codex alimentarius commission

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00100 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 4 **CX/MAS 02/4-Add.2**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

Twenty-fourth Session
Budapest, Hungary, 18-22 November 2002

CRITERIA FOR EVALUATING ACCEPTABLE METHODS OF ANALYSIS FOR CODEX PURPOSES - DISPUTE SITUATIONS

(Prepared by France)

The 23rd Session of the Committee on Methods of Analysis and Sampling discussed dispute situations in the framework of the application of the criteria approach.

The Committee agreed that the Delegation of France with the assistance of Australia, Belgium, Brazil, Canada, Denmark, Egypt, Greece, the Netherlands, Norway, South Africa, the United Kingdom and the United States would prepare a discussion paper addressing dispute situations for consideration by the next session. Governments would be requested by Circular Letter to provide information on the current practices in this regard in member countries, in order to facilitate the preparation of the above discussion paper (ALINORM 01/23, para. 35).

The document was considered by a group of French experts working in the framework of the French Standardization Organization, AFNOR. The document is based on ISO 4259 of December 1995 (NF EN ISO 4259 de 1995): Petroleum products – Determination and application of precision data in relation to methods of test. Only paragraphs 7,8,9,10 Application have been used and are included.

It should be noted that this procedure (pages 1-3) is applied in the absence of references to specific rules in the specification or the test method.

The Delegation of France prepared the attached discussion paper and circulated it to the countries mentioned in the report in April 2002. The comments received from Brazil are attached as Annex 1.

Member countries were invited to provide comments and information on current practices for the selection of methods in dispute situations in CL 2001/5-MAS (ALINORM 01/23) in March 2001. The comments received from Thailand are attached as Annex 2.

Procedure for settling interlaboratory disputes

It must be highlighted that this procedure applies in the absence of specific rules set out in the specification or mentioned in the test method.

Preamble : Interpretation of the results obtained in a laboratory

Supplier (exporter) inspection

A supplier who has no other source of information regarding the value of a characteristic than an individual result shall consider that the product conforms to the specification limit, only if the result *X* meets the following condition(s) which correspond to a confidence level of not less than 95 % :

- In the case of a single upper limit *Max* (at the maximum),

  $X < Max - 1{,}64\ S_R$ (or $0{,}59R$) (1)

- In the case of a single lower limit *Min* (at the minimum),

  $X > Min + 1{,}64\ S_R$ (or $0{,}59R$) (2)

  S_R is the standard deviation of reproducibility of the method, worked out according to ISO 5725 for a fully validated method or established on the basis of the data from the in-house quality inspection for a method validated in a laboratory.

  R is the reproducibility ($= 2{,}8\ S_R$)

- In the case of a double limit (*Max* and *Min*), these two conditions shall be met.

The use of equations (1) and (2) constitutes a guide for the supplier and shall not be interpreted as an obligation. A value plotted between the specification value and the limit of equation 1/2 does not constitute a proof of nonconformity.

Consignee (importer) inspection

A consignee who has no other source of information regarding the true value of a characteristic than an individual result shall consider that that the product does not comply with the specification limit(s), only if the result *X* is such that the following condition (or one of the following conditions) is met, the corresponding confidence level being not less than 95 % :

- In the case of a single upper limit *Max* (at the maximum),

  $X > Max + 1{,}64\ S_R$ (3)

- In the case of a single lower limit *Min* (at the minimum),

  $X < Min - 1{,}64\ S_R$ (4)

- In the case of a double limit (*Max* and *Min*), both of these conditions shall be met.

The interlaboratory dispute appears if the product has been analysed in two laboratories and if the results lead to a disagreement

The supplier's (exporter) laboratory and the consignee's (importer) laboratory compare their results, the implemented analysis methods, the method of expression of the results, the method of use of the recovery coefficient and all the obtention conditions likely to influence the analytical results.

If it is impossible for the supplier and consignee to reach agreement regarding the appraisal of the quality of the product, on the basis of the results obtained, the procedure described below shall be applied.

THE INITIAL ANALYSES RESULTS ARE NO LONGER TAKEN INTO ACCOUNT

The parties conduct together a representative sampling operation, taking three samples from the batch forming the subject of the dispute. A sample is forwarded to each of the laboratories, the third one being kept.

The two laboratories compare their respective procedures and apparatus. After these investigations, a correlation test shall be conducted between both laboratories. The analysis may also be performed in a single laboratory in presence of a representative from the second laboratory.

On its sample, each laboratory shall obtain at least three results under repeatability conditions. Each laboratory calculates the mean of the results meeting the repeatability conditions (range test).

Let :

$\overline{X}_1$ be the supplier's mean ;

$\overline{X}_2$ be the consignee's mean ;

Max be the upper specification limit ;

Min be the lower specification limit.

$\overline{X}_1$ and $\overline{X}_2$ shall be compared as follows to *Max* and *Min.*

1 - $\frac{\overline{X}_1 + \overline{X}_2}{2} < Max$ or $> Min$

1.1 if $|\overline{X}_1 - \overline{X}_2| \leq 0{,}84R_2$ $\Rightarrow$ the product is accepted

$$R_2 = \sqrt{R^2 - r^2\left(1 - \frac{1}{2k_1} - \frac{1}{2k_2}\right)}$$

R is the reproducibility of the method ;
r is the repeatability of the method ;
k_1 is the number of results of the first laboratory ;
k_2 is the number of results of the second laboratory ;

1.2 if $|\overline{X}_1 - \overline{X}_2| > 0{,}84R_2$ $\Rightarrow$ possible disagreement

It cannot be declared with confidence that the product complies or does not comply with the specification limit ; consequently, the disagreement can be resolved by negotiation.

2 - $\frac{\overline{X}_1 + \overline{X}_2}{2} > Max$ or $< Min$ $\Rightarrow$ the product is declared non conform whatever the difference $|\overline{X}_1 - \overline{X}_2|$.

If the disagreement persists, a third laboratory (neutral, qualified and accepted by both parties) shall be requested to perform the test on a third sample. Let $\overline{X}_3$ be the mean of the acceptable results of this laboratory.

Check that the three mean results meet the reproducibility conditions.

Calculate the difference Δ between the most divergent laboratory mean and the mean of the means of the two other laboratories.

$$R_3 = \sqrt{\frac{R_1^2}{2} + \frac{R_4^2}{2*3}}$$

$$R_4 = \sqrt{R^2 - \frac{r^2}{3}\left(3 - \frac{1}{k_1} - \frac{1}{k_2} - \frac{1}{k_3}\right)}$$

$R_1 = \sqrt{R^2 - r^2\left(1 - \frac{1}{k}\right)}$ corresponds to the most divergent laboratory mean.

3 - If this difference Δ is equal to or less than R_3, in absolute value, all the results shall be considered as acceptable and their mean taken as an estimated value of the characteristic.

3.1 $\frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3}{3} \leq Max$ or $\geq Min$ $\Rightarrow$ the product is accepted

3.2 $\frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3}{3} > Max$ or $< Min \Rightarrow$ the product is refused.

4 - If the difference Δ between the most divergent laboratory mean and the mean $\overline{X}$ of the two other laboratories is greater than R3, the procedure described below shall be adopted.

4.1 $\overline{X} \leq Max$ or $\geq Min \Rightarrow$ the product is accepted.

4.2 $\overline{X} > Max$ or $< Min \Rightarrow$ the product is refused.

According to NF EN ISO 4259 Petroleum products – Determination and application of precision data in relation to methods of test

Signification of the repeatability *r* and of the reproducibility *R* exposed in the previous clauses

The repeatability and reproducibility values are estimated from analysis of variance (ANOVA) (a two factor factorial design with repetition) performed on the results of an interlaboratory programme organised for statistical purposes and in which different laboratories each conduct testing on a range of samples. The repeatability and reproducibility values shall be included in each published test method and it is to be noted that reproducibility is usually higher than repeatability if they are calculated in accordance with this international Standard.

See also annex H for an insight into the underlying statistical reasoning for the formulae of this clause.

REPEATABILITY *R*

For the needs of routine quality inspections, the majority of laboratories only perform a single test per sample, except in special cases, for instance in the case of dispute or where the operator wishes to check that his technicality is correct. In these cases, where several results are obtained, it is worthwhile checking the coherence of the repeated results with respect to the repeatability of the method ; the appropriate procedure is explained in 7.1.1. It is also useful to know the degree of confidence which can be granted to the results of the means and, to achieve this, the method is explained in 7.1.2.

Acceptability of the results

Where only two results are obtained under repeatability conditions and where their difference is below or equal to *r*, the operator can consider that his work is valid and can take the mean of the two results as the estimated value of the characteristic being measured.

If the two results differ by more than *r*, they shall be considered as suspect and it is necessary to obtain at least three additional results. The difference between the most divergent results and the mean of the others (the initial two included) shall be calculated and compared with a new value r_1 instead of *r*, given by the following formula :

$$r_1 = r\sqrt{\frac{k}{2(k-1)}}$$ where *k* is the total number of results obtained. (16)

If the difference is below or equal to r_1, all the results shall be accepted. If the difference is greater than r_1, the most divergent one shall be rejected and the procedure set out in this subclause shall be repeated until an acceptable set of results is obtained.

The mean of the acceptable results shall be taken as the estimated value of the characteristic. However, if two results or more, out of a total not exceeding 20, are rejected, it is necessary to verify both the procedure and the apparatus, and to undertake, if possible, a new series of tests.

Confidence intervals

Where a single operator, working within the limits of precision of the method, obtains under repeatability conditions a series of *k* results providing a mean $\overline{X}$, it can be accepted, at the 95 % confidence level, that the true value, μ, of the characteristic is situated within the following limits

(17)

$$\overline{X} - \frac{R_1}{\sqrt{2}} \leq \mu \leq \overline{X} + \frac{R_1}{\sqrt{2}}$$

where $R_1 = \sqrt{R^2 - r^2\left(1 - \frac{1}{k}\right)}$ (18)

Likewise, in the case of a single limit, where a single limit (upper or lower) is fixed, it can be accepted, at the 95 % confidence interval, that the true value, μ, of the characteristic is situated within the following limits :

$\mu \leq \overline{X} + 0{,}59R_1$ (upper limit) (19)

where $\mu \geq \overline{X} + 0{,}59R_1$ (lower limit) (20)

The factor 0,59 is the ratio $0{,}84/\sqrt{2}$, where 0,84 is calculated in annex H.

However, insofar as *r*, in the majority of test methods, is far smaller than *R*, a repetition of the tests, under repeatability conditions, only provides a small improvement to the confidence interval of the mean.

If the reproducibility *R* of a test method has been found to be far higher than the repeatability *r*, the reasons for the high value of the *R*/*r* ratio shall be analysed, and the method shall, as far as possible, be improved.

REPRODUCIBILITY *R*

Acceptability of the results

The procedure described in this subclause is intended for appraising, in terms of the reproducibility of the test method, the compatibility of the results obtained by several laboratories in routine operations and during transactions. In the event of disagreement between a supplier and a consignee, the procedure set out in clauses 8 to 10 shall be adopted.

Where the difference between two individual results obtained in two laboratories is below or equal to R, both results shall be considered as compatible, and their mean – and not one or the other separately – shall be taken as the estimated value of the characteristic being measured.

If the difference between both results is greater than *R*, they shall both be considered as suspect. Each laboratory shall then obtain at least three other acceptable results (see 7.1.1).

In this case, the difference between the mean of all the acceptable results of each laboratory shall be appraised using, instead of *R*, a new value R_2, given by the following formula :

$$R_2 = \sqrt{R^2 - r^2\left(1 - \frac{1}{2k_1} - \frac{1}{2k_2}\right)} \quad (21)$$

where

R is the reproducibility of the method ;

r is the repeatability of the method ;

k_1 is the number of results of the first laboratory ;

k_2 is the number of results of the second laboratory ;

If the difference between the means is below or equal to R_2, these means are then acceptable and their mean shall be considered as the estimated value of the characteristic being measured. If the difference between the means is greater than R_2, the procedure set out in clauses 8 to 10 shall be adopted.

However, if two or more laboratory means, out of a total not exceeding 20, have been rejected, it is necessary to verify both the procedure and the apparatus, and to undertake, if possible, a new series of tests.

Confidence intervals

Where N laboratories obtain, under repeatability and reproducibility conditions, one or more results giving a laboratory mean $\overline{X}$, it is perhaps accepted, at the 95 % confidence interval, that the true value, μ, of the characteristic is situated within the following limits :

$$\overline{X} - \frac{R_4}{\sqrt{2N}} \leq \mu \leq \overline{X} + \frac{R_4}{\sqrt{2N}} \qquad R_4 = \sqrt{R^2 - \frac{r}{N}^2 (N - \frac{1}{k_1} - \frac{1}{k_2} - \cdots - \frac{1}{k_N})} \tag{24}$$

Likewise, in the case of a single limit, where a single upper or lower limit is fixed, it can be accepted, at the 95 % confidence interval, that the true value, μ, of the characteristic is situated within the following limits :

$$\mu \leq \overline{X} + 0{,}59\frac{R_4}{\sqrt{N}} \text{ (upper limit)} \tag{25}$$

where

$$\mu \geq \overline{X} - 0{,}59\frac{R_4}{\sqrt{N}} \text{ (lower limit)} \tag{26}$$

These formulae also allow a given laboratory (N = 1) to determine the confidence level which can be assigned to an individual result in comparison with the true value.

Specifications

PURPOSE OF THE SPECIFICATIONS

The purpose of a specification is to fix a limit (or limits) to the true value for the relevant characteristic. In practice, however, this true value can never be obtained exactly. The characteristic is measured in a laboratory by applying a standardised test method, the results of which can have a dispersion defined by its repeatability and reproducibility. Consequently, there exists some uncertainty regarding the true value of each measured characteristic.

The specifications for petroleum products are inspected in accordance with clauses 9 and 10. On the basis of a previous agreement, a supplier and a consignee can use the alternative procedures described in annex J.

DRAWING UP OF THE SPECIFICATIONS

The specifications normally deal with the limits of the values of the characteristics. In order to avoid confusion, such limits shall normally be expressed in the form « not lower than » or « not higher than ». There are two types of limits :

— the double limit, upper and lower, e.g. a viscosity which shall not be lower than 5 mm2/s and shall not be higher than 10 mm2/s ; a boiling point of 100 °C ± 0,5 °C ;

— the single limit, lower or upper, e.g. a sulphur content which shall not be higher than 2 % ; a lead content which shall not be higher than 3,0 g/l ; a bitumen solubility which shall not be lower than 99 %.

The single limit case amounts to the double limit case when, as in the majority of cases, there exists a second implicit limit which refers to the first case. This is illustrated by the previous examples in which the additional implicit limit is 0 %, 0 g/l and 100 % in each case. Where it is a matter of a veritable single limit, e.g. a flash point which shall not be lower than 60 °C, the following considerations do not apply.

In this clause and in those which follow, A_1 designates the upper limit and A_2 designates the lower limit.

The value chosen for a limit specification shall take into account the reproducibility of the test method as follows :

— in the case of a double limit (A_1 or A_2), the specified range (expressed or implicit) shall not be less than four times the reproducibility R, i.e. :

$$(A_1 - A_2) \geq 4R$$

— in the case of a single limit (A_1 or A_2), the specified limit shall not be less than twice the reproducibility R, i.e. :

$A_1 \geq 2R$ or $A_2 \geq 2R$

The requirements of this International Standard apply to specifications drawn up in accordance with these principles.

In the case where, for practical reasons, the value of (A_1 - A_2) is less than 4R, the obtained results will be uncertain regarding their capacity to determine whether a sample meets or does not meet the requirements of the specification. According to statistical reasoning, it is desirable that (A_1 - A_2) be far greater than 4R. In such a case, one or both of the following two solutions shall be adopted :

a) the specification limits shall be examined in order to know whether they can be extended in order to match the precision of the test method ;

b) the test method shall be examined in order to see whether the precision can be improved or whether another method having a better precision can be adopted in order to match the desired specification limits.

Quality inspection and specifications

This clause provides general information allowing suppliers and consignees to appraise the quality of a product with respect to the specifications, when a single result is available. If the consignee is obliged, after having examined this result, go further, the procedure described in clause 10 shall be adopted.

SUPPLIER'S MARGIN OF INSPECTION

A supplier who has no other source of information regarding the true value of a characteristic than an individual result shall consider that the product conforms to the specification limit, only if the result X meets the following condition or conditions which correspond to a confidence level of not less than 95 % :

in the case of a single upper limit A_1, $X \leq A_1 - 0{,}59R$ (27)

in the case of a single lower limit A_2, $X \geq A_2 + 0{,}59R$ (28)

in the case of a double limit (A_1 and A_2), these two conditions shall be met (see 7.1.2).

The use of equations (27) and (28) constitutes a guide for the supplier and shall not be interpreted as an obligation. A value plotted between the value and the specification and the limit of equation 27/28 does not constitute a proof of nonconformity.

CONSIGNEE'S MARGIN OF INSPECTION

A consignee who has no other source of information regarding the true value of a characteristic than an individual result shall consider that the product does not comply with the specification limit(s), only if the result X is such that the following condition (or one of the following conditions) is achieved, the corresponding confidence level being not less than 95 % :

in the case of a single upper limit A_1,

$X > A_1 + 0{,}59R$ (29)

in the case of a single lower limit A_2,

$X < A_2 - 0{,}59R$ (30)

in the case of a double limit (A_1 and A_2), both of these conditions shall be achieved.

Conformity and nonconformity rules in event of disagreement

If it is impossible for the supplier and consignee to reach an agreement regarding the appraisal of the quality of the product, on the basis of the obtained results, the procedure described below shall be adopted.

10.1 In event of disagreement, both laboratories shall take up contact and compare their respective procedure and apparatus. After these investigations, a correlation test between both laboratories shall be carried out on the two inspection samples. The mean of at least three acceptable results shall be calculated in each laboratory and compared as indicated in 10.2.

10.2 Each laboratory shall reject its own results and obtain at least three other acceptable results on the inspection sample in order to be certain that the work has been performed under repeatability conditions. The mean of the acceptable results of each laboratory shall then be calculated, the divergent results being rejected as indicated in 7.1.1.

Let :

$\overline{X}_S$ be the supplier's mean ;

$\overline{X}_R$ be the consignee's mean ;

A_1 be the upper specification limit ;

A_2 be the lower specification limit.

where

$\overline{X}_S \leq A_1 \ \overline{X}_R$

$\overline{X}_S \geq A_2 \ \overline{X}_R$

This signifies that $\overline{X}_S$ and $\overline{X}_R$ shall be compared as follows with A_1 and A_2.

10.2.1 If :

$$\frac{\overline{X}_S + \overline{X}_R}{2} \leq A_1 \text{ or } \geq A_2$$

the product is accepted if $|\overline{X}_S - \overline{X}_R| \leq 0{,}84R_2$ (for R_2 see 7.2.1).

possible disagreement if $|\overline{X}_S - \overline{X}_R| > 0{,}84R_2$.

In the last case, it cannot be declared with confidence that the product conforms or not to the specification limit ; consequently, the disagreement can be resolved by negotiation.

10.2.2 If :

$$\frac{\overline{X}_S + \overline{X}_R}{2} > A_1 \text{ or } < A_2$$ disagreement whatever the difference $\overline{X}_S - \overline{X}_R$.

10.3 If the disagreement persists, a third laboratory (neutral, qualified and accepted by both parties) shall be requested to conduct the test on a third sample. Let $\overline{X}_E$ be the mean of the three acceptable results of this laboratory. If the difference between the most divergent laboratory mean and the mean of the means of the two other laboratories is below or equal to R_3 (see 7.2.1), the procedure described below shall be adopted.

10.3.1 If : $\dfrac{\overline{X}_S + \overline{X}_R + \overline{X}_E}{3} \leq A_1 \text{ or } \geq A_2$ the product is accepted.

10.3.2 If : $\dfrac{\overline{X}_S + \overline{X}_R + \overline{X}_E}{3} > A_1 \text{ or } < A_2$ the product is refused.

10.4 If the difference between the most divergent laboratory mean and the mean $\overline{X}$ of the two other laboratories is greater than R_3, the procedure described below shall be adopted.

10.4.1 If : $\overline{X} \leq A_1 \text{ or } \geq A_2$ the product is accepted.

10.4.2 If : $\overline{X} > A_1$ or $< A_2$ the product is refused.

ANNEX 1

COMMENTS OF BRAZIL ON THE DISCUSSION PAPER

1- We suggest in second page after "third one being kept" to add "*this sample would be analysed in the same laboratory that refused the sample in presence of a representative from the second laboratory."*

2- We suggest in the second page between:

The parties conduct together a representative sampling operation, taking three samples from the batch forming the subject of the dispute. A sample is forwarded to each of the laboratories, the third one being kept.

and

The two laboratories compare their respective procedures and apparatus. After these investigations, a correlation test shall be conducted between both laboratories. The analysis may also be performed in a single laboratory in presence of a representative from the second laboratory.

To add "*Both of the parties shall agree with analytical method to be utilized."*

ANNEX 2

COMMENTS IN REPLY TO CL 2001/5-MAS (March 2001)

THAILAND

Criteria for Evaluating Acceptable Methods of Analysis : Dispute Situations.

In case of dispute situation, Thailand proposes information on the practical point of view as follows:

1) Selected methods of analysis should be suitable to detection limit.
2) Laboratories should practice quality assurance system on the basis of international guideline.
3) Laboratories should be cooperated in proficiency testing at international or national level, where available.
4) Whenever dispute situations occur, it is necessary to carry on an interlaboratory comparison of export and import laboratories. The new sample must be taken. It is homogenized and separated for both laboratories to analyse in the same period by the use of accepted method.
5) If interlaboratory comparison can not be agreeable. The analysis should be managed by the accredited and accepted laboratory as the third party to finalize on the results.