0.0019
```
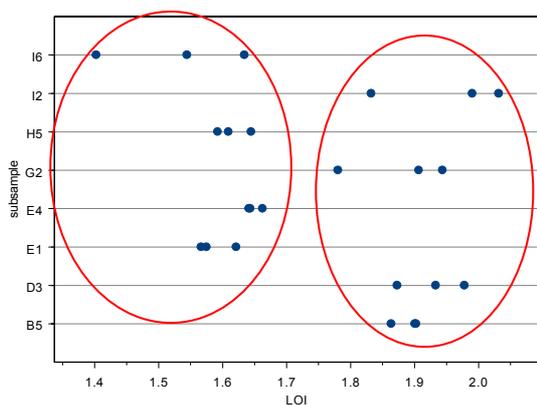
```
*** Analysis of Variance Model Sample B***
Short Output:
Call:
   aov(formula = TC ~ Subsample, data = homogeniteitsdata.TC.B, na.action = na.exclude)
Terms:
                Subsample Residuals
 Sum of Squares   1545766    8455505
Deg. of Freedom        4         22
Residual standard error: 620
Estimated effects may be unbalanced
         Df Sum of Sq Mean Sq F Value Pr(F)
Subsample  4   1545766  386441    1.01 0.426
Residuals 22   8455505  384341
```

For the measurement of Kjeldahl N of sample A and B, one can say that a single factor ANOVA with unequal replications confirms homogeneity in sample B but not in sample A. Multiple pairwise comparisons show that subsample G6 was significantly larger than B5, D3 and I5.

Homogeneity test based on 'Loss on ignition'

Eight subsamples of each of the sample A, B and C were analysed for loss on ignition in triplicate. The results are shown in the plots below (Figure 8 till 10).

Figure I.8 : Homogeneity test of sample A: Loss on ignition
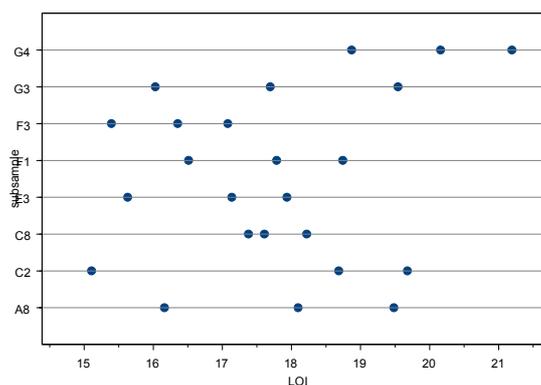
Figure I.9 : Homogeneity test of sample B: loss on ignition
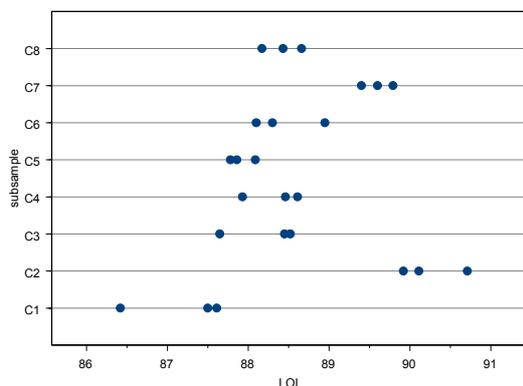
Figure I.10: Homogeneity test on Sample C: Loss on ignition

After normality check and check for homogenous distribution of the residues, analysis of variance (ANOVA) was conducted.

Sample A was log linear transformed. Significant differences between the groups were seen.

```
  *** Analysis of Variance Model Sample A***
Call:
   aov(formula = logLOI ~ subsample, data = homotest.loi.A, na.action = na.exclude)
Terms:
              subsample Residuals
 Sum of Squares    0.210      0.025
Deg. of Freedom        7         16
Residual standard error: 0.0398
Estimated effects are balanced
         Df Sum of Sq Mean Sq F Value     Pr(F)
subsample  7     0.210  0.0300       19 1.22e-006
Residuals 16     0.025  0.0016
```

The multiple comparisons revealed significant differences between two groups of samples:

```
E1, E4 ,H5, I6  <  B5, D3, I2, G2
```

For sample B, the ANOVA did not show any significant differences between the subsamples.

```
*** Analysis of Variance Model Sample B***
Short Output:
Call:
   aov(formula = LOI ~ subsample, data = homotest.loi.B, na.action = na.exclude)
Terms:
              subsample Residuals
 Sum of Squares       25        33
Deg. of Freedom        7        16

Residual standard error: 1.44
Estimated effects are balanced
         Df Sum of Sq Mean Sq F Value Pr(F)
subsample  7       25    3.57    1.73 0.172
Residuals 16       33    2.07
```

For sample C, the ANOVA reports differences between the subsamples.

```
*** Analysis of Variance Model Sample C***
Call:
   aov(formula = LOI ~ subsample, data = homotest.loi.C, na.action = na.exclude)
Terms:
              subsample Residuals
 Sum of Squares     19.4       2.6
Deg. of Freedom        7        16
Residual standard error: 0.401
Estimated effects are balanced
         Df Sum of Sq Mean Sq F Value     Pr(F)
subsample  7     19.4    2.77    17.2 2.38e-006
Residuals 16      2.6    0.16
```

The multiple comparisons revealed the next significant differences:

```
C1 < C2, C4, C6, C7 and C8
C2 > C3, C4, C5, C6 and C8
C7 > C3, C4, C5 and C8
```

In conclusion one could say for loss on ignition that there are significant differences between the subsamples of sample A and B BUT one need to consider the magnitude of the differences. If this variance is small compared to the differences seen between the laboratories in the ring test, it is possible that the differences between the subsamples become insignificant (which is the case in the 3ʳᵈ ring test). On the other hand, these analyses indicate that a better homogenisation of the samples, which means riffling twice, is absolutely required when conducting a next ring test.
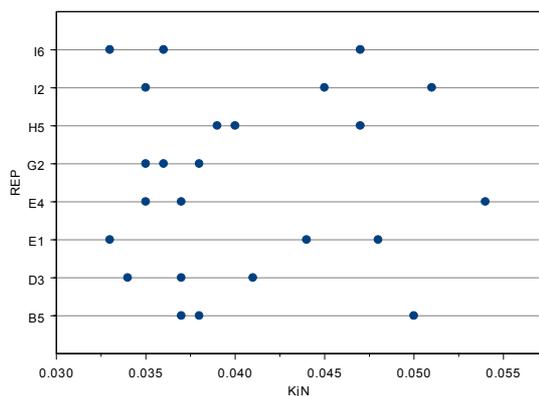
Homogeneity test based on Kjeldahl N

Figure I.11: Homogeneity test on sample A: Kjeldahl N
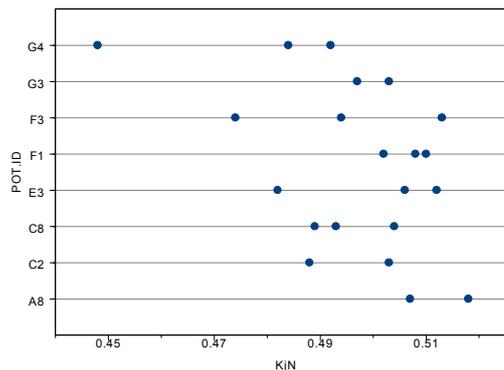
Figure I.12: Homogeneity test on sample B: Kjeldahl N
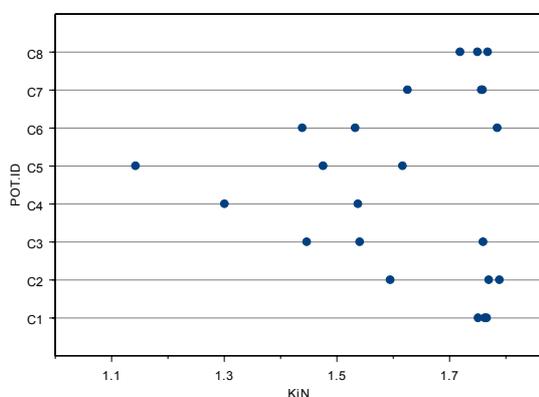
Figure I.13: Homogeneity test on sample C: Kjeldahl N

After the ANOVA, the next results were obtained:

```
*** Analysis of Variance Model Sample A***
Short Output:
Call:
   aov(formula = KjN ~ POT.ID, data = homotestdata.loi.KjN.A, na.action = na.exclude)
Terms:
                 POT.ID Residuals
 Sum of Squares 0.000144  0.000750
Deg. of Freedom        7        16
Residual standard error: 0.00685
Estimated effects are balanced
         Df Sum of Sq    Mean Sq F Value Pr(F)
   POT.ID  7  0.000144 0.0000205   0.438 0.864
Residuals 16  0.000750 0.0000469
```

Sample A, B and C are homogeneous at the 0.01 level but sample C showed significant differences between the subsamples at the 0.05 significance level.

```
** Analysis of Variance Model Sample B***
Short Output:
Call:
   aov(formula = KjN ~ POT.ID, data = homotestdata.loi.KjN.B, na.action = na.exclude)
Terms:
                 POT.ID Residuals
 Sum of Squares 0.00276   0.00274
Deg. of Freedom       7        15
Residual standard error: 0.0135
Estimated effects may be unbalanced
```

```
          Df Sum of Sq  Mean Sq F Value  Pr(F)
   POT.ID  7   0.00276 0.000394    2.16 0.0997
Residuals 15   0.00274 0.000182
```

```
** Analysis of Variance Model Sample C***
Short Output:
Call:
   aov(formula = KjN ~ POT.ID, data = homotestdata.loi.KjN.C, na.action = na.exclude)
Terms:
                POT.ID Residuals
 Sum of Squares  0.379     0.298
Deg. of Freedom      7        15
Residual standard error: 0.141
Estimated effects may be unbalanced
          Df Sum of Sq Mean Sq F Value Pr(F)
   POT.ID  7     0.379  0.0541    2.72 0.049
Residuals 15     0.298  0.0199
```

Annex II: Standard form for the input of the results (Sample A)

FSCC interlaboratory comparison, July 2002

Lab name	
Country	

Form for the collection of the data of sample A

Parameter	Unit	Replicate 1	Replicate 2	Replicate 3
Particle size: clay	%			
Particle size: silt	%			
Particle size: sand	%			
pH(CaCl$_2$)	-			
pH(H$_2$O)	-			
Carbonates	g/kg			
Organic carbon	g/kg			
Total N	g/kg			
Exchangeable acidity	cmol+/kg			
Exchangeable Al	cmol+/kg			
Exchangeable Ca	cmol+/kg			
Exchangeable Fe	cmol+/kg			
Exchangeable K	cmol+/kg			
Exchangeable Mg	cmol+/kg			
Exchangeable Mn	cmol+/kg			
Exchangeable Na	cmol+/kg			
Free H$^+$ acidity	cmol+/kg			
Extracted Al	mg/kg			
Extracted Ca	mg/kg			
Extracted Cd	mg/kg			
Extracted Cr	mg/kg			
Extracted Cu	mg/kg			
Extracted Fe	mg/kg			
Extracted Hg	mg/kg			
Extracted K	mg/kg			
Extracted Mg	mg/kg			
Extracted Mn	mg/kg			
Extracted Na	mg/kg			
Extracted Ni	mg/kg			
Extracted P	mg/kg			
Extracted Pb	mg/kg			
Extracted S	mg/kg			
Extracted Zn	mg/kg			
Total Al	mg/kg			
Total Ca	mg/kg			
Total Fe	mg/kg			
Total K	mg/kg			
Total Mg	mg/kg			
Total Mn	mg/kg			
Total Na	mg/kg			
Reactive Al	mg/kg			
Reactive Fe	mg/kg			

Annex III : Laboratory Questionnaire

Guidelines : **Please fill in all questions.**
For some questions more than one answer is allowed.
In case of subcontracted laboratories, each laboratory has to fill the General Information form.

A. General Information

Name of Laboratory:
Address
        Street :
        Number :
        PO Box :
        Postal Code :
        City :
        Country :

Telephone Number :
Fax Number :
Website :

Contact person
        Name :
        Email :

Responsible person (laboratory)
        Name :
        Email :

Statute (belonging to) : ☐ University, Name :
                               ☐ State Institute, Name :
                               ☐ Private Institute, Name :
                               ☐ Private Lab
                               ☐ Other, Specify :

Type of laboratory: ☐ Soil (Forest)
                           ☐ Soil (General)
                           ☐ Plant & Soil (Forest)
                           ☐ Plant & Soil (General)
                           ☐ General

Actual Number of laboratory personnel:

Lab is working since (year) :

Certificates of laboratory :

The results of the ring test will be published in a report. Do you agree that your laboratory is mentioned with full name in this evaluation report ? Laboratories indicating 'no' will be given a code only. In that case only the NFC will have knowledge of the laboratory codes of its country.
        ☐ Yes
        ☐ No

B. Analysis Information

Analysis : Determination of Particle Size Distribution

Laboratory that performed analysis
☐ Own laboratory
☐ Subcontracted laboratory, Specify
      (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 11277) ?
☐ Yes
☐ No, Lab method :
- Has the lab experience with reference method ?
☐ High
☐ Normal
☐ Little
- Does the lab possess any accreditation for this method ?
☐ Yes, delivered by :
☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Method detection limit

- What is the experience of the lab with the used equipment ?
☐ High
☐ Normal
☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
☐ Yes
☐ No
- What is the experience of the lab personnel with this method
☐ < 1 year
☐ 1 - 3 years
☐ > 3 years

Quality Assurance
- Does the lab use for this method
☐ International Reference Material (IRM), Provider/Code :
☐ National Reference Material (NRM), Provider/Code :
☐ Local Reference Material (LRM)
☐ No reference material
- Which reference material does the lab use ?
☐ Matrix reference material
☐ Method reference material
- Does the lab use Calibration Standards ?
☐ Yes, Provider/Code :
☐ No
- Does the lab use Control Charts ?
☐ Yes (if possible, please provide)
☐ No

Analysis : Determination of Soil pH

Laboratory that performed analysis
☐ Own laboratory
☐ Subcontracted laboratory, Specify
(subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 10390) ?
☐ Yes
☐ No, Lab method :
- Has the lab experience with reference method ?
☐ High
☐ Normal
☐ Little
- Does the lab possess any accreditation for this method ?
☐ Yes, delivered by :
☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Method detection limit
pH			

- What is the experience of the lab with the used equipment ?
☐ High
☐ Normal
☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
☐ Yes
☐ No
- What is the experience of the lab personnel with this method
☐ < 1 year
☐ 1 - 3 years
☐ > 3 years

Quality Assurance
- Does the lab use for this method
☐ International Reference Material (IRM), Provider/Code :
☐ National Reference Material (NRM), Provider/Code :
☐ Local Reference Material (LRM)
☐ No reference material
- Which reference material does the lab use ?
☐ Matrix reference material
☐ Method reference material
- Does the lab use Calibration Standards ?
☐ Yes, Provider/Code :
☐ No
- Does the lab use Control Charts ?
☐ Yes (if possible, please provide)
☐ No

Analysis : Determination of Carbonate Content

Laboratory that performed analysis
☐ Own laboratory
☐ Subcontracted laboratory, Specify
(subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 10693) ?
☐ Yes
☐ No, Lab method :
- Has the lab experience with reference method ?
☐ High
☐ Normal
☐ Little
- Does the lab possess any accreditation for this method ?
☐ Yes, delivered by :
☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Method detection limit
Carbonate content			

- What is the experience of the lab with the used equipment ?
☐ High
☐ Normal
☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
☐ Yes
☐ No
- What is the experience of the lab personnel with this method
☐ < 1 year
☐ 1 - 3 years
☐ > 3 years

Quality Assurance
- Does the lab use for this method
☐ International Reference Material (IRM), Provider/Code :
☐ National Reference Material (NRM), Provider/Code :
☐ Local Reference Material (LRM)
☐ No reference material
- Which reference material does the lab use ?
☐ Matrix reference material
☐ Method reference material
- Does the lab use Calibration Standards ?
☐ Yes, Provider/Code :
☐ No
- Does the lab use Control Charts ?
☐ Yes (if possible, please provide)
☐ No

Analysis : Determination of Organic Carbon Content

Laboratory that performed analysis
- ☐ Own laboratory
- ☐ Subcontracted laboratory, Specify
    (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 10694) ?
    - ☐ Yes
    - ☐ No, Lab method :
- Has the lab experience with reference method ?
    - ☐ High
    - ☐ Normal
    - ☐ Little
- Does the lab possess any accreditation for this method ?
    - ☐ Yes, delivered by :
    - ☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?
    - No

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Method detection limit
Organic Carbon			

- What is the experience of the lab with the used equipment ?
    - ☐ High
    - ☐ Normal
    - ☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
    - ☐ Yes
    - ☐ No
- What is the experience of the lab personnel with this method
    - ☐ < 1 year
    - ☐ 1 - 3 years
    - ☐ > 3 years

Quality Assurance
- Does the lab use for this method
    - ☐ International Reference Material (IRM), Provider/Code :
    - ☐ National Reference Material (NRM), Provider/Code :
    - ☐ Local Reference Material (LRM)
    - ☐ No reference material
- Which reference material does the lab use ?
    - ☐ Matrix reference material
    - ☐ Method reference material
- Does the lab use Calibration Standards ?
    - ☐ Yes, Provider/Code : Fisons Instruments, ThermoQuest Italia s.p.a.
    - ☐ No
- Does the lab use Control Charts ?
    - ☐ Yes (if possible, please provide)
    - ☐ No

Analysis : Determination of total Nitrogen Content

Laboratory that performed analysis
- ☐ Own laboratory
- ☐ Subcontracted laboratory, Specify
   (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 13878) ?
- ☐ Yes
- ☐ No, Lab method :
- Has the lab experience with reference method ?
- ☐ High
- ☐ Normal
- ☐ Little
- Does the lab possess any accreditation for this method ?
- ☐ Yes, delivered by :
- ☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Method detection limit
Total Nitrogen Content			

- What is the experience of the lab with the used equipment ?
- ☐ High
- ☐ Normal
- ☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
- ☐ Yes
- ☐ No
- What is the experience of the lab personnel with this method
- ☐ < 1 year
- ☐ 1 - 3 years
- ☐ > 3 years

Quality Assurance
- Does the lab use for this method
- ☐ International Reference Material (IRM), Provider/Code :
- ☐ National Reference Material (NRM), Provider/Code :
- ☐ Local Reference Material (LRM)
- ☐ No reference material
- Which reference material does the lab use ?
- ☐ Matrix reference material
- ☐ Method reference material
- Does the lab use Calibration Standards ?
- ☐ Yes, Provider/Code :
- ☐ No
- Does the lab use Control Charts ?
- ☐ Yes (if possible, please provide)
- ☐ No

Analysis : Determination of Exchangeable Acidity, Exchangeable Cations (Al, Ca, Fe, K, Mg, Mn, Na) and Free H⁺ Acidity

Laboratory that performed analysis
- ☐ Own laboratory
- ☐ Subcontracted laboratory, Specify
    (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 11260 and ISO 14254) ?
    - ☐ Yes
    - ☐ No, Lab method :
- Has the lab experience with reference method ?
    - ☐ High
    - ☐ Normal
    - ☐ Little
- Does the lab possess any accreditation for this method ?
    - ☐ Yes, delivered by :
    - ☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Wavelength	Method detection limit
Exchangeable acidity				
Free H⁺ acidity				
Exchangeable Al				
Exchangeable Ca				
Exchangeable Fe				
Exchangeable K				
Exchangeable Mg				
Exchangeable Mn				
Exchangeable Na				

- What is the experience of the lab with the used equipment ?
    - ☐ High
    - ☐ Normal
    - ☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
    - ☐ Yes
    - ☐ No
- What is the experience of the lab personnel with this method
    - ☐ < 1 year
    - ☐ 1 - 3 years
    - ☐ > 3 years

Quality Assurance
- Does the lab use for this method
    - ☐ International Reference Material (IRM), Provider/Code :
    - ☐ National Reference Material (NRM), Provider/Code :
    - ☐ Local Reference Material (LRM)
    - ☐ No reference material
- Which reference material does the lab use ?
    - ☐ Matrix reference material
    - ☐ Method reference material
- Does the lab use Calibration Standards ?
    - ☐ Yes, Provider/Code :

☐ No

- Does the lab use Control Charts ?
    ☐ Yes (if possible, please provide)
    ☐ No

Analysis : Aqua Regia Extractant Determinations (P, Ca, K, Mg, Mn, Cu, Pb, Cd, Zn, Al, Fe, Cr, Ni, S, Hg, Na)

Laboratory that performed analysis
☐ Own laboratory
☐ Subcontracted laboratory, Specify
(subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 11466) ?
☐ Yes
☐ No, Lab method :
- Has the lab experience with reference method ?
☐ High
☐ Normal
☐ Little
- Does the lab possess any accreditation for this method ?
☐ Yes, delivered by :
☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
- What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Wavelength	Method detection limit
Extracted Al				
Extracted Ca				
Extracted Cd				
Extracted Cr				
Extracted Cu				
Extracted Fe				
Extracted Hg				
Extracted K				
Extracted Mg				
Extracted Mn				
Extracted Na				
Extracted Ni				
Extracted P				
Extracted Pb				
Extracted S				
Extracted Zn				

- What is the experience of the lab with the used equipment ?
☐ High
☐ Normal
☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
☐ Yes
☐ No
- What is the experience of the lab personnel with this method
☐ < 1 year
☐ 1 - 3 years
☐ > 3 years

Quality Assurance
- Does the lab use for this method
☐ International Reference Material (IRM), Provider/Code :
☐ National Reference Material (NRM), Provider/Code :
☐ Local Reference Material (LRM)
☐ No reference material

- Which reference material does the lab use ?

☐ Matrix reference material
☐ Method reference material
- Does the lab use Calibration Standards ?
☐ Yes, Provider/Code :
☐ No
- Does the lab use Control Charts ?
☐ Yes (if possible, please provide)
☐ No

Analysis : Determination of Total Elements (Ca, Mg, Na, K, Al, Fe, Mn)

Laboratory that performed analysis
 ☐ Own laboratory
 ☐ Subcontracted laboratory, Specify
   (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISO 14869 or Michopoulos, 1995) ?
 ☐ Yes
 ☐ No, Lab method :
- Has the lab experience with reference method ?
 ☐ High
 ☐ Normal
 ☐ Little
- Does the lab possess any accreditation for this method ?
 ☐ Yes, delivered by :
 ☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Wavelength	Method detection limit
Total Al				
Total Ca				
Total Fe				
Total K				
Total Mg				
Total Mn				
Total Na				

- What is the experience of the lab with the used equipment ?
 ☐ High
 ☐ Normal
 ☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
 ☐ Yes
 ☐ No
- What is the experience of the lab personnel with this method
 ☐ < 1 year
 ☐ 1 - 3 years
 ☐ > 3 years

Quality Assurance
- Does the lab use for this method
 ☐ International Reference Material (IRM), Provider/Code :
 ☐ National Reference Material (NRM), Provider/Code :
 ☐ Local Reference Material (LRM)
 ☐ No reference material
- Which reference material does the lab use ?
 ☐ Matrix reference material
 ☐ Method reference material
- Does the lab use Calibration Standards ?
 ☐ Yes, Provider/Code :
 ☐ No
- Does the lab use Control Charts ?
 ☐ Yes (if possible, please provide)
 ☐ No

Analysis : Acid Oxalate Extractable Fe and Al

Laboratory that performed analysis
- ☐ Own laboratory
- ☐ Subcontracted laboratory, Specify
  (subcontracted laboratory should fill in General Information and Analysis Information Forms)

Method
- Is the lab working according to the reference method (ISRIC 1992) ?
  - ☐ Yes
  - ☐ No, Lab method :
- Has the lab experience with reference method ?
  - ☐ High
  - ☐ Normal
  - ☐ Little
- Does the lab possess any accreditation for this method ?
  - ☐ Yes, delivered by :
  - ☐ No
- Does the lab encounters any specific problems analysing the samples according to this method ?

Equipment
What is the used equipment for this analysis ?

Analysis	Analytical technique	Analytical instrument (+manufacturer)	Wavelength	Method detection limit
Reactive Al				
Reactive Fe				

- What is the experience of the lab with the used equipment ?
  - ☐ High
  - ☐ Normal
  - ☐ Little

Personnel
- Has the lab personnel specifically trained for this analysis ?
  - ☐ Yes
  - ☐ No
- What is the experience of the lab personnel with this method
  - ☐ < 1 year
  - ☐ 1 - 3 years
  - ☐ > 3 years

Quality Assurance
- Does the lab use for this method
  - ☐ International Reference Material (IRM), Provider/Code :
  - ☐ National Reference Material (NRM), Provider/Code :
  - ☐ Local Reference Material (LRM)
  - ☐ No reference material
- Which reference material does the lab use ?
  - ☐ Matrix reference material
  - ☐ Method reference material
- Does the lab use Calibration Standards ?
  - ☐ Yes, Provider/Code :
  - ☐ No
- Does the lab use Control Charts ?
  - ☐ Yes (if possible, please provide)
  - ☐ No

Annex IV: Accompanying letter for the sending of the samples (customs)

To whom it may concern,

I hereby declare that the parcel with invoice number EE 476 548 583 BE, sent to Dr. Marina Nadporozskaja, Laboratory of Soil Biochemistry, Biological Research Institute of Sankt-Petersburg State University, contains 3 soil samples, including a CD-ROM with the analysis methodologies, without any commercial value.
These soil samples form part of a study on the quality of the laboratories that participate in the ICP Forests Programme and have no commercial value. They are freely distributed to the participating laboratories. The CD-ROM contains information on the analysis methodologies. The documents on this CD-ROM can be freely downloaded on the FSCC website (www.ibw.vlaanderen.be/fscc/) as well.

Unfortunately the accompanying invoice of the parcel is only in Dutch and French. I hereby give a copy of the invoice with a translation of the part (CN 22) destined for customs.

CN 22 May be opened officially	VALUE	I declare that this shipment does not contain any hazardous material, prohibited by the postal regulations.
Detailed description of content		
	WEIGHT	I have taken note of the general transport conditions of EMS-TAXIPOST®.
Tick ☐ Documents to be validated before export ☐ Gift ☐ Sample of no commercial value		**Signature**

The parcel contains only items with no commercial value and that are not for sale. The estimated value of the shipment is therefore not determined by its contents, but it is only determined by the value of the carrying objects, *in casu* the CD-ROM (1 Euro) and the soil recipients (3 Euro).

I hereby do declare that this explanation is stated according to the truth.

Dr. Xavier Scheldeman
FSCC
Belgium