

INTERNATIONAL  
STANDARD

ISO  
21543

IDF  
201

First edition  
2006-09-01

---

---

**Milk products — Guidelines for the  
application of near infrared spectrometry**

*Produits laitiers — Lignes directrices pour l'application de la  
spectrométrie dans le proche infrarouge*

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[ISO 21543:2006](https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006)

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>



Reference numbers  
ISO 21543:2006(E)  
IDF 201:2006(E)

© ISO and IDF 2006

**PDF disclaimer**

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. Neither the ISO Central Secretariat nor the IDF accepts any liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies and IDF national committees. In the unlikely event that a problem relating to it is found, please inform the ISO Central Secretariat at the address given below.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

ISO 21543:2006

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>

© ISO and IDF 2006

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO or IDF at the respective address below.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

International Dairy Federation  
Diamant Building • Boulevard Auguste Reyers 80 • B-1030 Brussels  
Tel. + 32 2 733 98 88  
Fax + 32 2 733 04 13  
E-mail [info@fil-idf.org](mailto:info@fil-idf.org)  
Web [www.fil-idf.org](http://www.fil-idf.org)

Published in Switzerland

## Contents

Page

Foreword.....	iv
<b>1 Scope .....</b>	<b>1</b>
<b>2 Terms and definitions.....</b>	<b>1</b>
<b>3 Principle.....</b>	<b>1</b>
<b>4 Reagents .....</b>	<b>1</b>
<b>5 Apparatus .....</b>	<b>2</b>
<b>6 Calibration and initial validation .....</b>	<b>2</b>
6.1 Selection of calibration samples.....	2
6.2 Reference analyses and NIR measurements.....	3
6.3 Calibration .....	3
6.4 Outliers in calibration.....	4
6.5 Validation of calibration models .....	5
6.6 Changes in measuring and instrument conditions.....	6
6.7 Outlier detection .....	6
<b>7 Statistics for performance measurement.....</b>	<b>6</b>
7.1 Standard error of prediction (SEP) and bias.....	6
7.2 Root mean square error of prediction (RMSEP).....	7
7.3 Root mean square error of cross validation (RMSECV) .....	7
<b>8 Sampling.....</b>	<b>7</b>
<b>9 Procedure .....</b>	<b>8</b>
9.1 Preparation of test sample.....	8
9.2 Measurement.....	8
9.3 Evaluation of results.....	8
<b>10 Checking instrument stability .....</b>	<b>9</b>
10.1 Control sample.....	9
10.2 Instrument diagnostics .....	9
<b>11 Running performance check of calibration.....</b>	<b>9</b>
<b>12 Precision and accuracy.....</b>	<b>10</b>
12.1 Repeatability.....	10
12.2 Intralaboratory reproducibility .....	10
12.3 Accuracy.....	11
<b>13 Test report .....</b>	<b>11</b>
<b>Annex A (informative) Examples of SEP and RMSEP values .....</b>	<b>12</b>
<b>Annex B (informative) Examples of figures .....</b>	<b>14</b>
<b>Bibliography .....</b>	<b>22</b>

## 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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

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.

ISO 21543|IDF 201 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.

ITEH STANDARD PREVIEW  
(standards.iteh.ai)

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>

## Foreword

**IDF (the International Dairy Federation)** is a worldwide federation of the dairy sector with a National Committee in every member country. Every National Committee has the right to be represented on the IDF Standing Committees carrying out the technical work. IDF collaborates with ISO in the development of standard methods of analysis and sampling for milk and milk products.

Draft International Standards adopted by the Action Teams and Standing Committees are circulated to the National Committees for voting. Publication as an International Standard requires approval by at least 50 % of the IDF National Committees casting a vote.

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.

ISO 21543|IDF 201 was prepared by the International Dairy Federation (IDF) and Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products*. It is being published jointly by IDF and ISO.

All work was carried out by Joint ISO-IDF Action Team on *Automated methods*, of the Standing Committee on *Quality assurance, statistics of analytical data and sampling*, under the aegis of its project leader, Mr L.K. Sørensen (DK).

(standards.iteh.ai)

[ISO 21543:2006](https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006)

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

ISO 21543:2006

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>

# Milk products — Guidelines for the application of near infrared spectrometry

## 1 Scope

This International Standard provides guidance on use of near infrared spectrometry in the determination of

- the total solids, fat and protein contents in cheese,
- the moisture, fat, protein and lactose contents in dried milk, dried whey and dried butter milk, and
- the moisture, fat, non-fat solids and salt contents in butter.

## 2 Terms and definitions

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

### 2.1

**near infrared instrument**

ISO 21543:2006

**NIR instrument**

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347409512138-2006>

proprietary apparatus which, when used under the conditions defined in this International Standard, estimates the mass fractions of the substances specified in Clause 1

### 2.2

**total solids, moisture, non-fat solids, fat, protein, lactose and salt contents**

mass fraction of substances determined using the method specified in this International Standard

NOTE These contents are expressed as mass fractions in percent.

## 3 Principle

The sample is pretreated 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. The absorbance at wavelengths in the near infrared region is measured and the spectral data are transformed to constituent concentrations by calibration models developed on representative samples from the population to be tested.

## 4 Reagents

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

**4.1 Ethanol**, or other appropriate solvent or detergent mixture, for cleaning re-usable sample cups.

## 5 Apparatus

**5.1 Near-infrared (NIR) instrument**, based on diffuse reflectance or transmittance measurement in the whole near infrared wavelength region of 700 nm to 2 500 nm or segments of this or at selected wavelengths.

The optical operation principle may be dispersive (e.g. grating monochromators), interferometric or non-thermal (e.g. light-emitting diodes, laser diodes and lasers). The instrument should be provided with a diagnostic test system for testing photometric instrument noise, wavelength accuracy and wavelength precision (for scanning spectrophotometers). The wavelength accuracy should be better than 0,5 nm and the repeatability standard deviation better than 0,02 nm.

The instrument should be equipped with a sample holder, which allows measurement of a sufficiently large sample volume or surface to eliminate any significant influence of 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, a quartz window or other appropriate material to eliminate drying effects should preferably cover the interacting sample surface layer.

The sample cup (cuvette) may be re-usable or made of disposable material.

**5.2 Grinding or grating device**, appropriate for preparing the sample (e.g. a food processor for semi-hard cheese).

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

ITeH STANDARD PREVIEW  
(standards.iteh.ai)

## 6 Calibration and initial validation

ISO 21543:2006

### 6.1 Selection of calibration samples

<https://standards.iteh.ai/catalog/standards/sist/1b4c1c96-8deb-4b3f-b66e-347fa86b572f/iso-21543-2006>

The instrument should be calibrated before being used. Because of the complex nature of near infrared spectral data, which are mainly 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).

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 domain 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: analytes (e.g. total solids, fat and protein) and non-analytes;
- b) seasonal, geographic and genetic effects on milk composition;
- c) processing techniques and conditions;
- d) ripening stages of cheeses;
- e) storage and storage conditions.



The accuracy of calibration is influenced by the extent of variation in the sample material and the analyte concentration range. A moderate variation is usually easier to fit than a large variation. If the required accuracy cannot be obtained by a single calibration, then the application area should be split up into static or dynamic sub-areas, each with an associated calibration, in order to fulfil the requirements. Dynamic sub-areas are used in locally weighted regression algorithms where calibration samples close in spectral space to the actual prediction sample are selected from a larger population to create a local calibration equation.

It is generally preferable that the whole calibration range be 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. Furthermore, the samples should be collected and measured over a certain period of time to ensure 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 [1], [2] may be used as a tool in the selection of samples 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. More formal cluster analyses may be obtained using techniques based on distance measurements [2]. Further samples may be added over a period of time to this pool of selected samples using PCA space or distance measurement to identify differing samples.

## iTeh STANDARD PREVIEW (standards.iteh.ai)

### 6.2 Reference analyses and NIR measurements

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 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 built into the calibration model. An important factor is sample temperature.

The sampling procedure used and the sample size measured by NIR spectroscopy may be critical for the accuracy obtained [3]. 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. The optimal sample size should be determined from experiments where the prepared sample material (see 9.1) is measured repeatedly after repacking of the sample cup.

Special care should be taken to avoid surface drying effects, particularly in reflectance measurements.

The NIR measurements and reference analyses should preferably be performed on the same test sample in order to eliminate effects related to sampling uncertainty. The NIR measurements and the initiation of reference analyses 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.

### 6.3 Calibration

Because NIR instruments apply 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.

The calibration may be performed using different techniques [e.g. multiple linear regression (MLR), multivariate algorithms such as partial least-squares regression (PLS), locally weighted regression (LWR) or artificial neural networks (ANN)]. 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 ANN are used for calibration, a substantially higher number of samples is required to avoid over fitting of data because ANN are very flexible functions with many parameters to be determined. Three different data sets are normally required for determining the architecture, fitting the parameters and validating the network. The concept of LWR also requires a considerably larger database from which local calibration samples can be selected.

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. Often used treatments are multiplicative scatter correction (MSC) [4], standard normal variate (SNV) [5], de-trending [5] and first or second derivatives [2]. The optimal transformation and other pretreatments of spectra (e.g. smoothing) should be determined from trials. Several techniques often give equivalent results. The optimal techniques should be assessed from cross validation where models are subsequently developed on parts of the data and tested on other parts [6]. Additional information may be obtained from testing on an independent test set.

An important issue is 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 solution is obtained, which means that the model is not large enough to capture the important variability in the data. If too many variables or factors are used, an over-fitted solution may be obtained where much of the redundancy in the NIR data is modelled. Both cases can result in poor predictions on future samples. The optimal number can be determined by plotting RMSECV (see Clause 7) obtained from cross validation or RMSEP (see Clause 7) obtained from an independent test set versus the number of variables or factors (Figure B.1). Typically RMSECV (RMSEP) is large for small numbers of factors and decreases as the number increases, before it increases again when the number becomes too large. Generally, the best solution is the one giving the lowest RMSECV (RMSEP) with the fewest variables or factors.

The reference results should be plotted against predicted values obtained by cross validation. The plot should be examined for outliers. The plot should 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 region.

## 6.4 Outliers in calibration

### 6.4.1 General

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

### 6.4.2 $x$ -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 projections of the five first PCA axes can be useful to reveal  $x$ -outliers either globally outside the population or falling in a gap in the PCA space. A more formal identification of outliers may be performed using, for example, the principle of Mahalanobis distance applied on PCA reduced data [7] or the so-called leverage [8].

Figure B.2 shows a case from practice without outliers. In Figure B.3, an  $x$ -outlier is present.

### 6.4.3 $y$ -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 has to be performed at a later

stage than reference analyses, which may make it impossible to repeat analyses because of sample instability. There is no correct way to treat  $y$ -outliers, but outliers should generally be removed if the difference between NIR and reference results in cross validation exceeds three times the RMSECV (see Clause 7).

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 robustness of the calibration and overestimate the accuracy. Care should be taken to preserve the optimum distribution of the calibration set when outliers are removed.

Figure B.2 shows a case from practice without outliers. In Figure B.4, a  $y$ -outlier is present.

#### 6.4.4 Combined $x$ - and $y$ -outliers

Samples which are both  $x$ - and  $y$ -outliers (influential outliers) have a very strong effect on the regression equation and can be very harmful. Such outliers may give slope effects and increase the prediction error considerably.

Figure B.2 shows a case from practice without outliers. In Figures B.5 and B.6, a sample is present which is both an  $x$ - and a  $y$ -outlier.

### 6.5 Validation of calibration models

When calibration equations have been developed, they should be validated on an independent test set, preferably sampled after the calibration period. The test set should cover all variations in the sample population and should contain at least 25 samples. The use of cross validation in the calibration process, where subsequent parts of the calibration set are reserved for validation, can give a good estimate of the uncertainty of the method when the calibration samples are properly selected.

However, the potential risk is that cross validation may underestimate the ruggedness of the calibration and the predicted uncertainty because cross validation samples are taken from the pool of samples used for calibration.

The results obtained on the independent test set are plotted, reference against NIR and residuals against reference, to give a visual impression of the performance of the calibration. The SEP is calculated (see Clause 7) and the residual plot of data corrected for mean systematic error (bias) is examined for outliers; i.e. samples with a residual exceeding  $3 \times \text{SEP}$ . If an outlier occurs and this cannot be classified as an  $x$ -outlier and re-analysis of the sample by NIR and reference methods confirms the result, the outlier should not be removed.

In this case, the ruggedness of the calibration is not sufficient and the calibration set should be expanded. The next step is to fit NIR and reference data by linear regression (reference =  $b \times \text{NIR} + a$ ) to support the visual impression. If the slope ( $b$ ) is significantly different from 1, the calibration is skewed. Adjusting the slope of the calibration is generally not recommended. If a re-investigation of the calibration does not detect outliers, especially influential outliers, it is preferable to expand the calibration set to include more samples. However, if the slope is adjusted, the calibration should be tested on a new independent test set. The data are also examined for a bias between the methods. An intercept ( $a$ ) significantly different from 0 indicates that the calibration is biased. A bias may be removed by adjusting the constant term in the calibration equation. However, if the accuracy of the bias-adjusted calibration is significantly poorer than expected from cross validation on the calibration set, i.e. SEP is significantly larger than RMSECV, the calibration set should be expanded to include more samples. In all cases when a new calibration is developed on an expanded calibration set, the validation process should be repeated on a new independent test set. If necessary, expansion of the calibration set should be repeated until acceptable results are obtained on an independent test set.

## 6.6 Changes in measuring and instrument conditions

Unless additional validation is performed, a local validation of an NIR method stating the accuracy of the method may generally not be considered valid if the test conditions are changed.

For example, the calibrations developed for a certain population of samples may not be valid for samples outside this population, although the analyte concentration range is unchanged. A calibration developed on cheeses from one dairy may not give the same accuracy on cheeses produced in another dairy if the processing and ripening parameters are different.

Changes in the sample presentation technique or the measuring conditions (e.g. temperature) not included in the calibration set may also influence the analytical results.

Furthermore, calibrations developed on a certain instrument cannot always be transferred directly to an identical instrument operating under the same principle. It may be necessary to perform bias and slope adjustments to calibration equations. In some cases, it may even be necessary to standardize the two instruments against each other by mathematical procedures before calibration equations can be transferred [2]. Standardization procedures may be used to transfer calibrations between instruments of different types provided that samples are measured in the same way (reflectance, transmittance) in similar cups and that most of the spectral region is common. Adding a few samples scanned with the second instrument in the database can contribute to the transfer.

If the conditions are changed, a supplementary validation should be performed.

The calibrations should be checked whenever any major part of the instrument (optical system, detector) has been changed or repaired.

ITeH STANDARD PREVIEW  
(standards.iteh.ai)

## 6.7 Outlier detection

ISO 21543:2006

Use of NIR methods is generally limited to samples in the population covered by the calibration set with respect to sample material characteristics and analyte concentration. An outlier detection system should accompany the NIR method to reduce the risk of unintentional use of NIR spectrometry on samples outside this population. The system should be able to detect  $x$ -outliers and samples falling outside the concentration range. The principle of Mahalanobis distance [7] applied to PCA-reduced data or "leverage" [8] may be used for the detection of  $x$ -outliers. If a sample is detected as an outlier, the sample should be re-analysed by reference methods to obtain the final result.

## 7 Statistics for performance measurement

### 7.1 Standard error of prediction (SEP) and bias

The standard error of prediction (SEP), which expresses the accuracy of routine results corrected for the mean difference between routine and reference methods (bias), can be calculated by using the following equation:

$$SEP = \sqrt{\frac{1}{N-1} \sum (x_i - y_i - B)^2}$$