Accuracy of industrial methods for quality measurements

Eli Vibeke Olsen **Danish Meat Research Institute, Roskilde, Denmark, E-mail: evo@dmri.dk**

Introduction

Measuring equipment for documentation of the quality of meat products is increasingly applied in slaughterhouses. Starting with the requirements for laboratory methods, an example of the documentation procedure for accuracy of industrial measurement methods is given below. New measuring methods typically replace expensive and time-consuming laboratory measurements which then become reference measurements. However, if the reference measurements are not precise it is difficult to estimate the accuracy of the new measuring methods.

Documentation of measuring equipment

In addition to a manual, the documentation must include a description of the measuring properties. These are laid down on basis of experiments which include all facets of the application of the instruments and should include the following aspects:

Accuracy
The accuracy of the method includes trueness and precision. *Trueness* refers to the agreement between the measuring result and an accepted reference value, and is normally expressed as the bias. *The precision* refers to the agreement between the measuring results divided into *repeatability* and *reproducibility*. The two measures express the lowest and the highest variation of the results and are indicated by the dispersions σ_r and σ_R. Finally the reliability of the method is relevantly defined as the ratio $\sigma^2_D/(\sigma^2_D+\sigma^2_R)$ where σ_D indicates the natural variation of the characteristic. As a rule-of-thumb the reliability should be at least 80%.

Robustness
It is essential for determination of the accuracy that the sources of measuring variations are known, and thereby a measure of the robustness the method towards external factors. The influence of external factors (temperature, light etc.) should be limited by determination of a tolerance field for these factors.

Reference
If the reference of the measuring method is another measuring method it should be described by its precision. As far as possible certified reference materials or measurements from accredited laboratories should be applied as absolute references.

Example 1: Accuracy of pH-meters

The precision of the measurement of ultimate pH in pork (longissimus dorsi muscle (LD)) is determined. Random samples from 2x25 carcasses were measured over two days. Two operators measured all carcasses using four different electrodes. In all cases the measurements were independent duplicates.
Data are analysed by means of the following model: $pH_{ijkl} = No_i + Operator_j + Electrode_k + e_{ijkl}$, where No_i indicates the true pH value in carcass number i (i=1,..,50). $Operator_j \sim N(0,\sigma^2_{op})$ and $Electrode_k \sim N(0,\sigma^2_e)$ indicate normally distributed, independent random effects and e_{ijkl} the random error or the repeatability $\sim N(0,\sigma^2_r)$. The reproducibility is defined by $\sigma_R=\sqrt{[\sigma^2_{op}+\sigma^2_e+\sigma^2_r]}$.

The repeatability forms the main part of the precision, which in LD is estimated to s_r=0.067 while s_R=0.072. It means that the average of the duplicates will lie within $\pm 2\times\sqrt{[0.072^2 - 0.067^2/2]} = \pm 2\times0.054$ with 95% certainty and therefore, the result of the pH-measurements should be stated with only one decimal. The precision s_R=0.054 of duplicate pH-measurements is just acceptable with respect to reliability, as a typical variation of the ultimate pH in pork (LD) is 0.1-0.2 (=σ_D).

The electrodes are the most variable part of the pH-meter. An investigation concerning 12 electrodes, all complying with the producer's quality requirements, showed that the electrodes could be divided into two groups with different reproducibility in the aqueous buffers. The reproducibility s_R at 6 °C was determined to 0.016 and 0.037, respectively. It was not possible to characterize the two groups of electrodes by age, sensitivity, type or the like. *The trueness* of the pH-meters has been investigated by testing the linearity between two calibration values. All electrodes were calibrated in certified aqueous buffers with the values 4.01 and 7.00. In the control buffers, 5.00 and 5.50, a bias of 0.07-0.11 was demonstrated. The consequence of these investigations is that a regular (every 6 months) determination of the accuracy of all electrodes in stock in the slaughterhouse should be performed. Furthermore the calibration must be controlled frequently during the measuring process.

Example 2: Robustness by objective measurement of meat colour

The colour of pork is an important quality parameter - for instance for the Japanese meat processing industry. The colour varies very little and an objective measurement of colour should therefore be performed with great accuracy. In order to achieve this accuracy the influence of external factors on the measurements should be uncovered.
The chosen colour measuring equipment is a Minolta CR 300, which complies with the functional requirements. The meat colour is measured on a meat surface and it is only relevant to measure on oxygenated meat with a stable colour.

The measurement consists of an average of four point measurements dispersed on a cut surface of a muscle, so that the colour variation of the surface is compensated for. Seven factors are analysed on two levels in a fractional factorial design (2^{7-4} design) with eight runs.

Factors	"Low" level (–)	"High" level (+)	Combination of level
Thickness	15 mm	35 mm	+ + + + – – – –
Surface	film	no film	+ + – – + + – –
Temperature	2 °C	7 °C	+ – + – + – + –
Time of expos.	1 hour	2 hours	+ + – – – – + +
Pressure	own weight	+1 kg	+ – – + + – – +
Position	template	subj. choice	+ – – + + – – +
Dabbing method	swab	blotting paper	+ – – + – + + –

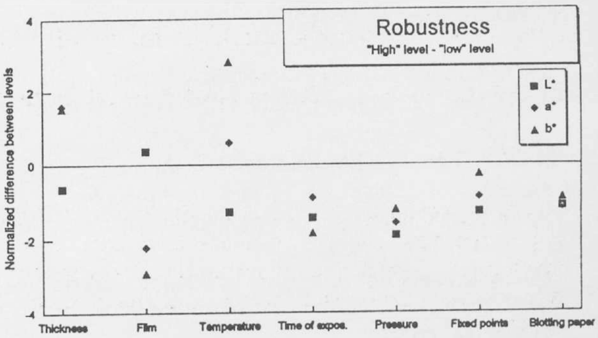

*Figure 1 The effect of 7 factors at L*a*b* –values*

The colour measurements are indicated in the L*a*b* system. The figure shows the normalized differences between average colour value by "Low" and "High" level for each of the seven factors. The results are based on 12 repetitions of each experiment.

The investigation had immediately the following consequences for the measuring procedure: The film – which is put on the surface of the meat during oxygenation – should be removed before measurement, and the temperature must be standardised.

Example 3: Evaluation of the level of the accuracy

The choice of instrument may be between a cheap instrument with inferior accuracy, and an expensive one with good accuracy. It is difficult to argue in favour of the expensive instrument based on reliability alone. The following – realistic – examples show how to use Bayes's decision theory. A random loin belongs to one of three categories which are determined from the weight of the longissimus dorsi muscle (LD). The price of a loin depends on LD's proportion of the loin, and apart from that the price is relative to the size. The application possibilities for loins are: export to Japan, bacon or unmanufactured loin. The production planning principle is that a loin is used for the most profitable end product. It is possible to predict the size of LD from the calculated percentage of lean meat in the carcass before cutting. By means of Bayes's equation $P(\theta_i| z_j) = P(\theta_i) P(z_j| \theta_i)/P(z_j)$ the probability of a loin to belong to each of the quality categories seen from a prediction point of view is calculated – and the average, expected price of sorting category can be calculated. In an experiment the actual category of quality is determined, and *the value of complete information* is calculated as well. Before that the carcass is measured with different instruments – giving information about the weight of the LD – and doing so it is possible to determine *the value of different accuracies.*

Table 1 Value of information

Level of information	*Error of prediction*	*Value, DKK/carcass*
Complete information	0 g	7.48
Max. information about measurements on the slaughter line	± 2 × 206 g	2.06
Information about meat% (high level of accuracy) and carcass weight	± 2 × 227 g	1.80
Information about meat% (low level of accuracy) and carcass weight	± 2 × 249 g	1.55

By improving the error of prediction from 249 g til 206 g, a selection can increase the value per carcass with DKK 0.51. But the value might increase by another DKK 5.42/carcass, if the error of prediction for information by measurement could be improved to a value close to 0.

Example 4: Evaluation of the precision of the reference

The *reliability* defined by Cronbach's coefficient $\alpha_N = \sigma^2_D /[\sigma^2_D + (\sigma^2/N)]$, is an adequate way to assess a sensory panel with N members. Further, the *individual reliability* can be assessed by $\alpha_{(i)} = \rho^2_{obs}/\alpha_{N-1}$, where ρ_{obs} is the observed correlation between the specific member *i* and the total score of the panel, and α_{N-1} is the reliability of the panel without member *i*. σ is the repeatability of an average panellist. As an example a test of new members for a sensory panel regarding evaluation of meat colour resulted in the following:

New member No. A: $\alpha_{N-1} = 0.95$, $\rho_{obs} = 0.85$, $\alpha_{(i)} = 0.76$ New member No. B: $\alpha_{N-1} = 0.96$, $\rho_{obs} = 0.69$, $\alpha_{(i)} = 0.51$

The reliability of the panel with all members is α_N=0.96. The reliability of the panel is unchanged by omission of person B, who does not contribute positively to the panel. It corresponds with a low individual reliability for B. Tests have shown that – with this reliability – the precision contributes approx. 15% of the residual variation when calibrating an objective colour measuring equipment.

References

International Standard ISO 5725 (1994), Accuracy (trueness and precision) of measurement methods and results.

Carmines, E.G., and Zeller, R.A. (1979), Reliability and Validity Assessment

Jannik Godt (1995), Definition, measurement and value of the colour of pork (28 December, 1995, 03463 (in Danish)