INTERNATIONAL STANDARD

ISO
21543

IDF 201

Second edition
2020-10

Milk and milk products — Guidelines for the application of near infrared spectrometry

Lait et produits laitiers — Lignes directrices pour l'application de la spectrométrie dans le proche infrarouge

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 21543:2020
https://standards.iteh.ai/catalog/standards/sist/bf845e6a-2e35-462b-9080-
f28358157506/iso-21543-2020

Contents

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products*, and the International Dairy Federation (IDF). It is being published jointly by ISO and IDF.

This second edition cancels and replaces the first edition (ISO 21543 | IDF 201:2006), which has been technically revised. The main changes compared with the previous edition are as follows:

— the measurement principles "Transmittance" and "Transflectance" have been added and defined;

— all sample types have been covered: liquids, solids and semi-solids;

— the calibration and validation sections have been reviewed and updated;

— the outlier section has been revised and the plots renewed;

— the procedures for sample handling and measurement have been expanded to include liquid samples and other examples;

— Annex A has been expanded to include raw milk analysis references.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

IDF (the International Dairy Federation) is a non-profit private sector organization representing the interests of various stakeholders in dairying at the global level. IDF members are organized in National Committees, which are national associations composed of representatives of dairy-related national interest groups including dairy farmers, dairy processing industry, dairy suppliers, academics and governments/food control authorities.

ISO and IDF collaborate closely on all matters of standardization relating to methods of analysis and sampling for milk and milk products. Since 2001, ISO and IDF jointly publish their International Standards using the logos and reference numbers of both organizations.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. IDF shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

This document was prepared by the IDF *Standing Committee on Statistics and Automation* and ISO Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products*. It is being published jointly by IDF and ISO.

The work was carried out by the IDF/ISO Action Team (S16) of the *Standing Committee on Statistics and Automation* under the aegis of its project leader, Mr A. Niemoeller (DE).

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 21543:2020
https://standards.iteh.ai/catalog/standards/sist/bf845e6a-2e35-462b-9080-
f28358157506/iso-21543-2020

Milk and milk products — Guidelines for the application of near infrared spectrometry

1 Scope

This document gives guidelines for the use of near infrared (NIR) spectrometry in the analysis of milk and milk products in liquid, semi-solid or solid form. Depending on the sample form and application, different instrument setups for transmittance, diffuse reflectance or transflectance can be applied.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
near infrared spectroscopy
NIR spectroscopy
spectroscopic method for measuring spectra using the NIR region of the electromagnetic spectrum (from 780 nm to 2 500 nm, 12 800 cm^{-1} to 4 000 cm^{-1})

3.2
near infrared spectrometry
NIR spectrometry
application of *near infrared spectroscopy* (3.1) yielding quantifiable results or other application-related evaluations of NIR spectra

3.3
parameter
constituent
proximate (e.g. total solids, moisture, non-fat solids, fat, protein, lactose and salt contents) or other property (e.g. pH value) of samples applicable for NIR analysis methods

Note 1 to entry: Most parameters are measured in mass fractions expressed in per cent, but some parameters will be determined in percentage by volume or with specific units, e.g. pH and freezing point depression.

4 Principle

4.1 General

A NIR method contains the following steps:

— a sample is presented to a NIR instrument;

— if required, it is preprocessed in advance to obtain a homogeneous test sample representing the chemical composition of the sample material;

— it is loaded into the sample holder of the NIR spectrometer and, during the spectroscopic measurement, the absorbance at wavelengths in the NIR region is measured;

— the spectral data are transformed to constituent concentrations by calibration models developed on samples from the population to be tested (see 7.1).

4.2   Measurement principles

4.2.1   Transmittance (transmission)

In transmittance measurements, light is directed at a sample with a focused or parallel beam. The transmitted part is detected after passing through the sample. This setup is not only used for clear liquids, but also scattering samples in liquid form (e.g. milk, cream) or semi-solid form (e.g. yoghurt, cheese) can be analysed by diffuse transmittance with a suitable optical setup. The optical path length is an important parameter and should be optimized and maintained according to the manufacturer's recommendations.

4.2.2   Diffuse reflectance (reflection)

When light is directed onto a sample, it is reflected and transmitted at differing amounts depending on the bulk properties of the material. Solid and semi-solid samples reflect light in diffuse reflection, i.e. reflecting in all angles equally or near-equally distributed. There are different instrumental setups used to collect and detect diffuse reflected light. Instruments with two principles for illuminating the sample in a container and collecting the reflected light either from above (top-down set-up) or from the bottom through a window (bottom-up) are common.

4.2.3   Transflectance (transflection)

When partially transparent samples are analysed, the reflectance and transmittance properties are not sufficient for using one of these dedicated measuring principles. By selecting a suitable path length and having a reflecting surface (reflector, mirror) on the other side of the sample, the incident light is transmitted through the sample, reflected and transmitted backwards to be detected. Thus, the diffuse reflected and the transmitted light is measured in a ration depending on the optical properties of the sample.

In cases where an NIR instrument provides reflectance measurements, even transparent or turbid liquids can be analysed in transflectance.

5   Reagents

Use only reagents of recognized analytical grade, unless otherwise specified, and distilled or demineralized water or water of equivalent purity.

5.1   Ethanol, or another appropriate solvent or detergent mixture, for cleaning re-usable sample containers.

5.2   Cleaning solutions.

For analysis of liquids in flow cells, appropriate cleaning solutions and maintenance protocols should be used according to the manufacturer's recommendations.

6 Apparatus

6.1 Near infrared (NIR) instrument.

NIR instruments are based on diffuse reflectance, transmittance or transflectance measurement principle by NIR spectroscopy in the whole NIR wavelength region of 780 nm to 2 500 nm (12 800 cm^{-1} to 4 000 cm^{-1}) or segments of this or at selected wavelengths.

The optical operation principle may be dispersive (e.g. grating monochromators), interferometry or non-thermal (e.g. light-emitting diodes, laser diodes, lasers). The instrument should be provided with a diagnostic test system, e.g. for testing photometric instrument noise, wavelength/wavenumber accuracy and wavelength/wavenumber precision/repeatability (for scanning spectrophotometers) according to the manufacturer's recommendations.

The instrument should be equipped with sample containers (cups, glass or polystyrene Petri dishes, flow cells, etc.), which allow repeatable filling of the sample. It is important that the window, Petri dish or flow cell material that allows the NIR light to interact with the sample is manufactured to a consistent standard by the manufacturer or supplier to avoid biased results or even re-calibration work. A sufficient sample volume or surface is required to eliminate any influence from inhomogeneity derived from the chemical composition or physical properties of the test sample. The sample path length (sample thickness) in transmittance measurements should be optimized according to the manufacturer's recommendations with respect to signal intensity for obtaining linearity and maximum signal/noise ratio. In reflectance measurements in a top-down instrument set-up, a quartz window or other appropriate material to eliminate drying effects should preferably cover the interacting sample surface layer. In a bottom-up instrument setup, the container and the sample are covered and protected against drying by the measurement window.

The sample containers may be re-usable or made of disposable material.

6.2 Grinding or grating device.

An appropriate device for preparing the sample (e.g. a food processor for semi-hard cheese) can be required. A procedure should be established to ensure consistency over time and with different operators.

Changes in grinding or grating conditions may influence the NIR measurements and results.

7 Calibration and initial validation

7.1 Selection of calibration samples

The instrument should be calibrated before being used or installed pre-calibrated. Because of the complex nature of NIR spectral data, which consists mainly of overtones and combination bands of fundamental vibrations in the mid-infrared region, the instrument should be calibrated using a series of natural samples (often at least 120 samples). No specific procedure for the calibration of a specific application can be given due to the wide variations over all aspects in sample types, constituents, concentration ranges, instrument setups and possible calibration algorithms. Examples and results are published in an IDF bulletin[1].

The accuracy and robustness of calibration models are dependent on the strategies used for sample selection and calibration. Developed calibration models are only valid for samples covered by the range and variations of the calibration samples. The first step in calibration development is therefore to define the application (e.g. sample types and concentration ranges). When calibration samples are selected,

care should be taken to ensure that all major factors affecting the accuracy of calibration are covered within the limits of the defined application area. These factors include the following:

a) combinations and composition ranges of major and minor sample components to include:

   1) range of analytes (e.g. total solids, fat and protein) and non-analytes; or

   2) recipes with dairy and non-dairy ingredients and additives (herbs, nuts, spices, chocolate, etc.);

b) seasonal, geographic and genetic effects on milk composition and other raw materials;

c) processing techniques and conditions;

d) ripening stages of cheeses;

e) storage and storage conditions;

f) sample and instrument temperatures and changes in them;

g) instrument variations.

The accuracy of calibrations is influenced by the extent of variation in the sample material and the analyte concentration range. A moderate variation can be easier to fit than a large variation but this depends on the analyte and sample matrices. If the required accuracy cannot be obtained by a single calibration, then stepwise evaluation is an option or the calibration data set should be split up into static or dynamic sub-areas, each with an associated calibration.

It is generally preferable that the whole calibration range is covered in a uniform way, with samples from low to high concentrations of analytes. The sample spread should also be as uniform as possible with respect to the other variables, including those mentioned above. In cases where there are gaps in some areas of the calibration range, all possible actions should be taken to fill these areas, e.g. mixing of samples or using samples from beginning or ending of production. Furthermore, the samples should be collected and measured over a certain period of time to ensure the inclusion of time-dependent effects. This design will improve the ruggedness and give a more even performance of the calibration over the entire analyte concentration range.

Multivariate methods[2][3] may be used as a tool in the selection of samples before the calibration process to ensure a homogeneous calibration set covering all variation in spectroscopic data induced by chemical, biological and physical factors without duplication of samples with similar information. In practice, a larger sample population is measured by NIR spectroscopy for collection of NIR data only. Then samples differing in spectral information are selected for reference analyses. Identification of differing samples may be obtained from inspection of score plots from principal component analysis (PCA) using, for example, the first three components. This may be less practical in the case of many samples. However, it is recommended always to perform a PCA and inspect score plots to obtain a visual overview of the sample set. It is important to check which spectral variations are influencing the loadings and therefore the PCA score plots, i.e. high variation in water or fat content can dominate the PCA results whereas protein variation could be lower but is of importance as well. More formal cluster analyses may be obtained using techniques based on distance measurements[3]. Further samples may be added over a period of time to this pool of selected samples using PCA score space or distance measurement to identify differing samples.

7.2 Reference analyses and NIR instruments

Internationally accepted reference methods for the determination of analytes should be used. The reference method used for calibration should be in statistical control, i.e. the variability should consist of a constant system of random variations. To support the assessment of outliers, it may be useful to perform replicate analyses in independent series (different analysts, different equipment, etc.).

All major variations in NIR measuring conditions that may appear in practice should be considered and built into the calibration model. An important factor is sample temperature, especially for transmittance measurements of aqueous liquids such as milk, etc.

The sampling procedure and sample presentation at the instrument as well as the actual test sample measured by NIR spectroscopy may be critical for the accuracy obtained[4]. The test sample volume or surface interacting in measurements should be large enough to avoid sample inhomogeneity having a significant influence. Reflectance measurement at higher wavelengths normally requires a larger sample surface than transmittance measurement at shorter wavelengths because the light penetration is much less. If possible with the NIR instrument, the optimal sample size should be determined from experiments where the prepared sample material (see 10.1) is measured repeatedly after the refilling of the sample container. Rotating of solid and semi-solid samples helps in acquiring an averaged NIR result for reflectance and diffuse transmittance.

The NIR measurements and reference analyses should preferably be performed on the same sample in order to eliminate effects related to sampling uncertainty. The sample can be divided for NIR measurements and the reference analyses, which should also be performed with a minimum time lag (preferably less than one day). It is good practice to randomize the order in which the samples are presented for both the reference analysis and NIR measurement.

The chemical and physical effects used for the analysis of a certain parameter shall be considered for the reference method and the corresponding NIR method which is calibrated by the reference values. The reference and NIR results may have a bias that is due to differing measurement principles.

7.3 Calibration

Because NIR instruments are applied utilizing different calibration systems, no specific procedure can be given for calibration. However, the person performing the calibration should be familiar with the statistical principles behind the calibration algorithm used.

There are commercially available calibration packages delivered with instruments as part of turn-key solutions. The same principles as described need to be considered when applying such calibrations. An initial validation step is required to check if the commercial methods are applicable for the samples analysed. In some cases, a need for a bias adjustment may occur and this will help to adapt to the current needs. Because the samples and the conditions can be specific an adaptation or expansion of calibrations might be required.

The calibration may be performed using different techniques, e.g. multiple linear regression (MLR), multivariate algorithms such as partial least-squares regression (PLS)[2], locally weighted regression (LWR)[5][6][7] or artificial neural networks (ANNs)[7][8]. The latter techniques are recommended if linearity problems between the spectral response and the constituent occur. Typically, at least 120 calibration samples are needed to obtain rugged calibrations with MLR and PLS. When ANNs are used for calibration, a substantially higher number of samples is required to avoid over-fitting of data because ANNs are very flexible functions with many parameters that shall be determined. Three different data sets (calibration, test or training and validation set) are normally required for determining the architecture, fitting the parameters and validating the ANN calibration. The concept of LWR also requires a considerably larger database from which local calibration samples can be selected for the prediction step.

Spectra should normally be preprocessed prior to calibration to remove or reduce the weighting of effects which are not related to the chemical absorption of light. Frequently used pretreatments are multiplicative scatter correction (MSC)[9], standard normal variate (SNV)[10], de-trending[10] and first or second derivatives[3]. The optimal transformation and other pretreatments of spectra (e.g. smoothing) should be determined by trials. Several techniques may often give equivalent results. The optimal techniques should be assessed from cross validation or independent test sets where models are subsequently developed on parts of the data and tested on other parts[11]. The preprocessing step is important to ensure the robustness of a calibration and one aspect to be checked and considered for the validation process.

An important issue is the selection of the optimal number of variables (in MLR) or factors (in multivariate calibrations). If too few variables or factors are used, an under-fitted model is obtained, which means that the model is not covering enough of important and analyte related variability in the data. If too many variables or factors are used, an over-fitted model may be obtained where noise and minor or

irrelevant variations are included in the model. Both cases can result in poor predictions on future samples. The optimal number can be determined by plotting a prediction error [root mean square error of prediction (RMSEP), see 8.2, obtained from an independent test set, or root mean square error of cross validation (RMSECV), see 8.3] versus the number of variables or factors (see Figure B.1). Typically, RMSEP or RMSECV are high for small numbers of variables or factors and decreases as the number increases, before it increases again when the number becomes too high. Generally, the best solution is the one giving the lowest RMSEP or RMSECV with the fewest variables or factors.

The predicted values obtained by cross or test set validation should be plotted against reference results (see Figure B.2). The plot should be examined for outliers and also be investigated for regions with different levels of prediction accuracy, random or systematic, which may indicate the need for more calibration samples or a segmentation of the calibration range.

7.4 Outliers in calibration

7.4.1 General

Outliers may be related to NIR spectra (x-outliers) or errors in reference data or samples with a different relationship between reference data and NIR spectra (y-outliers).

7.4.2 x-outliers

A homogeneous calibration set of spectrally similar samples is required for a robust predictive model. This can also form the basis of an outlier warning system. Any x-outliers should thus be removed before calibration. The scores of, for example, the five first PCA or PLS factors can be useful to reveal x-outliers either globally outside the population or falling in a gap in the score spaces. A more formal identification of outliers may be performed using, for example, the principle of Mahalanobis distance applied on PCA or PLS reduced data[2][12] or the so-called leverage[13].

It is important to check whether such x-outlier candidates are due to extreme high or low concentrations or other reasons which can be related to their identification as outliers. In such a case, the sample is probably important and should be added to the calibration set to widen the range, thus incorporating a different process situation, other batches of ingredients or a changed matrix. In cases where a deviating spectrum is attributed to an insufficient sample preparation or presentation, it shall be removed.

Figure B.2 shows an example plot of NIR prediction versus reference without outliers. In the PCA scores plot in Figure B.3, an x-outlier is present.

7.4.3 y-outliers

When a y-outlier is observed in the calibration set, the reference data should be checked for errors in sample identification, reference analyses, computations, data transfer, etc. However, it may be difficult to relate outliers to errors in reference analyses because the calibration step usually is performed at a later stage than reference analyses, which may make it impossible to repeat analyses because of sample instability. There is no standard way to treat y-outliers, but outliers should generally be removed if the difference between NIR predictions and reference results in cross or test set validation exceeds three times the RMSECV or RMSEP (see Clause 8).

It is important to note that the removal of outliers can influence the future prediction of similar samples. Outliers should be removed as a batch before a new calibration model is created. The outlier removal step should only be performed one or two times in order not to reduce the coverage and robustness of the calibration and overestimate the accuracy. Care shall be taken to preserve the optimum distribution of the calibration set when outliers are removed.

Figure B.2 shows an example plot of NIR prediction versus reference without outliers. In Figure B.4, a y-outlier is present.