

Measurement uncertainty in wine appreciation

CARSTEN FAUHL

Bundesinstitut für Risikobewertung
D-14195 Berlin, Thielallee 88-92
E-Mail: c.fauhl@bfr.bund.de

In an accredited laboratory measurement uncertainty must be considered within the evaluation of the assessment of relevant parameters. The crucial point as to determine for wine appreciation is whether a value does or does not differ significantly - with due regard being given to measurement uncertainty - from the pre-set limit that is regulated by law or acknowledged by experience. If the measured analytical value is really exceeding this crucial point - the decision limit -, this may give rise to complaints against the product. Several concepts for estimating the measurement uncertainty and its further consideration in wine appreciation by defining the decision limit are discussed, e.g. the expanded measurement uncertainty, the decision limit $CC\alpha$ according to Commission decision 2002/657/EC and the critical difference according to the ISO 5725 norm. In principle, decrease of measurement uncertainty will be obtained by performing repetitive successive measurements. The recommendations of ISO 5725 specify the role of the critical difference for the determination of the decision limit. The examination of the critical difference depending on the ratio of repeatability/reproducibility limit and on the number of successive measurements shows that distinct reduction of the critical difference can be obtained by carrying out the n -repetition of measurement with approximating precision limits only. In the area of wine analysis, however, the accentuation of measurement uncertainty proved to be slight in fact, persistently slight even in the case of multiple measurement repetition. The differences found among decision limits comparison of the various concepts having been used respectively to pay tribute to measurement uncertainty, could be considered as relatively minor shown by the chosen examples. From the point of view of wine control, the application of a one-tailed distribution approach for instance results in accentuating the limits, but the basic problem formulation and research hypothesis, includes specific requirements for the analysis of natural / native components must be taken into account.

Key words: measurement uncertainty, wine analysis, appreciation, decision limit

Messunsicherheit in der Weinbeurteilung. Die Messunsicherheit ist bei der Bewertung von beurteilungsrelevanten Parametern in einem akkreditierten Prüflaboratorium unbedingt zu berücksichtigen. Der Übergang, ab wann ein Parameter unter Berücksichtigung der Messunsicherheit von einem bestimmten Limit, das gesetzlich festgelegt oder auf Erfahrungswerten basieren kann, signifikant abweicht, ist für die Beurteilung eines Weines besonders kritisch. Ab diesem Wert, der Entscheidungsgrenze, überschreitet der ermittelte Analysenwert den Grenzwert tatsächlich und gegebenenfalls sind weitere Schritte, wie zum Beispiel eine Beanstandung des Produktes, notwendig. Verschiedene Möglichkeiten für die Abschätzung der Messunsicherheit, die Festlegung der Entscheidungsgrenze sowie die entsprechenden Auswirkungen bei der Weinbeurteilung werden diskutiert, unter anderem die erweiterte Messunsicherheit, die Entscheidungsgrenze $CC\alpha$ nach der Kommissionsentscheidung 2002/657/EC und die kritische Differenz nach der ISO-Norm 5725. Prinzipiell kann die Messunsicherheit mittels Durchführung von Wiederholmessungen verkleinert werden. Für eine entsprechende Berücksichtigung bei der Festlegung der Entscheidungsgrenze ist die Anwendung der kritischen Differenz nach ISO 5725 beschrieben. Die Untersuchung der kritischen Differenz in Abhängigkeit des Quotienten aus Wiederholgrenze und Vergleichsgrenze und der Anzahl von Wiederholmessungen zeigt, dass die n -fache Wiederholung eines Messwertes nur bei angenäherten Präzisionsgrenzen zu einer deutlichen Reduktion der kritischen Differenz führt. Beispiele aus der Weinanalytik zeigen, dass die tatsächliche Verschärfung der Messunsicherheit auch bei einer vielfachen Wiederholung des Messwertes gering ist. Die Unterschiede zwischen den Entscheidungsgrenzen bei Anwendung der verschiedenen Konzepte zur Berücksichtigung der Messunsicherheit sind bei den ausgewählten Beispielen als eher geringfügig einzustufen. Die Verwendung von einseitigen

Verteilungsansätzen führt zu einer Verschärfung der Grenzen aus der Sicht der Weinkontrolle, allerdings sind bei der Analyse von natürlichen Inhaltsstoffen besondere Anforderungen hinsichtlich der zugrunde liegenden Hypothese bzw. Fragestellung zu berücksichtigen.

Schlagwörter: Messunsicherheit, Weinanalyse, Beurteilung, Entscheidungsgrenze

Imprécision des mesures dans l'évaluation des vins. Lors de l'évaluation de paramètres importants, il faut absolument tenir compte de l'imprécision des mesures effectuées dans un laboratoire d'essais accrédité. Le seuil qui peut être fixé par la loi ou basé sur des valeurs empiriques, à partir duquel un paramètre s'écarte significativement d'une valeur limite, est particulièrement critique pour l'évaluation d'un vin. À partir de cette valeur, le seuil de décision, soit la valeur d'analyse déterminée, dépasse effectivement la valeur limite et, le cas échéant, d'autres mesures, telles qu'une réclamation concernant le produit, sont nécessaires. Les différentes possibilités d'estimer l'imprécision des mesures, la détermination du seuil de décision ainsi que les conséquences correspondantes pour l'évaluation du vin font l'objet du présent article et, entre autres, l'imprécision des mesures élargie, le seuil de décision CC α conformément à la Décision 2002/657/CE de la Commission et la différence critique selon la norme ISO 5725. En principe, l'imprécision des mesures peut être réduite en effectuant des mesures répétées. L'application de la différence critique selon ISO 5725 est décrite pour qu'elle puisse être prise en compte lors de la détermination du seuil de décision. L'examen de la différence critique en fonction du quotient de la limite de reproductibilité et du seuil de décision, ainsi qu'en fonction du nombre des mesures répétées montre que la répétition n fois d'une valeur mesurée ne conduit à une réduction sensible de la différence critique que dans le cas de limites de précision rapprochées. Il ressort d'exemples d'analyses des vins que l'augmentation de l'imprécision des mesures est faible, également dans le cas d'une répétition multiple de la valeur de mesure. Dans les exemples choisis, les différences entre les seuils de décision lors de l'application des différents concepts destinés à tenir compte de l'imprécision des mesures peuvent être considérées comme plutôt insignifiantes. L'utilisation d'approches de répartition unilatérales conduit à un renforcement des limites du point de vue du contrôle des vins; lors de l'analyse des composants naturels, il faut cependant tenir compte des exigences particulières en ce qui concerne les hypothèses ou bien les questions sur lesquelles ils se fondent.

Mots clés : Imprécision des mesures, analyse du vin, évaluation, seuil de décision

Not only since accreditation became essential in official control laboratories it is a demand of reliable analysis that if a parameter is determined, its particular measurement uncertainty should be directly stated with the actual value or full information about the precision, repeatability and reproducibility of the applied method should be available. Otherwise the quantity of the determined parameter can hardly be assessed correctly. The definition of measurement uncertainty supposes that the stated value varies within a certain range due to the precision of the analytical method and the width of the range depends on the probability level considered. Apart from the so-called component by component approach in which the error of each relevant analytical step is considered and summed up according to the rules of error propagation, it is generally accepted that accurate estimation for the uncertainty derives from a collaborative trial (Eurachem, 2000). Presuming conduction and evaluation followed international recommendations (ISO, 1994; HORWITZ, 1995). Measurement uncertainty is usually defined by the standard deviation of the re-

producibility, optionally with a stated factor for increasing of the probability level defining the so-called expanded measurement uncertainty. Official wine control within Europe is highly unified and validated methods shall be applied. Therefore precision characteristics deriving from a collaborative trial for the relevant methods are generally available and uncertainty can be assessed.

The comparison of an actual measurement value with a reference value, that can be in the case of wine a regulatory limit at European level, national level, a recommendation by the O.I.V. or a so-called experience value for certain parameters is one major piece of the control process. Of course the overall process of wine appreciation is more complex because a lot of parameters of the analysed product contribute to the total impression of an expert, but this paper will focus on the „simple“ single parameter proof only. This part of the appreciation always relies back on whether the analytical parameter determined, considering also the measurement uncertainty, is violating the established limit or if it is still in compliance with it. The paper will

highlight some different situations, which may occur during the analysis and appreciation of wine.

Wine analysis in Europe

In the European Union the analysis of wines is a special case because methods of analysis are laid down in Annex of Regulation (EEC) No 2676/1990 (EU, 1990). The determination of 45 analytical parameters is described and for some of them more than one method is stated, the usual and the reference method. For substances for which reference methods and usual methods are prescribed, the results obtained by the use of the reference methods shall prevail. For almost all reference methods precision parameters were determined in collaborative trials. Table 1 shows the summary of all methods and their precision performance parameter, the limits of repeatability (r) and reproducibility (R). Apart from the method collection in Annex of Regulation (EEC) No 2776/1990 Article 46 of the superior basic Council Regulation (EC) No 1493/1999 (EU, 1999) regulates the application of O.I.V. methods for other parameters than the provisioned ones in Regulation (EEC) No 2776/1990. The O.I.V. provides a comprehensive compilation of internationally accepted methods, the Compendium of International Methods of Analysis (O.I.V., 2006), for many parameters. Further Article 46 states that if parameters have to be analysed for which no methods are available in the mentioned sources methods complying the standards of the International Organisation for Standardization (ISO) shall be used, or in case of absence and by reason of their repeatability, reproducibility and accuracy methods allowed at national level or any other appropriate method can be used.

Starting situation

The European Norm which sets the standards and requirements for accreditation for testing and calibration laboratories is the EN ISO/IEC 17025 (ISO, 2005). According to Regulation (EC) No 882/2004 (EU, 2004a) accreditation will be required for official control laboratories in food and feed analysis from the 1. January 2006 on. It is evident that the requirements of accreditation set by EN ISO/IEC 17025 (ISO, 2005) are fundamental in wine control as well. According to this norm the test result of an analytical measurement must be stated with an estimate of its uncertainty in particular when it is relevant to the application of the test results or when uncertainty affects compliance to a specification limit. For most of the pa-

rameters which are relevant for the wine appreciation measurement uncertainty shall therefore be stated and it must be considered for drawing conclusions, respectively.

From a statistical point of view five different situations, graphically presented in Figure 1, must be considered when a test result was determined in a laboratory for wine appreciation and control. The circles present the test result and the error bars are defining the measurement uncertainty of the value.

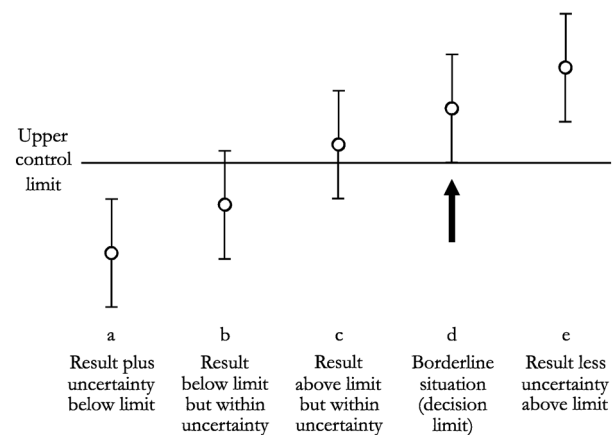


Fig. 1: Statistically different situations during appreciation.

Situation

- result plus uncertainty below limit
- result below limit, but limit within uncertainty
- result above limit, but limit within uncertainty
- borderline situation (decision limit)**
- result minus uncertainty above limit

Situations **a**) and **b**) do not cause any consequences since the results are in compliance with the control limit. Situation **c**) may already lead to different interpretations, when uncertainty is taken into account the result is compliant with the limit, but when measurement uncertainty is completely ignored by the analyst, the result is interpreted to be non-compliant. In this case the terminus „beyond reasonable doubt“ - that should be demonstrated in case of official objection - does not fit perfectly and also conventions of accreditation are mistreated. On the other hand situation **e**) clearly presents a violation.

The particular situation **d**) is of utmost importance in laboratory praxis because it defines if a measurement value violates a limit truly or if it is still in compliance with the control limit and depending on the decision

Tab. 1: Summary of methods in Annex of Regulation (EEC) No 2676/1990 with precision parameters

Method name		r	R	r/R	Method
1 Density and specific gravity at 20°C	dry, semi-sweet w.	0,00010	0,00037	0,27	ref. pycnometer
	sweet w.	0,00018	0,00045	0,40	u hydrometry
					u densimetry using hydrostatic balance
2 Refraction index in grape musts					only refractometry
3 Alcoholic strength by volume		0,10 %vol	0,19 %vol	0,53	ref. pycnometer
		0,074 %vol	0,229 %vol	0,32	ref. hydrostatic balance
		0,061 %vol	0,174 %vol	0,35	ref. electronic densimeter
4 Total dry extract					u hydrometry
					u refractometry
					only densimetry
5 Reducing sugars		0,015 x _i	0,058 x _i	0,26	ref. (ion exchange/Pb-acatate) iodometrically
6 Sucrose					u (Pb-acatate/Zn2-hexacyanoferrate) iodometrically
					u thin-layer chromatography
7 Glucose and fructose		0,056 x _i	0,12 + 0,076 x _i		ref. enzymatic
8 SNIF-NMR					only SNIF - NMR
9 Ash content					only
10 Alkalinity of the ash					only
11 Chlorides		1,2 mg/l	4,1 mg/l	0,29	only potentiometry
12 Sulphates	< 1000 mg/l	27 mg/l	51 mg/l	0,53	ref. gravimetrically
	~ 1500 mg/l	41 mg/l	81 mg/l	0,51	
13 Total acidity	g tartaric acid /l	0,07 g/l	0,3 g/l	0,23	titration with indicator
			0,4 g/l (r.w.)	0,18	
14 Volatile acidity	g acetic acid /l	0,04 g/l	0,08 g/l	0,50	only potentiometric titration
15 Fixed acidity					distillation and titration difference
16 Tartaric acid					ref. gravimetrically
					u colorimetrically
17 Citric acid	< 400 mg/l	14 mg/l	39 mg/l	0,36	only enzymatic
	> 400 mg/l	28 mg/l	65 mg/l	0,43	
18 Lactic acid		0,02 + 0,07 x _i	0,05 + 0,125 x _i		ref. enzymatic
19 L-Malic acid		0,03 + 0,034 x _i	0,05 + 0,071 x _i		only enzymatic
20 D-Malic acid		11 mg/l	20 mg/l	0,55	ref. enzymatic
	< 50 mg/l	12,4 mg/l (w.w.)			
21 Total malic acid	< 2 g/l	0,1 g/l	0,3 g/l	0,33	only colorimetrically
	> 2 g/l	0,2 g/l		0,67	
22 Sorbic acid					u spectrophotometry
					u GC
					u thin-layer chromatography
23 L-Ascorbic acid					ref. fluorimetry
24 pH					
25 Sulphur dioxid	≤ 50 mg	1 mg/l	9 mg/l	0,11	ref. oxidation and titration
	> 50 mg	6 mg/l	15 mg/l	0,40	
26 Sodium		1 + 0,024 x _i	2,5 + 0,05 x _i		ref. AAS
		1,4 mg/l	4,7 + 0,08 x _i		u flame photometry
	liqueur wine	2,0 mg/l			
27 Potassium		35 mg/l	66 mg/l	0,53	ref. AAS
		17 mg/l	66 mg/l	0,26	u flame photometry
28 Magnesium		3 mg/l	8 mg/l	0,38	only AAS

Tab. 1 (cont.): Summary of methods in Annex of Regulation (EEC) No 2676/1990 with precision parameters

Method name		r	R	r/R	Method	
29	Calcium	< 60 mg > 60 mg	2,7 mg/l 4 mg/l	0,114 x _i - 0,5	only AAS	
30	Iron				ref. AAS u photometry	
31	Copper				only AAS	
32	Cadmium				only AAS	
33	Silver				only AAS	
34	Zinc				only AAS	
35	Lead				only AAS	
36	Fluorides				only Ion-selective electrode	
37	Carbon dioxide				ref. titration	
37a	Excess pressure				only aphrometer	
38	Cyanide derivatives	white wine red wine	3,1 µg/l 6,4 µg/l	12 µg/l 23 µg/l	0,26 0,28	ref. colorimetry
39	Allylthiocyanate				only GC	
40	Chromatic properties				ref. spectrophotometry	
41	Folin-Ciocalteu-Index		≤ 1		only spectrophotometry	
42	Special methods for rectified concentrated grape must					
43	¹⁸ O/ ¹⁶ O ratio of wine water		0,24 ‰	0,50 ‰	0,48	ref. IRMS
	wine > 14 %vol	µg/l	µg/l			
	40	1,59	4,77	0,33		
	80	3,32	7	0,47		
	162	8,2	11,11	0,74		
44	Ethyl carbamate	wine > 14 %vol	µg/l	µg/l		ref. GC/MS
	11	0,43	2,03	0,21		
	25	1,67	2,67	0,63		
	48	1,97	4,25	0,46		
45	¹³ C/ ¹² C ratio in ethanol		0,24 ‰	0,6 ‰	0,40	ref. IRMS

x_i = concentration of the components per sample

ref. = reference method

u = usual method

further action (e.g. official objection) may become necessary. Therefore this crucial point will be highlighted in the following. Two factors possibly affecting the limit, measurement uncertainty and the character of the control limit itself, are considered in more detail.

Measurement uncertainty (error bar)

Different fundamental papers published by standardisation organisations cover the expression and estimation of measurement uncertainty which alternatively could be described with analytical variability. When a sample is analysed, the obtained results vary according to the analytical precision and measurement uncertainty attempts to define the degree of variation. In 1993 the ISO Guide (ISO, 1995) introduced the idea of uncertainty and distinguished it from „error“. The definition of uncertainty given in the EURACHEM Guide (Eurachem, 2000) is: „a parameter associated with the result of a measurement that characterises the dispersion of the values that could reasonably be attributed to the

measurand“. The ISO Guide (ISO, 1995) and also the EURACHEM Guide (Eurachem, 2000) place emphasis on the so called component-by-component approach in which the analytical procedure of the applied method is dissected and the combined measurement uncertainty is calculated by incrementing the relevant steps. Every single analytical step that might affect the final uncertainty must be detected and its uncertainty, expressed as standard deviation contributes to the total uncertainty. Typical steps are for example weighing, pipetting, calibrant uncertainty, signal measurement, extraction variability and so on. However this approach also called bottom-up, is theoretically and often criticised to be not practicable to analytical chemistry.

On the other hand the Eurachem Guide (2000) states also that results obtained from a collaborative study - that followed international rules - may deliver an acceptable estimation of uncertainty unless a relevant factor of possibly measurement error was not covered by

the conducted study. Then, eventually further factors should be included into the estimate of uncertainty.

This anticipates that detailed information on the collaborative trial is available and can be critically reviewed. Particularly in the case of traditional, classical wine analysis, which are defined in Regulation (EEC) No. 2676/1990 exact information on the actual trial is hardly be assessable. However, the conduction of a good validation study provides all necessary data and it is possible to justify the use of an appropriate statistic, such as standard deviation of the reproducibility (S_R), to determine uncertainty. This approach - the use of S_R obtained in a collaborative study as estimate for measurement uncertainty - was internationally confirmed and accepted in the field of wine analysis by the O.I.V. with Resolution OENO 09/2005 (O.I.V., 2005) recently, which recommends the use of S_R if available. Nevertheless it must be underpinned that it is a demand of accreditation rules (ISO, 2005) that if a method, which was validated previously in a collaborative trial, is applied in a laboratory at least basic validation experiments should be performed in order to confirm that the established precision parameters are also achieved in the actual laboratory.

A very pragmatic approach was proposed by HORWITZ (2003) which described the application of the Horwitz-Formula as rough calculation of the expected uncertainty for the anticipated concentration of the ana-

lyte. From experience this gives an estimate of the expected uncertainty. However the method considered should perform „normative“ for this assumption and surely not all developments will fit this prerequisite.

A recently published EU report presented a comprehensive overview about the different approaches and reflects problematic effects of measurement uncertainty on provisions in food and feed legislation (EU, 2004b). From a practical point of view and taking into account the recommendation of the O.I.V. (O.I.V., 2005) for wine analysis measurement uncertainty often relies on S_R obtained in a collaborative study. For a better understanding the term standard deviation can be visualised in a frequency distribution. If a sample is analysed with a certain method very often the obtained results vary according to the precision and can be illustrated in a frequency distribution. If the measurement is carried out infinitely often and normal distribution is assumed the frequency distribution will result in the so called Gauß-distribution. Looking at the normally distributed frequency distribution, the arithmetic mean \pm standard deviation defines the limits in which 68.3 % of the measurement values will fall. The conditions of the analysis - during repetition - are very important and more deeply defined. By convention the circumstances for the measurements are generally categorised into repeatability and reproducibility conditions, resulting in S_r (standard deviation of the repeatability) and

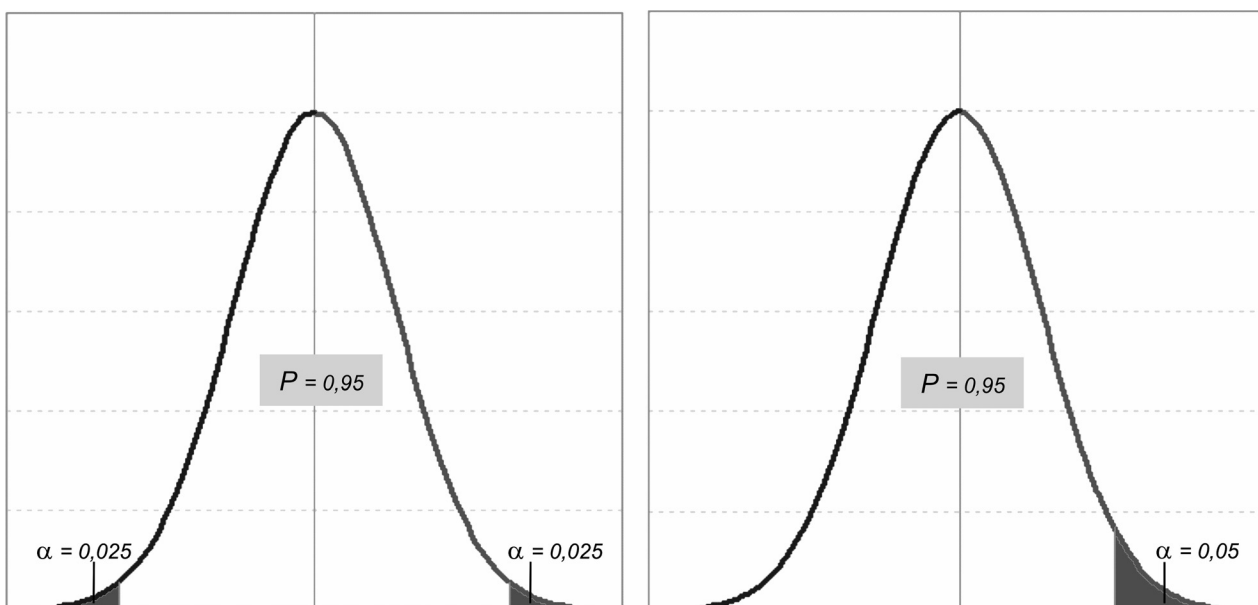


Fig. 2: One-tailed and two-tailed distribution

S_R (standard deviation of the reproducibility) (ISO, 1994). Measurement uncertainty defines the interval of $x \pm S_R$ around the stated value (x). When the expanded measurement uncertainty is used the interval is expanded to the desired or required probability by multiplication of the standard deviation by a coverage factor, then it usually assigns the 95% probability interval. If the expanded uncertainty is stated it should be clarified to which level of probability it is anticipated, that is however not always the case. Even if extension is defined it already provides the possibility of misunderstandings, because the multiplication with a factor results in a certain probability for a defined situation only. Multiplication with 1.96, typically done for the probability level of 95%, is valid for a two-tailed situation. In case of a one-tailed distribution only a factor of 1.64 is necessary to cover also 95% probability. This is shown in Figure 2 where the difference from the centre to the limit in the upward direction is higher for the two sided distribution - because 2.5% are covered on one side - than for the one-tailed distribution.

In many fields of analysis the „classical“ decision limit derives from the actual measurement value reduced by two times the measurement uncertainty for the relevant concentration. If this value still exceeds the specification then non-compliance is given. This approach is discussed at a European level for contaminants (EU, 2004b) and is a common practice for parts of residual analysis such as pesticides.

In case of residue analysis of veterinary drugs (EU, 2002) the decision limit has been defined with the one-tailed distribution. The concept of performance criteria for analytical methods qualified for the analysis of veterinary drug residues defines the decision limit $CC\alpha$ (see Equation 1), from which value on true violation of the limit is given. One-tailed distribution is considered only because violation of a provision will always be an exceeding of a certain limit for a veterinary drug residue.

$$CC\alpha = MRL + 1.64 \times S_R$$

Equation 1: $CC\alpha$ for Maximum Residue Level (MRL) substances (EU, 2002)

Application of one-tailed distributions or definitions of one-tailed error bars in wine analysis is clearly indicated for contaminants or undesired ingredients such as lead or volatile acidity for example.

On the other hand naturally distributed variables like e.g. mineral contents in wine would in principle follow a two-tailed distribution, because the variation can fol-

low two directions. From a statistical point of view the application of one-tailed distributions requires the theoretical exclusion of one of the two directions. The question of the hypothesis to be answered defines the appropriate use of one-tailed or two-tailed distributions. For isotopic parameters the consideration of one- or two-tailed distributions depends on the suspect and information about the wine. If the addition of sugar is the clear assumption the proof of chaptalisation by the use of the D/H ratios and the $\delta^{13}C$ value of wine ethanol would allow the application of a one-sided distribution. In case that water addition is the clear suspect in the expert opinion for the explanation of a low $\delta^{18}O$ value found - supported by other indications drawn from further parameters - one-sided distribution probability limits are in principle applicable. Although the sharpness of the evaluation is better if one-tailed distribution is applied from a control point of view, reasonable exclusion of one direction of distribution remains and its justification is left to the expert. Application of two-tailed distribution probabilities seems to be the most appropriate if no specific information is available, because this surly covers all doubts.

Concept of critical difference (effect of repeatability)

The extensive application of precision parameters, determined in a collaborative trial for method validation is described in the ISO 5725 (Part 6) (ISO, 1994) and was adapted to different statistical documentation, e.g. the Swiss Food Law (Schweizerisches Lebensmittelbuch, 2005). The concept was originally developed before the term measurement uncertainty became established in analytical chemistry. It is important to note that it is based on the assumption that laboratories are operating at the same level of quality as those which participated in the original collaborative trial to validate the method. But referring to wine analysis according to unified methods, e.g. of Regulation (EEC) No 2676/1990, it must be assumed that the control laboratories are performing at a similar level, and therefore application of this concept appears to be appropriate in the sector of wine control.

The most easiest and known cases of the critical difference are the repeatability limit (r) and reproducibility limit (R), which describe the comparison of two single measurements a) in one laboratory (r) or b) in two laboratories (R). Apart from these well established norm situations more complex situations comparing values of two laboratories with n and n_1 determinations in

each laboratory are defined in ISO 5725 (ISO, 1994). In addition different situations for the comparison with limits are defined. In this case the critical difference defines from which distance to the comparative value the measurement value is truly violating the limit with a given probability.

$$CrD95 = \frac{0.84}{\sqrt{2}} \times \sqrt{R^2 - \left(r^2 \times \frac{(n-1)}{n} \right)}$$

Equation 2: Calculation of Critical Difference (95% significance, one-tailed)

For $n = 1$ $CrD95$ in equation 2 equals $CC\alpha$ (Equation 1).

$$CrD95_{n=1} = \frac{0.84}{\sqrt{2}} \times R = \frac{0.84}{\sqrt{2}} \times 1.96 \times \sqrt{2} \times S_R = 1.64 \times S_R$$

Equation 3: $CrD95$ for $n = 1$

Looking at Equation 2 it is obvious that repeatability affects the critical difference if more than one determination is carried out in the laboratory.

It is likely assumed that with the number of determinations the analytical variation is reduced and sometimes it is even supposed that it can be totally neglected if the determination has been repeated often within one laboratory but in fact the improvement in terms of the reduction of the critical difference depends on the ratio

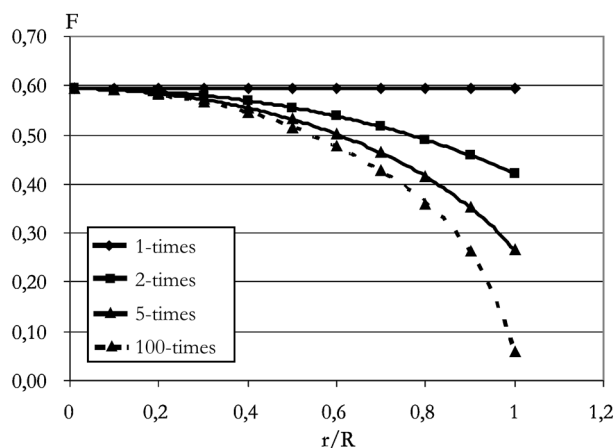


Fig. 3: Influence of the number of repetitions on the critical difference and ratio of r/R (with $CrD = F \times R$, so F is the factor by which R must be multiplied in order to obtain CrD , displayed for different ratio r/R and number of repetitions, for $n = 1$ the factor F becomes 0.59)

of reproducibility and repeatability. Figure 3 shows the behaviour of the critical difference for n -repetitions in one laboratory and the ratio of r/R . Only if this ratio is close to 1 a considerable improvement is observed.

On the other hand it should be noted that the critical difference is significantly reduced if a second laboratory analyses the relevant parameter in the suspicious sample and the mean value of the two laboratories is then compared with the specification (ISO, 1994) for appropriate equation. This is of course not always applicable in practice although in crucial cases the benefit of information regarding the sharpness of decision is valuable.

Control limit

The nature of the control limit can also be different. The upper-limit in Figure 1 can be a limit set by legislation, a recommended value or a comparative limit, what means that the limit is set by the analyst e.g. with help of data bases. Of course the situation shown in Figure 1 is valid for lower control limit in the opposite way, too.

In the case of wine analysis parameter limits are fixed in EU regulations, other limits are defined by the O.I.V. and additional limits have been established on a national level. For several relevant analytical parameters no limits are laid down by law or established by international organisation in wine appreciation. This situation occurs e.g. often in the case of authentication with stable isotope analysis but also in several other cases such as mineral content analysis and evaluation. For these parameters the experience of the analyst and availability of comparison or reference values is important, because the effectiveness of control depends on it.

The selection of reference collectives from data bases is a very crucial and delicate matter and should therefore left to experts only. Representativeness and comparability are only two points of importance. Apart from the process of selecting of relevant reference data - that is not discussed in this paper - the process after selection of the reference data set is highlighted from a statistical point of view. After the selection of comparison samples these values are evaluated regularly by calculating the arithmetic mean, median, standard deviation and their confidence limits. For these operations ideally at least 30 reference samples should be available, what is not always the case, of course. Calculation of the confidence limits by consideration of the Student-Factor (t -distribution) for a low number of reference samples was described for authenticity proof in wine analysis previously (CHRISTOPH et al., 2003; OTTENEDER et al.,

2004) and is accepted to be valuable for data interpretation. The t-distribution for a two-tailed question delivers an upper or a lower limit and defines the range which is used for the comparison with the actual value of the wine in suspect. If the actual measurement value falls in the range then it fits statistically into the reference collective with a given probability.

$$C_{Upper-Limit} = X + (S \times t)$$

$$C_{Lower-Limit} = X - (S \times t)$$

Equation 4: Upper and lower confidence limit (with X = arithmetic mean of reference samples, S = standard deviation and t = Student-Factor for n = values and certain probability)

Again what must be selected by the expert for each individual case is what kind of distribution - one- or two-tailed - is appropriate for interpretation. If a one-tailed distribution is considered, only one limit - a lower or upper - is the result for the confidence limit calculation and consequently only exceeding or falling below is decisive. CHRISTOPH et al. (2003) depicted the data handling and interpretation with an example for stable isotope analysis.

It is sometimes discussed that the analytical variation must also be considered additionally if a set of reference samples was used for the calculation of the confidence intervals. However MARTIN et al. (1996) have described concretely that the analytical variation can be neglected when the natural variation is considered because the analytical variation is already included in the natural one. Therefore the limits obtained from such a calculation are comparable with regulatory limits from a statistical point of view.

Examples in wine analysis

Volatile acidity. Specifications for this quality criteria are laid down in Regulation (EC) No. 1493/1999. The example presented here is based on the case of red wine for which a maximum content of 20 milli equivalents volatile acidity per liter, corresponding to 20 meq/l is fixed. The method to be used for official control purposes is Method 14 of Annex of Regulation (EEC) No 2676/1990 „Volatile acidity“, with r = 0.7 meq/l and R = 1.3 meq/l. Figure 4 shows the decision limits calculated with the different concepts, ranging from 20.93 meq/l for calculating the decision limit (DL) by subtracting 2 x the measurement uncertainty

(MU) from the actual measurement value (X), down to 20.7 meq/l for application of the critical difference with 100 repetitions.

All limits are relatively close together, and the improvement from the control point of view obtained by the repetition of the measurement is not that effective even when the measurement would have been carried out 100 times, what is of course absolutely unrealistic but for demonstration of the behaviour this number was implemented.

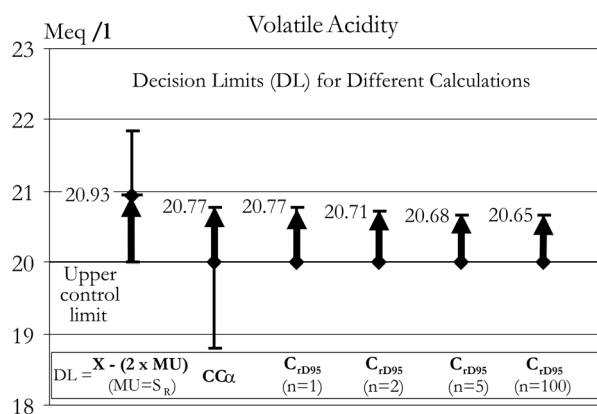


Fig. 4: Volatile acidity - Effect of different concepts for calculating the decision limit

δ¹⁸O-Value. The official method for the measurement of the δ¹⁸O-value of wine water is defined in method 43 in the Annex of Regulation (EEC) No 2676/1990. The precision parameters r and R are given with 0.24 ‰ and 0.5 ‰ δ¹⁸O, respectively. Limits for δ¹⁸O-value are not laid down in any provision. The example considered in Figure 5 is based on fictive data set for a country or region for which the statistical characteristics are given with Table 2.

Tab. 2: Fictive data set for calculation δ¹⁸O-value confidence limits

δ ¹⁸ O- values in ‰	Two-tailed	One-tailed
n	44	
Minimum	-1,26	
Max	2,14	
Mean	0,58	
Standard deviation	0,89	
Median	0,82	
Student-Factor	2,02	1,69
95% C _{Lower limit (-)}	-1,22	-0,92
95% C _{Upper limit (+)}	2,37	

In the case of water addition only the lower limit is relevant because the tap water added usually is characterized by a very negative $\delta^{18}\text{O}$ -value and its addition causes a reduction of the initial value in wine water. Figure 5 shows that the decision limit varies only from -1.5 and -1.6 ‰ between the different calculations. Even the conduction of 100 repetitions reduces the decision limit not significantly. However it should be noted that if a one-tailed distribution is assumed for the data given in Table 2 the lower limit would be -0.91 ‰ instead of -1.22 ‰ leading to decision limits between -1.2 ‰ and -1.3 ‰, so 0.3 ‰ lower than using the two-tailed distribution.

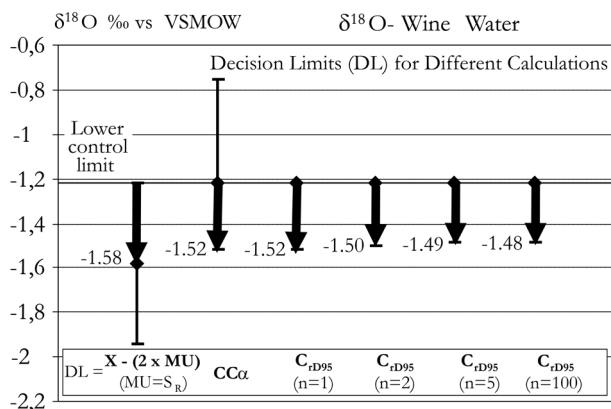


Fig. 5: $\delta^{18}\text{O}$ -value - Effect of different concepts for calculating the decision limit

^2H -NMR. In case of the deuterium (^2H)-NMR method for the determination of chaptalisation no precision data are stated in method 8 in the Annex of Regulation (EEC) No 2676/1990. However, the method has been validated comprehensively for fruit juice, AOAC 995.17 Method (AOAC, 2000), in which the standard deviation of repeatability (S_r) is 0.25 ppm and 0.37 ppm for standard deviation of reproducibility (S_R) were obtained for the worst sample ($r = 0.69$ and $R = 1.02$) that is taken for the example presented here. CHRISTOPH et al. (2003) discussed in detail one wine sample coming from Franconia, that was found to be suspicious in terms of chaptalisation with beet sugar showing a very low (D/H)I-value of 98.4 ppm. For this sample minimum (D/H)I-values were calculated taking into account different reference samples collectives. The lower control limits obtained for a probability level of 95% were: 99.2, 99.1, 98.9 and 99.0 ppm depending on the reference data set selected. For each data set Student-

factors for the relevant number of samples were used. In Figure 6 the decision limit for a sample with (D/H)I-value of 98.4 ppm is displayed applying the different concepts of respecting the measurement uncertainty. Not for all considerations the lowest control limit of 98.9 ppm is violated, but anyway for three out of four reference sample collectives non-compliance is evident, employing the decision limit of 99.01 ppm calculated by the critical difference with two measurements. This clearly underpins the importance of reference sample selection on one hand and on the other hand the importance of measurement uncertainty related to the measurement of the suspicious sample itself which should not be neglected during evaluation.

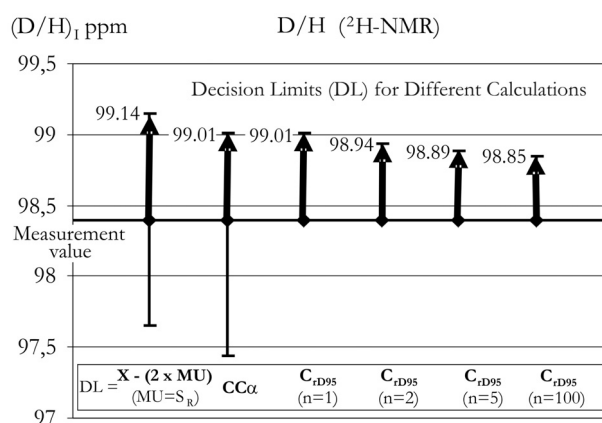


Fig. 6: (D/H)I-value - Effect of different concepts for calculating the decision limit

Conclusions

Apart from the complexity of the only partially discussed statistical approaches it was underlined that the consideration of measurement uncertainty in data interpretation is an essential point and a demand in accredited laboratories. Different methods for using - and previously even the estimation of - measurement uncertainty are available for the calculation of decision limits, unless the examples in study demonstrated that the final differences were relatively low. It should have been pointed out that the appropriate consideration of measurement uncertainty is mandatory but the exact proceeding is individually left to the expert in wine analysis and appreciation. Although wine analysis is highly unified at an analytical level further harmonisation is needed within data interpretation process.

Acknowledgement

The author would like to thank BARBARA BRAKOWIECKA-SASSY for the thorough compilation of figures and tables.

References

- AOAC (2000): Official methods of analysis, Method 995.17: Beet sugar in fruit juices. 17th Edition (Ed. by HORWITZ, W.). - Gaithersburg, Maryland: Association of Official Analytical Chemists International, 2000
- CHRISTOPH, N., ROSSMANN, A. and VOERKELIUS S. (2003): Possibilities and limitations of wine authentication using stable isotope and meteorological data, data banks and statistical tests. Part 1: Wines from Franconia and Lake Constance 1992 to 2001. Mitt. Klosterneuburg 53: 23-40
- EU (1990): Regulation (EEC) No 2676/1990 of 17 September 1990, amended with Regulation (EC) No 1293/2005 of 5. August 2005
- EU (1999): Council Regulation (EC) No 1493/1999 of 17 May 1999, amended with Regulation (EC) No 1795/2003 from 13. October 2003
- EU (2002): Commission Decision 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results
- EU (2004a): Regulation (EC) No 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules
- EU (2004b): Report on the relationship between analytical results, measurement uncertainty, recovery factors and the provisions of EU food and feed legislation, with particular reference to community legislation concerning - contaminants in food (council regulation (EEC) No 315/93 of 8 February 1993 laying down community procedures for contaminants in food) - undesirable substances in feed (directive 2002/32/EC of the European parliament and of the council of 7 May 2002 on undesirable substances in animal feed) (http://europa.eu.int/comm/food/food/chemicalsafety/contaminants/sampling_en.htm)
- Eurachem (2000): EURACHEM/CITAC Guide : Quantifying measurement uncertainty in analytical measurement. 2nd Edition (Ed. by ELLISON, S.L.R., ROSSLEIN, M. and WILLIAMS). (<http://www.measurementuncertainty.org>)
- HORWITZ, W. (1995): Protocol for the design, conduct and interpretation of method-performance studies (Technical Report). Pure & Appl. Chem 67: 331-343
- HORWITZ, W. (2003): The uncertainty of measurement uncertainty. Journal of AOAC International 86: 109-111
- ISO (1994): ISO 5725: Accuracy (trueness and precision) of measurement methods and results (Part 1-6). - Geneva: International Organisation for Standardization, 1994
- ISO (1995): Guide to the expression of uncertainty in measurement. 1st Edition. - Geneva: International Organisation for Standardization, 1995
- ISO (2005): DIN ISO/IEC 17025. General Requirements for the Competence of Testing and Calibration Laboratories. - Geneva: International Organisation for Standardization, 2005
- MARTIN, G.G., MARTIN, Y.L. and HERTZ, J. 1996: Statistische Konzepte in der Fruchtsaftanalytik. Flüss. Obst 63: 647-654
- O.I.V. (2006): O.I.V. Compendium of International Methods of Analysis of Wine and Musts. - Paris: Organisation Internationale de la Vigne et du vin, 2006 (www.oiv.int)
- O.I.V.: (2005): Resolution (2005): OENO 09/2005. Recommendations on measurement uncertainty. - Paris: Organisation Internationale de la Vigne et du vin, 2005
- OTTENEDER, H., MARX, R. and ZIMMER, M. 2004: Analysis of the anthocyanin composition of Cabernet Sauvignon and Portugieser wines provides an objective assessment of the grape varieties. Austral. J. Grape Wine Res. 10(1): 3-7
- Schweizerisches Lebensmittelbuch (2005): Kapitel 60: Statistik und Ringversuche. (<http://www.bag.admin.ch/slmb/d/>)

Received March, 13, 2006