

团 体 标 准

T/CAGDRS XX—2026

土壤含盐量实验室高光谱监测技术规程

Technical specification for laboratory hyperspectral monitoring
of soil salt content

征求意见稿

2026-XX-XX 发布

2026-XX-XX 实施



中国农业绿色发展研究会 发布

目 次

前 言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 监测技术流程 2

5 土壤样品采集 2

6 土壤样品制备 3

7 土壤含盐量测定 3

8 土壤含盐量高光谱测量 3

9 土壤含盐量高光谱监测模型 4

10 试验报告 5

附 录 A （资料性附录） 光谱预处理与数学变换方法 6

附 录 B （资料性附录） 竞争性自适应重加权采样（CARS）特征筛选方法 9

附 录 C （资料性附录） 土壤含盐量监测模型 10

附 录 D （资料性附录） 样品采样记录表示例 12

参考文献 13

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。

文件由中国农业科学院农业资源与农业区划研究所提出。

本文件由中国农业绿色发展研究会归口。

本文件起草单位：中国农业科学院农业资源与农业区划研究所，国家盐碱地综合利用技术创新中心，北京市农林科学院信息技术研究中心，自然资源部国土整治中心。

本文件主要起草人：陆苗，吴文斌，查燕，田恬，李明丽，王虹扬，宋英强，李存军，钱建平，李玉义，温彩运，武雪萍，朴艺璇，农吉静，王宇文，张建财。

土壤含盐量实验室高光谱监测技术规程

1 范围

本文件确立了受控实验室条件下利用高光谱反射数据监测土壤含盐量的技术流程，规定了资料收集与整理、土壤样品采集与制备、土壤含盐量测定、高光谱数据测量、光谱数据处理、特征筛选、监测模型建立、模型验证等技术要求。

本文件适用于土壤含盐量的实验室高光谱快速监测及盐渍化监测工作。

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 36199 土壤质量 土壤采样程序设计指南

GB/T 42363 土壤质量 土壤理化分析样品的预处理

NY/T 1121.1 土壤检测 第1部分：土壤样品的采集、处理和贮存

NY/T 1121.16 土壤检测 第16部分：土壤水溶性盐总量的测定

TD/T 1043.1 暗管改良盐碱地技术规程 第1部分：土壤调查

3 术语和定义

下列术语和定义适用于本文件。

3.1

高光谱反射光谱 hyperspectral reflectance spectrum

在一定波长范围内，由连续或近连续窄波段反射率构成，能够表征样品光谱特性随波长变化规律的数据。

3.2

土壤含盐量 soil salt content

土壤中水溶性盐的总量，以每千克干土中水溶性盐的克数表示，单位为克每千克(g/kg)。

[来源：TD/T 1043.1—2013]

3.3

监测模型 estimation model

利用具有土壤含盐量参考值的样品，建立高光谱反射数据与土壤含盐量之间定量关系的数学模型。

3.4

训练集 training set

用于光谱预处理方法选择、特征筛选、模型参数优化和监测模型建立的样品集合。

3.5

独立验证集 independent validation set

在模型建立和参数确定过程中保持封存，仅用于评价锁定模型预测性能的样品集合。

4 监测技术流程

土壤含盐量实验室高光谱监测技术流程包括：

- a) 土壤样品采集；
- b) 土壤样品制备；
- c) 土壤含盐量测定；
- d) 土壤含盐量高光谱测量；
- e) 土壤含盐量高光谱监测模型；
- f) 试验报告。

5 土壤样品采集

5.1 资料收集

5.1.1 采样前应收集研究区土壤含盐量空间分布相关资料，包括土壤类型、土地利用现状、盐渍化程度、地形地貌、水文地质、地下水埋深、灌排条件及历史监测资料等。

5.1.2 根据土壤含盐量空间分布资料划分采样单元，保证采样点具有代表性，能够反映研究区土壤含盐量空间差异。

5.2 采样单元划分和样点布设

5.2.1 根据土壤类型、土地利用方式、盐渍化程度、地形和灌排条件划分采样单元。采样单元应覆盖拟应用模型范围内的主要土壤类型和含盐量梯度。

5.2.2 采样点布设及采样程序设计应符合 GB/T 36199 的有关规定。采样点宜采用网格法均匀布设，并结合土壤类型、土地利用方式、盐渍化程度、地形及灌排条件等因素合理调整，使采样点覆盖主要土壤类型和不同含盐量水平，保证样品具有代表性。

5.3 采集工具

采集工具包括土钻或土铲、洁净样品袋、防水标签、记号笔、全球导航卫星系统(GNSS)定位设备和野外记录表。工具与容器应洁净、干燥，不应引入可溶性盐污染。样品采样记录表示例见附录 D。

5.4 采样深度

一般采集 0 cm~10 cm 土层样品。开展耕层或剖面研究时，应根据研究目的预先规定土层，各样点采用相同的分层规则。地表存在盐壳时，应明确盐壳是否纳入样品并在记录中说明。

5.5 混合采样

5.5.1 混合采样应符合 NY/T 1121.1 中有关混合采样的原则。在能够代表目标空间尺度的范围内采集不少于 5 个子样，各子样采样深度和采样量应保持一致。

5.5.2 各子样充分混匀后采用四分法缩分，缩分后样品质量宜为 500 g。

5.6 样品标识、运输和贮存

样品装袋后应立即标识，标签信息应与采样记录一致。样品应避免雨淋、渗漏、交叉污染和高温暴晒，并按照 NY/T 1121.1 的规定运输和贮存。

6 土壤样品制备

6.1 土壤样品预处理应符合 GB/T 42363 的有关规定，风干样品的处理和贮存应符合 NY/T 1121.1 的规定。

6.2 制备后的土壤样品应充分混匀并分为两份，分别用于土壤含盐量测定和高光谱数据采集，保证实测含盐量数据与光谱数据来源一致。

7 土壤含盐量测定

7.1 土壤含盐量测定应按照 NY/T 1121.16 的规定执行。

7.2 称取通过 2 mm 筛孔的风干土壤样品 50 g（精确至 0.01 g），加入无二氧化碳蒸馏水 250 mL（水土比 5:1），振荡后过滤，制备土壤浸提液。

7.3 弃去初滤液，取 20 mL~50 mL 清亮滤液于已知质量的蒸发皿中蒸干，必要时进行有机质去除处理。

7.4 根据烘干残渣质量计算土壤水溶性盐总量，测定结果作为土壤含盐量实测值，单位为克每千克（g/kg）。

8 土壤含盐量高光谱测量

8.1 测量环境

光谱测量宜在常温暗室环境中进行。测量期间应保持环境条件稳定，并控制外部杂散光、振动和气流等因素对测量结果的影响。

8.2 仪器与设备

光谱仪测量波长范围宜覆盖 350 nm~2500 nm，光谱分辨率应不大于 10 nm，并配备视场角为 1°~25° 的探头、稳定人工光源、标称反射率不低于 95% 的标准白板、光谱采集软件和计算机等设备。

8.3 仪器启动与预热

测量前应连接光谱仪、计算机及相关设备，启动光谱仪和光源，并按照仪器使用说明进行预热。无明确规定时，预热时间不应少于 30 min。预热完成后，应检查仪器连接状态和光谱信号稳定性。

8.4 光源与测量几何

8.4.1 宜采用卤素灯或仪器配套光源作为测量光源，根据光照强度调整光源与样品之间的距离和入射角度，使光线均匀覆盖样品有效测量区域。同一测量批次内，光源位置、入射角度以及光源与样品之间的距离应保持一致。

8.4.2 探头应垂直于样品表面，并根据探头视场角设置探头与样品表面的距离。探头视场应完全覆盖于样品有效测量区域内，且测量视场内不应包含样品容器边缘、杂物或阴影。同一测量批次内，探头位置和测量高度应保持一致。

8.5 样品装填

将制备好的土壤样品均匀装入样品皿。样品皿宜采用直径约 10 cm、深度约 2 cm 的培养皿。装样后应轻敲样品皿，使样品自然沉实，并将样品表面刮平。所有样品应采用一致的装样量、装样厚度和表面处理方式。

8.6 暗电流与标准白板校正

8.6.1 样品测量前，应按照仪器操作要求采集暗电流，并使用标称反射率达到 95% 及以上的标准白板采集参考光谱。标准白板与土壤样品应采用相同的光照条件和测量几何，通过参考光谱校正消除光源强度及传感器响应差异，获得样品反射率光谱。

8.6.2 测量过程中，每测量 5 个样品后，应重新进行标准白板校正。光源位置、探头高度或环境条件发生变化时，应重新进行暗电流和标准白板校正。

8.7 重复测量与数据处理

为降低样品表面不均匀性对光谱测量结果的影响，每个样品应按照 0° 、 90° 、 180° 和 270° 四个方向进行重复测量，每个方向采集的有效光谱曲线数量不应少于 5 条。完成光谱质量检查后，应对同一样品的有效光谱曲线进行算术平均，获得该样品的反射率光谱数据。

9 土壤含盐量高光谱监测模型

9.1 数据集建立

应根据土壤样品含盐量的分布特征，采用分层随机划分方法建立训练集和独立验证集，训练集与独立验证集的样品比例应为 7:3。训练集用于光谱特征选择与模型构建，独立验证集用于模型精度验证。

9.2 数据预处理

9.2.1 原始光谱反射率数据宜进行预处理，以降低仪器噪声、随机波动及非目标因素对光谱信息的影响，提高光谱数据质量。

9.2.2 光谱预处理宜采用 Savitzky-Golay (SG) 平滑方法对原始光谱反射率进行处理，以减少随机噪声并保持光谱曲线特征。根据仪器信噪比和测量条件，可剔除边缘低信噪比波段。

9.2.3 经预处理后的光谱数据，可根据建模需求进行数学变换，以增强光谱特征信息和光谱反射率与土壤含盐量之间的响应关系。常用数学变换方法包括一阶导数 (FD)、二阶导数 (SD)、标准正态变量变换 (SNV)、多元散射校正 (MSC)、对数变换 (Log) 和倒数变换 (REC) 等，数据预处理及数学变换方法见附录 A。

9.3 光谱特征选择

9.3.1 经光谱预处理和数学变换后的光谱数据，应进行特征波段筛选，以减少冗余光谱变量和无效信息，提高模型构建效率和稳定性。

9.3.2 特征波段筛选宜采用竞争性自适应重加权采样（competitive adaptive reweighted sampling, CARS）等变量筛选方法，筛选与土壤含盐量监测相关的特征波段组合。采用 CARS 方法进行特征波段筛选时，应基于偏最小二乘回归模型，通过蒙特卡洛采样和变量权重评价确定特征波段组合，并以交叉验证均方根误差（RMSECV）等模型评价指标确定筛选结果。

9.3.3 完成特征波段筛选后，应利用筛选得到的特征波段数据进行后续土壤含盐量监测模型构建。CARS 方法具体流程见附录 B。

9.4 监测模型构建

9.4.1 根据训练集样品数据，以筛选得到的土壤含盐量敏感光谱特征为自变量，以土壤含盐量实测值为因变量，建立土壤含盐量高光谱监测模型。

9.4.2 可采用偏最小二乘回归（Partial Least Squares Regression, PLSR）、随机森林回归（Random Forest Regression, RF）、支持向量回归（Support Vector Regression, SVR）和极端随机树回归（Extremely Randomized Trees Regression, ExtraTrees）等方法构建土壤含盐量监测模型。土壤含盐量监测模型见附录 C。

9.5 模型验证与评价

9.5.1 模型建立后，利用独立验证集样本对模型监测效果进行评价。以决定系数（ R^2 ）和均方根误差（RMSE）作为评价指标。按照公式（1）计算决定系数，按照公式（2）计算均方根误差。

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

9.5.2 式中， R^2 表示决定系数；RMSE表示均方根误差，单位为 g/kg；n表示验证样品数量； y_i 表示第 i 个验证样品的土壤含盐量实测值，单位为 g/kg； $\hat{y}_i$ 表示第 i 个验证样品的土壤含盐量监测值，单位为 g/kg； $\bar{y}$ 表示验证样品土壤含盐量实测值的平均值，单位为 g/kg。

9.5.3 模型精度应结合 R^2 和 RMSE 指标综合评价，判断模型是否满足应用要求。当模型精度不能满足应用要求时，可通过增加样品数量、优化光谱预处理方法、改进特征筛选方法、优化模型参数或调整建模算法等措施，提高模型精度。

10 试验报告

试验报告至少应包括以下内容：

- a) 样品编号、来源、采样日期、采样深度和样品制备方法；
- b) 参考值测定方法及质量控制结果；
- c) 光谱仪型号、波长范围、测量几何、积分时间、扫描次数和校正记录；
- d) 光谱预处理方法、特征选择方法、模型算法和模型版本；
- e) 土壤含盐量监测结果、单位、精度；
- f) 异常情况、复核结果、试验人员和日期。

附 录 A

（资料性附录）

光谱预处理与数学变换方法

A.1 概述

土壤高光谱数据在采集过程中容易受到仪器噪声、样品颗粒散射以及基线漂移等因素影响。为降低非目标因素对光谱信息的影响，提高光谱特征提取能力，在模型构建前可对原始光谱进行预处理和数学变换。

本文件涉及的光谱预处理与数学变换方法主要包括 Savitzky-Golay（SG）平滑、标准正态变量变换（Standard Normal Variate, SNV）、多元散射校正（Multiplicative Scatter Correction, MSC）、一阶导数（First Derivative, FD）、二阶导数（Second Derivative, SD）、对数变换（Log）和倒数变换（Reciprocal Transformation, REC）。

A.2 Savitzky-Golay（SG）平滑

Savitzky-Golay（SG）平滑是一种基于局部多项式拟合的光谱平滑方法。该方法通过对一定窗口范围内的光谱数据进行多项式拟合，降低随机噪声影响，同时保持光谱曲线的主要变化趋势和特征峰位置。

设原始光谱反射率为 R_i ，经 SG 平滑后的光谱值为：

$$R_i^{SG} = \sum_{j=-m}^m C_j R_{i+j} \quad (\text{A.1})$$

式中：

R_i^{SG} ——第 i 个波段 SG 平滑后的反射率；

R_{i+j} ——第 $i+j$ 个波段原始反射率；

C_j ——SG 平滑系数；

m ——半窗口宽度。

A.3 一阶导数变换（First Derivative, FD）

一阶导数变换用于增强光谱曲线变化特征，减弱基线漂移对光谱信息的影响。

计算公式如下：

$$FD_i = \frac{R_{i+1} - R_i}{\lambda_{i+1} - \lambda_i} \quad (\text{A.2})$$

式中：

FD_i ——第 i 个波段的一阶导数光谱值；

R_i 、 R_{i+1} ——相邻波段光谱值；

λ_i 、 λ_{i+1} ——对应波长。

A.4 二阶导数变换（Second Derivative, SD）

二阶导数变换通过进一步计算光谱变化率，增强光谱局部变化特征。
计算公式如下：

$$SD_i = \frac{R_{i+1} - 2R_i + R_{i-1}}{(\Delta\lambda)^2} \quad (\text{A.3})$$

式中：

SD_i ——第 i 个波段的二阶导数光谱值；

$\Delta\lambda$ ——光谱采样间隔。

A.5 对数变换 (Log)

对数变换通过对光谱反射率进行对数转换，压缩高反射率区域差异，增强低反射区域光谱响应
计算公式如下：

$$R_i^{Log} = \log(R_i) \quad (\text{A.4})$$

式中：

R_i^{Log} ——第 i 个波段对数变换后的光谱值；

A.6 倒数变换 (Reciprocal Transformation, REC)

倒数变换通过计算光谱反射率倒数，提高光谱变化敏感性，增强部分吸收特征信息。
计算公式如下：

$$R_i^{REC} = \frac{1}{R_i} \quad (\text{A.5})$$

式中：

R_i^{REC} ——第 i 个波段倒数变换后的光谱值。

A.7 标准正态变量变换 (Standard Normal Variate, SNV)

标准正态变量变换通过对单条光谱进行中心化和标准化处理，降低由于样品粒径、表面粗糙度及散射效应引起的光谱差异。

计算公式如下：

$$R_i^{SNV} = \frac{R_i - \bar{R}}{s} \quad (\text{A.6})$$

式中：

R_i^{SNV} ——SNV 处理后的光谱值；

R_i ——原始光谱值；

$\bar{R}$ ——单条光谱平均值；

s ——单条光谱标准差。

A.8 多元散射校正 (Multiplicative Scatter Correction, MSC)

多元散射校正通过建立样品光谱与参考光谱之间的线性关系，对不同样品间由于散射效应造成的光谱差异进行校正。

其计算模型为：

$$R_i = a + bR_i^{ref} \quad (\text{A.7})$$

校正后的光谱值为：

$$R_i^{MSC} = \frac{R_i - a}{b} \quad (\text{A.8})$$

式中：

R_i^{ref} ——参考光谱值；

a ——线性拟合截距；

b ——线性拟合斜率；

R_i^{MSC} ——MSC 校正后的光谱值。

A.9 光谱变换方法选择

在土壤含盐量高光谱监测模型构建过程中，应根据光谱数据质量和模型性能选择适宜的光谱预处理与数学变换方法。

不同光谱处理方法的适用效果可采用训练集内部交叉验证指标进行评价，并据此选择适宜的光谱处理方式。独立验证集不应参与光谱处理方法选择和模型参数优化。

- a) SG 平滑，用于降低原始光谱噪声；
- b) 一阶导数（FD）和二阶导数（SD）变换，用于增强光谱曲线变化特征；
- c) 对数变换（Log）和倒数变换（REC），用于增强光谱响应差异，提高光谱特征表达能力；
- d) 标准正态变量变换（SNV）和多元散射校正（MSC），用于降低样品散射效应对光谱数据的影响。

附 录 B (资料性附录)

竞争性自适应重加权采样 (CARS) 特征筛选方法

B.1 概述

竞争性自适应重加权采样算法 (competitive adaptive reweighted sampling, CARS) 是一种基于蒙特卡洛采样和自适应加权策略的变量筛选方法。该方法通过筛选对模型贡献较大的光谱变量, 剔除冗余波段和无效信息, 提高模型稳定性和预测性能。

B.2 CARS算法流程

CARS 算法以偏最小二乘回归模型中变量回归系数绝对值作为变量重要性评价依据, 通过蒙特卡洛采样方法建立不同变量组合模型, 并利用指数衰减函数逐步减少变量数量, 结合自适应重加权采样保留贡献较大的变量。CARS 特征筛选主要包括以下步骤:

- a) 利用蒙特卡洛采样方法随机选择样本子集, 建立偏最小二乘回归模型;
- b) 计算各光谱变量的回归系数绝对值, 并根据变量权重评价其重要性;
- c) 依据指数衰减函数逐步剔除贡献较低的变量;
- d) 采用自适应重加权采样保留重要变量;
- e) 建立不同变量组合模型, 并评价模型性能;
- f) 根据模型评价结果确定最优特征波段组合。

B.3 参数设置与特征波段确定

CARS 算法主要参数包括蒙特卡洛采样次数 (迭代次数)、蒙特卡洛采样比例、交叉验证方式和最小保留变量数等。蒙特卡洛采样次数用于控制变量筛选过程的重复程度和搜索范围; 蒙特卡洛采样比例用于确定每次迭代参与模型构建的样本数量; 交叉验证方式用于评价不同变量组合的模型性能; 最小保留变量数用于控制特征筛选过程中保留变量数量的下限。

参数应根据样本数量、光谱变量数量及模型稳定性合理设置。在每次迭代过程中, 根据偏最小二乘回归模型中变量回归系数的绝对值评价变量重要性, 并结合指数衰减函数和自适应重加权采样逐步筛选变量。对不同迭代过程获得的特征变量组合建立偏最小二乘回归模型, 并采用交叉验证评价模型性能, 以交叉验证均方根误差 (RMSECV) 最小时对应的变量组合作为最优特征波段组合。

RMSECV 计算公式如下:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (B.1)$$

式中:

$RMSECV$ ——交叉验证均方根误差;

y_i ——第 i 个样本实测含盐量;

$\hat{y}_i$ ——第 i 个样本预测含盐量;

n ——样本数量。

附录 C

(资料性附录)

土壤含盐量监测模型

C.1 偏最小二乘回归模型 (Partial Least Squares Regression, PLSR)

偏最小二乘回归 (Partial Least Squares Regression, PLSR) 是一种适用于高维变量和多重共线性数据分析的线性回归方法。土壤高光谱数据波段数量较多, 相邻波段之间通常具有较强相关性。PLSR 通过从光谱变量中提取能够表征光谱信息并与土壤含盐量具有较强相关性的潜变量, 在降低变量维度和多重共线性影响的基础上建立光谱特征与土壤含盐量之间的回归关系。

设光谱特征矩阵为 X , 土壤含盐量为 Y , PLSR 将光谱特征矩阵分解为:

$$X = TP^T + E \quad (C.1)$$

土壤含盐量可表示为:

$$Y = Tq + F \quad (C.2)$$

式中:

X ——光谱特征矩阵;

Y ——土壤含盐量;

T ——潜变量得分矩阵;

P ——光谱特征载荷矩阵;

q ——土壤含盐量对应的回归系数向量;

E 、 F ——残差矩阵。

通过确定适宜的潜变量数量建立 PLSR 模型, 并利用建立的回归关系对待测样品的土壤含盐量进行监测。

C.2 随机森林模型 (Random Forest, RF)

随机森林 (Random Forest, RF) 是一种基于决策树的集成学习回归方法。该方法通过对训练样本进行随机抽样, 并在决策树节点分裂过程中随机选择部分特征建立多棵回归树, 从而降低单一决策树模型的随机性及过拟合风险, 提高模型的稳定性和泛化能力。

对于待测样品, 各回归树分别给出预测结果, 随机森林回归模型的最终预测结果为各回归树预测结果的平均值, 计算公式如下:

$$\hat{Y} = \frac{1}{N} \sum_{i=1}^N \hat{Y}_i \quad (C.3)$$

式中:

$\hat{Y}$ ——随机森林回归模型的土壤含盐量最终预测结果;

N ——回归树数量;

$\hat{Y}_i$ ——第 i 棵回归树的土壤含盐量预测结果。

随机森林能够处理复杂的非线性关系, 适用于土壤高光谱变量与含盐量之间复杂映射关系的建模。

C.3 支持向量回归 (Support Vector Regression, SVR)

支持向量回归（Support Vector Regression, SVR）是一种基于统计学习理论的回归方法。该方法通过引入核函数，将原始特征空间中的光谱变量映射到高维特征空间，并在允许一定预测误差的条件下寻找具有较好泛化能力的回归函数，从而建立光谱特征与土壤含盐量之间的关系。

SVR 回归函数可表示为：

$$\hat{Y}(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (\text{C.4})$$

式中：

$\hat{Y}$ ——待测样品的土壤含盐量预测值；

n ——训练样品数量；

x_i ——第 i 个训练样品的光谱特征；

x ——待测样品的光谱特征；

α_i 、 α_i^* ——模型训练得到的拉格朗日乘子；

$K(x_i, x)$ ——核函数；

b ——回归函数的偏置项。

核函数可用于描述样品之间的非线性关系。模型构建时应根据训练数据对核函数及相关参数进行优化，以提高土壤含盐量监测模型的预测性能。

C.4 极端随机树回归（Extremely Randomized Trees Regression, ExtraTrees）

极端随机树回归（Extremely Randomized Trees Regression, ExtraTrees）是一种基于多棵随机化决策树的集成学习回归方法。该方法在构建回归树时随机选择候选特征，并在候选特征范围内随机生成分裂阈值，以增加各回归树之间的差异性，降低模型方差，提高模型的稳定性和泛化能力。

对于待测样品，ExtraTrees 模型分别利用各回归树进行预测，并对所有回归树的预测结果取平均值作为最终监测结果，计算公式如下：

$$\hat{Y} = \frac{1}{M} \sum_{m=1}^M f_m(x) \quad (\text{C.5})$$

式中：

$\hat{Y}$ ——待测样品的土壤含盐量预测值；

M ——回归树数量；

$f_m(x)$ ——第 m 棵回归树对待测样品 x 的土壤含盐量预测结果；

x ——待测样品的光谱特征。

ExtraTrees通过增强决策树节点分裂过程的随机性，能够降低单棵决策树对训练数据局部特征的依赖，适用于土壤高光谱特征与土壤含盐量之间复杂非线性关系的监测。

附 录 D
(资料性附录)
样品采样记录表示例

样品采样记录表示例见表 D.1。

表 D.1 样品采样记录表示例

序号	记录项	内容示例
1	样点编号	A1-1
2	位置	山东省东营市
3	采集日期	2023.10.12
4	经度 (°)	118.659685E
5	纬度 (°)	37.383468N
6	土地利用类型	耕地
7	采样深度	0-10cm
8	作物或植被类型	大豆
9	土壤颜色	黄褐色
10	土壤湿度	较干
11	地表盐碱状况	无明显盐斑，盐壳

参 考 文 献

- [1] 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
- [2] 史舟 等, 2014. 土壤地面高光谱遥感原理与方法[M]. 北京：科学出版社.
- [3] 郭艳萍, 王雪梅, 赵枫, 李平平, 2025. 基于最优数学和小波变换的绿洲耕层土壤盐分RF高光谱反演. 农业工程学报, 41(3): 83-93.
- [4] 李明丽, 温彩运, 马东豪, 李存军, 王宇文, 康璐, 陆苗, 2025. 基于土壤高光谱特征波段优选的盐分预测模型构建. 中国农业科学, 58(20): 4054-4069.
- [5] 王涛, 喻彩丽, 姚娜, 张楠楠, 白铁成, 2018. MLR和PLSR的沙壤土盐分含量光谱检测对比研究. 干旱区地理, 41(6): 1295-1302.
- [6] Lin N, Ma X, Sui Y, Zhu R, Liu H, Wu M, Jiang R, 2026. Soil salinity estimation based on satellite hyperspectral and synthetic aperture radar remote sensing image fusion. Agricultural Water Management, 323: 110076.
- [7] Mahajan G R, Das B, Gaikwad B, Murgaonkar D, Desai A, Morajkar S, Patel K P, Kulkarni R M, 2021. Monitoring properties of the salt-affected soils by multivariate analysis of the visible and near-infrared hyperspectral data. Catena, 198: 105041.
- [8] Qiu Y, Xu J, Zhu L, Chen Z, Lou F, Chen S, Tan S, 2026. Moisture interference-free estimation of soil salinity using proximal hyperspectral sensing through external parameter orthogonalization correction, feature extraction and SVM-AdaBoost modelling. Soil Use and Management, 42(2): e70208.
- [9] Wu S, Wang C, Liu Y, Li Y, Liu J, Xu A, Pan K, Li Y, Pan X, 2018. Mapping the salt content in soil profiles using Vis-NIR hyperspectral imaging. Soil Science Society of America Journal, 82(5): 1259-1269.
-