

ICS 65.120

CCS B 20

NY

中华人民共和国农业行业标准

NY/T ××××-202×

饲料的近红外光谱分析应用指南

Guidelines for the application of near infrared spectrometry for feed

(ISO 12099:2017, Animal feeding stuffs, cereals and milled cereal products—Guidelines for the application of near infrared spectrometry, MOD)

(公开征求意见稿)

202×-××-××发布

202×-××-××实施

中华人民共和国农业农村部 发布

目 次

前言.....	III
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 原理.....	1
5 仪器设备.....	2
5.1 近红外光谱分析仪.....	2
5.2 样品制备设备.....	2
6 样品.....	2
6.1 总则.....	2
6.2 定标样品和验证样品.....	2
6.3 试样.....	2
6.4 质控样品.....	3
6.5 监控样品.....	3
7 定标模型的建立与验证.....	3
7.1 总则.....	3
7.2 参考值.....	3
7.3 异常值.....	3
7.4 定标模型的验证.....	3
7.5 检测条件和仪器状况的变化.....	4
8 定标模型性能统计学评价.....	5
8.1 总则.....	5
8.2 结果散点图.....	5
8.3 偏差.....	6
8.4 预测误差均方根 (s_{RMSEP}).....	6
8.5 预测标准误差 (s_{SEP}).....	7
8.6 斜率.....	8
9 试样的测定.....	9
9.1 试样近红外光谱的测量.....	9
9.2 试样近红外光谱预测值的测定.....	9
9.3 试样近红外光谱预测结果的表述.....	9
10 仪器稳定性检查.....	9
10.1 用质控样品检查仪器稳定性.....	9
10.2 仪器诊断.....	10
10.3 仪器联网.....	10
11 定标模型运行性能监控.....	10
11.1 总则.....	10
11.2 参考值和近红外预测值差异控制图.....	10
12 精密度和准确度.....	11
12.1 重复性.....	11

12.2	重现性.....	11
12.3	准确度.....	11
12.4	不确定度.....	11
13	检测报告.....	11
附录 A (资料性)	具体成分或技术指标的近红外分析标准制定指南.....	13
附录 B (资料性)	异常值和控制图示例.....	14
附录 C (资料性)	补充性条款及定义.....	18
附录 D (资料性)	本文件与 ISO 12099:2017 相比的结构变化情况.....	23
附录 E (资料性)	本文件与 ISO 12099:2017 的技术差异及其原因.....	24
	参考文献.....	27

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件使用重新起草法修改采用ISO 12099:2017《动物饲料、谷物及谷物精制料的近红外光谱分析应用指南》。

本文件与ISO 12099:2017相比在结构上有较多调整，附录D中列出了本标准与ISO 12099:2017的章条编号对照一览表。

本文件与ISO 12099:2017相比存在技术性差异，这些差异涉及的条款已通过在其外侧页边空白位置的垂直单线（|）进行了标示，附录E中给出了相应技术性差异及其原因的一览表。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中华人民共和国农业农村部畜牧兽医局提出。

本文件由全国饲料工业标准化技术委员会归口。

本文件起草单位：

本文件主要起草人：

饲料近红外光谱分析应用指南

1 范围

本文件规定了饲料成分如水分、粗脂肪、粗蛋白、淀粉、粗纤维含量以及消化率等技术指标的近红外光谱分析应用指南。

本文件适用于可采用近红外光谱分析（见附录 C.1.3）的饲料。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 10647 饲料工业术语

GB/T 14699.1 饲料 采样

GB/T 29858 分子光谱多元校正定量分析通则

3 术语和定义

GB/T 10647 和 GB/T 29858 界定的以及下列术语与定义适用于本文件。

3.1

成分含量 **constituent content**

用适当的、标准化的或验证过的化学方法检测的饲料成分的含量。

示例：水分、粗脂肪、粗蛋白、淀粉、粗纤维、中性洗涤纤维及酸性洗涤纤维。

3.2

技术指标 **technological parameter**

用适当的、标准化的或验证过的方法检测的饲料属性或功能指标。

示例：消化率。

4 原理

采集近红外光谱（NIRS）数据，并利用以所分析产品的代表性样本为基础开发出的定标模型，给出所分析产品的成分含量和技术指标值。

5 仪器设备

5.1 近红外光谱分析仪

仪器基于近红外光谱区的漫反射（见附录 C.1.4）或/和透射（见附录 C.1.5）进行检测，其光谱区包括：近红外全谱区 770 nm-2500 nm（12900 cm^{-1} -4000 cm^{-1} ），或全谱区内的某段光谱范围，或是选择的波长或波数。光学原理可是色散型（如：光栅单色仪），干涉型或非热型（如：发光二极管、激光二极管以及激光）等。仪器应具备自我诊断系统，用于检测仪器的噪音、重现性、波长/波数准确度和波长/波数精密度（对扫描型光谱仪）。

仪器应采集足够样品量（包括样品厚度和样品表面积）的光谱信息，以消除待检样品化学组成和物理性质不均一的影响。对于透射扫描方式，应依据仪器制造商关于信号强度的推荐值进行光程（样品厚度）的优化，以获得最佳的线性和最大的信噪比。

5.2 样品制备设备

如需对样品进行粉碎等预处理，应选择适当的研磨或粉碎设备。

注：研磨或粉碎条件的改变会影响近红外光谱的测量，例如，发热会导致像水分这类挥发性成分发生变化。

6 样品

6.1 总则

样品包括对近红外仪器进行定标的定标样品、对定标模型进行验证的验证样品、待测的试样、对仪器进行稳定性检查的质控样品、和对定标模型运行性能进行日常监控的监控样品。

样品的采集可参照 GB/T 14699.1。

应取用对所分析成分具有代表性的样品。样品在运输和贮存过程中不应受到损坏或发生变化。应确保样品从采样到进行分析的整个过程中其成分含量保持不变。

6.2 定标样品和验证样品

对于定标样品和验证样品应有足够多的代表性样品，并且覆盖样品的各种影响因素，例如：

- a) 应综合考虑样品的主要成分与次要成分及其成分含量范围；
- b) 对饲草和饲料原料样品应考虑季节、地域和遗传等因素的影响；
- c) 应考虑加工技术和条件的不同；
- d) 应考虑不同的储存条件；
- e) 应考虑样品和仪器温度变化；
- f) 应考虑仪器变化（例如：仪器间差异）。

定标样品和验证样品可从同一样品池中，按照轮流或取第 n 个样品（按分析成分含量排序）的方法选择。

6.3 试样

试样的制备应与验证样品的制备保持一致。

6.4 质控样品

质控样品应能够保持长时间的稳定，并尽可能与所分析样品类似。质控样品的分析参数同样需要保持稳定，并等同于或至少从生物化学性质的角度尽可能的接近样品分析物。质控样品应以能够使其存储效期最大化的方式进行存储，且其稳定性应根据实际情况予以检测。

6.5 监控样品

定标运行性能监控样品应从待测定的试样中随机选取。宜采用一些采样策略以确保样品在整个定标范围内分布的均匀性，例如，按含量高低进行分段然后在每段中随机选择样品，或确保样品能够涵盖常规范围。

7 定标模型的建立与验证

7.1 总则

在进行试样的测定之前，近红外光谱分析仪应预先定标，即：采用定标方法（见附录 C.2）将样品的近红外光谱吸光度值与样品的成分含量/技术指标的参考值进行关联，建立定标模型。对近红外仪器的定标有很多不同的方法，在此不推荐具体的定标方法。关于定标方法的选择可参考文献[1]、GB/T 29858 或所使用仪器所附的使用手册。

7.2 参考值

样品水分、粗脂肪、粗蛋白、淀粉、粗纤维等成分含量以及消化率等技术指标的参考值，应使用标准方法或国际公认的参考方法（见附录 C.1.1）进行赋值，参考值的精度应在参考方法要求的精度范围内，并应记录参考值的具体精度值。

7.3 异常值

在定标和验证时可能会出现统计异常值（见附录 C.2.7）。异常值的出现可能是光谱数据异常（光谱异常值，简称为-异常值）（见附录 C.2.8）或参考值存在误差或参考值与近红外光谱回归关系异常（简称为-异常值）（见附录 C.2.9），示例见附录 B 中的图 B.1-B.5。

判断一个样品是否异常，应首先判断它是不是-异常值，如果超出了定标确定的-异常值限值，则该样品需要被剔除。如果判断不是-异常值，则需要检查参考值和近红外预测值是不是异常，例如通过重复分析检查：如果参考值和近红外预测值确证无误，则该样品应保留并参与验证的统计分析；如果重复测量值证明原始参考值或近红外预测值有误，那么该样品应使用新测定值进行赋值。

为了达到定标模型验证的目的，下列情况不能作为异常值：

- a) 样品成分含量/技术指标在定标样品含量范围内；
- b) 样品光谱在定标样品光谱的变异范围内，例如可以用马氏距离（见附录 C.2.11）判断；
- c) 样品光谱残差（见附录 C.2.14）在定标样品确定的残差范围内；
- d) 样品预测残差（见附录 C.2.13）在定标样品确定的残差范围内。

7.4 定标模型的验证

7.4.1 总则

定标模型在分析试样之前应采用具有广泛代表性的独立验证样品进行验证。验证至少需要 20 个样

品以确定偏差、斜率和预测标准误差（见 8.5）。验证应充分考虑样品种类、成分含量/技术指标的范围、温度和其它已知有影响或预计有影响的因素，只有验证样品包含这些影响因素时才能证明定标的有效性。

使用线性回归（）拟合近红外预测值（）和参考值（）形成描述验证结果的统计指标。

如果验证过程显示定标模型不能获得可接受的统计值，那么该定标模型就无法使用。

注：什么是可接受的统计值由建立定标模型者决定，例如：可取决于参考值的精度、覆盖的含量范围和分析的目的等。

独立验证样品的参考值和近红外预测值的散点图和与残差（见附录 C.2.13）的散点图可对定标模型的效果进行直观显示。使用偏差（bias）矫正后的残差散点图和计算出的（见 8.5）可检验-异常值，如：残差大于 ± 3 的即可判断为异常值。

7.4.2 偏差校正（截距调整）

近红外预测值和参考值之间的偏差应被检查。近红外预测值和参考值的平均值的差值与 0 有显著性差异表明定标存在偏差，应通过调整定标模型的常数项（截距（见附录 C.3.11））来消除偏差（见 8.3）的影响。

7.4.3 斜率调整

斜率（见附录 C.3.10）与 1 存在显著性差异说明定标出现偏斜。

通常不推荐对斜率进行调整，除非定标模型应用于新样品类型或新仪器。如果对定标模型的验证没有发现异常值，尤其是没有发现高杠杆值（见 C.2.10）的异常样品，应扩充现有定标样品使之包含更多有代表性的样品。但是，斜率一旦被调整，那么定标模型应使用独立验证样品进行重新验证。

含量范围太窄或分布不均匀都会导致斜率校正无用。仅当验证样品覆盖了定标样品大部分含量范围时，方可对斜率进行调整。

7.4.4 定标样品的扩充

在定标模型的准确度未达到预期效果时，应扩充定标样品，或建立新的定标模型。定标样品扩展后建立的新模型应采用新验证样品进行重新验证，直至验证所获得的结果满足要求为止。

7.5 检测条件和仪器状况的变化

检测条件改变定标模型可能失效，应根据新的检测条件进行重新定标。

例如，尽管成分含量的范围一致，为某个样品集所开发的定标模型对该样品集之外的样品仍可能是无效的。例如，遗传因素、生长条件和加工工艺参数等不同，针对某个地区的青贮饲草建立的定标模型不能对其它地区的样品给出相同准确度的预测。

未在定标样品中包含的样品制备技术或测量条件的改变同样会影响分析结果，例如样品研磨粒度、温度等。

在某一仪器上开发的定标模型并非总是可以直接转移到另一台相同原理、相同型号的仪器上。定标模型转移之前，必须对定标模型的斜率/截距进行相应调整。在多数情况下，在进行定标模型的转移之前，必须进行两台仪器之间的标准化（见附录 C.1.6，参考文献[1]）。经过标准化后不同类型仪器之间可以进行定标模型转移，但要求样品检测采用相同方式（反射、透射），且波长范围拥有重叠波段。

检测条件和仪器状况的改变，都应对定标模型进行重新验证。

仪器任何主要部件（光学系统、检测器）更换或维修时，也应对定标模型进行重新验证。

8 定标模型性能统计学评价

8.1 总则

定标模型的性能应使用验证样品来检验。验证样品应独立于定标样品。

验证样品应该严格地采用参考方法进行参考值赋值，控制验证样品参考值的精度非常重要，甚至比控制定标样品参考值的精度更重要。

为便于进行可信的数据统计，验证样品数量应不少于 20 个。

对定标模型性能的评价所用的 NIRS 检测方案应与试样测定时所用的方案相同（一次分析或两次分析）。

8.2 结果散点图

可视化绘图是定标性能统计学评价的重要方法之一。例如参考值-近红外预测值或残差-近红外预测值的散点图。

残差按式（1）计算：

$$\dots\dots\dots (1)$$

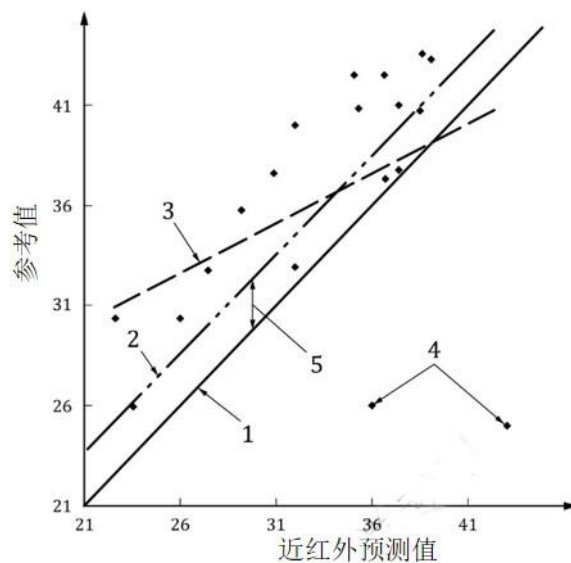
式中：

——第个参考值（）；

——第个近红外预测值（）。

与参考值相比，近红外预测值偏大时偏差为负，近红外预测值偏小时偏差为正。

数据散点图可以清晰的表现相关关系、偏差、斜率和异常值的情况。（见图 1）。



说明:

- 1——45°线 (偏差为 0、斜率为 1 的理想线); 3——线性回归线; 5——偏差。
 2——带偏差的 45°线 (45°线减去偏差得到的线) 4——异常值;

注: 异常值 (图中 4) 对斜率有很大的影响, 如果该结果是用于斜率调整的, 则异常值需要剔除。

图 1 定标样品的散点图,)

8.3 偏差

偏差产生的主要原因包括: 未被包含在定标模型中的新样品类型、仪器的漂移、参考值的漂移、过程的变化、样品制备的变化等。

偏差是残差的平均值, 个独立样品的偏差按式 (2) 计算:

$$\dots\dots\dots (2)$$

式中:

——式 (1) 中定义的残差, 由此可以得到式 (3):

$$\dots\dots\dots (3)$$

式中:

- 第个参考值;
- 第个近红外预测值;
- 参考值的平均值;
- 近红外预测值的平均值。

偏差的显著性用 *t*-test 进行检验。偏差置信区间 (BCL) (见附录 C.3.2) 的计算见式 (4):

$$\dots\dots\dots (4)$$

式中:

- 发生 I 类错误的概率, 即检验的显著性水平;
- 双尾检验学生氏 *t* 值, 其值大小取决于与 SEP 有关的自由度和 I 类错误的概率, 可使用 Excel 中的 TINV 函数计算得到;
- 独立样品数;
- 预测标准误差 (见 8.5)。

示例: 当 $n = 20$, 且, BCL 按式 (5) 计算:

$$\dots\dots\dots (5)$$

这意味着, 当偏差高于预测标准误差的 48%时, 该 20 个样品检测的偏差与 0 有显著性差异。

8.4 预测误差均方根 ()

预测误差均方根 () 按式 (6) 计算:

..... (6)

式中：

——第 i 个样品的残差；

——独立样品数。

该值可以与和进行比较（见附录 C.3.3 和 C.3.4）。

包含随机误差（即预测标准误差）和系统误差（即偏差），同时也包含参考方法的误差。计算公式见式（7）：

..... (7)

式中：

——独立样品数；

——预测标准误差（见 8.5）；

——偏差或系统误差。

对不进行统计检验，而是分别对系统误差（即：偏差/）和随机误差（即：）进行相应的统计检验。

8.5 预测标准误差（）

预测标准误差（），也就是残差的标准差，计算公式见式（8）：

..... (8)

式中：

——独立样品数；

——第 i 个样品的残差；

——偏差或系统误差。

与或的比较（见附录 C），可用于检查定标模型相对于所选验证集的有效性。

随机误差置信区间（UECL），（见附录 C.3.8），用 F 检验计算（参考文献[2]），见式（9）：

..... (9)

式中：

——定标标准误差（见附录 C.3.3）；

—— F 累积分布的概率值，大小取决于 I 类错误概率、自由度 ν 和，可使用 Excel 中的 FINV 函数计算；

——发生 I 类错误的概率，即检验的显著性水平；

ν ——验证样品相对应的自由度，即，其中是验证样品数；

——定标标准误差相对应的自由度，即。其中，是定标样品数；是模型的因子数或 PLS 主成分数，或人工神经网络模型中的神经元连接的权重（见附录 C.2）。

注：由于过于理想化，故（见附录 C.3.7）是比更好的统计量，经常用于取代。

示例：,,,，计算。

20 个验证样品，最大允许大于的 30%。

F 检验需要两个独立的样品集，所以不能用于比较基于同一样品集的两种定标方法的检验。

8.6 斜率

回归方程中的斜率和截距的计算根据最小二乘拟合，计算公式分别见式（10）和（11）：

..... (10)

式中：

——参考值和近红外预测值的协方差；

——一个近红外预测值的方差。

..... (11)

式中：

——参考值的平均值；

——斜率；

——近红外预测值的平均值。

可用 t 检验对进行假设检验，如式（12）所示：

..... (12)

式中：

——一个近红外预测值的方差；

——独立样品数；

——残差标准偏差。

残差标准偏差，按式（13）计算：

..... (13)

式中：

——独立样品数；

——截距（见式 11）；

——斜率（见式 10）；

——第 i 个参考值；

——第 i 个近红外预测值。

当满足式（14）时，斜率与 1 有显著性差异：

..... (14)

式中：

——由式（12）计算的 t 值；

——双尾检验学生氏 t 值，其值大小取决于自由度和 I 类错误的概率；可使用 Excel 中的 TINV 函数计算得到。

示例：样品数量=20，残差标准偏差（见公式（13））为 1，近红外预测值的标准偏差=2，并且计算的斜率=1.2，则得到的值为 1.7，由于对应 20 个样品的为 2.09（），故斜率与 1 之间的差异不显著。若斜率为 1.3，则得到的值为 2.6，那么斜率与 1 之间存在显著性差异。

当计算的斜率用于校正近红外预测结果时，斜率的计算必须以参考值作为因变量，以近红外预测值作为自变量。

9 试样的测定

9.1 试样近红外光谱的测量

参照近红外光谱仪器制造商或供应商的仪器使用说明进行。

试样的温度应在验证样品温度范围之内。近红外光谱分析时，样品温度不可过高或过低，应在常温范围内。

9.2 试样近红外光谱预测值的测定

应检查定标模型对试样的适用性。可采用在定标和验证过程中所使用的异常值剔除方法，如马氏距离（见附录 C.2.11）和光谱残差（见附录 C.2.14）等。这在大多数仪器中可自动执行。

如果试样没有通过适用性检验，则不能用现有的定标模型进行检测，需对定标模型进行修正或升级。如果定标模型能够适用于被分析样品，则该光谱可采用该定标模型进行预测。

试样的测定结果应在所使用的定标模型成分含量范围内方可有效结果。

若试样的光谱异常，则近红外光谱测定结果不可采纳。

异常样品的检测可用于选择哪些样品用参考方法赋值后作为升级定标模型之用。

9.3 试样近红外光谱预测结果的表述

若对同一样品有多次测量，且满足重复性要求的（见 12.1），计算其算术平均值。

对结果的表述可参照附录 A。

10 仪器稳定性检查

10.1 用质控样品检查仪器稳定性

为确保仪器硬件的稳定性及预防仪器故障，每天至少进行一次质控样品的检测。质控样品应该与试样测定时间同步，以保障不间断的质量控制。

在质控图上，应对每天质控结果的波动进行记录，并应对有显著性变化趋势的质控图进行研究。

10.2 仪器诊断

对扫描型光谱仪，应对其波长/波数（见 5.1）的准确度和精密度进行至少每周一次的检查，或者如果仪器商推荐，其检查频度应该更高。并且其结果需要与相应技术规格及要求进行对比（见 5.1）。

应至少每周或按照厂商推荐的时间间隔对仪器的噪音进行检查。

10.3 仪器联网

如果几台仪器联网使用，需根据厂商的推荐对仪器之间的标准化（见附录 C.1.6）给予特别关注。

11 定标模型运行性能监控

11.1 总则

定标模型运行性能应持续选用根据参考方法赋值的监控样品进行验证，以确保定标处于稳定的最优状态并满足准确性要求。对近红外定标模型运行性能的验证频度，应足以确保该方法运行的稳定性受控于参考方法的系统偏差和随机误差。该频度尤其依赖于每天分析的样品数量和样品集的波动。

定标模型运行性能监控样品数量应能满足统计分析需要。可靠的监控过程需要至少 20 个样品（以满足变量的正态分布）。独立验证样品的结果可以用于定标模型运行性能的首次监控，接下来每周 5-10 个样品即可满足对定标模型性能的监控。如果监控样品数量太少，一旦有结果超出了控制限则很难做出正确判断。

11.2 参考值和近红外预测值差异控制图

应采用控制图对结果进行评估，控制图是以分析监控样品数作为横坐标，参考值和近红外预测值的差作为纵坐标；（95%的概率）和（99.8%的概率）可用作报警限和行动限，采用独立于定标样品的验证样品获得。

在定标和参考值的分析均按要求执行的条件下，则应只有 5%的样品在报警限之外，只有 2%的样品在行动限之外。

对控制图应检查系统偏差、偏差的正负和极端异常值。可使用 Shewart 控制图的一般规则对结果进行评估。然而，过多规则同时使用可导致虚假报警。

以下规则的组合应用已被证明具备发现问题的实用价值：

- a) 1 个点超过了行动限；
- b) 连续 3 个点中有 2 个在报警限之外；
- c) 连续 9 个点在 0 线的同一侧。

在对结果的评估中，应注意和所测得的近红外预测值与参考值的差异也包含参考值的不确定度。如果参考值的不确定度小于的 1/3，则该影响可忽略不计^[3]。

为减少虚假报警的风险，监控样品的近红外光谱分析和参考值分析应以不同的顺序分别独立进行，以避免（如在参考方法中）日间系统偏差的影响。

如果样品经常超出报警限，并且控制图显示都是随机误差（无趋势性或系统偏差），那可能是在确定行动限和报警限时用的太过理想。此时，通过频繁调整定标以勉强将结果限于控制限范围内并不可取，而应对进行重新评估。

稳定运行一段时间后的定标模型可能会开始出现失控情况,此时应对定标模型进行升级。升级之前,首先要评估该失控是否是因为参考值的变化、测量条件预料之外的变化(如更换新的操作者所致)、仪器漂移或发生故障等。在有些情况下,只通过调整定标模型的常数项(即截距)就能满足要求(示例见图 B.6)。在其它情况下,可能需要重新运行定标程序。重新定标时,定标样品应包含对定标模型运行性能进行监控的监控样品,亦可使用专门为重新定标选择的样品(示例见图 B.7)。

鉴于参考值已满足参考方法精密度要求,且近红外分析测量条件与仪器性能可保持不变,偏差显著或值增大可能是由于样品相比于定标样品的化学、生物学或物理性质发生了变化。

其它控制图[例如,得分(见附录 C.1.8)图]和其它控制原则也可使用,以强化评估效果。

12 精密度和准确度

12.1 重复性

重复性是对完全相同的试样,采用相同的检测方法,在同一实验室,相同的操作人员,用相同的操作仪器,在短时间内获得的两次独立测量结果的差异,该差异不应该大于 5%。重复性的大小依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略等。所有情况下都应进行重复性的检测。

12.2 重现性

重现性是对完全相同的试样,不同的实验室、不同的操作者、在不同的时间获得的两次独立测定结果的差异,该差异不应该大于 5%。重现性大小依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略。所有情况下都应进行重现性的检测。

12.3 准确度

准确度包含了偏离样品参考值的系统偏差的不确定度(准确度)和随机误差的不确定度(精密度),尤其依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略。所有情况下都应进行精确度的检测。报告的和值也包括了参考值的不确定度。

12.4 不确定度

不确定度(是用来描述合理地分布于结果附近的测量值的离散度的参数。对于近红外预测的结果,不确定度通常用式(15)表示:

$$\dots\dots\dots (15)$$

如果乘数为 2,可理解为真值有 95%的可能性落在之间。

13 检测报告

检测报告应包括以下信息:

- a) 完整识别样品所需的全部信息;
- b) 参考值赋值采用的参考方法;
- c) 所有该标准未明确或被视为可选的所有操作条件;
- d) 任何可能影响结果的环境因素;

e) 获得的试样的测定结果；

f) 目前的和偏差（有统计差异时），用最少 20 个监控样品进行定标模型运行性能测试得到（见本文件第 11 章）。

附录 A
(资料性)

具体成分或技术指标的近红外分析标准制定指南

具体成分或技术指标的近红外分析标准,可依据通过近红外光谱法测定饲料中具体成分或技术指标的专用定标模型进行制定。

这些标准应具备以下信息:

- 在验证过程中,遵循本标准所规范的内容;
- 针对浓度、准确度与精密度,给出需满足的具体参数范围;
- 指定在开发模型时应考虑到的变量的来源;
- 应具体规定分析流程、计算方法和结果的表达。

具体成分或技术指标的近红外分析标准应符合标准文本的格式,并给出以下相关信息:

- 在标准名称或范围中应给出:样品类型、通过近红外光谱仪检测的成分/技术指标和定标模型;
- 在“规范性引用文件”中列出验证用的参考方法;
- 光谱的原理(如近红外反射 NIR、近红外透射 NIT)和定标原理(如偏最小二乘方法 PLS、人工神经网络 ANN);
- 测试样品的制备、测定和质量控制过程;
- 依据准确度、精密度及范围制定需满足的具体值;表 A.1 是一个验证统计值的示例。

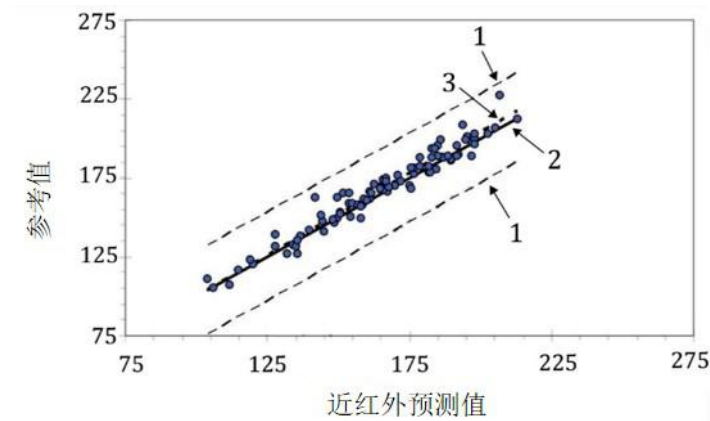
表 A.1 验证统计值的示例

成分	模型	样品数 <i>N</i>	精密度	最小含量 %质量比	最大含量 %质量比	<i>R</i> ²
粗脂肪	ANN	183	0.50	2.8	12.9	0.94
水分	ANN	183	0.47	9.2	12.3	0.83
粗蛋白	ANN	179	0.72	11.0	29.1	0.96
粗纤维	ANN	123	1.11	0.5	18.0	0.90
淀粉	PLS	113	1.80	7.8	50.2	0.92

附录 B
(资料性)

异常值和控制图示例

图 B.1 显示的是饲草中粗蛋白含量的测定，这是一个没有异常值的例子。使用已建立的定标模型分析 95 个独立验证样品的结果：；；斜率。

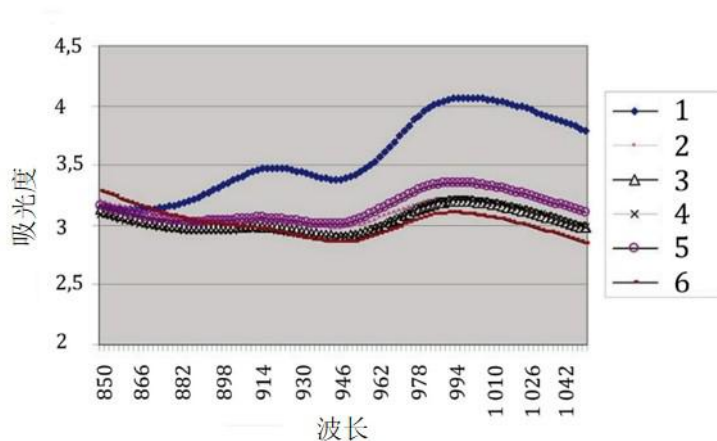


说明：

- 1——临界线
- 2——45° 线（斜率为 1，截距为 0 的理想线）
- 3——回归线

图 B.1 无异常值

图 B.2 显示的是存在-异常值的吸收光谱。光谱 1（图 B.2 中最上方的光谱）是异常光谱。

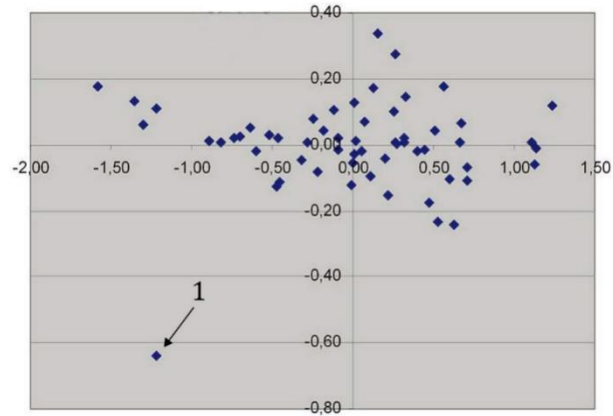


说明：

- 1——光谱 1
- 2——光谱 2
- 3——光谱 3
- 4——光谱 4
- 5——光谱 5
- 6——光谱 6

图 B.2 存在-异常值的吸收光谱

图 B.3 显示的是存在-异常值的主成分得分图（图 B.3 中 1 为-异常值）。

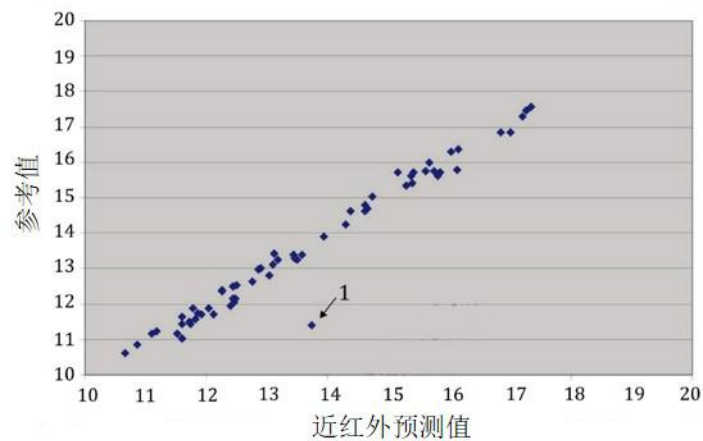


说明:

1——异常值

图 B.3 存在-异常值的主成分分析得分图

图 B.4 显示的是存在-异常值的散点图（图 B.4 中 1 为-异常值）。参考值 vs 预测值的散点图（反之亦然）显示出与其它样品严重偏离的一个样品（图 B.4 中 1 所示）。若其偏离的原因与 NIRS 数据（ x -异常值）无关，则该样品是一个由于错误的参考值或参考值与光谱值之间存在不同的相关关系所导致的-异常值。

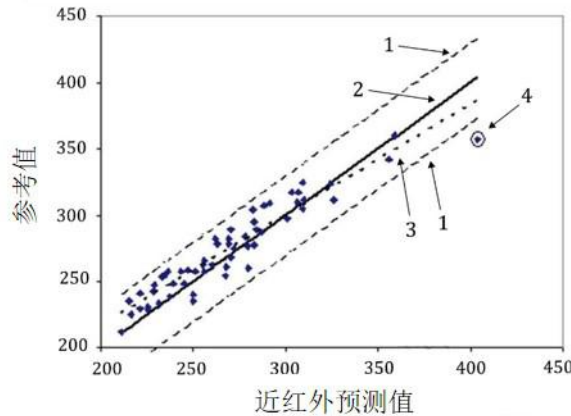


说明:

1——异常值

图 B.4 存在-异常值的散点图

图 B.5 显示的是存在-异常值（图 B.5 中 1 所示）的饲草中酸性洗涤纤维 ADF 含量的测定。



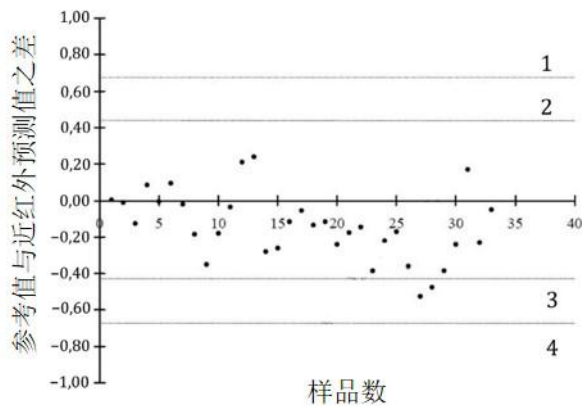
说明:

- 1——临界线
- 2——45°线
- 3——回归线
- 4——异常值

图 B. 5 存在-异常值的饲草 ADF 的测定

图 B.6 显示的是谷物饲料原料中粗脂肪含量的测定。在行动限上限 UAL 和行动限下限 LAL 之外没有异常值点。但有连续的 9 个点（如 14-22）在零线同一侧，这意味着偏差问题的存在。另有连续 3 点中的 2 点（27 和 28）超出了报警限下限 LWL，但没有出现超出报警限上限 UWL 的点，这同样表明偏差问题的存在。未观察到随机变化增加的情况，分布依然小于。

总之，该定标模型应进行偏差校正。



说明:

- 1——行动限上限 (UAL,)
- 2——报警限上限 (UWL,)
- 3——报警限下限 (LWL, -)
- 4——行动限下限 (LAL, -)

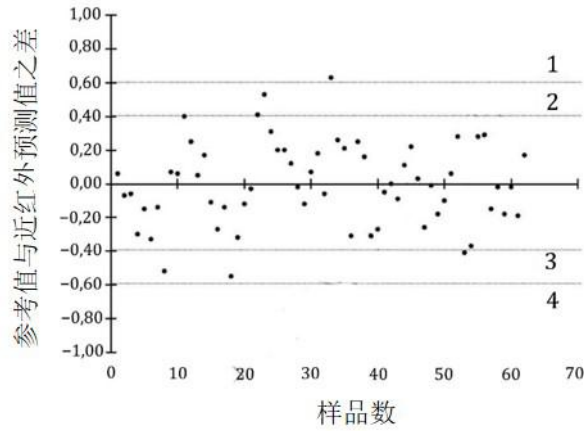
图 B. 6 谷物饲料原料中粗脂肪含量测定的控制图

图 B.7 为某模型中一项参数（范围：44%~57%）测定的控制图，在第 35 点处进行了重新定标。

观察前 34 个检测点，其中有 1 点超出了行动限上限 UAL，这意味着严重问题的出现。连续 3 点中的 2 点（22 和 23）超出了报警限上限 UWL，2 个不连续点同样超出了报警限下限 LWL。样品点均衡分布于零线附近（遵守 9 点规则），但 34 个点中有 5 个在 95%置信区间[UWL, LWL]之外，并且有 1 个点超出了 99.9%的置信区间[UAL, LAL]。这远远超出了期望。

出现这种结果的原因之一是计算临界值的值过于理想，这意味着临界值应该要适当放宽。另一种原

因可能是实测样品与定标样品有些不同。为检测这种可能，对定标样品进行扩展，使之包含质控样品，并生成新的定标模型。通过 35-62 质控样品的结果可以看到，新定标模型的性能得到了明显改善。



说明：

1——行动限上限 (UAL, +3)

3——报警限下限 (LWL, -2)

2——报警限上限 (UWL, +2)

4——行动限下限 (LAL, -3)

图 B. 7 某模型中一项参数测定的控制图 (范围：44%~57%)

附录 C
(资料性)
补充性条款及定义

C.1 总则

C.1.1 参考方法

获得国家标准化管理委员会或其它国际组织认可、经过验证的检测方法，能够给出待测参数的包括测量不确定度的“真值”或“设定值”。

C.1.2 间接法

通过检测与待测指标具备函数关系的某些特性来测定待测指标的方法，该方法所获取的信号与利用参考方法测定的“真值”相关。

C.1.3 近红外光谱分析 (NIRS)

近红外光谱，是指在 770 nm~2500 nm (12900 cm^{-1} ~ 4000 cm^{-1}) 的近红外光范围内，测量样品对光的吸收强度。NIRS 仪器使用该光谱范围的部分或整体，或更宽的能包含该范围的光谱区域（如 400 nm~2500 nm）进行检测，然后采用多元校正分析技术将吸光度值与样品成分含量或属性进行关联。

C.1.4 近红外漫反射 (NIR)

其检测基础是利用样品前侧传感器对所吸收到的样品表面漫反射回来的近红外光进行检测。

C.1.5 近红外透射 (NIT)

其检测基础是利用样品背侧传感器对所吸收到的透射过样品的近红外光进行检测。

C.1.6 仪器的标准化

对一组近红外仪器进行调整的过程。通过调整，当利用统一的定标模型时，所有设备对相同的样品获取相近的预测值。有许多技术可用，但若从广义的范畴进行定义，这些技术可以分为两种：

C.1.6.1 前预测法

即通过调整样品光谱使“主”仪器与该组内的每台仪器之间的差异最小。

C.1.6.2 后预测法

即使用线性回归分析调整每台仪器的预测值，使之尽可能接近“主”仪器的预测值。

C.1.7 NIRS 网络

使用相同定标模型的一定数量的近红外仪器。网内仪器通常需要标准化，从而能使标准样品预测值的差异最小。

C.1.8 z-得分

利用近红外预测值与参考值的差值除以标准偏差（例如 RMSEP）得到的一项性能指标。

C.2 定标方法

C.2.1 主成分分析 (PCA)

主成分分析是一种数据压缩方式。只对 X (光谱) 数据进行计算, 每个主成分都表达了光谱数据投影的最大方差, 并且这些主成分互不相关。第一主成分尽可能多的表达了原始数据的变化。然后将其影响从 X 数据中扣除, 新的主成分尽可能多的表达剩余数据的特征。能够得到和光谱数据点或数据集中样品个数相同的主成分数, 但光谱中主要影响集中在前几个主成分, 因此需要考虑的数据量就明显降低了。

在每一步计算中, 主成分分析 (PCA) 产生两个新的变量集: 主成分得分代表了每个样品在每个主成分上的响应; 主成分载荷代表了原始光谱中每个数据点对主成分的相对重要性。

主成分分析有很多用途, 例如光谱解析, 但是最广泛的用途是用于光谱异常值的判别。

C.2.2 主成分回归 (PCR)

主成分回归是在多元线性回归中, 以每个主成分作为回归量、样品的成分含量作为 y 值的一种回归分析技术。因为主成分是相互正交的, 因此主成分得分可形成性能优于原始光谱数据的互不相关的数据集。尽管能够以每个主成分与目标成分含量相关性的优劣为依据选择回归的主成分组合, 但大多数商业软件会强制采用为模型所选择的最大主成分数进行回归。

当分析近红外数据时, 通常将主成分空间的回归系数转换回波长空间中用所有数据点建立的预测模型。

C.2.3 偏最小二乘回归 (PLS)

偏最小二乘回归是一种分析技术, 像 PCA 一样, 也是一种数据压缩形式。其提取因子的规则是每个因子能够依次实现 Y 值和 X 数据所有可能线性组合之间最大的协方差。PLS 在计算主成分时, 在考虑所计算的主成分方差尽可能最大的同时, 还使主成分与浓度最大程度地相关。因此与主成分相比, PLS 因子与 Y 值的变化更直接相关。PLS 生成三个新变量, 载荷权重 (之间不是正交关系)、载荷和得分 (相互正交的)。

PLS 模型通过回归 Y 值和 PLS 得分获得。和主成分回归一样, 当分析近红外数据时, 通常将 PLS 空间的回归系数转换回波长空间中用所有数据点建立的预测模型。

C.2.4 多元线性回归 (MLR)

多元线性回归是联合几个 X 变量预测一个 Y 变量的回归技术。对近红外光谱技术, X 变量既可以是选择波长处的吸光度值, 又可以是像主成分得分和偏最小二乘得分等类似的派生变量。

C.2.5 人工神经网络 (ANN)

人工神经网络是基于生物神经系统结构的非线性模型技术。该网络最初通过几个具有 X 值 (光谱数据或如主成分得分等派生变量) 和 Y 参考值的数据集进行“训练”。在训练过程中, 为产生最好的参数值预测, 对输入变量和输出变量, 可能需要修改网络结构, 以及重新分配神经元加权系数。

训练时神经网络需要大量数据。对 ANN 的一些优化方案, 需要设定一个停止点以终止模型的优化 (例如: 权重)。

C.2.6 多变量模型

多变量模型是用一组 X 变量预测一个或多个 Y 变量的模型。

C.2.7 异常值

从统计的角度讲，异常值是数据集内位于预期分布之外的那些点。对 NIRS 数据来说，异常值通常分为-异常值（光谱异常）或-异常值（参考值异常）。

C.2.8 -异常值

-异常值是与近红外光谱相关的异常值。-异常值可能由以下原因引起：仪器故障、完全不同于其他样品的样品类型、或预测的样品类型不包含在原始定标样品中。

C.2.9 -异常值

-异常值是与参考数据中的错误相关的异常值，例如参考数据誊写错误或由于参考实验室获得的参考值错误。

C.2.10 杠杆值

杠杆值是一个度量样品与模型所定义的群体空间中心距离的测量值。具有高杠杆值的样品对模型有非常大的影响。杠杆值通过计算一个投影点和模型中心的距离得到。

C.2.11 马氏距离

马氏距离是数据点和主成分空间中心在主成分空间的距离（参见下面的 h 值），这是一个线性测量值。在主成分空间，一组样品通常形成一个曲线形状的分布。最能代表数据集概率分布的椭圆柱体可以通过构建样本的协方差矩阵进行估计。简单说，马氏距离通过测试点与数据集中心的距离除以测试点方向椭圆柱体的宽度得到。

C.2.12 h 值

在一些软件中，马氏距离被称为“全局 h 值”，并且通过样品距数据集中心 h 值的标准偏差来检测异常值。另一个是“邻近 h 值”，它是一个数据点和离它最近 n 个数据点在主成分空间的距离，它表明该样品是孤立的、还是隶属于某一集中分布的。

C.2.13 残差

残差是参考值和回归模型预测值之间的差异，用于回归统计计算。

C.2.14 光谱残差

化学计量学（如 PCA、PLS 等）处理光谱后的残差，是模型未描述的光谱变异。

C.2.15 交互验证

计算预测统计值的方法，重复从定标样品中移出不同的子集，用剩余的定标集样品建立定标模型，并用移出的样品作为验证样品计算残差；当这个过程运行多次后，用所有的残差计算预测统计值。全交互验证每次移出一个样品，共运行 n 次（ n 是定标样品数）。如需选择移出更大的子集，则应保证在计算预测统计值之前运行交互验证循环 8 次以上。最终，使用所有的定标样品计算出一个定标模型。

需谨慎使用交互验证。首先，交互验证统计值与独立验证统计值相比，更趋于给出偏乐观的结果。第二，如果在定标样品中有重复样品（例如：相同的样品在几台仪器上扫描或扫描不同的次数），要确保每次把该样品所有的重复数据一起移出作为验证样品，否则将会给出过于理想的统计结果。

C.2.16 过拟合

在一个多元线性回归中使用了太多的回归变量。当对不在定标样品中的样品进行预测时，过拟合模型的预测结果会比期望值差得多，像或统计值。

C.2.17 得分/得分图

一个主成分或偏最小二乘因子的得分与另一主成分或偏最小二乘因子得分的散点图。标注了样品编号或成分含量值的得分图更加实用，能够容易的识别出原始数据中不明显的规律。

C.3 统计表述

也可参见本文件第 8 章。

C.3.1 偏差 ()

参考值平均值和近红外预测值平均值的差异。

C.3.2 偏差置信区间 (BCL,)

位于置信区间之外，表明偏差与 0 存在给定置信水平上的显著差异（参见 7.3）。

C.3.3 定标标准误差 (SEC,)

对任一定标模型，定标标准误差 (SEC) 是建模样品的预测值和参考值的平均差异。在此及后续统计中，该平均差异是指残差值平方和除以自由度的平方根，68% 的误差小于该值。

C.3.4 交互验证标准误差 (SECV,)

对给定的定标模型，交互验证标准误差 (SECV) 是指，交互验证（参见 C.2.19）过程中，选择定标样品的子集作为预测样品，其预测值和参考值经偏差校正过的平均差异。

C.3.5 预测标准误差 (SEP,)

预测标准误差 (SEP) 是指，选择不包含于定标过程中的样品作为预测样品，其回归模型预测值和参考值经偏差校正过的平均差异。SEP 涵盖了 68% 的置信区间（该值乘以 1.96 就是 95% 的置信区间）。

C.3.6 预测误差均方根 (RMSEP,)

预测误差均方根 (RMSEP) 是指，选择不包含在定标过程中的样品作为预测样品，其参考值和回归模型预测值间的平均差异。

注：RMSEP 包含了预测过程中所有偏差。

C.3.7 交互验证误差均方根 (RMSECV,)

交互验证误差均方根 (RMSECV) 是指，交互验证（参见 C.2.19）过程中，选择定标样品的子集作为预测样品，其预测值和参考值的平均差异。

注：RMSECV 包含了预测过程中所有偏差。

C.3.8 随机误差置信区间 (UECL,)

随机误差置信区间是指，在给定置信水平上与存在显著性差异的界限。

C.3.9 相关系数的平方 (RSQ,)

预测值和参考值多重相关系数的平方。当用百分比表示时，代表回归模型可以解释的变量的比例。

C.3.10 斜率 ()

对任一回归线，斜率代表 X 每增加 1 个单位对应的 Y 的增量。

C.3.11 截距

对任一回归线，截距代表 X 等于 0 时 Y 的值。

C.3.12 残差标准偏差 ()

残差标准偏差是指，参考值和经斜率和截距校正后的拟合值之间差异的平均尺寸。

C.3.13 协方差 ()

协方差是测量两个随机变量一起变化的相关程度。对于样本总体，如果 y 随 x 增大而增大，那么这两个变量的协方差是正值。如果 y 随 x 增大而减小，那么这两个变量的协方差是负值。当两个变量不相关时，其协方差为 0。

C.3.14 不确定度 ()

不确定度是用来描述合理地分布于结果附近的测量值的离散度的参数。

附录 D

(资料性)

本文件与 ISO 12099:2017 相比的结构变化情况

本文件与 ISO 12099:2017 相比，结构上的修改具体对照情况见表 D.1。

表 D.1 本文件与 ISO 12099:2017 的章条编号对照情况

本标准章条编号	对应的 ISO 12099:2017 章条编号
7 定标模型的建立与验证	6 定标模型与初始验证
8 定标模型性能统计学评价	7 模型性能统计学评价
9 试样的测定	9 检测过程
C.1	附录 C 表中 General 子表
C.1.1 到 C.1.5	附录 C 表中 General 子表下第 1-5 行
C.1.6	附录 C 表中 General 子表下第 7 行
C.1.7	附录 C 表中 General 子表下第 6 行
C.1.8	附录 C 表中 General 子表下第 8 行
C.2	附录 C 表中 Calibration techniques 子表
C.2.1 到 C.2.14	附录 C 表中 Calibration techniques 子表下第 1-14 行
C.2.15 到 C.2.16	附录 C 表中 Calibration techniques 子表下第 19-20 行
C.2.17	附录 C 表中 Calibration techniques 子表下第 22 行
C.3	附录 C 表中 Statistical expressions 子表
C.3.1 到 C.3.14	附录 C 表中 Statistical expressions 子表下第 1-14 行

附录 E

(资料性)

本文件与 ISO 12099:2017 的技术差异及其原因

表 E.1 给出了本文件与 ISO 12099:2017 的技术性差异及其原因。

表 E.1 本文件与 ISO 12099:2017 的技术性差异及其原因

本标准章条编号	技术性差异	原因
封面	标准名称修改为“饲料的近红外光谱分析应用指南”，删除谷物及谷物精制料。	本标准为饲料行业标准，对适应对象进行了限定。
1	删除适用对象中的谷物及谷物精制料。	本标准为饲料行业标准，对适应对象进行了限定。
1	对使用范围进行了明确“本标准适用于可采用近红外光谱分析（见附录 C.1.3）的饲料”。	根据标准撰写规范，明确了使用范围的说明。
2	增加：规范性引用文件。	将标准中使用的相关标准文件：GB/T 10647 饲料工业术语，GB/T 14699.1 饲料 采样和 GB/T 29858 分子光谱多元校正定量分析通则列为规范性引用文件。
3	删除：ISO 和 IEC 的术语数据库资源，替换为：GB/T 10647 和 GB/T 29858 界定的以及下列术语与定义适用于本标准。	本标准用到的所有术语在 GB/T 10647 和 GB/T 29858 和本标准附录 C 中均已列出。
3	删除：ISO 标准中对近红外光谱仪和动物饲料两个术语定义。	附录 C 中对近红外光谱分析进行了定义；动物饲料术语见 GB/T 10647。
6	增加：第 6 章 样品	在该标准中，所用的样品比较复杂，包括对近红外仪器进行定标的定标样品、对定标模型进行验证的验证样品、待测的试样、对仪器进行稳定性检查的质控样品、和对定标模型运行性能进行日常监控的监控样品。ISO 标准是在用到相关样品时随带进行解释，这样样品的描述比较混乱且不符合国内标准的逻辑，因此，将标准中用到的所有样品进行了统一的描述，增加了本章。
6.1	用等同采用国际标准的 GB/T 14699.1，代替了 ISO 6497	引用 GB/T 14699.1，便于标准使用者使用中文版标准。
7.1	增加对定标模型的解释，删除关于验证样品的相关描述。	增加定标模型的解释更便于读者理解和使用；验证样品的相关描述已经在 6.2 中统一说明。

表 E.1 本文件与 ISO 12099:2017 的技术性差异及其原因 (续)

本标准章条编号	技术性差异	原因
7.3	调整段落顺序。	为便于读者理解和使用。
7.4.1	调整段落顺序。	为便于读者理解和使用。
7.4.3	对 7.4.3 通常不推荐对其斜率或截距进行调整, 除非定标模型应用于新类型样品或仪器, 删除了其中的截距。	本标准中 7.4.2 是截距调整, 7.4.3 是斜率调整, 将 7.4.3 中的截距删除逻辑上更清晰, 且生产实际应用中证明: 对一些模型进行截距调整对改善系统偏差是比较常用的方法, 并且效果也很好。
7.4.3	将 8.6 中关于斜率校正的描述移到此处: 含量范围太窄或分布不均匀都会导致斜率校正无用。仅当验证样品覆盖了定标样品大部分含量范围时, 方可对斜率进行调整。	为便于读者理解和使用。
7.5	未在定标样品中包含的样品制备技术或测量条件的改变同样会影响分析结果, 例如样品研磨粒度、温度等。举例增加了样品研磨粒度。	技术实际应用中发现样品研磨粒度的影响非常大, 影响因素举例中把样品研磨粒度也列出, 可引起使用者的注意, 便于更好的利用该技术。
8.3	对公式 (4) 中参数的解释和示例计算做了具有技术性差异的调整, 如下: ——公式 (4) 中在参数的解释后增加了利用 Excel 计算的函数; ——删除了表 1。	——增加利用 Excel 计算的函数, 便于使用者使用。 ——表 1 仅给出了时, 部分对应的 t 值, 使用时意义不大, 并且 t 可用 Excel 中的函数计算, 此表只是一个示例, 对于整个标准没有必要。
8.5	技术性差异调整如下: ——公式 (9) 中增加了参数的解释: 大小取决于 I 类错误概率和自由度 v 和, 可使用 Excel 中的 FINV 函数计算; ——删除了表 2。	——界定符号的名称定义, 使公式计算更清晰; ——表 2 仅给出了时, 部分自由度对应的 F 值, 使用时意义不大, 并且 F 可用 Excel 中的函数计算, 此表只是一个示例, 对于整个标准没有必要。
8.6	对公式 (14) 中的解释做了调整。	与公式 (4) 的说法保持一致。
8.6	对示例中的下述表述做了调整, 具体调整如下: 由于对应 20 个样品的为 2.09 (。	增加了, 更易于使用者理解和使用。
9	将待测试样的描述移到第 6 章样品中 6.3 中说明。	为便于读者理解和使用。

表 E.1 本文件与 ISO 12099:2017 的技术性差异及其原因 (续)

本标准章条编号	技术性差异	原因
9.1	增加：近红外分析时，样品温度不可过高或过低，应在常温范围内。	ISO 标准文本中只提到所制备样品的温度应在验证时样品温度的范围之内。但未对温度提出具体要求，增加此说明，更便于标准的实施。
9.2	增加：试样近红外光谱预测值的测定相关描述。	将 ISO 标准 11.1 中与样品测试原则和方法相关叙述移至此条描述，更便于读者理解和使用。
10.1	将质控样品的描述移到第 6 章样品中 6.4 中说明。	为便于读者理解和使用。
11.1	将监控样品的描述移到第 6 章样品中 6.5 中说明。	为便于读者理解和使用。
附录 A	删除谷物及谷物精制料的描述。	本标准为饲料行业标准，对适应对象进行了限定。
附录 C	C.1.1 参考方法将“获得 ISO 或其他国际认可和经过验证的检测方法”改为“获得国家标准化管理委员会或国际标准化组织认可和经过验证的检测方法”	适应中国标准情况。
附录 C	C.1.3 近红外光谱技术 (NIRS) 将“近红外光谱，是指在 700 nm~2500 nm (14300 cm ⁻¹ ~4000 cm ⁻¹) 的近红外光范围内，测量样品对光的吸收强度。”改为“近红外光谱，是指在 770 nm~2500 nm (12900 cm ⁻¹ ~4000 cm ⁻¹) 的近红外光范围内，测量样品对光的吸收强度。”	光谱范围与标准正文 5.1 保持一致，统一为 770 nm~2500 nm。

参考文献

- [1] NÆS T., ISAKSSON T., FEARN T., DAVIES T. A user-friendly guide to multivariate calibration and classification. NIR Publications, Chichester, 2002.
- [2] SØRENSEN, L.K. Use of routine analytical methods for controlling compliance of milk and milk products with compositional requirements. IDF Bull. 2004, (390), pp. 42-49
- [3] SHENK, J.S., WESTERHAUS, M.O., ABRAMS, S.M. Protocol for NIRS calibration: Monitoring analysis results and recalibration. In: Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality. USDA ARS Handbook 643, (Marten, G.C., Shenk, J.S., Barton, F.E., eds.). US Government Printing Office, Washington, DC, 1989, pp.104-110.
-

饲料的近红外光谱分析应用指南

编制说明

（公开征求意见稿）

起草单位：中国农业大学

中国农业科学院农业质量标准与检测技术研究所

一、标准制定背景及任务来源

1.1 任务来源

2018年7月农业农村部下发了2018年农业行业标准制定和修订项目“制定《饲料近红外光谱应用指南》标准”任务，由中国农业大学和中国农业科学院农业质量标准与检测技术研究所[国家饲料质量监督检验中心（北京）]承担该标准的制定工作，技术归口单位为全国饲料工业标准化技术委员会。具体参照农财发[2018]46号《农业农村部关于支付2018年农产品质量安全监管专项经费等项目资金的通知》。

1.2 标准制定的意义和必要性

饲料是动物生产的最主要投入品，是动物性食品生产的源头。确保饲料优质安全，是实现动物性食品优质安全的关键。当前全球工业领域的质量管理，已提升到以“原料控制”及“生产过程质量控制”等预防性的质量控制和检验手段为主。要满足上述要求，须有快速、适宜现场及在线检测的检验手段作为支撑。

与其他分析技术尤其是传统的实验室化学分析技术相比，近红外光谱（Near Infrared Reflectance Spectroscopy, NIRS）分析技术在分析速度、检测成本、可同时检测多种理化性质、易操作性等主要检测性能方面具有显著优势。在全球饲料行业，NIRS技术的优势已经获得了极大的认可和广泛的应用。在ISO 12099:2010 Animal feeding stuffs, cereals and milled cereal products-Guidelines for the application of near infrared spectrometry 标准颁布之前，国际上的近红外光谱技术在饲料行业中并没有通行的、普适性的国际标准。2010年6月15日，ISO 12099的颁布实施在动物饲料行业树立了行业公认的交流准则，从而让不同NIRS光谱用户实现了结果的互认与交流。该标准在2017年进行了修订。

在我国，大量的科研机构及企事业单位越来越重视并充分挖掘和利用着NIRS分析的优势。在饲料行业，我国从20世纪90年代中期开始引进近红外饲料分析仪器，到2002年底，正式颁布了饲料行业近红外分析的国家标准GB/T18868-2002《饲料中水分、粗蛋白

质、粗纤维、粗脂肪、赖氨酸、蛋氨酸快速测定近红外光谱法》、2007年颁布实施了 NY/T 1423-2007《鱼粉和反刍动物精料补充料中肉骨粉快速定性检测 近红外反射光谱法》，2012年颁布了地方标准 DB21/T 2048-2012《饲料中粗蛋白、粗脂肪、粗纤维、水分、钙、总磷、粗灰分、水溶性氯化物、氨基酸的测定 近红外光谱法》。这些标准的颁布实施，标志着这项检测新技术在我国的饲料检测方面受到了广泛的关注和认可。

国内许多大型饲料企业和科研院所均在饲料的 NIRS 检测软硬件方面投入了巨大的财力物力，既促进了饲料行业的飞速发展，又提升了 NIRS 技术的普及与推广。但与飞速发展的 NIRS 饲料分析技术以及对应的国际标准方面的发展相比，我国针对 NIRS 技术在饲料检测方面的标准制订还有待完善，这对我国 NIRS 技术在饲料行业的全面、稳健、规范的发展形成了制约。故此，亟需推出针对饲料行业检测具有指导性质的、能适应 NIRS 检测技术发展态势的指南标准。

ISO 12099 为使用近红外光谱进行动物饲料的成分如水分、脂肪、蛋白、淀粉、粗纤维含量以及相关性能参数如消化率等的检测提供了综合性指南。ISO 12099 国际标准的引用，为各项技术环节提供了非常细致的指导基础，是后续开发和应用具体 NIRS 解决方案的重要基石。

二、主要工作过程

2018年7月任务下达之后，中国农业大学工学院和中国农业科学院农业质量标准与检测技术研究所[国家饲料质量监督检验中心（北京）]成立了标准编制小组。

2018年7-12月，标准编制组根据标准制定项目计划的要求，开展了广泛的国内外调研与技术资料收集整理工作。并在此基础上制定了标准的制定原则，即修改采用国际标准 ISO 12099:2017

《Animal feeding stuffs, cereals and milled cereal products-Guidelines for the application of near infrared spectrometry 动物饲料、谷物及谷物精制料的近红外光谱分析应用指南》。

2019年1-3月，完成国际标准 ISO 12099:2017 《Animal feeding stuffs, cereals and milled cereal products-Guidelines for the application of near infrared spectrometry 动物饲料、谷物及谷物精制料的近红外光谱分析应用指南》的翻译与校对。

2019年4-5月，依据中华人民共和国国家质量监督检验检疫总局令第10号《采用国际标准管理办法》，并按照 GB/T 20000.2-2009《标准化工作指南 第2部分：采用国际标准》规定，使用重新起草法修改采用 ISO 12099:2017《动物饲料、谷物及谷物精制料的近红外光谱分析应用指南》，形成农业行业标准《饲料近红外光谱应用指南（征求意见稿）》。

2019年6-7月标准编制组进行了领域内同行专家的意见定向征集，共发出征求意见稿25份，收到回函18份其中含有建议或意见回函18份，意见116条。

2019年8月，认真开展定向征求意见回函的分析、处理与汇总，对提出的116条意见，采纳88条，部分采纳或不采纳28条。形成农业行业标准《饲料近红外光谱应用指南（预审稿）》。

2019年9月27日，中国农业大学工学院组织专家对农业行业标准《饲料近红外光谱应用指南（预审稿）》进行了认真的审查。与会专家认为：标准内容详实、完整，标准预审稿可行。并对进一步提升标准质量提出了修改建议。

2019年10-12月，标准起草组根据预审专家组提出的修改意见对标准进行了修改完善，形成农业行业标准《饲料近红外光谱应用指南（送审稿）》。

受疫情影响，2020年8月提交送审稿电子版。

2021年4-5月，根据 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的新要求对标准文本送审稿的用语和规范性进行了修改。

三、标准编制原则和主要技术内容确定的依据

本标准修改采用 ISO 12099:2017《动物饲料、谷物及谷物精制

料的近红外光谱分析应用指南》。ISO 12099:2017 标准原文见附件 1，翻译稿见附件 2。

本标准与 ISO 12099:2017 相比在结构上基本保持一致，表 1 为本标准与 ISO 12099:2017 的章条编号对照一览表。

表 1 本标准与 ISO 12099:2017 的章条编号对照情况

本标准章条编号	对应的 ISO 12099:2017 章条编号
7 定标模型的建立与验证	6 定标模型与初始验证
8 定标模型性能统计学评价	7 模型性能统计学评价
9 试样的测定	9 检测过程
C.1	附录 C 表中 General 子表
C.1.1 到 C.1.5	附录 C 表中 General 子表下第 1-5 行
C.1.6	附录 C 表中 General 子表下第 7 行
C.1.7	附录 C 表中 General 子表下第 6 行
C.1.8	附录 C 表中 General 子表下第 8 行
C.2	附录 C 表中 Calibration techniques 子表
C.2.1 到 C.2.14	附录 C 表中 Calibration techniques 子表下第 1-14 行
C.2.15 到 C.2.16	附录 C 表中 Calibration techniques 子表下第 19-20 行
C.2.17	附录 C 表中 Calibration techniques 子表下第 22 行
C.3	附录 C 表中 Statistical expressions 子表
C.3.1 到 C.3.14	附录 C 表中 Statistical expressions 子表下第 1-14 行

本标准与 ISO 12099:2017 相比存在技术性差异，这些差异涉及的条款已通过在其外侧页边空白位置的垂直单线（|）进行了表示，表 2 给出了相应技术性差异及其原因的一览表。

表 2 本标准与 ISO 12099:2017 的技术性差异及其原因

本标准章条编号	技术性差异	原因
封面	标准名称修改为“饲料的近红外光谱分析应用指南”，删除谷物及谷物精制料。	本标准为饲料行业标准，对适应对象进行了限定。
1	删除适用对象中的谷物及谷物精制料。	本标准为饲料行业标准，对适应对象进行了限定。

表 2 本标准与 ISO 12099:2017 的技术性差异及其原因 (续)

本标准章条编号	技术性差异	原因
1	对使用范围进行了明确“本标准适用于可采用近红外光谱分析(见附录 C.1.3)的饲料”。	根据标准撰写规范,明确了使用范围的说明。
2	增加:规范性引用文件。	将标准中使用的相关标准文件:GB/T 10647 饲料工业术语,GB/T 14699.1 饲料采样和 GB/T 29858 分子光谱多元校正定量分析通则列为规范性引用文件。
3	删除:ISO 和 IEC 的术语数据库资源,替换为:GB/T 10647 和 GB/T 29858 界定的以及下列术语与定义适用于本标准。	本标准用到的所有术语在 GB/T 10647 和 GB/T 29858 和本标准附录 C 中均已列出。
3	删除:ISO 标准中对近红外光谱仪和动物饲料两个术语定义。	附录 C 中对近红外光谱分析进行了定义;动物饲料术语见 GB/T 10647。
6	增加:第 6 章 样品	在该标准中,所用的样品比较复杂,包括对近红外仪器进行定标的定标样品、对定标模型进行验证的验证样品、待测的试样、对仪器进行稳定性检查的质控样品、和对定标模型运行性能进行日常监控的监控样品。ISO 标准是在用到相关样品时随带进行解释,这样样品的描述比较混乱且不符合国内标准的逻辑,因此,将标准中用到的所有样品进行了统一的描述,增加了本章。
6.1	用等同采用国际标准的 GB/T 14699.1, 代替了 ISO6497	引用 GB/T 14699.1, 便于标准使用者使用中文版标准。
7.1	增加对定标模型的解释,删除关于验证样品的相关描述。	增加定标模型的解释更便于读者理解和使用;验证样品的相关描述已经在 6.2 中统一说明。
7.3	调整段落顺序。	为便于读者理解和使用。
7.4.1	调整段落顺序。	为便于读者理解和使用。

表 2 本标准与 ISO 12099:2017 的技术性差异及其原因 (续)

本标准章条编号	技术性差异	原因
7.4.3	对 7.4.3 通常不推荐对其斜率或截距进行调整, 除非定标模型应用于新类型样品或仪器, 删除了其中的截距。	本标准中 7.4.2 是截距调整, 7.4.3 是斜率调整, 将 7.4.3 中的截距删除逻辑上更清晰, 且生产实际应用中证明: 对一些模型进行截距调整对改善系统偏差是比较常用的方法, 并且效果也很好。
7.4.3	将 8.6 中关于斜率校正的描述移到此处: 含量范围太窄或分布不均匀都会导致斜率校正无用。仅当验证样品覆盖了定标样品大部分含量范围时, 方可对斜率进行调整。	为便于读者理解和使用。
7.5	未在定标样品中包含的样品制备技术或测量条件的改变同样会影响分析结果, 例如样品研磨粒度、温度等。举例增加了样品研磨粒度。	技术实际应用中发现样品研磨粒度的影响非常大, 影响因素举例中把样品研磨粒度也列出, 可引起使用者的注意, 便于更好的利用该技术。
8.3	对公式 (4) 中参数的解释和示例计算做了具有技术性差异的调整, 如下: ——公式 (4) 中在参数 t 的解释后增加了利用 Excel 计算 t 的函数 ——删除了表 1	——增加利用 Excel 计算 t 的函数, 便于使用者使用 ——表 1 仅给出了 $\alpha = 0.05$ 时, 部分 n 对应的 t 值, 使用时意义不大, 并且 t 可用 Excel 中的函数计算, 此表只是一个示例, 对于整个标准没有必要。
8.5	技术性差异调整如下: ——公式 (9) 中增加了参数 F 的解释: $F_{(\alpha, \nu, M)}$ 大小取决于 I 类错误概率 α 和自由度 ν 和 M , $F_{(\alpha, \nu, M)}$ 可使用 Excel 中的 FINV 函数计算; ——删除了表 2	——界定符号的名称定义, 使公式计算更清晰 ——表 2 仅给出了 $\alpha = 0.05$ 时, 部分自由度对应的 F 值, 使用时意义不大, 并且 F 可用 Excel 中的函数计算, 此表只是一个示例, 对于整个标准没有必要。
8.6	对公式 (14) 中 $t_{(1-\frac{\alpha}{2})}$ 的解释做了调整。	与公式 (4) 的说法保持一致。
8.6	对示例中的下述表述做了调整, 具体调整如下: 由于对应 20 个样品的 $t_{(1-\frac{\alpha}{2})}$ 为 2.09 ($\alpha = 0.05$)。	增加了 $\alpha = 0.05$, 更易于使用者理解和使用。

表 2 本标准与 ISO 12099:2017 的技术性差异及其原因 (续)

本标准章条编号	技术性差异	原因
9	将待测试样的描述移到第 6 章样品中 6.3 中说明。	为便于读者理解和使用。
9.1	增加: 近红外分析时, 样品温度不可过高或过低, 应在常温范围内。	ISO 标准文本中只提到所制备样品的温度应在验证时样品温度的范围之内。但未对温度提出具体要求, 增加此说明, 更便于标准的实施。
9.2	增加: 试样近红外光谱预测值的测定相关描述。	将 ISO 标准 11.1 中与样品测试原则和方法相关叙述移至此条描述, 更便于读者理解和使用。
10.1	将质控样品的描述移到第 6 章样品中 6.4 中说明。	为便于读者理解和使用。
11.1	将监控样品的描述移到第 6 章样品中 6.5 中说明。	为便于读者理解和使用。
附录 A	删除谷物及谷物精制料的描述。	本标准为饲料行业标准, 对适应对象进行了限定。
附录 C	C.1.1 参考方法将“获得 ISO 或其他国际认可和经过验证的检测方法”改为“获得国家标准化管理委员会或国际标准化组织认可和经过验证的检测方法”	适应中国标准情况。
附录 C	C.1.3 近红外光谱技术 (NIRS) 将“近红外光谱, 是指在 700 nm~2500 nm (14300 cm ⁻¹ ~4000 cm ⁻¹) 的近红外光范围内, 测量样品对光的吸收强度。”改为“近红外光谱, 是指在 770 nm~2500 nm (12900 cm ⁻¹ ~4000 cm ⁻¹) 的近红外光范围内, 测量样品对光的吸收强度。”	光谱范围与标准正文 5.1 保持一致, 统一为 770 nm~2500 nm。

四、采用国际标准

本标准修改采用 ISO 12099:2017 《动物饲料、谷物及谷物精制料的近红外光谱分析应用指南》。

五、与现行法律法规和强制性标准的关系

与现行法律、法规和强制性国家标准没有冲突。

六、重大分歧意见的处理经过和依据

无重大分歧意见。

七、标准作为强制性或推荐性标准的意见

标准作为推荐性标准发布。

八、贯彻标准的要求和措施建议

该标准为在饲料分析领域应用近红外光谱分析方法的指南性标准，无特殊实施条件。

九、废止现行有关标准的建议

无。

十、其他应予说明的事项

无。

附件 1

ISO 12099:2010 Animal feeding stuffs, cereals and milled cereal products-Guidelines for the application of near infrared spectrometry

**Animal feeding stuffs, cereals and
milled cereal products — Guidelines
for the application of near infrared
spectrometry**

*Aliments des animaux, céréales et produits de mouture des céréales —
Lignes directrices pour l'application de la spectrométrie dans le
proche infrarouge*

用户名称: 中国农业大学
订单号: 01020180816162041017
购买日期: 2018-08-16
销售机构: 国家标准化管理委员会标准信息中心
联系电话: 400
网址: <http://www.sacinfo.cn>



用户名称：中国农业大学
订单号：01020180816162041017
购买日期：2018-08-16
销售机构：国家标准化管理委员会标准信息中心
联系电话：400
网址：<http://www.sacinfo.cn>



版权声明

经ISO、IEC授权，国家标准化管理委员会负责ISO、IEC正版标准在中国的发行销售的监督管理工作以及ISO、IEC标准版权的保护工作。

国家标准化管理委员会标准信息中心 是国家标准化管理委员会授权的在中国的国际标准销售点。

未经授权，严禁对ISO、IEC标准出版物及相关技术文件进行复制、销售、传播和翻译出版，严禁将国际标准的任何部分通过电子信息网络或制作成标准数据库用于传播。

国家标准化管理委员会标准信息中心
电话：400 650 6190 网址：www.gb638.cn

ISO 12099:2017(E)

中国农业大学专用

防伪查询：KNqV-GPIN-2sTv-Xigl-wYFZ

网址：<http://www.sacinfo.cn>

销售机构：国家标准化管理委员会标准信息中心



COPYRIGHT PROTECTED DOCUMENT

© ISO 2017, Published in Switzerland

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents		Page
Foreword		iv
Introduction		v
1 Scope		1
2 Normative references		1
3 Terms and definitions		1
4 Principle		2
5 Apparatus		2
6 Calibration and initial validation		2
6.1 General.....		2
6.2 Reference methods.....		3
6.3 Outliers.....		3
6.4 Validation of calibration models.....		3
6.4.1 General.....		3
6.4.2 Bias correction.....		4
6.4.3 Slope adjustment.....		4
6.4.4 Expansion of calibration set.....		4
6.5 Changes in measuring and instrument conditions.....		4
7 Statistics for performance measurement		5
7.1 General.....		5
7.2 Plot the results.....		5
7.3 Bias.....		6
7.4 Root mean square error of prediction (s_{RMSEP}).....		8
7.5 Standard error of prediction (s_{SEP}).....		8
7.6 Slope.....		10
8 Sampling		12
9 Procedure		12
9.1 Preparation of test sample.....		12
9.2 Measurement.....		12
9.3 Evaluation of result.....		12
10 Checking instrument stability		13
10.1 Control sample.....		13
10.2 Instrument diagnostics.....		13
10.3 Instruments in a network.....		13
11 Running performance check of calibration		13
11.1 General.....		13
11.2 Control charts using the difference between reference and NIR results.....		14
12 Precision and accuracy		15
12.1 Repeatability.....		15
12.2 Reproducibility.....		15
12.3 Accuracy.....		15
12.4 Uncertainty.....		15
13 Test report		15
Annex A (informative) Guidelines for specific NIR standards		16
Annex B (informative) Examples of outliers and control charts		17
Annex C (informative) Supplementary terms and definitions		23
Bibliography		28

ISO 12099:2017(E)

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 on 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 the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 10, *Animal feeding stuffs*.

This second edition cancels and replaces the first edition (ISO 12099:2010), which has been technically revised.

Introduction

This document has been drafted using, as a basis, ISO 21543 | IDF 201, which was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 5, *Milk and milk products*, and the International Dairy Federation (IDF).

防伪查询：KNqV-GpFN-2sTv-Xigl-wYFZ

网址：<http://www.sacinfo.cn>

销售机构：国家标准化管理委员会标准信息中心

中国农业大学专用

中国农业大学专用

Animal feeding stuffs, cereals and milled cereal products — Guidelines for the application of near infrared spectrometry

1 Scope

This document gives guidelines for the determination by near infrared spectroscopy of constituents such as moisture, fat, protein, starch and crude fibre and parameters such as digestibility in animal feeding stuffs, cereals and milled cereal products.

The determinations are based on spectrometric measurement in the near infrared spectral region.

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:

- IEC Electropedia: available at <http://www.electropedia.org/>
- ISO Online browsing platform: available at <http://www.iso.org/obp>

3.1

near infrared instrument NIR instrument

apparatus which, when used under the conditions defined in this document, predicts *constituent contents* (3.3) and *technological parameters* (3.4) in *animal feeding stuffs* (3.2), cereals and milled cereal products through relationships to absorptions in the near infrared range

3.2

animal feeding stuffs

substance or product, including additives, whether processed, partially processed or unprocessed, intended to be used for oral feeding to animals

EXAMPLE Raw materials, fodder, meat and bone meal, mixed feed and other end products, pet food, etc.

3.3

constituent content

mass fraction of substances determined using the appropriate, standardized or validated chemical method

Note 1 to entry: The mass fraction is often expressed as a percentage.

Note 2 to entry: For examples of appropriate methods, see References [1] to [12].

EXAMPLE Moisture, fat, protein, crude fibre, neutral detergent fibre and acid detergent fibre.

ISO 12099:2017(E)

3.4 technological parameter

property or functionality of *animal feeding stuffs* (3.2), cereals and milled cereal products that can be determined using the appropriate, standardized or validated method(s)

Note 1 to entry: It is possible to develop and validate NIR methods for other parameters and sample types than listed above, as long as the procedure from this document is observed. The measuring units of the parameters determined follow the units used in the reference methods.

EXAMPLE Digestibility.

4 Principle

Spectral data in the near infrared region are collected and transformed to constituent or parameter concentrations by calibration models developed on representative samples of the products concerned.

5 Apparatus

5.1 Near infrared instruments.

Instruments based on diffuse reflectance or transmittance measurement covering the near infrared wavelength region of 770 nm to 2 500 nm ($12\,900\text{ cm}^{-1}$ to $4\,000\text{ cm}^{-1}$) or segments of this or at selected wavelengths or wavenumbers. The optical 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 noise and reproducibility, wavelength/wavenumber accuracy and wavelength/wavenumber precision (for scanning spectrophotometers).

The instrument should measure a sufficiently large sample volume or surface to eliminate any significant influence of inhomogeneity derived from 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 recommendation with respect to signal intensity for obtaining linearity and maximum signal/noise ratio.

5.2 Appropriate milling or grinding device, for preparing the sample (if needed).

NOTE Changes in grinding or milling conditions can influence NIR measurements due, for example, to heating which can drive off volatile components such as water.

6 Calibration and initial validation

6.1 General

The instrument shall be calibrated before use. Calibration involves the comparison with a reference and adjustment processes to the instrument. Because a number of different calibration systems can be applied with NIR instruments, no specific procedure can be given for calibration.

For an explanation of methods for calibration development, see Reference [16] and the respective manufacturer's manual. For the validation, it is important to have a sufficient number of representative samples, covering variations such as the following:

- a) combinations and composition ranges of major and minor sample components;
- b) seasonal, geographic and genetic effects on forages, feed raw material and cereals;
- c) processing techniques and conditions;
- d) storage conditions;

- e) sample and instrument temperature;
- f) instrument variations (i.e. differences between instruments).

NOTE For a solid validation, at least 20 samples are needed.

6.2 Reference methods

Internationally accepted reference methods for determination of moisture, fat, protein and other constituents and parameters should be used. See References [1] to [12] for examples.

The reference method used for calibration should be in statistical control. It is essential to know the precision of the reference method.

Where possible, references that provide measurement traceability to the SI (International system of units), such as certified reference materials, should be used.

6.3 Outliers

In many situations, statistical outliers are observed during calibration and validation. Outliers may be related to NIR data (spectral outliers, hereafter referred to as “x-outliers”) or errors in reference data or samples with a different relationship between reference data and NIR data (hereafter referred to as “y-outliers”); see [Figures B.1](#) to [B.5](#) for examples.

For the purpose of validation, samples are not to be regarded as outliers if they fulfil the following conditions:

- a) if they are within the working range of the constituents/parameters in the calibration(s);
- b) if they are within the spectral variation of the calibration samples, as, for example, estimated by Mahalanobis distance;
- c) if the spectral residual is below a limit defined by the calibration process;
- d) if the prediction residual is below a limit defined by the calibration process.

If a sample appears as an outlier, then it should be checked initially to see if it is an x-outlier. If it exceeds the x-outlier limits defined for the calibration, it should be removed. If it is not an x-outlier, then both the reference value and the NIR predicted value should be checked, e.g. by repeated measurements. If these confirm the original values, then the sample should not be deleted and the validation statistics should include this sample. If the repeat values show that either the original reference values or the NIR predicted ones were in error, then the new values should be used.

6.4 Validation of calibration models

6.4.1 General

Before use, calibration equations shall be validated locally on an independent test set that is representative of the sample population to be analysed. For the determination of bias, slope and for the determination of standard error of prediction (SEP, see [7.5](#)), at least 20 samples are needed. Validation shall be carried out for each sample type, constituent/parameter, temperature and other factors known to affect or expected to have an effect the measurement. The calibration is valid only for the variations, i.e. sample types, range and temperature, used in the validation.

NOTE 1 Calibration models can only be used in the range they have been validated.

Results obtained on the independent test set are plotted, reference against NIR, and residuals against reference results, to give a visual impression of the performance of the calibration. The SEP is calculated (see [7.5](#)) and the residual plot of data corrected for mean systematic error (bias) is examined for outliers, i.e. samples with a residual exceeding $\pm 3 s_{SEP}$.

ISO 12099:2017(E)

If the validation process shows that the model cannot produce acceptable statistics, then it should not be used.

NOTE 2 What will be acceptable will depend, for example, on the performance of the reference method, the covered range, the purpose of the analysis, etc., and is up to the parties involved to decide.

Where available and suitable, reference materials or certified reference materials can be used as part of validation of calibration models.

The next step is to fit NIR, y_{NIRS} , and reference data, y_{ref} , by linear regression ($y_{\text{ref}} = a + b \times y_{\text{NIRS}}$) to produce statistics that describe the validation results.

6.4.2 Bias correction

The data are also examined for a bias between the methods. If the difference between means of the NIR predicted and reference values is significantly different from zero, then this indicates that the calibration is biased. A bias may be removed by adjusting the constant term (see 7.3) in the calibration equation.

6.4.3 Slope adjustment

If the slope, b , is significantly different from 1, the calibration is skewed.

Adjusting the slope/intercept of the calibration is generally not recommended unless the calibration is applied to new types of samples or instruments. If a reinvestigation of the calibration does not detect outliers, especially outliers with high leverage, it is preferable to expand the calibration set to include more samples. However, if the slope is adjusted, the calibration should then be tested on a new independent test set.

6.4.4 Expansion of calibration set

If the accuracy of the calibration does not meet expectations, the calibration set should be expanded to include more samples or a new calibration should be made. In all cases when a new calibration is developed on an expanded calibration set, the validation process should be repeated on a new validation set. If necessary, expansion of the calibration set should be repeated until acceptable results are obtained on a validation set.

6.5 Changes in measuring and instrument conditions

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

For example, 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 grass silages from one area may not give the same accuracy on silages from another area if the genetic, growing and processing 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.

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 or slope / intercept adjustments to calibration equations. In many cases, it will be necessary to standardize the two instruments against each other before calibration equations can be transferred^[16]. Standardization procedures can be used to transfer calibrations between instruments of different types provided that samples are measured in the same way (reflectance, transmittance) and that the spectral region is common.

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.

7 Statistics for performance measurement

7.1 General

The performances of a prediction model shall be determined by a set of validation samples. This set consists of samples which are independent of the calibration set. In a plant, it will be new batches; in agriculture, it will be a new crop or a new experiment location.

This set of samples shall be carefully analysed following the reference methods. The care to analyse validation samples shall be emphasized and the precision of these results is more important for the validation set than for the samples used at the calibration phase.

The number of validation samples shall be at least 20 to compute the statistics with some confidence.

The NIR protocol used for the determination of the performances of the prediction model shall be the same as that used in routine (one measurement or two measurements).

7.2 Plot the results

It is important to visualize the results in plots, i.e. reference vs. predicted values or residuals vs. predicted values.

The residuals are defined by [Formula \(1\)](#):

$$e_i = y_i - \hat{y}_i \quad (1)$$

where

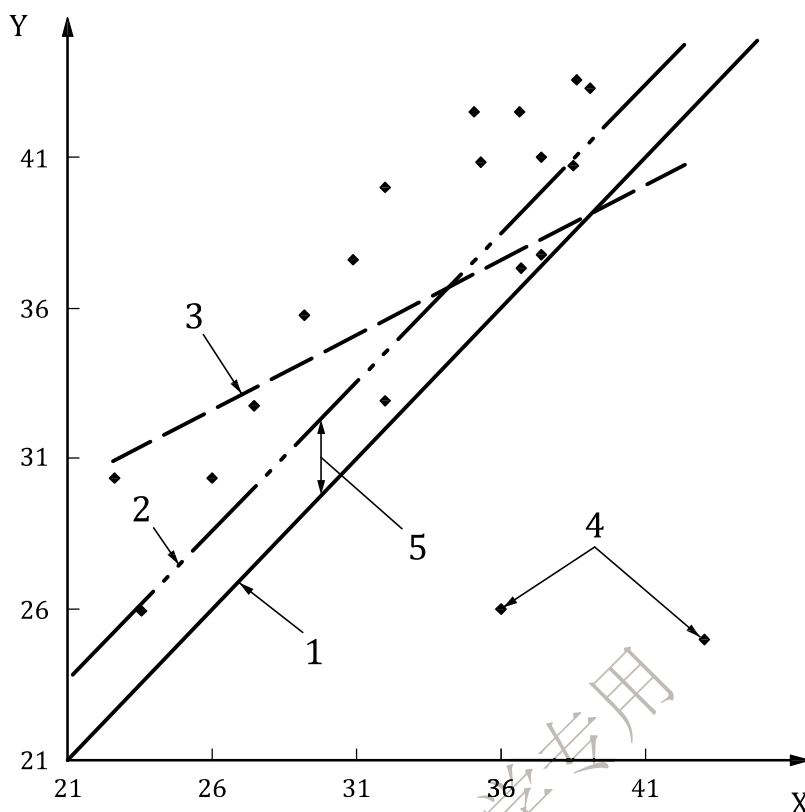
y_i is the i^{th} reference value (y_{ref});

\hat{y}_i is the i^{th} predicted value (y_{NIRS}) obtained when applying the multivariate NIR model.

The way the differences are calculated will give a negative bias when the predictions are too high and a positive one when the predictions are too low compared to the reference values.

A plot of the data immediately gives an overview of the correlation, the bias, the slope and the presence of obvious outliers (see [Figure 1](#)).

ISO 12099:2017(E)



Key

- 1 45° line (ideal line with bias = 0 and slope = 1)
- 2 (45° - bias) line
- 3 linear regression line
- 4 outliers
- 5 bias

NOTE The outliers (key 4) have a strong influence on the calculation of the slope and should be removed if the results are to be used for adjustments.

Figure 1 — Scatter plot for a validation set, $y_{ref} = f(a + b \times y_{NIRS})$

7.3 Bias

Most of the time, a bias or systematic error is observed with NIR models. Bias can occur due to several causes: new samples of a type not previously seen by the model, drift of the instrument, drift in wet chemistry, changes in the process, in the sample preparation, etc.

With n , the number of independent samples, the bias (or offset) is the mean difference and can be defined by Formula (2):

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i \tag{2}$$

where e_i is the residual as defined by Formula (1) resulting in Formula (3):

$$\bar{e} = \frac{1}{n} \left(\sum_{i=1}^n y_i - \sum_{i=1}^n \hat{y}_i \right) = \bar{y} - \bar{\hat{y}} \tag{3}$$

where

- y_i is the i^{th} reference value;
- \hat{y}_i is the i^{th} predicted value;
- \bar{y} is the mean of the reference values;
- $\bar{\hat{y}}$ is the mean of the predicted values.

The significance of the bias is checked by a t test. The calculation of the bias confidence limits (BCLs), T_b , determines the limits for accepting or rejecting formula performance on the small set of samples chosen from the new population; see [Formula \(4\)](#):

$$T_b = \pm \left[t_{(1-\alpha/2)} s_{\text{SEP}} \right] / \sqrt{n} \tag{4}$$

where

- α is the probability of making a type I error;
- t is the appropriate student's t value for a two-tailed test with degrees of freedom associated with s_{SEP} and the selected probability of a type I error;
- n is the number of independent samples;
- s_{SEP} is the standard error of prediction (defined in [7.5](#)).

As an example, with $n = 20$, and $s_{\text{SEP}} = 1$, the BCLs are as in [Formula \(5\)](#):

$$T_b = \pm (2,09 \times 1) / \sqrt{20} = \pm 0,48 \tag{5}$$

This means that the bias tested with 20 samples shall be higher than 48 % of the standard error of prediction to be considered as different from zero.

Table 1 — Values of the t distribution with a probability $\alpha = 0,05$ (5 %)

n	t values	n	t values
10	2,23	75	1,99
15	2,13	100	1,98
20	2,09	200	1,97
30	2,04	500	1,96
40	2,02	1000	1,96
50	2,01	—	—

NOTE The Excel function "TINV" can be used.^a

^a Excel is the trade name of a product supplied by Microsoft. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

ISO 12099:2017(E)

7.4 Root mean square error of prediction (s_{RMSEP})

The s_{RMSEP} is defined by [Formula \(6\)](#):

$$s_{RMSEP} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (6)$$

where

- e_i is the residual of the i^{th} sample;
- n is the number of independent samples.

This value can be compared with s_{SEC} and s_{SECV} (see [Annex C](#)).

s_{RMSEP} includes the random error (s_{SEP}) and the systematic error (bias). It includes also the error of the reference methods (as do s_{SEC} and s_{SECV}); see [Formula \(7\)](#):

$$s_{RMSEP} = \sqrt{\frac{(n-1)}{n} s_{SEP}^2 + \bar{e}^2} \quad (7)$$

where

- n is the number of independent samples;
- s_{SEP} is the standard error of prediction (defined in [7.5](#));
- \bar{e} is the bias or systematic error.

There is no direct test for s_{RMSEP} . This is the reason to separate the systemic error (bias or \bar{e}) and the random error s_{SEP} .

7.5 Standard error of prediction (s_{SEP})

The standard error of prediction (s_{SEP}), i.e. standard deviation of the residuals, which expresses the accuracy of routine NIR results corrected for the mean difference (bias) between routine NIR and reference method, can be calculated by using [Formula \(8\)](#):

$$s_{SEP} = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-1}} \quad (8)$$

where

- n is the number of independent samples;
- e_i is the residual of the i^{th} sample;
- \bar{e} is the bias or systematic error.

The s_{SEP} should be related to the s_{SEC} (respectively, s_{SECV} ; see [Annex C](#)) to check the validity of the calibration model for the selected validation set.

The unexplained error confidence limits (UECLs), T_{UE} , are calculated from an F -test (ratio of two variances) (see Reference [18] and Table 2). See Formula (9):

$$T_{UE} = s_{SEC} \sqrt{F_{(\alpha, v, M)}} \quad (9)$$

where

- s_{SEC} is the standard error of calibration (see Annex C);
- α is the probability of making a Type I error;
- v is $n - 1$ numerator degrees of freedom associated with s_{SEP} of the test set;
- n is the number of samples in the validation process;
- M is $n_c - p - 1$ denominator degrees of freedom associated with s_{SEC} (standard error of calibration) [n_c is the number of calibration samples, p is the number of terms or PLS factors of the model or weights in the case of ANN (see Annex C). In ANN, weights are all unknown parameters in the model].

NOTE 1 s_{SEC} can be replaced by s_{SECV} , which is a better statistic than s_{SEC} ; very often, s_{SEC} is too optimistic; ($s_{SEP} > s_{SECV} > s_{SEC}$).

EXAMPLE Where $n = 20$, $\alpha = 0,05$, $M = 100$, and $s_{SEP} = 1$, gives the following value: $T_{UE} = 1,30$.

With 20 samples, a s_{SEP} that is up to 30 % larger than the s_{SEC} can be accepted.

NOTE 2 The Excel¹⁾ function "FINV" can be used.

The F -test cannot be used to compare two calibrations on the same validation set. It needs (as here) two independent sets to work. Another test shall be used to compare two or more models on the same data set.

1) Excel is the trade name of a product supplied by Microsoft. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO of the product named. Equivalent products may be used if they can be shown to lead to the same results.

ISO 12099:2017(E)

Table 2 — F values and squared root of the F values in function of the degrees of freedom of the numerator associated with s_{SEP} and of the denominator associated with s_{SEC}

		$F(\alpha, \nu, M)$							$\sqrt{F(\alpha, \nu, M)}$				
		df (s_{SEC})							df (s_{SEC})				
df (s_{SEP})		50	100	200	500	1 000	df (s_{SEP})		50	100	200	500	1 000
10		2,03	1,93	1,88	1,85	1,84	10		1,42	1,39	1,37	1,36	1,36
11		1,99	1,89	1,84	1,81	1,80	11		1,41	1,37	1,36	1,34	1,34
12		1,95	1,85	1,80	1,77	1,76	12		1,40	1,36	1,34	1,33	1,33
13		1,92	1,82	1,77	1,74	1,73	13		1,39	1,35	1,33	1,32	1,32
14		1,89	1,79	1,74	1,71	1,70	14		1,38	1,34	1,32	1,31	1,30
15		1,87	1,77	1,72	1,69	1,68	15		1,37	1,33	1,31	1,30	1,29
16		1,85	1,75	1,69	1,66	1,65	16		1,36	1,32	1,30	1,29	1,29
17		1,83	1,73	1,67	1,64	1,63	17		1,35	1,31	1,29	1,28	1,28
18		1,81	1,71	1,66	1,62	1,61	18		1,30	1,31	1,29	1,27	1,27
19		1,80	1,69	1,64	1,61	1,60	19		1,34	1,30	1,28	1,27	1,26
29		1,69	1,58	1,52	1,49	1,48	29		1,30	1,26	1,23	1,22	1,22
49		1,60	1,48	1,42	1,38	1,37	49		1,27	1,22	1,19	1,17	1,17
99		1,53	1,39	1,32	1,28	1,26	99		1,24	1,18	1,15	1,13	1,12
199		1,48	1,34	1,26	1,21	1,19	199		1,22	1,16	1,12	1,10	1,09
499		1,46	1,31	1,22	1,16	1,13	499		1,21	1,14	1,11	1,08	1,07
999		1,45	1,30	1,21	1,14	1,11	999		1,20	1,14	1,10	1,07	1,05

NOTE 1 See explanation to [Formula \(9\)](#).
 NOTE 2 df is the degree of freedom; $n - 1$ for s_{SEP} ; and $n_c - p - 1$ for s_{SEC} .

7.6 Slope

The slope, b , of the simple regression $y = a + b \cdot \hat{y}$ is often reported in the NIR reports and publications.

Notice that the slope shall be calculated with the reference values as the dependent variable and the predicted NIR values as the independent variable, if the calculated slope is intended to be used for adjustment of NIR results (like in the case of the inverse multivariate regression used to build the prediction model).

From the least squares fitting, slope and intercept are calculated by [Formula \(10\)](#) and [Formula \(11\)](#), respectively:

$$b = \frac{s_{\hat{y}y}}{s_y^2} \tag{10}$$

where

$s_{\hat{y}y}$ is the covariance between reference and predicted values;

s_y^2 is the variance of the n predicted values.

$$a = \bar{y} - b\bar{\hat{y}} \quad (11)$$

where

\bar{y} is the mean of the reference values;

b is the slope;

$\bar{\hat{y}}$ is the mean of the predicted values.

As for the bias, a t test can be calculated to check the hypothesis that $b = 1$ as in [Formula \(12\)](#):

$$t_{\text{obs}} = |b - 1| \cdot \sqrt{\frac{s_y^2 \cdot (n - 1)}{s_{\text{res}}^2}} \quad (12)$$

where

s_y^2 is the variance of the n predicted values;

n is the number of independent samples;

s_{res} is the residual standard deviation.

The residual standard deviation, s_{res} , is defined in [Formula \(13\)](#):

$$s_{\text{res}} = \sqrt{\frac{\sum_{i=1}^n [y_i - (a + b\hat{y}_i)]^2}{n - 2}} \quad (13)$$

where

n is the number of independent samples;

a is the intercept (see [Formula \(11\)](#));

b is the slope (see [Formula \(10\)](#));

y_i is the i^{th} reference value;

\hat{y}_i is the i^{th} predicted value obtained when applying the multivariate NIR model.

NOTE s_{res} is like s_{SEP} when the predicted values are corrected for slope and intercept. Be aware to not confuse bias and intercept. See also [Figure 1](#).

The bias equals the intercept only when the slope is exactly one.

The slope, b , will be considered as different from 1 when [Formula \(14\)](#) applies:

$$t_{\text{obs}} \geq t_{(1-\alpha/2)} \quad (14)$$

ISO 12099:2017(E)

where

t_{obs} is the observed t value, calculated according to [Formula \(12\)](#);

$t_{(1-\alpha/2)}$ is the t value obtained from [Table 1](#) for a probability of $\alpha = 0,05$ (5 %).

A too-narrow range or uneven distribution will lead to useless correction of the slope even when the s_{SEP} is correct. The slope can only be adjusted when the validation set covers a large part of the calibration range.

EXAMPLE For $n = 20$ samples with a residual standard deviation [see [Formula \(13\)](#)] of 1, a standard deviation of the predicted values of $S_{\hat{y}} = 2$ and a calculated slope of $b = 1,2$, the observed t_{obs} value is 1,7 and then the slope is not significantly different from 1 as the t value (see [Table 1](#)) for $n = 20$ samples is 2,09. If the slope is 1,3, the t_{obs} value is 2,6 and then the slope is significantly different than 1.

8 Sampling

Sampling is not part of the method specified in this document.

NOTE Recommended sampling procedures are given in ISO 6497 and ISO 24333.

It is important that the laboratory receives a sample which is truly representative and has not been damaged or changed during transport or storage.

9 Procedure

9.1 Preparation of test sample

All laboratory samples should usually be kept under conditions that will not change the composition of the sample from the time of sampling to the time of commencing the procedure.

The preparation of samples for routine measurements needs to be made in the same way as the preparation of the validation samples. It is necessary to apply standard conditions.

Before the analysis, the sample should be taken in such a way as to obtain a sample representative of the material to be analysed.

For specific procedures, see specific NIR standards.

Guidelines for specific NIR standards are given in [Annex A](#).

9.2 Measurement

Follow the instructions of the instrument manufacturer/supplier.

The prepared sample should reach a temperature within the range included in the validation.

9.3 Evaluation of result

For the routine results to be valid, they shall be within the range of the calibration model used.

Results obtained on samples detected as spectral outliers cannot be regarded as reliable.

If multiple measurements are made on the same sample, calculate the arithmetic mean if the repeatability conditions (see [12.1](#)) are observed.

For the expression of results, refer to specific NIR standards.

10 Checking instrument stability

10.1 Control sample

At least one control sample should be measured at least once per day to check instrument hardware stability and to detect any malfunction. Knowledge of the true concentration of the analyte in the control sample is not necessary. The sample material should be stable and, as far as possible, resemble the samples to be analysed. The parameter(s) measured should be stable and, as far as possible, identical to or at least biochemically close to the sample analyte. A sample is prepared as described in 9.1 and stored in such a way as to maximize the storage life. These samples are normally stable for lengthy periods but the stability should be tested in the actual cases. Control samples should be overlapped to secure uninterrupted control.

The recorded day-to-day variation should be plotted in control charts and investigated for significant patterns or trends.

10.2 Instrument diagnostics

For scanning spectrophotometers the wavelength/wavenumber (see 5.1) accuracy and precision should be checked at least once a week, or more frequently if recommended by the instrument manufacturer, and the results should be compared to specifications and requirements (see 5.1).

A similar check of the instrument noise shall be carried out weekly or at intervals recommended by the manufacturer.

10.3 Instruments in a network

If several instruments are used in a network, special attention shall be given to standardization of the instruments according to the manufacturer's recommendations.

11 Running performance check of calibration

11.1 General

The suitability of the calibration for the measurement of individual samples should be checked. The outlier measures used in the calibration development and validation can be applied, e.g. Mahalanobis distance and spectral residuals. In most instruments, this is done automatically.

If the sample does not pass the test, i.e. the sample does not fit into the population of the samples used for calibration and/or validation, it cannot be determined by the prediction model, unless the model is changed. Thus, the outlier measures can be used to decide which samples should be selected for reference analysis and included in a calibration model update.

If the calibration model is found to be suitable for the measured sample, the spectrum is evaluated according to the validated calibration model.

NIR methods should be validated continuously against reference methods to secure steady optimal performance of calibrations and observance of accuracy. The frequency of checking the NIR method should be sufficient to ensure that the method is operating under steady control with respect to systematic and random deviations from the reference method. The frequency depends *inter alia* on the number of samples analysed per day and the rate of changes in sample population.

The running validation should be performed on samples selected randomly from the pool of analysed samples. It may be necessary to resort to some sampling strategy to ensure a balanced sample distribution over the entire calibration range, e.g. segmentation of concentration range and random selection of test samples within each segment or to ensure that samples with a commercially important range are covered.

ISO 12099:2017(E)

The number of samples for the running validation should be sufficient for the statistics used to check the performance. For a solid validation, at least 20 samples are needed (to expect a normal distribution of variance). One can fill in the results of the independent validation set for starting the running validation. To continue with about 5 to 10 samples every week is sufficient to monitor the performance properly. Using fewer samples it is hard to take the right decision in case one of the results is outside the control limits.

11.2 Control charts using the difference between reference and NIR results

Results should be assessed by control charts, plotting running sample numbers on the abscissa and the difference between results obtained by reference and NIR methods on the ordinate; $\pm 2 s_{SEP}$ (95 % probability) and $\pm 3 s_{SEP}$ (99,8 % probability) may be used as warning and action limits where the s_{SEP} has been obtained on a test set collected independently of calibration samples.

If the calibration and the reference laboratories are performing as they should, then only 1 point in 20 points should plot outside the warning limits and 2 points in 1 000 points outside the action limits.

Control charts should be checked for systematic bias drifts from zero, systematic patterns and excessive variation of results. General rules applied for Shewart control charts may be used in the assessment. However, too many rules applied simultaneously may result in too many false alarms.

The following rules used in combination have proved to be useful in detection of problems:

- a) one point outside either action limit;
- b) two out of three points in a row outside a warning limit;
- c) nine points in a row on the same side of the zero line.

Additional control charts plotting other features of the running control (e.g. mean difference between NIR and reference results) and additional rules may be applied to strengthen decisions.

In the assessment of results, it should be remembered that s_{SEP} and measured differences between NIR and reference results also include the imprecision of reference results. This contribution can be reduced to a negligible part if the imprecision of reference results is reduced to less than one third of the s_{SEP} ^[17].

To reduce the risk of false alarms, the control samples should be analysed independently (in different series) by both NIR spectrometry and reference methods to avoid the influence of day-to-day systematic differences in, for example, reference analyses.

If the warning limits are often exceeded and the control chart only shows random fluctuations (as opposed to trends or systematic bias), the control limits may have been based on a too optimistic s_{SEP} value. An attempt to force the results within the limits by frequent adjustments of the calibration will not improve the situation in practice. The s_{SEP} should instead be re-evaluated using the latest results.

If the calibration equations after a period of stability begin to move out of control, the calibration should be updated. Before this is done, an evaluation should be made of whether the changes could be due to changes in reference analyses, unintended changes in measuring conditions (e.g. caused by a new operator), instrument drift or malfunction, etc. In some cases, a simple adjustment of the constant term in the calibration equation may be sufficient (an example is shown in [Figure B.6](#)). In other cases, it may be necessary to run a complete re-calibration procedure, where the complete or a part of the basic calibration set is expanded to include samples from the running validation, and perhaps additional samples selected for this purpose (an example is shown in [Figure B.7](#)).

Considering that the reference analyses are in statistical control and the measuring conditions and instrument performance are unchanged, significant biases or increased s_{SEP} values can be due to changes in the chemical, biological or physical properties of the samples compared to the underlying calibration set.

Other control charts, for example, using z-scores, may be used.

12 Precision and accuracy

12.1 Repeatability

The repeatability, i.e. the difference between two individual single test results, obtained with the same method on identical test material in the same laboratory by the same operator using the same equipment within a short interval of time, which should not be exceeded in more than 5 % of cases, depends on the sample material, the analyte, sample and analyte variation ranges, method of sample presentation, instrument type and the calibration strategy used. The repeatability should be determined in each case.

12.2 Reproducibility

The reproducibility, i.e. the difference between two individual single test results, obtained on identical test material by different laboratories and by different operators at different times, which should not be exceeded in more than 5 % of cases, depends on the sample material, the analyte, sample and analyte variation ranges, method of sample presentation, instrument type and the calibration strategy used. The reproducibility should be determined in each case.

12.3 Accuracy

The accuracy, which includes uncertainty from systematic deviation from the true value on the individual sample (trueness) and uncertainty from random variation (precision), depends *inter alia* on the sample material, the analyte, sample and analyte variation ranges, method of sample presentation, instrument type and the calibration strategy used. The accuracy should be determined in each case. The reported s_{SEP} and s_{RMSEP} values also include uncertainty of reference results which may vary from case to case.

12.4 Uncertainty

Uncertainty, U_e , is a parameter characterizing the dispersion of values that can reasonably be attributed to the result. For NIR predicted results, the uncertainty is usually expressed by [Formula \(15\)](#):

$$U_e = \pm 2s_{RMSEP} \quad (15)$$

If a multiplier of 2 is used, this can be understood to indicate limits corresponding to 95 % probability of the true value lying within the range $\pm U_e$.

The s_{RMSEP} shall be determined locally.

13 Test report

The test report shall specify:

- a) all information necessary for complete identification of the sample;
- b) the test method used, with reference to the relevant International Standard;
- c) all operating conditions not specified in this document, or regarded as optional;
- d) any circumstances which may have influenced the results;
- e) the test result(s) obtained;
- f) the current s_{SEP} and bias (if statistically significant), estimated from running a performance test on at least 20 test samples (see [Clause 11](#)).

Annex A (informative)

Guidelines for specific NIR standards

Specific NIR standards may be developed for specific calibrations for the determination of specific constituents and parameters in animal feeding stuff, cereals and milled cereal products by NIR spectrometry.

A specific standard should

- respect the recommendation of this document for the process of validation,
- indicate particular specifications to be met as range of concentration, accuracy and precision,
- fix the sources of variation that shall be taken into account in the development of the model, and
- specify the procedure, the calculation and the expression of results.

A specific standard is not made for a particular apparatus and calibration.

These standards should follow the ISO format and provide information regarding

- the type of samples and constituents/parameters determined followed by “near infrared spectrometry” and the calibration model(s) used in the title and the scope,
- the reference methods used for the validation under “Normative references”,
- the spectroscopic principle (e.g. NIR, NIT) and calibration principle (e.g. PLS, ANN),
- the procedure(s) including preparation of the test sample(s), measurement and quality control, and
- specifications to be followed in terms of accuracy, precision and range; for example, as in [Table A.1](#) for a validation set.

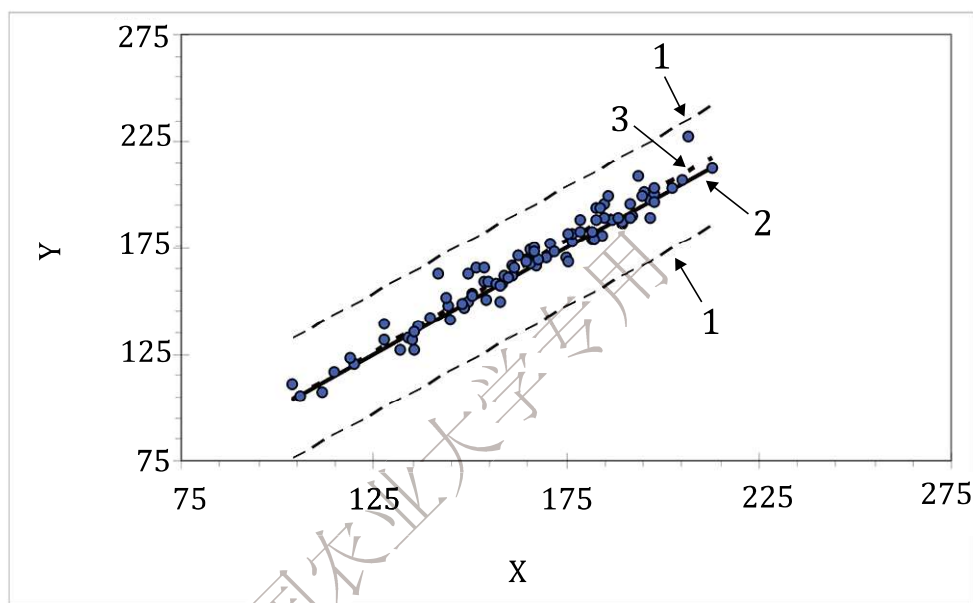
Table A.1 — Example for specifications given for a validation set

Component	Model	N	Precision (<i>s</i> _{SEP})	Min %	Max %	RSQ
Fat	ANN	183	0,50	2,8	12,9	0,94
Moisture	ANN	183	0,47	9,2	12,3	0,83
Protein	ANN	179	0,72	11,0	29,1	0,96
Fibre	ANN	123	1,11	0,5	18,0	0,90
Starch	PLS	113	1,80	7,8	50,2	0,92

Annex B (informative)

Examples of outliers and control charts

[Figure B.1](#) shows the determination of crude protein in forages, which is an example with no outliers. The results were obtained on an independent test set of 95 samples using the developed calibration equation: $s_{SEP} = 4,02$; $s_{RMSEP} = 6,05$; Slope, $b = 1,04$.



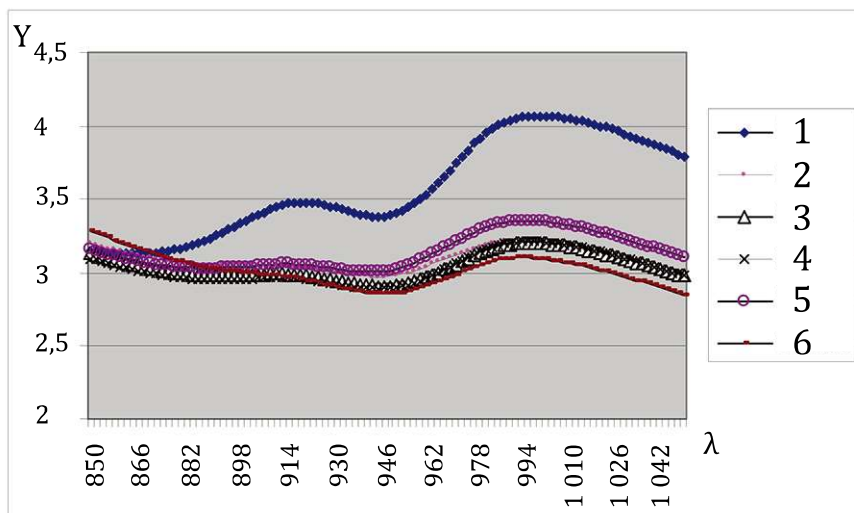
Key

- | | | | |
|---|---|---|-----------|
| 1 | $\pm 3 s_{SEP}$ limits, | X | NIR |
| 2 | 45 degree line (ideal line with slope = 1 and bias = 0) | Y | reference |
| 3 | regression line | | |

Figure B.1 — No outliers

ISO 12099:2017(E)

Figure B.2 shows an absorbance spectra with X-outlier. The series 1 (upper) spectrum indicates a spectral outlier.

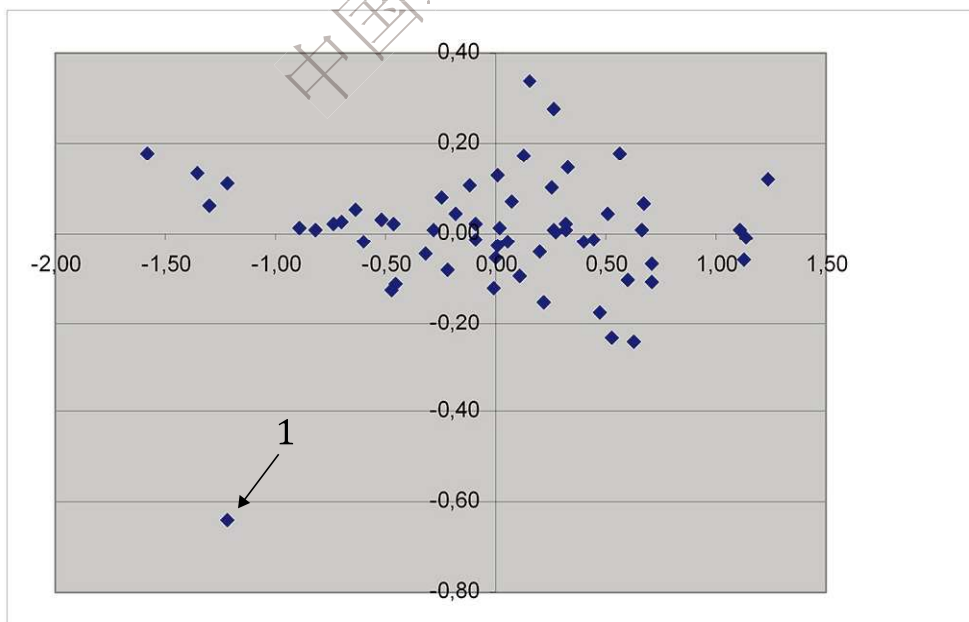


Key

- | | |
|------------|-------------------|
| 1 series 1 | 5 series 5 |
| 2 series 2 | 6 series 6 |
| 3 series 3 | λ wavelength (nm) |
| 4 series 4 | Y absorbance |

Figure B.2 — Absorbance spectra with X-outlier

Figure B.3 shows a PCA score plot with X-outlier (1).

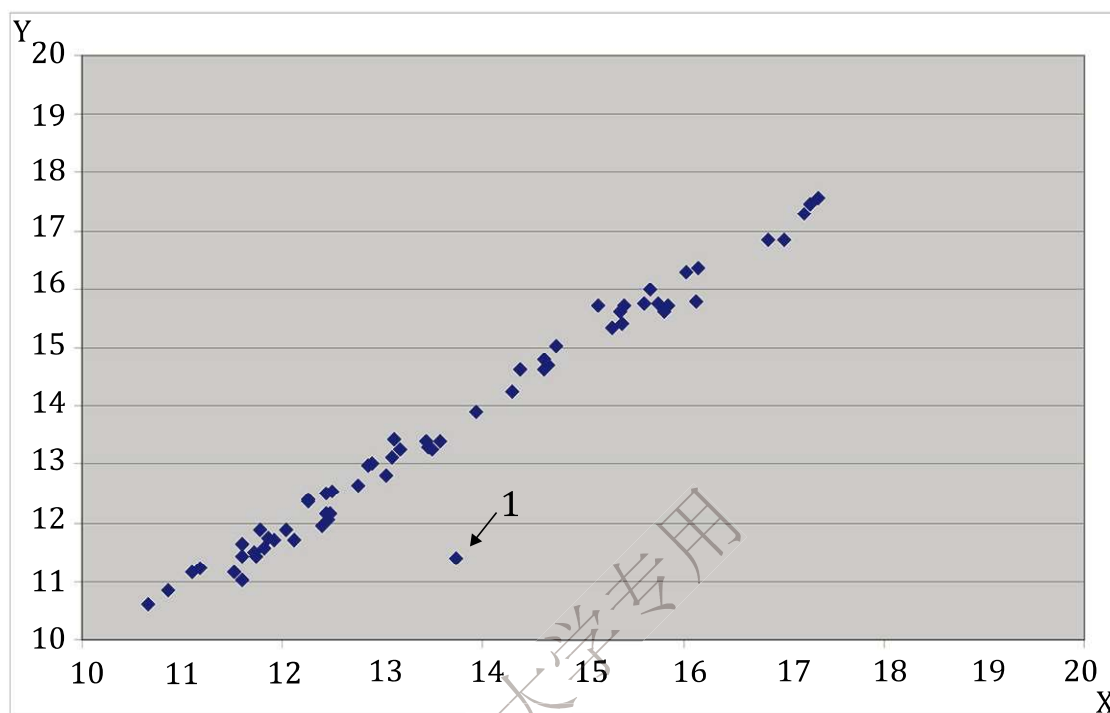


Key

- 1 outlier

Figure B.3 — PCA score plot with X-outlier (1)

Figure B.4 shows a scatter plot with Y-outlier (1). The plot of reference vs. predicted values (or vice versa) shows one sample (1) that strongly deviates from the other samples. If the reason for this deviation is not related to NIR data (X-outlier), this sample will be a Y-outlier due to erroneous reference data or a different relationship between reference data and spectral data.

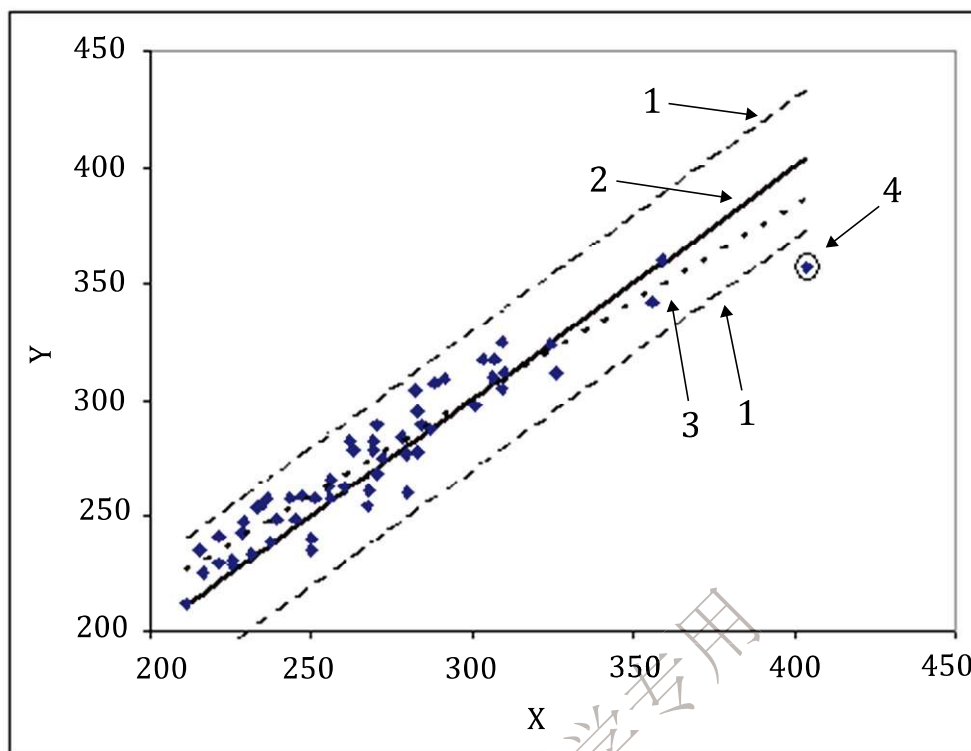
**Key**

- 1 outlier
- X NIR
- Y reference

Figure B.4 — Scatter plot with Y-outlier (1)

ISO 12099:2017(E)

Figure B.5 shows the determination of ADF in forages with Y-outlier (1).



Key

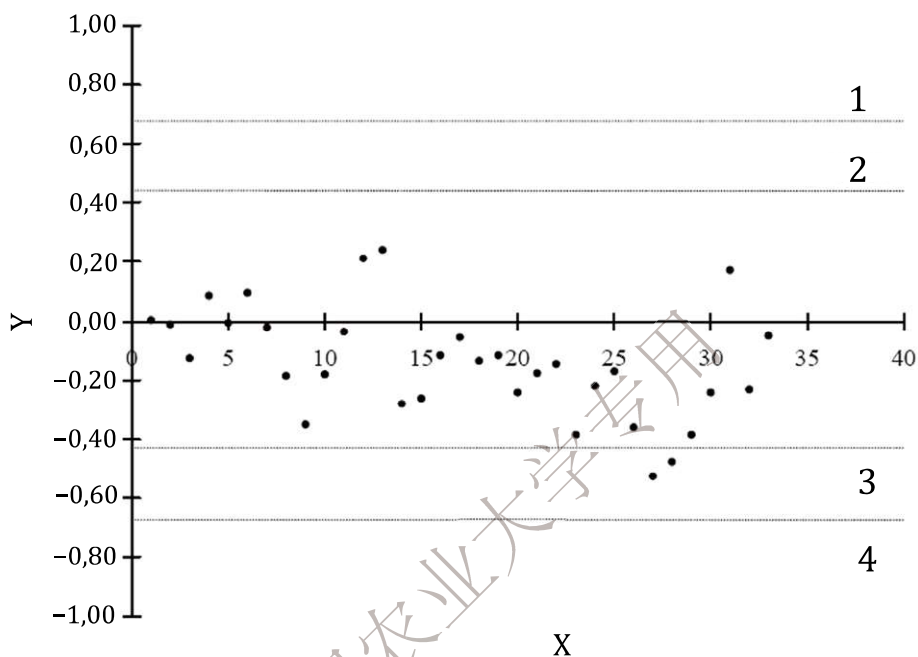
- 1 ±3 s_{SEP} limits
- 2 45 degree line
- 3 regression line
- 4 outlier

Figure B.5 — Determination of ADF in forages with Y-outlier (1)

Figure B.6 shows a control chart for determination of percent fat in cereals. No points are outside the upper action limit (UAL) or the lower action limit (LAL). However, 9 points in row (e.g. 14 to 22) are on the same side of the zero line. That indicates a bias problem. Two points (27 and 28) out of 3 points are outside the lower warning limit (LWL) but none are outside the upper warning limit (UWL). This also indicates a bias problem. No increase in random variation is observed. The spread is still less than 3 s_{SEP} .

In conclusion, the calibration should be bias adjusted.

(Difference Reference - NIR is plotted. UAL/LAL = 3 s_{SEP} and UWL/LWL = 2 s_{SEP})



Key

- | | |
|-----------------------------|-------------------|
| 1 upper action limit (UAL) | X run number |
| 2 upper warning limit (UWL) | Y reference - NIR |
| 3 lower warning limit (LWL) | |
| 4 lower action limit (LAL) | |

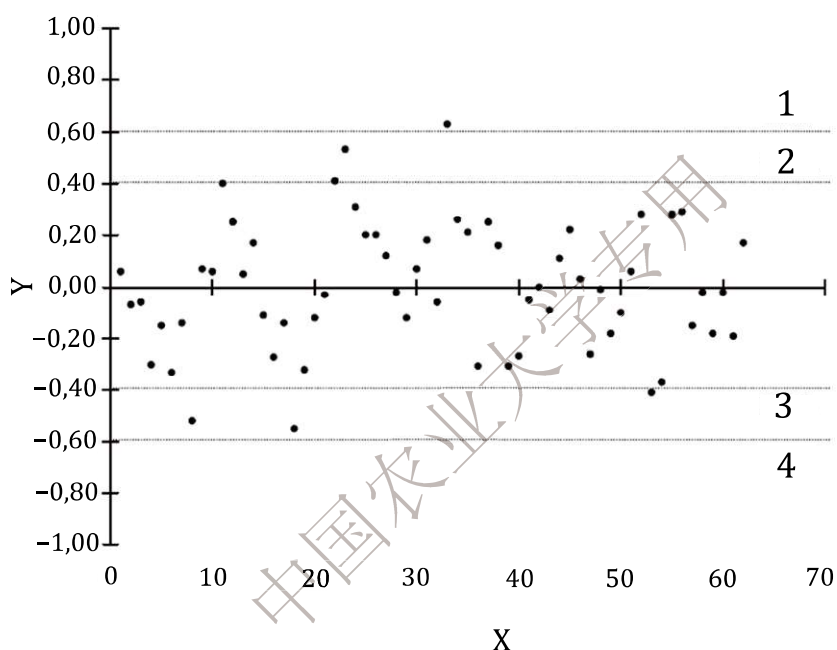
Figure B.6 — Control chart for determination of percent fat in cereals

ISO 12099:2017(E)

Figure B.7 shows a control chart for determination of XXXX in YYY in the range 44 % to 57 %, where recalibration was performed at point 35.

Viewing the first 34 points, one point is outside the UAL. This indicates a serious problem. Two points (22 and 23) out of 3 points are outside the UWL. Two separate points are also outside the LWL. The spread is uniform around the zero line (the 9-points rule is obeyed) but 5 point out of 34 points are outside the 95 % confidence limits (UWL, LWL) and 1 point out of 34 points is outside the 99,9 % confidence limits (UAL, LAL). This is much more than expected.

One reason for this picture could be that the s_{SEP} value behind calculation of the limits is too optimistic. This means the limits should be widened. Another reason could be that the actual samples are somewhat different from the calibration samples. To test this possibility, the calibration set was extended to include the control samples and a new calibration was developed. The performance of this calibration was clearly better, as shown by the control samples number 35 to 62.



Key

- | | | | |
|---|---------------------------|---|-----------------|
| 1 | upper action limit (UAL) | X | run number |
| 2 | upper warning limit (UWL) | Y | reference - NIR |
| 3 | lower warning limit (LWL) | | |
| 4 | lower action limit (LAL) | | |

Figure B.7 — Control chart for determination of XXXX in YYY (range 44 % to 57 %)

Annex C (informative)

Supplementary terms and definitions

General	
Reference method	ISO or otherwise internationally recognized and validated method, which gives the “true” or “assigned” value of the parameter to be measured, including the measurement uncertainty.
Indirect method	Method that measures properties that are functionally related to the parameter(s) to be determined. The obtained signal is related to the “true” value(s) as determined by the reference method(s).
Near infrared spectroscopy (NIRS)	NIR spectroscopy is the measurement of the intensity of the absorption of near infrared light by a sample within the range 700 nm to 2 500 nm (14 300 to 4 000 cm ⁻¹). NIR instruments use either part of, the whole, or ranges that include this region (e.g. 400 nm to 2 500 nm). Multivariate calibration techniques are then used to relate a combination of absorbance values either to composition or to some property of the samples.
Near infrared reflectance (NIR)	NIR spectroscopy where the basic measurements are the absorption of near infrared light diffusely reflected back from the surface of a sample to be collected by a detector in front of the sample.
Near Infrared transmittance (NIT)	NIR spectroscopy where the basic measurements are the absorption of near infrared light that has travelled through a sample and is then collected by a detector behind the sample.
NIR network(s)	A number of near infrared instruments operated using the same calibration models. Instruments in a network are usually standardized so that the differences in predicted values for a set of standard samples are minimized.
Standardization (of instruments)	A process whereby a group of near infrared instruments are adjusted so that they predict similar values when operating the same calibration model. A number of techniques can be used but these can be broadly defined as either pre-prediction methods where the spectra of samples are adjusted to minimize the differences between the response of a “master” instrument and each instrument in the group and “post-prediction” methods where linear regression is used to adjust the predicted values produced by each instrument to make them as similar as possible to those from a “master” instrument.
Z-Score	Performance criterion calculated by dividing the difference between the NIR predicted result and the reference value by a target value for the standard deviation, e.g. the RMSEP.

ISO 12099:2017(E)

Calibration techniques	
PCA	<p>Principle component (PC) analysis is a form of data compression. For a set of samples, it works solely with the X (spectral) data and finds principal components (factors) according to a rule that says that each PC expresses the maximum variation in the data at any time and is uncorrelated with any other PC. The first PC expresses as much as possible of the variability in the original data. Its effect is then subtracted from the X data and a new PC derived again expressing as much as possible of the variability in the remaining data. You can derive as many PCs as there are either data points in the spectrum or samples in the data set but the major effects in spectra can be shown to be concentrated in the first few PCs and therefore the amount of data that needs to be considered is dramatically reduced.</p> <p>PCA produces two new sets of variables at each stage: PC Scores represent the response of each sample on each PC; PC Loadings represent the relative importance of each data point in the original spectra to the PC.</p> <p>PCA has many uses, for example, in spectral interpretation, but is most widely used in the identification of spectral outliers.</p>
PCR	<p>Principle component regression uses the scores on each PC as regressors in a MLR against Y values representing composition of samples. As each PC is orthogonal to every other PC, the scores form an uncorrelated data set with better properties than the original spectra. While it is possible to select a combination of PCs for regression based on how well each PC correlates to the constituent of interest, most commercial software forces the regression to use all PCs up to the highest PC selected for the model (“the top down approach”).</p> <p>When used in NIRS, the regression coefficients in PC space are usually converted back to a prediction model using all the data points in wavelength space.</p>
PLS	<p>Partial least square regression is an analytical technique which, like PCA, is a form of data compression. With PLS, the rule used to derive the factors is that each factor in turn maximizes the covariance between the Y data and all possible linear combinations of the X data. PLS is therefore a balance between variance and correlation with each factor being influenced by both effects. PLS factors are therefore more directly related to variability in Y values than are principal components. PLS produces three new variables, loading weights (which are not orthogonal to each other), loadings, and scores which are both orthogonal.</p> <p>PLS models are produced by regressing PLS scores against Y values. As with PCR, when used in NIRS, the regression coefficients in PLS space are usually converted back to a prediction model using all the data points in wavelength space.</p>
MLR	<p>Multiple linear regression uses a combination of several X variables to predict a single Y variable. In NIRS, the X values are either absorbance values at selected wavelengths in the near infrared or derived variables such as PCA or PLS scores.</p>
ANN	<p>Artificial neural networks describes a nonlinear modelling technique loosely based on the architecture of biological neural systems. The network is initially “trained” by supplying a data set with several X (spectral or derived variables such as PCA scores) values and reference Y values. During the training process, the architecture of the network may be modified and the neurons assigned weighting coefficients for both inputs and outputs to produce the best possible predictions of the parameter values.</p> <p>Neural networks require a lot of data in training. Some optimization schemes for ANN require a stop set to end the optimization of the model (e.g. weights).</p>
Multivariate model	<p>Any model where a number of X values are used to predict one or more Y variables.</p>

防伪查询：KNqV-GpFN-zsTv-Xigl-wYFZ

网址：http://www.sacinfo.cn

销售机构：国家标准化管理委员会标准信息中心

Calibration techniques	
Outliers	Outliers are points in any data set that can be shown statistically to have values that lie well outside the expected distribution. For NIRS data, outliers are normally classified as either X (spectral) outliers or Y (reference data) outliers.
X-outliers	Outlier related to the NIR spectrum. An X-outlier can be a spectrum with instrumental faults or from a sample type that is radically different from the other samples or in prediction, a sample type not included in the original calibration set.
Y-outliers	Outliers related to errors in reference data, e.g. a transcription error or in the value obtained by the reference laboratory.
Leverage	A measure of how far a sample lies from the centre of the population space defined by a model. Samples with high leverage have high influence on the model. Leverage is calculated by measuring the distance between a projected point and the centre of the model.
Mahalanobis distance	The distance in PC space between a data point and the centre of the PC space (see h-value below). This is a linear measurement. In PC space, a set of samples usually form a curve shaped distribution. The ellipsoid that best represents the set's probability distribution can be estimated by building the covariance matrix of the samples. The Mahalanobis distance is simply the distance of the test point from the centre of mass divided by the width of the ellipsoid in the direction of the test point.
h-value	In some software, the Mahalanobis distance is referred to as the "Global h-value" and outlier detection depends upon how many standard deviations of h a sample is from the centre. A second measure "Neighbourhood h" is the distance in PC space between a data point and its "n" nearest neighbours and indicates whether a sample is isolated or in a well-populated part of the distribution.
Residual(s)	The difference between a reference value and the value predicted by a regression model. Used in the calculation of regression statistics.
Spectral residual	Residual after chemometric treatment (e.g. PCA, PLS, etc.) of a spectrum. Spectral variation not described by the model.
Test set	When testing a regression model, any set of samples that excludes those used to develop the calibration.
Independent test set	A test set that consists of samples that are from a different geographical region, a new plant (in industrial terms), or have been collected at a later time (e.g. from a different harvest) than those used to create and validate a regression model. These samples form a "true" test of a prediction model.
Validation set	Samples used to validate or "prove" a calibration. Usually samples having the same characteristics as those selected for calibration. Often alternate or "n th " samples (ranked in order of the constituent of interest) are allocated to the calibration and validation data sets from the same pool of samples.
Monitoring set	Set of samples that is used for the routine control of calibration models.

ISO 12099:2017(E)

Calibration techniques	
Cross validation	<p>A method of generating prediction statistics where, repeatedly, a subset of samples are removed from a calibration population, a model being calculated on the remaining samples and residuals calculated on the validation subset. When this process has been run a number of times, prediction statistics are calculated on all the residuals. Full cross validation omits one sample at a time and is run n times (where there are n calibration samples). Where a larger subset is removed, the cross validation cycle is usually run at least eight times before the statistics are calculated. Finally, a model is calculated using all the calibration samples.</p> <p>Cross validation should be used with caution. Firstly, cross validation statistics tend to be optimistic when compared with those for an independent test set. Secondly, care should be taken if there is any duplication in the calibration data (e.g. the same sample scanned on several instruments or at different times) to always assign all copies of the same sample to the same cross validation segment; otherwise, very optimistic statistics are produced.</p>
Overfitting	The addition of too many regression terms in a multiple linear regression. As a result, when samples not in the calibration set are predicted, statistics such as s_{RMSEP} or s_{SEP} are much poorer than expected.
PLS factors	See PLS above.
Scores/ score plots	Plots where the scores on one PC or PLS factor are plotted against those of another PC or PLS factor. Most useful if sample ID or concentration values are used to identify each point in the plot. Patterns in the data can then be seen which are not obvious from the raw data.

Statistical expressions	
	(see also Clause 7)
Bias, \bar{e}	The difference between the mean reference value \bar{y} and the mean value predicted by the NIR model $\bar{\hat{y}}$.
BCL, T_b	The bias confidence limit is the value greater than which a bias is significantly different from zero at the confidence level specified (see 7.3).
SEC, s_{SEC}	For any calibration model, the standard error of calibration (SEC) is an expression of the average difference between predicted and reference values for samples used to derive the model. In this and subsequent statistics, this expression of the average difference refers to the square root of the sum of squared residual values divided by the number of values corrected for degrees of freedom, where 68 % of the errors are below this value.
SECV, s_{SECV}	For a calibration model, the standard error of cross validation (SECV) is an expression of the bias corrected average difference between predicted and reference values for the subset of samples selected as prediction samples during the cross validation process (see “cross validation” above).
SEP, s_{SEP}	<p>The standard error of prediction (SEP) is an expression of the bias corrected average difference between predicted and reference values predicted by a regression model when applied to a set of samples not included in the derivation of the model.</p> <p>The SEP covers a confidence interval of 68 % (multiplied with 1,96 an interval of 95 %).</p>
RMSEP, s_{RMSEP}	<p>The root mean square error of prediction is an expression of the average difference between reference values and those predicted by a regression model when applied to a set of samples not included in the derivation of the model.</p> <p>NOTE RMSEP includes any bias in the predictions.</p>

防伪查询：KNqV-GPIN-2sTv-Xigl-wYFZ

网址：http://www.sacinfo.cn

销售机构：国家标准化管理委员会标准信息中心

Statistical expressions	(see also Clause 7)
RMSECV, s_{RMSECV}	The root mean square error of cross validation (RMSECV) is an expression of the average difference between predicted and reference values for the subset of samples selected as prediction samples during the cross validation process (see "cross validation" above). NOTE s_{RMSECV} includes any bias in the predictions.
UECL, T_{UE}	Unexplained error confidence limit, i.e. the limit which a validation s_{SEP} exceeds in order to be significantly different from the s_{SEC} at the confidence level specified.
RSQ (R^2)	The square of the multiple correlation coefficient between predicted and reference values. When expressed as a percentage, it represents the proportion of the variance explained by the regression model.
Slope, b	For a regression line, it represents the amount Y increases for unit increase in X.
Intercept	For a regression line, the value of Y at X = zero.
Residual standard deviation, s_{res}	The residual standard deviation is an expression of the average size of the difference between reference and fitted values after a slope and intercept correction has been performed.
Covariance, $S_{\hat{y}y}$	Covariance is a measure of how much two random variables vary together. If, for a population of samples, an increase in X is matched by an increase in Y, then the covariance between the two variables will be positive. If an increase in X is matched by a decrease in Y, then the covariance will be negative. When values are uncorrelated, then the covariance is zero.
U_e	Measurement uncertainty, characterizing the dispersion of values that can reasonably be attributed to the result.

Bibliography

- [1] ISO 712, *Cereals and cereal products — Determination of moisture content — Reference method*
- [2] ISO 5983-2, *Animal feeding stuffs — Determination of nitrogen content and calculation of crude protein content — Part 2: Block digestion and steam distillation method*
- [3] ISO 6492, *Animal feeding stuffs — Determination of fat content*
- [4] ISO 6496, *Animal feeding stuffs — Determination of moisture and other volatile matter content*
- [5] ISO 6865, *Animal feeding stuffs — Determination of crude fibre content — Method with intermediate filtration*
- [6] ISO 9622 | IDF 141, *Milk and liquid milk products — Guidelines for the application of mid-infrared spectrometry*
- [7] ISO 11085, *Cereals, cereals-based products and animal feeding stuffs — Determination of crude fat and total fat content by the Randall extraction method*
- [8] ISO 13906, *Animal feeding stuffs — Determination of acid detergent fibre (ADF) and acid detergent lignin (ADL) contents*
- [9] ISO 16472, *Animal feeding stuffs — Determination of amylase-treated neutral detergent fibre content (aNDF)*
- [10] ISO 16634-1, *Food products - Determination of the total nitrogen content by combustion according to the Dumas principle and calculation of the crude protein content — Part 1: Oilseeds and animal feeding stuffs*
- [11] ISO 16634-2, *Food products — Determination of the total nitrogen content by combustion according to the Dumas principle and calculation of the crude protein content — Part 2: Cereals, pulses and milled cereal products*
- [12] ISO 20483, *Cereals and pulses — Determination of the nitrogen content and calculation of the crude protein content — Kjeldahl method*
- [13] ISO 6497, *Animal feeding stuffs — Sampling*
- [14] ISO 24333, *Cereals and cereal products — Sampling*
- [15] ISO 21543 | IDF 201, *Milk products — Guidelines for the application of near infrared spectrometry*
- [16] NÆS T., ISAKSSON T., FEARN T., DAVIES T. *A user-friendly guide to multivariate calibration and classification*. NIR Publications, Chichester, 2002
- [17] SHENK J.S., WESTERHAUS M.O., ABRAMS S.M. Protocol for NIRS calibration: monitoring analysis results and recalibration. In: *Near Infrared Reflectance Spectroscopy (NIRS): Analysis of forage quality. USDA ARS Handbook 643*, (MARTEN G.C., SHENK J.S., BARTON F.E. eds.). U.S. Gov. Print. Office, Washington, DC, 1989, pp. 104–110.
- [18] SØRENSEN L.K. Use of routine analytical methods for controlling compliance of milk and milk products with compositional requirements. *IDF Bulletin* **390**, 2004

防伪查询：KNqV-GPIN-2sTv-Xigl-wYFZ

网址：<http://www.sacinfo.cn>

销售机构：国家标准化管理委员会标准信息中心

中国农业大学专用

ISO 12099:2017(E)

中国农业大学专用

ICS 65.120

Price based on 28 pages

附件 2

ISO 12099:2010 Animal feeding stuffs, cereals and milled cereal products-Guidelines for the application of near infrared spectrometry 翻
译稿

国际
标准

ISO
12099

第二版

2017-08

动物饲料、谷物及谷物精制料的
近红外光谱分析应用指南

参考号码
ISO 12099:2017(E)

目次

前言.....	III
引言.....	IV
1 范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 原理.....	2
5 仪器.....	2
6 定标模型与初始验证.....	2
6.1 总则.....	2
6.2 参考方法.....	2
6.3 异常值.....	3
6.4 定标模型的验证.....	3
6.5 检测条件和仪器状况的变化.....	4
7 模型性能统计学评价.....	4
7.1 总则.....	4
7.2 结果散点图.....	4
7.3 偏差.....	5
7.4 预测误差均方根 (<i>sRMSEP</i>).....	6
7.5 预测标准误差 (<i>sSEP</i>).....	7
7.6 斜率.....	8
8 取样.....	10
9 检测过程.....	10
9.1 待检样品的制备.....	10
9.2 测量.....	10
9.3 结果的评价.....	10
10 仪器稳定性检查.....	11
10.1 质控样品.....	11
10.2 仪器诊断.....	11
10.3 仪器联网.....	11
11 定标运行性能检测.....	11
11.1 总则.....	11
11.2 参考值和近红外预测值差异控制图.....	11
12 精度和准确度.....	12
12.1 重复性.....	12
12.2 重现性.....	12
12.3 精确度.....	13
12.4 不确定度.....	13
13 检测报告.....	13
附录 A (资料性附录) 专用近红外标准指南.....	14
附录 B (资料性附录) 异常值和控制图示例.....	15

附录 C (资料性附录) 补充性条款及定义	19
参考文献.....	23

前言

国际标准化组织（ISO）是各个 ISO 成员国的国家标准化机构间的全球联盟机构，其负责起草国际标准的相关工作通常由 ISO 技术委员会执行。对任何一种已经被某一技术委员会创建的科目感兴趣的各成员机构均有权被陈列于该委员会的名单中。其它与 ISO 保持联络的各种国际标准组织，无论是政府性的还是非政府性的，同样可以参与该项工作。ISO 与国际电工委员会（IEC）之间在所有与电气技术标准相关的领域有密切的合作关系。

ISO/IEC 指导准则第 1 部分中描述了用于开发本文档的程序以及用于进一步维护本文档的程序。应当注意，不同类型 ISO 文件所需的批准标准不同。本文件是根据 ISO/IEC 指导准则第 2 部分的条款起草的（参见 www.iso.org/directives）。

请注意，本国际标准中的某些内容可能会涉及到一些专利权的保护范畴，但 ISO 不对鉴别任何类似专利权利负有责任。在文件开发过程中确定的任何专利权的详细信息都将出现在 ISO 的专利介绍或收到的声明列表中（参见 www.iso.org/patents）。

本文件所使用的商标名称均为为方便使用者提供的资料，并不构成背书。

为了解释标准的自愿性，与合格性评估相关的 ISO 特定术语和表达，以及 ISO 在 TBT（技术性贸易壁垒）中坚持世界贸易组织（WTO）原则的相关信息参见：www.iso.org/iso/foreword.html。

本文件由 ISO/TC 34 食品技术委员会、动物饲料分委会 SC 10 编写。

本版本取消并替换了第一个版本（ISO 12099: 2010），在技术上进行了修订。

引言

本国际标准是根据 ISO 21543 | IDF 201 标准起草的，该标准由食品技术委员会 ISO/TC 34、牛奶和奶制品委员会 SC 5 以及国际奶制品联合会(IDF)编制。

国际标准 ISO 12099:2017(E)

动物饲料、谷物及谷物精制料的近红外光谱分析应用指南

1 范围

本标准为使用近红外光谱进行动物饲料、谷物和谷物精制料成分如水分、脂肪、蛋白、淀粉、粗纤维含量以及相关参数如消化率等的检测提供指南。

上述检测是基于近红外光谱区的光谱测定进行的检测。

2 规范性引用文件

本标准无规范性引用文件。

3 术语和定义

下列术语与定义适用于本标准。

ISO 和 IEC 维护术语数据库的标准化使用参考如下网址：

IEC 电子百科全书：<http://www.electropedia.org/>

ISO 在线浏览平台：<http://www.iso.org/obp>

3.1 近红外光谱分析仪/NIR 分析仪

Near infrared instrument/NIR instrument

在本标准定义条件下，利用近红外范围的吸光度与成分含量（3.3）及技术指标（3.4）的关系模型，对动物饲料（3.2）、谷物和谷物精制料成分含量及技术指标进行预测的仪器。

3.2 动物饲料

Animal feeding stuffs

用于动物的经口饲喂的包括添加剂在内的成品、半成品或未加工过的物质或产品。

示例：原料、饲草、肉骨粉、混合饲料及其他终端产品，包括宠物饲料等。

3.3 成分含量

Constituent content

用适当的、标准化的或验证过的化学方法检测的成分的质量浓度。

注 1：质量浓度通常使用百分比表示。

注 2：示例中成分的分析方法请参考文献[1]至[12]。

示例：水分、脂肪、蛋白、粗纤维、中性洗涤纤维及酸性洗涤纤维。

3.4 技术指标

Technological parameter

用适当的、标准化的或验证过的方法检测的动物饲料 (3.2)、谷物和谷物精制料的属性或功能性。

注：只要遵循本标准的流程，亦有可能针对以上列举之外的参数和饲料种类进行近红外方法的开发及验证。所测得参数的计量单位应遵循参考方法中使用的单位。

示例：消化率。

4 原理

收集近红外 (NIR) 范围内的光谱数据，并利用以所分析产品的代表性样本为基础开发出的定标模型转换成组成成分或技术指标的浓度值。

5 仪器

5.1 近红外光谱分析仪

仪器基于近红外谱区的漫反射或透射进行检测，其光谱区包括：近红外全谱区 770nm-2500nm (12900cm^{-1} - 4000cm^{-1})，或全谱区内的某段光谱范围，或是选择的波长或波数。光学原理可是色散型(如：光栅单色仪)，干涉型或非热型(如发光二极管、激光二极管以及激光)。仪器应具备自我诊断系统，用于检测仪器的噪音、重现性、波长/波数准确度和波长/波数精度(对扫描型光谱仪)。

仪器应该分析足够多体积或表面积样品以消除待检样品化学和物理性质不均匀的影响。对于透射扫描方式，应依据仪器制造商的关于信号强度的推荐值进行光程(样品厚度)的优化，以获得最佳的线性和最大的信噪比。

5.2 适当的研磨或粉碎设备，对样品进行预处理(如需要)。

注：研磨或粉碎条件的改变会影响近红外的测量，例如，发热会影响像水分这类挥发性成分。

6 定标模型与初始验证

6.1 总则

在使用之前，仪器应经过定标。定标包括与参比的比较和仪器的调整过程。对近红外仪器的定标有很多不同的方法，在此不推荐具体的定标方法。

关于定标方法的介绍可以参照参考文献[7]或所使用仪器所附的使用手册。对于验证样品要有足够的代表性样品覆盖样品的各种影响因素，例如：

- a) 应综合考虑样品的常量成分与微量成分及其成分含量范围；
- b) 对饲草、饲料原料和谷物样品应考虑季节、地域和遗传等因素的影响；
- c) 应考虑加工技术和条件的不同；
- d) 应考虑不同的储存条件；
- e) 应考虑样品和仪器温度变化；
- f) 应考虑仪器变化(例如：仪器间差异)。

注：对于一项可靠的验证至少需要 20 个验证样本。

6.2 参考方法

水分、脂肪、蛋白和其它成分含量和技术指标的测定要使用国际公认的参考方法。具体实例可参考

文献[1]至[12]。

为定标集样本赋值的参考方法要在统计控制之内。清楚参考方法的精度至关重要。

在可能的情况下，应使用有标准赋值的标准物质作为参考。

6.3 异常值

许多情况下，在定标和验证时会出现统计异常值。异常值的出现可能是光谱数据异常（光谱异常值，接下来简称为 x -异常值）或参考值存在误差或参考值与近红外光谱回归关系异常（接下来简称为 y -异常值），示例见图 B1-B5。

为了达到验证的目的，下列情况的样品不能作为异常值：

- a) 样品成分含量/技术指标在定标含量范围内；
- b) 样品光谱在定标样品光谱的变异范围内，例如可以用马氏距离判断；
- c) 样品光谱偏差在定标确定的偏差范围内；
- d) 样品预测偏差在定标确定的偏差范围内。

如果一个样品看似异常，那需要首先判断它是不是 x -异常值，如果超出了定标确定的 x -异常值限值，那该样品需要被剔除。如果不是 x -异常值，则需要检查参考值和近红外预测值是不是异常，例如通过重复分析：如果参考值和近红外预测值确证无误，则该样品应保留并参与验证集的统计分析；如果重复测量值显示原始参考值或近红外预测值有误，那么该样品应该使用新的测定值。

6.4 定标模型的验证

6.4.1 总则

定标模型在使用之前应用具有广泛代表性的独立验证集进行验证。为确定偏差、斜率和预测标准误差（SEP，见 7.5）至少需要 20 个样品。验证应针对每类样品、每个成分含量/技术指标、温度和其它已知有影响或预计有影响的因素进行，只有这些变异，例如样品种类、含量范围和温度，包含在验证集中时才能证明定标的有效性。

注 1：定标模型只能在已被验证的范围内使用。

独立验证集样品参考值和近红外预测值的散点图和与参考值之间的残差能对定标模型的效果进行直观显示。首先需要计算出 SEP（见 7.5），然后使用平均系统误差（bias）修正后的残差图来检验异常值，如：残差大于 $\pm 3S_{SEP}$ 的即可判断为异常值。

如果验证过程显示定标模型不能获得可接受的统计值，那么该定标模型就无法使用。

注 2：什么是可接受的统计值由建模者决定，例如：可取决于参考方法的精度、覆盖的含量范围和分析的目的等。

如果可能，参考物质或标准物质可用作定标模型验证的一部分。

接下来是使用线性回归（ $y_{ref} = a + b \times y_{NIRS}$ ）拟合近红外预测值（ y_{NIRS} ）和参考值（ y_{ref} ）生成描述验证结果的统计指标。

6.4.2 偏差校正

方法之间的偏差也应该检查。如果近红外预测值的平均值和参考值的差值与 0 有显著性差异，这表明定标存在偏差。需要通过调整定标方程的常数项（见 7.3）来消除偏差的影响。

6.4.3 斜率调整

如果斜率 b 与 1 存在显著性差异，那么说明该定标模型出现偏斜。

除非定标模型应用于新类型样品或仪器，否则通常不推荐对其斜率或截距进行调整。如果对定标模型的重新检验没有发现异常值，尤其是没有发现高杠杆值的异常样品，那更可取的措施是扩充该定标集使之包含更多有代表性的样品。但是，斜率一旦被调整，那么定标模型必须用新的独立验证集进行检验。

6.4.4 定标集的扩充

如果定标集的精度未达到预期效果，应扩充定标集使其包含更多的样品，或重新执行新的定标。只要是以扩充后的定标集建立的新定标模型，应使用新的验证集重复验证过程。必要的话，需持续进行定标集的扩充，直至验证集所获得的结果满足要求为止。

6.5 检测条件和仪器状况的变化

除非进行了额外的定标流程，否则如果检测条件改变，声明方法准确度的近红外方法的本地验证通常不被视为有效。

例如，尽管分析成分含量的范围一致，为某个样本集所开发的定标模型对该样本集之外的样本仍可能是无效的。例如，如果遗传因素、生长条件和加工工艺参数等不同，针对某个地区的青贮饲草建立的定标方程也不能对其它地区的样本给出相同准确度的预测。

未在定标集中包含的样品制备技术或测量条件，例如温度，的改变同样会影响分析结果。

在某一仪器上开发的定标方程并非总是可以直接转移到另一台相同原理、相同型号的仪器上。定标方程转移之前，必须对定标方程的偏差或斜率/截距进行相应调整。在很多情况下，在能够进行定标方程的转移之前，必须进行两台仪器之间的标准化（参考文献 7）。标准化的过程可以用来在不同类型仪器之间进行定标模型转移，但要求样品检测采用相同方式（反射、透射），且波长范围拥有重叠波段。

任何一种条件被改变后应进行相应的补充验证。

仪器任何主要部件（光学系统、检测器）更换或维修时，都应重新检验定标模型。

7 模型性能统计学评价

7.1 总则

预测模型的性能应该用验证集样品来检验。验证集样品是独立于定标集的样品。在工厂，可以是新批次的产品；在农业中，可以是新的作物或新试验地点。

验证集的样品应该严格地采用参考方法分析。验证集样品的分析应更被重视，其结果的精度比定标集样品的精度更重要。

为便于进行可信的数据统计计算，验证集样品数最少要 20 个。

对预测模型性能的评价所用的 NIR 检测方案应与用在常规分析中的相同（一次分析或两次分析）。

7.2 结果散点图

对结果的可视化绘图是很重要的。例如参考值 vs 预测值或残差 vs 预测值的散点图。

残差按公式 (1) 计算：

$$e_i = y_i - \hat{y}_i \quad (1)$$

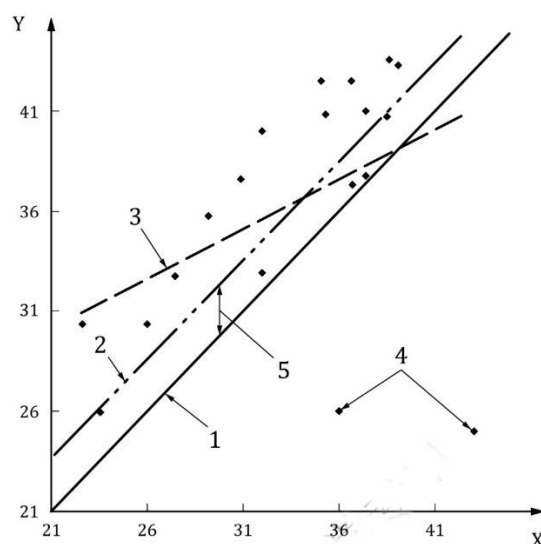
式中:

y_i 是第 i 个参考值 (y_{ref}):

\hat{y}_i 是第 i 个近红外预测值 (y_{NIRS})。

与参考值相比, 近红外预测偏高时偏差为负, 近红外预测偏低时偏差为正。

数据散点图可以清晰的表现出: 相关关系、偏差、斜率和显著异常值的存在。(见图 1)。



图中:

- | | | | |
|---|----------------------------|---|--------|
| 1 | 45° 线 (偏差为 0、斜率为 1 的理想线) | X | 近红外预测值 |
| 2 | 带偏差的 45° 线 (45° 线减去偏差得到的线) | Y | 参考值 |
| 3 | 线性回归线 | | |
| 4 | 异常值 | | |
| 5 | 偏差 | | |

注: 异常值 (图中 4) 对斜率有很大的影响, 如果结果是用于斜率调整的, 则异常值需要剔除。

图 1 定标集的散点图, $y_{ref} = f(a + by_{NIRS})$

7.3 偏差

大多数情况下, 偏差或系统误差是通过近红外模型观察到的。偏差的出现有几个主要原因: 未被包含在定标模型中的新样品类型、仪器的漂移、湿法化学值的漂移、过程的变化、样品制备的变化等。

偏差是残差的平均值 \bar{e} , 用 n 个独立样品按公式 (2) 计算:

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i \quad (2)$$

式中 e_i 是公式 (1) 中定义的残差, 由此可以得到公式 (3):

$$\bar{e} = \frac{1}{n} [\sum_{i=1}^n y_i - \sum_{i=1}^n \hat{y}_i] = \bar{y} - \bar{\hat{y}} \quad (3)$$

式中:

y_i 是第 i 个参考值；

\hat{y}_i 是第 i 个近红外预测值；

\bar{y} 是参考值的平均值；

$\bar{\hat{y}}$ 是近红外预测值的平均值。

偏差的显著性用 t-test 进行检验。新样本的小样本集偏差接受或拒绝的阈值，由偏差置信区间 (BCLs) T_b 确定，其计算见公式 (4)。

$$T_b = \pm \frac{t_{(1-\frac{\alpha}{2})} S_{SEP}}{\sqrt{n}} \quad (4)$$

式中：

α 是发生 I 类错误的概率；

t 是双尾检验学生氏 t 值，其值大小取决于与 SEP 有关的自由度和 I 类错误的概率；

n 是独立样本数；

S_{SEP} 是预测标准误差 (见 7.5)。

示例：当 $n = 20$ ，且 $S_{SEP} = 1$ ，BCLs 按公式 (5) 计算：

$$T_b = \pm \frac{2.09 \times 1}{\sqrt{20}} = \pm 0.48 \quad (5)$$

这意味着，当偏差高于预测标准误差的 48% 时，该 20 个样本检测的偏差与 0 有显著性差异。

表 1 — 给定概率的 t 分布值， $\alpha = 0.05 = 5\%$

n	t	n	t	n	t	n	t
5	2.57	11	2.20	17	2.11	50	2.01
6	2.45	12	2.18	18	2.10	75	1.99
7	2.36	13	2.16	19	2.09	100	1.98
8	2.31	14	2.14	20	2.09	200	1.97
9	2.26	15	2.13	30	2.04	500	1.96
10	2.23	16	2.12	40	2.02	1 000	1.96

注：可使用 Excel¹ 中的 TINV 函数计算。

¹ Excel 是微软 (Microsoft) 公司一项产品的商品名。这一信息仅为本文件的用户提供方便，而非通过 ISO 构成该指定产品的背书。其它等效产品，只要能够证明可生成具备可比性的结果，亦可以使用。

7.4 预测误差均方根 (S_{RMSEP})

S_{RMSEP} 按公式 (6) 计算：

$$S_{RMSEP} = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n}} \quad (6)$$

¹ Excel 是微软 (Microsoft) 公司一项产品的商品名。这一信息仅为本文件的用户提供方便，而非通过 ISO 构成该指定产品的背书。其它等效产品，只要能够证明可生成具备可比性的结果，亦可以使用。

式中:

e_i 是第 i 个样品的残差;

n 是独立样本数。

该值可以与 s_{SEC} 和 s_{SECV} 进行比较 (见附录 C)。

s_{RMSEP} 包含随机误差 (s_{SEP}) 和系统误差 (偏差)。同时也包含参考方法的误差 (如同 s_{SEC} 和 s_{SECV})。计算公式见式 (7):

$$s_{RMSEP} = \sqrt{\frac{(n-1)}{n} s_{SEP}^2 + \bar{e}^2} \quad (7)$$

式中:

n 是独立样本数;

s_{SEP} 是预测标准误差 (见 7.5);

\bar{e} 是偏差或系统误差;

对 s_{RMSEP} 没有直接的检验, 这就是将系统误差 (偏差/ \bar{e}) 和随机误差 (s_{SEP}) 分开的原因。

7.5 预测标准误差 (s_{SEP})

预测标准误差 (s_{SEP}), 也就是残差的标准差, 表达了通过 NIR 检测值和参考值之间的平均偏差 (bias) 折算的近红外分析结果的准确性, 计算公式见式 (8):

$$s_{SEP} = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-1}} \quad (8)$$

式中:

n 是独立样本数;

e_i 是第 i 个样品的残差;

\bar{e} 偏差或系统误差。

s_{SEP} 与 s_{SEC} 或 s_{SECV} 的关联 (见附录 C), 用于检查定标模型相对于所选验证集的有效性。

随机误差置信区间 (UECLs), T_{UE} , 用 F 检验 (两个方差的比值) 计算 (参考文献 [18] 和表 2), 见公式 (9):

$$T_{UE} = s_{SEC} \sqrt{F_{(\alpha, v, M)}} \quad (9)$$

式中:

s_{SEC} 是定标标准误差 (见附录 C);

α 是发生 I 类错误的概率;

v 等于 $n - 1$, 是与验证集 s_{SEP} 相对应的分子自由度;

n 是验证过程使用的样品数;

M 等于 $n_c - p - 1$, 是与定标标准误差 s_{SEC} 相对应的分母自由度 (其中, n_c 是定标样品数; p 是模型的因子数或 PLS 主成分数, 或神经网络模型中的神经元连接的权重 (见附录 C))。在神经网络中, 权重都是模型中未知的参数。

注 1: 由于 s_{SEC} 过于理想化, 故 s_{SECV} 是比 s_{SEC} 更好的统计量, 经常用于取代 s_{SEC} ($s_{SEP} > s_{SECV} > s_{SEC}$)。

示例: $n = 20$, $\alpha = 0.05$, $M = 100$, $s_{SEC} = 1$, 计算 $T_{UE} = 1.30$ 。

20 个验证样品, s_{SEP} 最大允许大于 s_{SEC} 的 30%。

注 2: 可使用 Excel² 中的 FINV 函数计算。

由于 F 检验需要两个独立的验证集, 所以不能用于比较基于同一个验证集的两种定标。基于相同数据集的两个或更多定标模型的比较需要采用另外的检验方法。

表 2 根据 SEP 自由度作为分子自由度和 SEC 自由度作为分母自由度的 F 值和 F 值的平方根
[请参照方程式 9 下面的相关定义]

$F_{(\alpha,v,M)}$						$\sqrt{F_{(\alpha,v,M)}}$					
SEP 自由度	SEC 自由度					SEP 自由度	SEC 自由度				
	50	100	200	500	1000		50	100	200	500	1000
5	2.4	2.31	2.26	2.23	2.22	5	1.55	1.52	1.5	1.49	1.49
6	2.29	2.19	2.14	2.12	2.11	6	1.51	1.48	1.46	1.45	1.45
7	2.20	2.10	2.06	2.03	2.02	7	1.48	1.45	1.43	1.42	1.42
8	2.13	2.03	1.98	1.96	1.95	8	1.46	1.43	1.41	1.4	1.4
9	2.07	1.97	1.93	1.90	1.89	9	1.44	1.41	1.39	1.38	1.37
10	2.03	1.93	1.88	1.85	1.84	10	1.42	1.39	1.37	1.36	1.36
11	1.99	1.89	1.84	1.81	1.80	11	1.41	1.37	1.36	1.34	1.34
12	1.95	1.85	1.80	1.77	1.76	12	1.40	1.36	1.34	1.33	1.33
13	1.92	1.82	1.77	1.74	1.73	13	1.39	1.35	1.33	1.32	1.32
14	1.89	1.79	1.74	1.71	1.7	14	1.38	1.34	1.32	1.31	1.30
15	1.87	1.77	1.72	1.69	1.68	15	1.37	1.33	1.31	1.30	1.29
16	1.85	1.75	1.69	1.66	1.65	16	1.36	1.32	1.30	1.29	1.29
17	1.83	1.73	1.67	1.64	1.63	17	1.35	1.31	1.29	1.28	1.28
18	1.81	1.71	1.66	1.62	1.61	18	1.30	1.31	1.29	1.27	1.27
19	1.80	1.69	1.64	1.61	1.60	19	1.34	1.30	1.28	1.27	1.26
29	1.69	1.58	1.52	1.49	1.48	29	1.30	1.26	1.23	1.22	1.22
49	1.60	1.48	1.42	1.38	1.37	49	1.27	1.22	1.19	1.17	1.17
99	1.53	1.39	1.32	1.28	1.26	99	1.24	1.18	1.15	1.13	1.12
199	1.48	1.34	1.26	1.21	1.19	199	1.22	1.16	1.12	1.10	1.09
499	1.46	1.31	1.22	1.16	1.13	499	1.21	1.14	1.11	1.08	1.07
999	1.45	1.30	1.21	1.14	1.11	999	1.20	1.14	1.10	1.07	1.05

7.6 斜率

² Excel 是微软 (Microsoft) 公司一项产品的商品名。这一信息仅为本文件的用户提供方便, 而非通过 ISO 构成该指定产品的背书。其它等效产品, 只要能够证明可生成具备可比性的结果, 亦可以使用。

在近红外分析报告和文献中，经常有关于简单的回归方程 $y = a + b\hat{y}$ 中斜率 b 的报道。

注意，如果计算的斜率用于修正近红外预测结果（例如用逆向多元回归建立预测模型），斜率的计算必须以参考值作为因变量，以近红外预测值作为自变量。

根据最小二乘拟合，斜率和截距的计算分别见公式（10）和（11）：

$$b = \frac{s_{\hat{y}y}}{s_{\hat{y}}^2} \quad (10)$$

式中：

$s_{\hat{y}y}$ 是参考值和预测值的协方差；

$s_{\hat{y}}^2$ 是 n 个预测值的方差。

$$a = \bar{y} - b\bar{\hat{y}} \quad (11)$$

式中：

\bar{y} 是参考值的平均值；

b 是斜率；

$\bar{\hat{y}}$ 是预测值的平均值。

可用 t 检验对 $b = 1$ 进行假设检验，如公式（12）所示：

$$t_{obs} = |b - 1| \sqrt{\frac{s_{\hat{y}}^2(n-1)}{s_{res}^2}} \quad (12)$$

式中：

$s_{\hat{y}}^2$ 是 n 个预测值的方差；

n 是独立样本数；

s_{res} 是残差标准偏差。

残差标准偏差 s_{res} 按公式（13）计算：

$$s_{res} = \sqrt{\frac{\sum_{i=1}^n [y_i - (a + b\hat{y}_i)]^2}{n-2}} \quad (13)$$

式中：

n 是独立样本数；

a 是截距（见公式 11）；

b 是斜率（见公式 10）；

y_i 是第 i 个参考值；

\hat{y}_i 是第 i 个近红外预测值。

注：当预测值用斜率和截距校正时， s_{res} 与 s_{SEP} 类似。不要把偏差和截距混淆（见图 1）。

只有斜率等于 1 时，偏差才等于截距。

当满足公式 (14) 时，斜率 b 与 1 有显著性差异：

$$t_{obs} \geq t_{(1-\frac{\alpha}{2})} \quad (14)$$

式中：

t_{obs} 是由公式 (12) 计算的 t 值；

$t_{(1-\frac{\alpha}{2})}$ 是置信度 $\alpha=0.05$ (5%) 时从表 1 中获取的双尾检验学生氏 t 值。

即使 s_{SEP} 正确，如果范围太窄或分布不均匀都会导致斜率的无用校正。仅当验证集覆盖了定标集大部分范围时，方可对斜率进行调整。

示例：样本数量 $n=20$ ，残差标准偏差（见公式 (13)）为 1，预测值的标准偏差 $S_y=2$ ，并且计算的斜率 $b=1.2$ ，则得到的 t_{obs} 值为 1.7，由于对应 20 个样本的 $t_{(1-\frac{\alpha}{2})}$ 为 2.09（见表 1），故斜率 b 与 1 之间的差异不显著。若斜率为 1.3，则得到的 t_{obs} 值为 2.6，那么斜率 b 与 1 之间存在显著性差异。

8 取样

取样不属于本标准所规范的内容。

注：推荐的取样过程详见 ISO 6497 和 ISO14333。

在实验室得到真正有代表性的样品非常重要。样品在运输和贮存过程中不应受到损坏或发生变化。

9 检测过程

9.1 待检样品的制备

所有试验样品应在特定的条件下保存，以确保样品从取样到进行分析的整个过程中其成分含量保持不变。

日常分析样品的制备需要与验证集样品的制备流程相同，应采用标准制备流程。

在分析之前，应取用对所分析成分具有代表性的样品。

对于具体样品制备过程，参见具体的近红外分析标准。

具体的近红外分析标准指南见附录 A。

9.2 测量

参照仪器制造商或供应商的仪器使用说明进行。

所制备样品的温度应该在验证时样品温度的范围之内。

9.3 结果的评价

日常分析结果应在所使用的定标模型范围内方可有效结果。

光谱异常样品的检测结果不可采信。

若对同一样本有多次测量，且满足重复性要求的（见 12.1），计算其算术平均值。

对结果的表达，参照具体的近红外标准。

10 仪器稳定性检查

10.1 质控样品

为确保仪器硬件的稳定性及预防仪器故障，每天至少进行一次质控样品的检测。不需要了解质控样品中分析物的真实含量。质控样品应足够稳定，并尽可能与所分析样本类似。质控样品的分析参数同样需要保持稳定，并等同于或至少从生物化学性质的角度尽可能的接近样品分析物。应按照 9.1 制备样品，并且以能够使其存储效期最大化的方式进行存储。这些样品通常能够保持长时间的稳定，但其稳定性应根据实际情况予以检测。质控样品应该与分析样品周期重叠，以保障不间断的质量控制。

在质控图上，应对每天质控结果的波动进行记录，并应对有显著性变化模式或趋势的质控图进行研究。

10.2 仪器诊断

对扫描型光谱仪，应对其波长/波数（见 5.1）的准确度和精度进行至少每周一次的检查，或者如果仪器厂商推荐，其检查频度应该更高。并且其结果需要与相应技术规格及要求对比（见 5.1）。

应该至少每周或按照厂商推荐的时间间隔对仪器的噪音进行类似的检查。

10.3 仪器联网

如果几台仪器联网使用，需根据厂商的推荐对仪器之间的标准化给予特别关注。

11 定标运行性能检测

11.1 总则

应检查定标模型对个体样品测试的适用性。可采用在定标和验证过程中所使用的异常值剔除方法，如马氏距离和光谱残差等。这在大多数仪器中可自动执行。

如果样品没有通过检验，说明该样品不能适应于定标和/或验证所用样品群体，除非对模型进行修正，否则该样品就不能用这一预测模型进行检测。由此，异常样品的检测可用于选择哪些样品用参考方法赋值并作为升级定标模型之用。

如果定标模型能够适用于被分析样品，则该光谱可采用验证过的定标模型进行预测。

近红外方法应持续使用参考方法进行验证，以确保定标处于稳定的最优状态并满足准确性要求。对近红外方法验证的频度，应足以确保该方法运行的稳定性受控于参考方法的系统和随机偏差。该频度尤其依赖于每天分析的样品数量和样品集的波动。

日常验证所用样品应从待分析的样品中随机选取。宜采用一些取样策略以确保样品在整个定标范围内分布的均匀性，例如，按含量高低进行分段然后在每段中随机选择样品，或确保样本能够涵盖常规范围。

日常验证样品数量应满足性能测试统计分析所需。可靠的验证过程需要至少 20 个样品（以满足变量的正态分布）。独立验证集的结果可以用于日常验证的启动，接下来每周 5-10 个样品即足够对定标性能的监控。如果采用的样品量太少，一旦有结果超出了控制限则很难做出正确判断。

11.2 参考值和近红外预测值差异控制图

应采用控制图对结果进行评估，控制图是以分析样品数作为横坐标，参考值和近红外预测值的差作为纵坐标； $\pm 2s_{SEP}$ （95%的概率）和 $\pm 3s_{SEP}$ （99.8%的概率）可用作报警限和行动限，此处的 s_{SEP} 采用独立于定标集的验证集获得。

如果定标和参考值的分析均按要求执行，则应只有 5% 的样品在报警限之外，只有 2% 的样品在行动限之外。

对控制图应检查系统偏差、偏差的正负和极端异常值。可使用 Shewart 控制图的一般规则对结果进行评估。然而，过多规则同时使用可导致虚假报警。

以下规则的组合应用已被证明具备发现问题的实用价值：

- a) 1 个点超过了行动限；
- b) 连续 3 个点中有 2 个在报警限之外；
- c) 连续 9 个点在 0 线的同一侧。

其它控制图（例如参考值和近红外预测值的平均差）和其它控制原则也可使用，以强化评估效果。

在对结果的评估中，应注意 s_{SEP} 和所测得的近红外预测值与参考值的差异也包含参考值的不确定度。如果参考值的不确定度小于 s_{SEP} 的 1/3，则该影响可忽略不计^[10]。

为减少虚假报警的风险，控制样品的近红外光谱分析和参考方法分析应以不同的顺序分别独立进行，以避免（如在参考方法中）日间系统偏差的影响。

如果样品经常超出报警限，并且控制图显示都是随机误差（无趋势性或系统偏差），那可能是控制限所用的 s_{SEP} 太过理想。此时，试图通过频繁调整定标以勉强将结果限于该控制限内并不能改善现实状况，反而应该用新的结果对 s_{SEP} 进行重新评估。

稳定运行一段时间后的定标方程可能会开始出现失控情况，此时应对定标进行升级。升级之前，首先要评估该失控是否因为参考值的变化、测量条件预料之外的变化（如更换新的操作者所致）、仪器漂移或发生故障等。在有些情况下，只通过调整定标方程的常数项就能满足要求（示例见图 B6）。在其它情况下，可能需要重新运行定标程序。重新定标时，定标集样品需扩展到能够包含日常验证的样品，亦可增加专门为重新定标选择的样本（示例见图 B7）。

鉴于参考值已满足参考方法精密度要求，且近红外分析测量条件与仪器性能可保持不变，偏差显著或 s_{SEP} 值增大可能是由于样品相比于定标集样品的化学、生物学或物理性质发生了变化。

其它控制图，例如z得分图也可以使用。

12 精度和准确度

12.1 重复性

重复性是对完全相同的测试原料，采用相同的检测方法，在同一实验室，相同的操作人员，用相同的操作仪器，在短时间内获得的两次独立测量结果的差异，该差异不应该大于 5%。重复性的大小依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略等。所有情况下都应进行重复性的检测。

12.2 重现性

重现性是对完全相同的测试原料，不同的实验室，不同的操作者，在不同的时间获得的两次独立测定结果的差异，该差异不应该大于 5%。重现性大小依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略。所有情况下都应进行重现性的检测。

12.3 精确度

精确度，包含了偏离样品真值的系统偏差的不确定度（准确度）和随机误差的不确定度（精密程度），尤其依赖于样品、分析成分、样品和分析成分的变化范围、样品制备方法、仪器类型及相应的定标策略。所有情况下都应进行精确度的检测。报告的 s_{SEP} 和 s_{RMSEP} 值也包括了参考值的不确定度，该不确定度会因情况而异。

12.4 不确定度

不确定度， U_e ，是用来描述合理地分布于结果附近的测量值的离散度的参数。对于近红外预测的结果，不确定度通常用式（15）表示：

$$U_e = \pm 2s_{RMSEP} \quad (15)$$

如果乘数为 2，可理解为真值有 95%的可能性落在 $\pm U_e$ 之间。

13 检测报告

检测报告应包括以下信息：

- a) 完整识别样本所需的全部信息；
- b) 使用的检测方法，参考相关的参考方法；
- c) 所有该标准未明确或被视为可选的所有操作条件；
- d) 任何可能影响结果的环境因素；
- e) 获得的检测结果；
- f) 目前的 s_{SEP} 和偏差（如果有统计差异），用最少 20 个样品进行性能测试得到（见第 11 章）。

附录 A

(资料性附录)

专用近红外标准指南

专用近红外标准，可依据通过近红外光谱法测定饲料、谷物和谷物精制料中具体成分或参数的专用定标模型进行制定。

一套专用标准应具备以下信息：

- 在验证过程中，遵循本标准所规范的内容；
- 针对浓度、精确度与精密度，给出需满足的具体参数范围；
- 指定在开发模型时应考虑到的变量的来源；
- 应具体规定分析流程、计算方法和结果的表达。

专用标准应符合标准文本的格式，并给出以下相关信息：

- 在标准名称或范围中应给出：样品类型、通过近红外光谱仪检测的成分/参数和定标模型；
- 在“规范性引用文件”中列出验证用的参考方法；
- 光谱仪的原理（如近红外反射 NIR、近红外透射 NIT）和定标原理（如偏最小二乘方法 PLS、人工神经网络 ANN）；
- 测试样品的制备、检测和质量控制过程；
- 依据精确度、精密度及范围制定需满足的具体值；表 A.1 是一个验证集参数的示例。

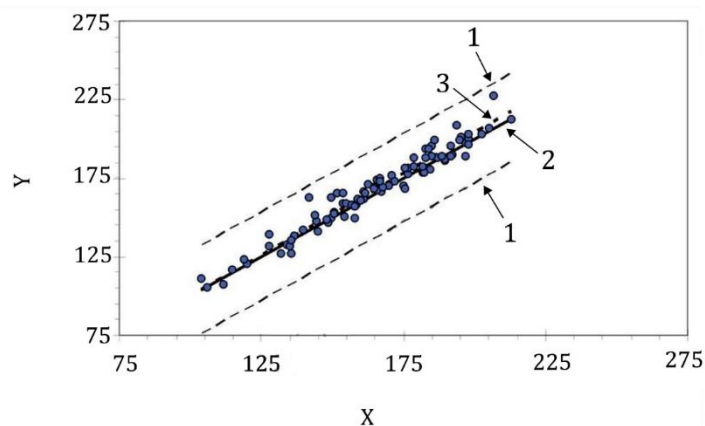
表 A.1 验证集参数的示例

成分	模型	样本数 <i>N</i>	精确度 <i>s_{SEP}</i>	最小含量 %质量比	最大含量 %质量比	R ²
脂肪	ANN	183	0.50	2.8	12.9	0.94
水分	ANN	183	0.47	9.2	12.3	0.83
蛋白	ANN	179	0.72	11.0	29.1	0.96
纤维	ANN	123	1.11	0.5	18.0	0.90
淀粉	PLS	113	1.80	7.8	50.2	0.92

附录 B
(资料性附录)

异常值和控制图示例

图 B.1 显示的是饲草中粗蛋白含量的测定，这是一个没有异常值的例子。使用已建立的定标方程分析包含 95 个样品独立验证集的结果： $s_{SEP} = 4.02$ ； $s_{RMSEP} = 6.05$ ；斜率 $b = 1.04$ 。

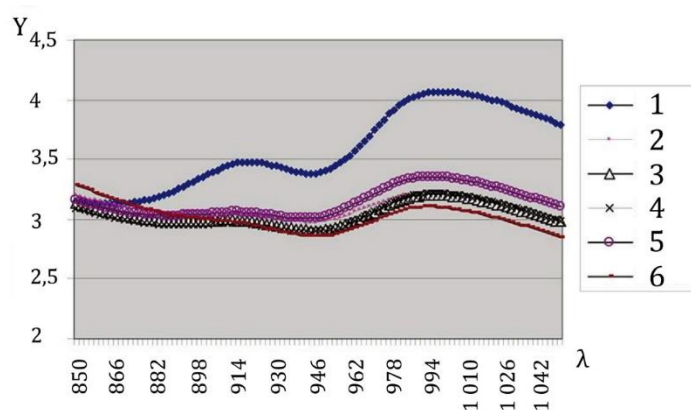


图中：

- | | | | |
|---|-------------------------|---|--------|
| 1 | $\pm 3s_{SEP}$ 临界线 | X | 近红外预测值 |
| 2 | 45° 线（斜率为 1，截距为 0 的理想线） | Y | 参考值 |
| 3 | 回归线 | | |

图 B.1 无异常值

图 B.2 显示的是存在 x -异常值的吸收光谱。光谱 1（图 B.2 中最上方的光谱）是异常光谱。

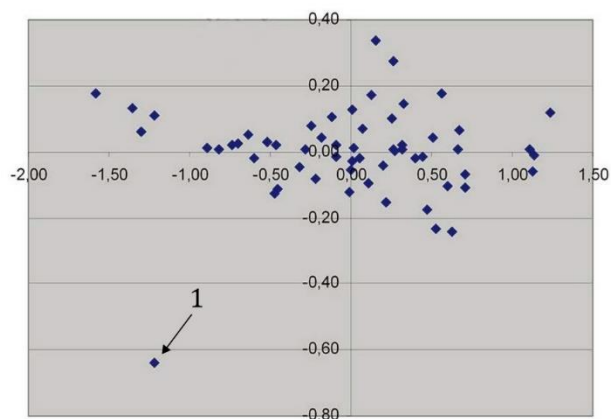


图中：

- | | | | |
|---|------|-----------|------|
| 1 | 光谱 1 | 5 | 光谱 5 |
| 2 | 光谱 2 | 6 | 光谱 1 |
| 3 | 光谱 3 | λ | 波长 |
| 4 | 光谱 4 | Y | 吸光度值 |

图 B.2——存在 x -异常值的吸收光谱

图 B.3 显示的是存在 x -异常值的主成分得分图（图 B.3 中 1 为 x -异常值）。

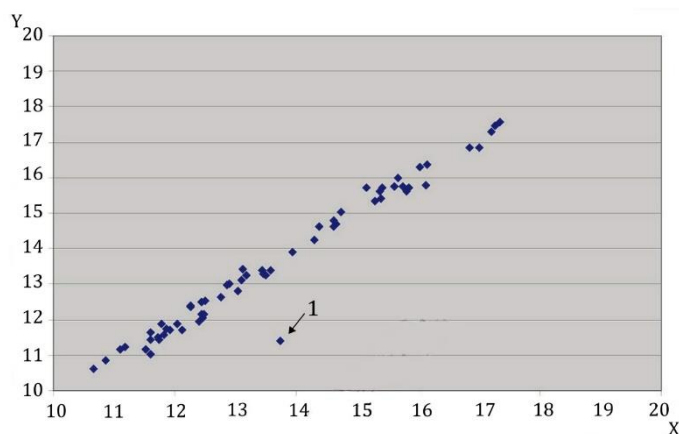


图中:

1 异常值

图 B.3——存在x-异常值的主成分分析得分图

图 B.4 显示的是存在y-异常值的散点图（图 B.4 中 1 为y-异常值）。参考值 vs 预测值的散点图（反之亦然）显示出与其它样品严重偏离的一个样品（图 B.4 中 1 所示）。如果其偏离的原因与 NIR 数据（x-异常值）无关，则该样品是一个由于错误的参考值或参考值与光谱值之间存在不同的相关关系所导致的y-异常值。

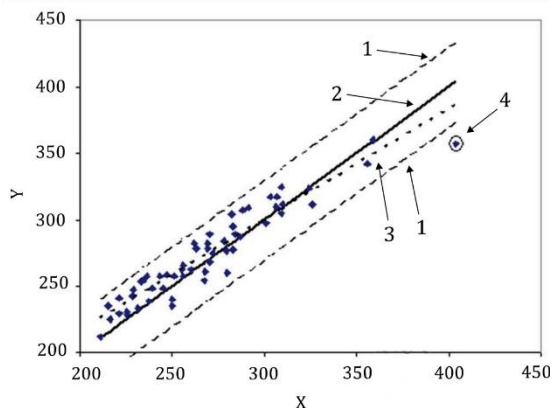


图中:

1 异常值
X 近红外预测值
Y 参考值

图 B.4 存在y-异常值的散点图

图 B.5 显示的是存在y-异常值（图 B.5 中 1 所示）的饲草中酸性洗涤纤维 ADF 含量的测定。



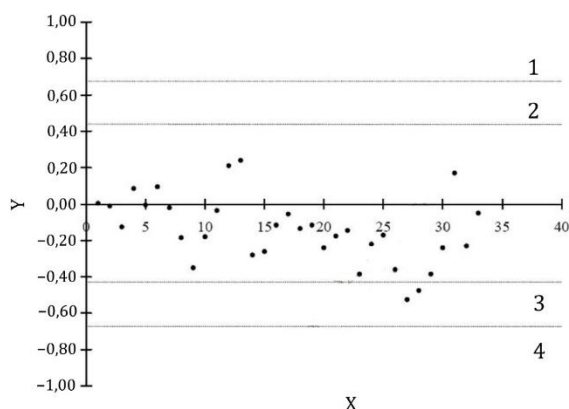
图中:

- | | | | |
|---|--------------------|---|--------|
| 1 | $\pm 3s_{SEP}$ 临界线 | X | 近红外预测值 |
| 2 | 45°线 | Y | 参考值 |
| 3 | 回归线 | | |
| 4 | 异常值 | | |

图 B.5 存在y-异常值的饲草 ADF 的测定

图 B.6 显示的是谷物饲料原料中脂肪含量的测定。在警戒线上限 UAL 和警戒线下限 LAL 之外没有异常值点。但有连续的 9 个点（如 14-22）在零线同一侧，这意味着偏差问题的存在。另有连续 3 点中的 2 点（27 和 28）超出了报警线下限 LWL，但没有出现超出报警线上限 UWL 的点，这同样表明偏差问题的存在。未观察到随机变化增加的情况，分布依然小于 $3s_{SEP}$ 。

总之，该定标模型应进行偏差校正。



图中:

- | | | | |
|---|---------------------------|---|--------------|
| 1 | 警戒线上限 (UAL, $+3s_{SEP}$) | X | 样品数 |
| 2 | 报警线上限 (UWL, $+2s_{SEP}$) | Y | 参考值与近红外预测值之差 |
| 3 | 报警线下限 (LWL, $-2s_{SEP}$) | | |
| 4 | 警戒线下限 (LAL, $-3s_{SEP}$) | | |

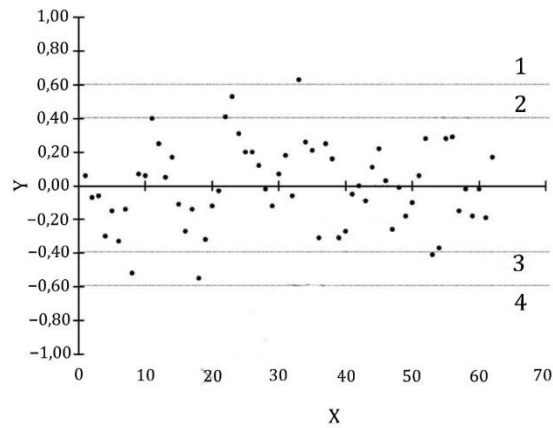
图 B.6 谷物饲料原料中脂肪含量测定的控制图

图 B.7 为某模型中一项参数（范围：44%-57%）测定的控制图，在第 35 点处进行了重新定标。

观察前 34 个检测点，其中有 1 点超出了警戒线上限 UAL，这意味着严重问题的出现。连续 3 点中的 2 点（22 和 23）超出了报警线上限 UWL，2 个不连续点同样超出了报警线下限 LWL。样品点均衡

分布于零线附近（遵守 9 点规则），但 34 个点中有 5 个在 95%置信区间[UWL, LWL]之外，并且有 1 个点超出了 99.9%的置信区间[UAL, LAL]。这远远超出了期望。

出现这种结果的原因之一是计算临界值的 s_{SEP} 值过于理想，这意味着临界值应该要适当放宽。另一种原因可能是实测样品与定标样本有些不同。为检测这种可能，对定标集进行扩展，使之包含质控样本，并生成新的定标模型。通过 35-62 质控样本的结果可以看到，新定标模型的性能得到了明显改善。



图中：

- | | | | |
|---|----------------------------|---|--------------|
| 1 | 警戒线上限 (UAL, +3 s_{SEP}) | X | 样品数 |
| 2 | 报警线上限 (UWL, +2 s_{SEP}) | Y | 参考值与近红外预测值之差 |
| 3 | 报警线下限 (LWL, -2 s_{SEP}) | | |
| 4 | 警戒线下限 (LAL, -3 s_{SEP}) | | |

图 B.7——某模型中一项参数测定的控制图（范围：44%-57%）

附录 C

(资料性附录)

补充性条款及定义

总则	
参考方法	获得 ISO 或其他国际认可和经过验证的检测方法，能够给出待测参数的包括测量不确定度的“真值”或“设定值”。
间接法	通过检测与待测指标具备函数关系的某些特性来测定待测指标的方法，该方法所获取的信号与利用参考方法测定的“真值”相关。
近红外光谱技术(NIRS)	近红外光谱，是指在 700 到 2500 nm (14300 至 4000 cm^{-1}) 的近红外光范围内，测量样品对光的吸收强度。NIRS 仪器使用该光谱范围的部分或整体，或更宽的能包含该范围的光谱区域（如 400 nm 至 2500 nm）进行检测，然后采用多元校正分析技术将吸光度值与样品成分含量或属性进行关联。
近红外反射 (NIR)	其检测基础是利用样本前侧传感器对所吸收到的样本表面漫反射回来的近红外光进行检测。
近红外透射 (NIT)	其检测基础是利用样本背侧传感器对所吸收到的透射过样本的近红外光进行检测。
NIRS 网络	使用相同定标模型的一定数量的近红外仪器。网内仪器通常需要标准化，从而能使标准样品预测值的差异最小。
仪器的标准化	对一组近红外仪器进行调整的过程。通过调整，当利用统一的定标模型时，所有设备对相同的样本获取相近的预测值。有许多技术可用，但若从广义的范畴进行定义，这些技术可以分为两种：1) “前预测法”，即通过调整样本光谱使“主”仪器与该组内的每台仪器之间的差异最小；2) “后预测法”，即使用线性回归分析调整每台仪器的预测值，使之尽可能接近“主”仪器的预测值。
z-得分	利用近红外预测值与参考值的差值除以标准偏差（例如 RMSEP）得到的一项性能指标。

定标技术	
主成分分析 (PCA)	<p>主成分分析是一种数据压缩方式。只对 x（光谱）数据进行计算，每个主成分都表达了光谱数据投影的最大方差，并且这些主成分互不相关。第一主成分尽可能多的表达了原始数据的变化。然后将其影响从 x 数据中扣除，新的主成分尽可能多的表达剩余数据的特征。能够得到和光谱数据点或数据集中样本个数相同的主成分数，但光谱中主要影响集中在前几个主成分，因此需要考虑的数据量就明显降低了。</p> <p>在每一步计算中，主成分分析 (PCA) 产生两个新的变量集：主成分得分代表了每个样本在每个主成分上的响应；主成分载荷代表了原始光谱中每个数据点对主成分的相对重要性。</p> <p>主成分分析有很多用途，例如光谱解析，但是最广泛的用途是用于光谱异常值的判别。</p>
主成分回归 (PCR)	主成分回归是在多元线性回归中，以每个主成分作为回归量、样品的成分含量作为 y 值的一种回归分析技术。因为主成分是相互正交的，因此

	<p>主成分得分可形成性能优于原始光谱数据的互不相关的数据集。尽管能够以每个主成分与目标成分含量相关性的优劣为依据选择回归的主成分组合,但大多数商业软件会强制采用为模型所选择的最大主成分数进行回归。</p> <p>当分析近红外数据时,通常将主成分空间的回归系数转换回波长空间中用所有数据点建立的预测模型。</p>
偏最小二乘回归 (PLS)	<p>偏最小二乘回归是一种分析技术,像 PCA 一样,也是一种数据压缩形式。其提取因子的规则是每个因子能够依次实现 y 值和 x 数据所有可能线性组合之间最大的协方差。PLS 在计算主成分时,在考虑所计算的主成分方差尽可能最大的同时,还使主成分与浓度最大程度地相关。因此与主成分相比,PLS 因子与 y 值的变化更直接相关。PLS 生成三个新变量,载荷权重(之间不是正交关系)、载荷和得分(相互正交的)。</p> <p>PLS 模型通过回归 y 值和 PLS 得分获得。和主成分回归一样,当分析近红外数据时,通常将 PLS 空间的回归系数转换回波长空间中用所有数据点建立的预测模型。</p>
多元线性回归 (MLR)	<p>多元线性回归是联合几个 X 变量预测一个 Y 变量的回归技术。对近红外光谱技术,X 变量既可以是选择波长处的吸光度值,又可以是像主成分得分和偏最小二乘得分等类似的派生变量。</p>
人工神经网络 (ANN)	<p>人工神经网络是基于生物神经系统结构的非线性模型技术。该网络最初通过几个具有 X 值(光谱数据或如主成分得分等派生变量)和 Y 参考值的数据集进行“训练”。在训练过程中,为产生最好的参数值预测,对输入变量和输出变量,可能需要修改网络结构,以及重新分配神经原加权系数。</p> <p>训练时神经网络需要大量数据。对 ANN 的一些优化方案,需要设定一个停止点以终止模型的优化(例如:权重)。</p>
多变量模型	<p>多变量模型是用一组 X 变量预测一个或多个 Y 变量的模型。</p>
异常值	<p>从统计的角度讲,异常值是数据集内位于预期分布之外的那些点。对 NIRS 数据来说,异常值通常分为 X 异常(光谱异常)或 Y 异常(参考值异常)。</p>
X-异常值	<p>X-异常值是与近红外光谱相关的异常值。X-异常值可能由以下原因引起:仪器故障、完全不同于其他样品的样品类型、或预测的样品类型不包含在原始定标集中。</p>
Y-异常值	<p>Y-异常值是与参考数据中的错误相关的异常值,例如参考数据誊写错误或由于参考实验室获得的参考值错误。</p>
杠杆值	<p>是一个度量样品与模型所定义的群体空间中心距离的测量值。具有高杠杆值的样品对模型有非常大的影响。杠杆值通过计算一个投影点和模型中心的距离得到。</p>
马氏距离	<p>马氏距离是数据点和主成分空间中心在主成分空间的距离(参见下面的 h 值),这是一个线性测量值。在主成分空间,一组样品通常形成一个曲线形状的分布。最能代表数据集概率分布的椭圆柱体可以通过构建样本的协方差矩阵进行估计。简单说,马氏距离通过测试点与数据集中心的距离除以测试点方向椭圆柱体的宽度得到。</p>
h 值	<p>在一些软件中,马氏距离被称为“全局 h 值”,并且通过样品距数据集中心 h 值的标准偏差来检测异常值。另一个是“邻近 h 值”,它是一个数据点和离它最近 n 个数据点在主成分空间的距离,它表明该样品是孤立的、还是隶属于某一集中分布的。</p>

残差	残差是参考值和回归模型预测值之间的差异，用于回归统计计算。
光谱残差	化学计量学（如 PCA、PLS 等）处理光谱后的残差，是模型未描述的光谱变异。
检验集	当进行一个回归模型检验时，除定标过程中用到的样本集之外的任意样品集。
独立检验集	与建立和验证回归模型的样品集相比，来自不同地理区域、新工业厂房或不同收集时间（例如不同的收获期）样品组成的检验集。这些样品构成了对预测模型的“真正”检验。
验证集	用于验证或“证明”定标的样品集。验证集通常包括和定标样品有相同属性的样品，经常从同一样品池中，按照轮流或取第 n 个样品（按分析成分含量排序）的方法选择定标集和验证集。
监控集	用于定标模型常规控制的样品集。
交互验证	<p>计算预测统计值的方法，重复从定标样品中移出不同的子集，用剩余的定标集样品建立定标模型，并用移出的样品作为验证集计算残差；当这个过程运行多次后，用所有的残差计算预测统计值。全交互验证每次移出一个样品，共运行 n 次（n 是定标样品数）。如需选择移出更大的子集，则应保证在计算预测统计值之前运行交互验证循环 8 次以上。最终，使用所有的定标样本计算出一个定标模型。</p> <p>需谨慎使用交互验证。首先，交互验证统计值与独立检验集相比，更趋于给出偏乐观的结果。第二，如果在定标集中有重复样品（例如：相同的样品在几台仪器上扫描或扫描不同的次数），要确保每次把该样品所有的重复数据一起移出作为验证集样品，否则将会给出过于理想的统计结果。</p>
过拟合	在一个多元线性回归中使用了太多的回归变量。当对不在定标集中的样品进行预测时，过拟合模型的预测结果会比期望值差的多，像 S_{RMSEP} 或 S_{SEP} 统计值。
PLS 因子	参见 PLS
得分/得分图	一个主成分或偏最小二乘因子的得分与另一主成分或偏最小二乘因子得分的散点图。标注了样品编号或成分含量值的得分图更加实用，能够容易的识别出原始数据中不明显的规律。

统计表述	也可参见第 7 章。
偏差 (\bar{e})	参考值平均值 \bar{y} 和近红外预测值平均值 $\bar{\hat{y}}$ 的差异。
偏差置信区间 (BCL, T_b)	位于置信区间之外，表明偏差与 0 存在给定置信水平上的显著差异(参见 7.3)。
定标标准误差 (SEC, s_{SEC})	对一定标模型，定标标准误差 (SEC) 是建模样品的预测值和参考值的平均差异。在此及后续统计中，该平均差异是指残差值平方和除以自由度的平方根，68% 的误差小于该值。
交互验证标准误差 (SECV, s_{SECV})	对给定的定标模型，交互验证标准误差 (SECV) 是指，交互验证 (参见 C.2.19) 过程中，选择定标样品的子集作为预测样品，其预测值和参考值经偏差校正过的平均差异。
预测标准误差 (SEP, s_{SEP})	预测标准误差 (SEP) 是指，选择不包含于定标过程中的样品作为预测样品，其回归模型预测值和参考值经偏差校正过的平均差异。SEP 涵盖了 68% 的置信区间 (该值乘以 1.96 就是 95% 的置信区间)。
预测误差均方根	预测误差均方根 (RMSEP) 是指，选择不包含在定标过程中的样品作

(RMSEP, s_{RMSEP})	为预测样品, 其参考值和回归模型预测值间的平均差异。 注: RMSEP 包含了预测过程中所有偏差。
交互验证误差均方根 (RMSECV, s_{RMSECV})	交互验证误差均方根 (RMSECV) 是指, 交互验证 (参加 C.2.19) 过程中, 选择定标样品的子集作为预测样品, 其预测值和参考值的平均差异。 注: RMSECV 包含了预测过程中所有偏差。
随机误差置信区间 (UECL, T_{UE})	随机误差置信区间是指, 在给定置信水平上 s_{SEP} 与 s_{SEC} 存在显著性差异的界限。
相关系数的平方 (RSQ, R^2)	预测值和参考值多重相关系数的平方。当用百分比表示时, 代表回归模型可以解释的变量的比例。
斜率 (b)	对任一回归线, 斜率代表 X 每增加 1 个单位对应的 Y 的增量。
截距	对任一回归线, 截距代表 X 等于 0 时 Y 的值。
残差标准偏差 (s_{res})	残差标准偏差是指, 参考值和经斜率和截距校正后的拟合值之间差异的平均尺寸。
协方差 (s_{yy})	协方差是测量两个随机变量一起变化的相关程度。对于样本总体, 如果 y 随 x 增大而增大, 那么这两个变量的协方差是正值。如果 y 随 x 增大而减小, 那么这两个变量的协方差是负值。当两个变量不相关时, 其协方差为 0。
不确定度 (U_e)	不确定度, 是用来描述合理地分布于结果附近的测量值的离散度的参数。

参考文献

- [1] ISO 712, Cereals and cereal products — Determination of moisture content — Reference method
- [2] ISO 5983-2, Animal feeding stuffs — Determination of nitrogen content and calculation of crude protein content — Part 2: Block digestion and steam distillation method
- [3] ISO 6492, Animal feeding stuffs — Determination of fat content
- [4] ISO 6496, Animal feeding stuffs — Determination of moisture and other volatile matter content
- [5] ISO 6865, Animal feeding stuffs — Determination of crude fibre content — Method with intermediate filtration
- [6] ISO 9622 | IDF 141, Milk and liquid milk products — Guidelines for the application of mid-infrared spectrometry
- [7] ISO 11085, Cereals, cereals-based products and animal feeding stuffs — Determination of crude fat and total fat content by the Randall extraction method
- [8] ISO 13906, Animal feeding stuffs — Determination of acid detergent fibre (ADF) and acid detergent lignin (ADL) contents
- [9] ISO 16472, Animal feeding stuffs — Determination of amylase-treated neutral detergent fibre content (aNDF)
- [10] ISO 16634-1, Food products — Determination of the total nitrogen content by combustion according to the Dumas principle and calculation of the crude protein content — Part 1: Oilseeds and animal feeding stuffs
- [11] ISO/TS 16634-2, Food products — Determination of the total nitrogen content by combustion according to the Dumas principle and calculation of the crude protein content — Part 2: Cereals, pulses and milled cereal products
- [12] ISO 20483, Cereals and pulses — Determination of the nitrogen content and calculation of the crude protein content — Kjeldahl method
- [13] ISO 6497, Animal feeding stuffs — Sampling
- [14] ISO 24333, Cereals and cereal products — Sampling
- [15] ISO 21543 | IDF 201, Milk products — Guidelines for the application of near infrared spectrometry
- [16] NÆS, T., ISAKSSON, T., FEARN, T., DAVIES, T. A user-friendly guide to multivariate calibration and classification. Chichester: NIR Publications, 2002. 344 p.
- [17] SHENK, J.S., WESTERHAUS, M.O., ABRAMS, S.M. Protocol for NIRS calibration: Monitoring analysis results and recalibration. In: Marten, G.C., Shenk, J.S., Barton, F.E., editors. Near infrared reflectance spectroscopy (NIRS): Analysis of forage quality, pp. 104-110. Washington, DC: US Government Printing Office, 1989. (USDA ARS Handbook 643.)
- [18] SØRENSEN, L.K. Use of routine analytical methods for controlling compliance of milk and milk products with compositional requirements. IDF Bull. 2004, (390), pp. 42-49