



中华人民共和国国家标准

GB/T □□□□□—202□

生态环境损害鉴定评估技术指南

总纲和关键环节 第 4 部分：土壤生态环境

基线调查与确定

**Technical guidelines for identification and assessment of environmental
damage—General principles and key components**

—Part 4: survey and determination of soil environmental baseline

(征求意见稿)

202□-□□-□□发布

202□-□□-□□实施

生态环境部 发布
国家市场监督管理总局

目 次

前 言.....	ii
1 适用范围.....	1
2 规范性引用文件.....	1
3 术语和定义.....	1
4 工作内容和程序.....	2
5 工作准备.....	4
6 基于历史数据和对照区调查确定土壤生态环境基线.....	5
7 基于其他方法确定土壤特征污染物生态环境基线.....	10
附录 A（资料性附录） 正态分布检验方法.....	11
附录 B（资料性附录） Walsh 检验法.....	13

前 言

为贯彻《中华人民共和国民法典》《中华人民共和国环境保护法》《中华人民共和国土壤污染防治法》和《生态环境损害赔偿制度改革方案》，规范土壤生态环境损害鉴定评估的土壤生态环境基线调查与确定工作，制定本标准。

本标准规定了土壤生态环境损害鉴定评估过程中土壤生态环境基线调查与确定的内容、工作程序、方法和技术要求。

本标准附录A和附录B为资料性附录。

本标准为首次发布。

本标准由生态环境部组织制定。

本标准主要起草单位：生态环境部环境规划院、中国科学院南京土壤研究所、北京农业信息技术研究中心。

本标准自20□□年□□月□□日起实施。

本标准由生态环境部解释。

生态环境损害鉴定评估技术指南 总纲和关键环节

第4部分：土壤生态环境基线调查与确定

1 适用范围

本标准规定了土壤生态环境损害鉴定评估过程中土壤生态环境基线调查与确定的内容、工作程序、方法和技术要求。

本标准适用于因环境污染或生态破坏导致的土壤生态环境损害鉴定评估的土壤生态环境基线调查与确定。

本标准不适用于核与辐射所致土壤生态环境损害鉴定评估的土壤生态环境基线调查与确定。

2 规范性引用文件

本标准引用下列文件或其中的条款。凡是注明日期的引用文件，仅注日期的版本适用于本标准。凡是未注日期的引用文件，其最新版本（包括所有的修改单）适用于本标准。

GB/T 4882	数据的统计处理和解释 正态性检验
GB/T 4883	数据的统计处理和解释 正态样本离群值的判断和处理
GB 15618	土壤环境质量 农用地土壤污染风险管控标准（试行）
GB 36600	土壤环境质量 建设用地土壤污染风险管控标准（试行）
GB/T 39791.2	生态环境损害鉴定评估技术指南 总纲和关键环节 第2部分：损害调查
GB/T 39792.1	生态环境损害鉴定评估技术指南 环境要素 第1部分：土壤和地下水
HJ 25.2	建设用地土壤污染风险管控和修复监测技术导则
HJ/T 166	土壤环境监测技术规范
HJ 710.10	生物多样性观测技术导则 中型土壤动物
HJ 1019	地块土壤和地下水中挥发性有机物采样技术导则
HJ 1185	区域性土壤环境背景含量统计技术导则（试行）

3 术语和定义

下列术语和定义适用于本标准。

3.1

土壤生态环境损害 environmental damage of soil

因污染环境、破坏生态造成土壤环境及土壤中生物要素的不利改变，及上述要素构成的生态系统的功能退化和服务减少。

3.2

土壤生态环境基线 environmental baseline of soil

污染环境或破坏生态未发生时评估区土壤的化学元素或化合物含量、理化性质以及生物物种、种群、群落、生态服务功能等的水平或状态。

3.3

对照区 reference area

具有与评估区相同或相似的地质地球化学特征和生态分区，但未受到评估区损害行为影响的区域。

3.4

历史数据 historical data

能表征损害发生前评估区土壤化学元素或化合物含量、理化性质、生物物种、种群、群落、生态服务功能水平或状态的相关数据。

4 工作内容和程序

4.1 工作程序

土壤生态环境基线调查与确定的工作程序如图 1 所示。

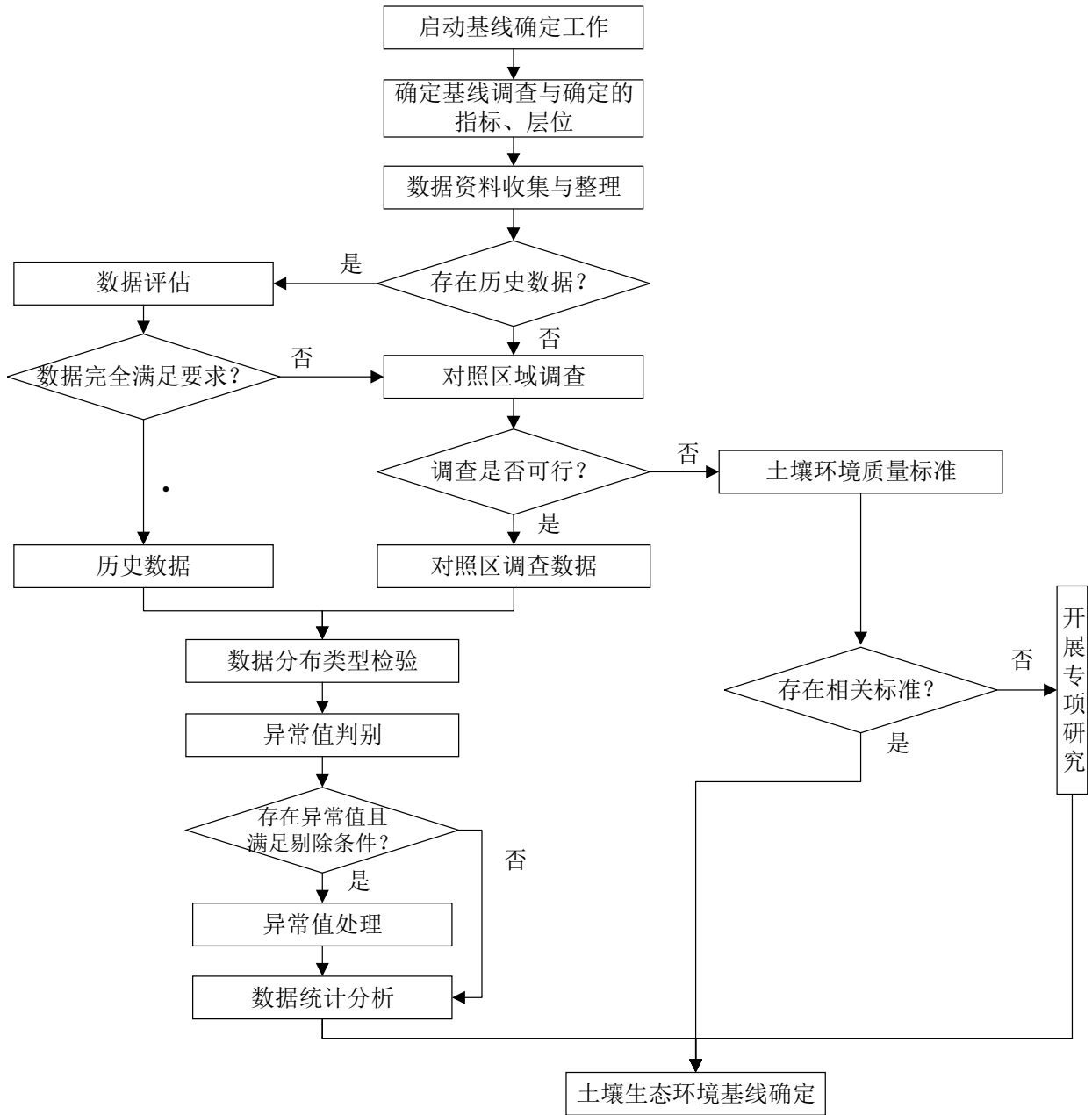


图1 土壤生态环境基线调查与确定工作程序

4.2 工作内容

4.2.1 工作准备

收集整理土壤生态环境基线调查与确定所需的相关数据资料，制定基线调查与确定工作方案。

4.2.2 基于历史数据和对照区数据确定土壤生态环境基线

- a) 当存在历史数据时，对历史数据的可用性进行评估；
- b) 当不存在可用的历史数据或历史数据不能完全满足要求时，应开展或结合历史数据开展对照区调查，获取数据；

c) 选择合适的对照区,开展对照区土壤生态环境调查,合理布设调查点位,采用规范的方法进行样品采集与分析监测,获取对照区土壤生态环境状况相关数据;

d) 综合可用的历史数据和对照区数据,通过数据分布类型检验、异常值判别处理、统计分析,确定土壤生态环境基线水平。

4.2.3 基于其它方法确定土壤生态环境基线

造成土壤环境质量下降的,当无法基于历史数据或对照区调查确定土壤生态环境基线,选择适用的土壤环境质量标准或基准确定土壤生态环境基线;当缺乏适用的标准或基准时,开展相关专项研究,确定土壤生态环境基线。

5 工作准备

通过资料收集分析、文献查阅、座谈走访、问卷调查、现场踏勘等方式,掌握评估区土壤生态环境损害的基本情况,了解评估区及周边的自然环境与社会状况,明确土壤生态环境基线调查与确定的工作内容,研究确定基线调查与确定的具体方法,编制基线调查与确定工作方案。

5.1 土壤生态环境损害相关信息调查

调查评估区土壤生态环境损害相关信息:

- a) 评估区土地利用变迁及相关影像资料(包括历史照片、遥感影像、航拍图片等);
- b) 评估区当前及历史上生产经营活动类型、时间、平面布置情况、地面硬化、地层人工改造、产排污、污染事故、生态破坏等情况;
- c) 评估区地层岩性、地下水埋深、地下水流向等地质和水文地质条件;
- d) 评估区土壤污染物类型、污染物空间分布等相关现状和历史监测数据;
- e) 评估区土壤理化性质,包括土壤类型、质地、土层厚度、砂砾含量、孔隙度、容重、养分含量、pH值、有机质含量等相关现状和历史监测数据;
- f) 评估区土壤生物,主要包括大型土壤动物种类组成、频度、密度、生物量等相关现状和历史监测数据。

调查周边区域土壤生态环境相关信息:

- a) 周边土地利用变迁资料及相关影像资料(包括历史照片、遥感影像、航拍图片等);
- b) 周边相关生产经营活动类型、时间、产排污情况;
- c) 评估区所在区域地层岩性、地下水埋深、地下水流向等地质和水文地质条件;
- d) 周边土壤、地表水及地下水污染等相关现状和历史监测数据,大气污染排放及沉降情况;
- e) 周边区域土壤理化性质相关现状和历史监测数据;
- f) 周边区域土壤生物,主要包括大型土壤动物种类组成、频度、密度、生物量等相关现状和历史监测数据。

5.2 自然环境与社会经济信息收集

收集区域的自然环境信息:

- a) 地理位置、地形地貌、水文、气候气象资料;
- b) 地质和水文地质资料;
- c) 土地利用的历史、现状、规划信息;
- d) 居民区、河流、饮用水水源地、生态保护红线、自然保护区、湿地、风景名胜区等环境敏感区分布信息以及主要生物资源的分布状况;

e) 厂矿、水库、构筑物、沟渠、地下管网、渗坑及其他面源污染等分布情况。

收集区域的社会经济信息：

- a) 经济和主要产业的现状和发展状况；
- b) 国家和地方相关法规、政策与标准等信息；
- c) 人口、交通、基础设施、能源和水资源供给等信息。

5.3 制定工作计划

根据所掌握的土壤生态环境损害基本情况和收集到的自然环境、社会经济信息，明确要开展土壤生态环境基线调查与确定的指标和层位，设计基线调查与确定工作程序，研究确定基线调查与确定方法，编制基线调查与确定工作计划。

土壤生态环境基线可能涉及的具体指标包括特征污染指标（重金属、有机物、硫酸盐、氯化物等元素或化合物含量等）、理化性质指标（pH 值、质地、容重、有机质含量、阳离子交换量等）、生物指标（土壤生物种类、密度、生物量、种群特征、群落特征等）。

6 基于历史数据和对照区调查确定土壤生态环境基线

6.1 历史数据收集整理

针对要进行土壤生态环境基线调查与确定的指标，补充收集和整理评估区及周边损害行为发生前的相关历史数据，包括政府机关和权威机构发布的环境调查监测数据（包括全国土壤普查、重点行业企业用地土壤污染状况调查、土壤背景值调查、常规土壤环境监测等）、环评等专项调查研究报告中的数据、文献中记载的数据、环境保护规划等历史档案中记录的监测数据等，明确数据来源，同时获取数据对应的采样时间、采样点位、采样深度、采样方法、保存流转方法、检测方法、质量控制措施及结果等信息。

6.2 历史数据评估

历史数据应满足以下要求：

a) 历史数据应具有较好的时间代表性，数据获取时间应在损害行为发生之前，且尽可能接近损害发生时间，数据获取时间到损害行为发生期间，不存在导致土壤物理、化学和生物学状态发生明显变化的因素；

b) 历史数据应具有较好的空间代表性，数据采集点位尽可能位于评估区内部或者周边，如果位于评估区周边，应确保与评估区为同一土地利用类型或没有导致其物理、化学和生物学状态与评估区存在明显差异的因素，具体参照 6.4 要求；

c) 历史数据对应的调查点位数量应满足 6.5.1 和 6.5.2 要求，如果数量不满足要求，应开展补充调查；

d) 历史数据对应的调查点位在评估区内部或周边区域原则上应均匀分布，确保空间上的均匀性，具体应满足 6.5.3 要求；

e) 历史数据对应的调查点原则上应与需要确定基线水平的土壤位于同一深度，且土壤类型相同，具体应满足 6.5.4 要求；

f) 历史数据的采样、保存流转、检测等方法与评估区生态环境损害调查过程所用的方法相同或具有等效性、可比性，具体参照 6.6 要求；

g) 历史数据获取过程的质量控制措施及结果符合相关标准规范要求，为有效数据，具体参照 6.7 要求；

h) 不同来源且符合上述要求的历史数据可以进行合并,但对于不同时间采集的历史数据,应进行适当分析和筛选。

6.3 启动对照区调查条件

当历史数据不满足或不完全满足要求时,可在符合要求的历史数据的基础上,开展对照区调查。

6.4 对照区要求

对照区应具有与评估区相似的生境特征,能代表评估区未受损害前的状态,未受到评估区所受生态环境损害行为的影响。

对照区的选择具体应遵循以下原则:

- a) 根据需要,可选择一个或多个对照区,对照区的总面积应满足 6.5.1 表 1 要求;
- b) 尽量选择受人为干扰最小的区域,且除未受评估区污染环境或破坏生态行为影响外,受到其它污染环境或生态破坏行为的影响与评估区相同;
- c) 选择评估区外部,且尽可能靠近评估区;
- d) 气象条件、地形地貌、生境特征、土地利用类型、水文地质条件、地表径流条件、降雨入渗条件、社会经济条件等应与评估区类似;
- e) 土壤类型、性质及成土母质应与评估区相同;
- f) 土壤 pH、质地、容重、有机质含量、阳离子交换量等理化性质应与评估区相似;
- g) 指示性生物种群特征(密度、性别比例、年龄组成等)、群落特征(如多度、密度、盖度、丰度等)或生态系统特征(如生物多样性、植被覆盖度等)、生态服务功能等应与评估区受损前相似;
- h) 如大气沉降是导致土壤生态环境损害的原因之一,应选在评估区常年主导风向上风向;
- i) 如地表径流是导致土壤生态环境损害的原因之一,应选在评估区周边地势较高区域;
- j) 如地下水污染是导致土壤生态环境损害的原因之一,应选在评估区地下水上游区域。

不适宜作为对照区的区域:

- a) 人工开挖区、施工区、回填区或填埋区;
- b) 处置、处理、储存危险废物、化学品、固体废物或废水的区域;
- c) 道路(公路、铁路等)周边临近区域;
- d) 建筑物附近区域,尤其是存在油漆碎片等可能导致污染的废弃物的区域;
- e) 停车场等地面硬化区域;
- f) 地表径流汇水区;
- g) 大气污染源周边,如铸造厂、燃煤发电厂等;
- h) 发生过历史污染事故或污染排放的区域。

6.5 点位和深度要求

6.5.1 点位数量

点位数量应按照以下方法确定:

- a) 结合评估区面积,确定对照区面积和点位数量(见表 1)。原则上对照区点位布设密度与评估区保持一致,在保证最少点位数量的前提下,根据情况适度增加采样点,最少点位数量参照表 1 确定;
- b) 基于已收集的数据资料和前期调查获取的数据,获得变异系数,或基于估计的变异系数,计算基础点位数量,计算方法参照 HJ 1185 中基础样本数量;
- c) 对照区土壤环境调查点位数量应同时满足 a) 和 b) 最小样本数量;
- d) 当涉及多类调查指标时,按照每类指标的要求分别布设调查点位,不同指标的调查点位可以协

调、合并。

表 1 对照区水平方向点位数量

序号	评估区面积/m ²	对照区面积/m ²	调查指标类型	最少点位数量/个
1	<64000	2000~8000	pH、有机质、氯化物、硫化物、氰化物、氟化物、有机物等	≥5
2	64000~128000	8000~16000	pH、有机质、氯化物、硫化物、氰化物、氟化物、有机物等	≥10
3	128000~256000	16000~32000	pH、有机质、氯化物、硫化物、氰化物、氟化物、有机物等	≥20
4	>256000	>32000	pH、有机质、氯化物、硫化物、氰化物、氟化物、有机物等	≥30
5	<32000	2000~8000	重金属	≥5
6	32000~64000	8000~16000	重金属	≥10
7	64000~128000	16000~32000	重金属	≥20
8	>128000	>32000	重金属	≥30
9			土壤理化性质指标	≥5

6.5.2 针对土壤生物调查的点位数量

土壤生物调查点位按照以下方法确定：

a) 参照 GB/T 39791.2，在每一损害类别区域及其对照区内至少各设置三个样方，如立地条件存在较大差异应适当增加样方数量；

b) 参照 HJ 710.10，使用随机数表和指南针选取样方位置，样方应设立在每类损害区域内的代表性区域中，每个样方面积为 25 m²（5 m×5 m）；

c) 在每个样方中设 2 个 30 cm×30 cm 均匀分布的样点。

6.5.3 布点方法

对照区土壤环境调查，一般采用系统布点法进行布点。将对照区划分成面积相等的网格，每个网格内布设 1 个采样点，结合对照区形状、面积、点位数量确定网格间距。根据实际情况，结合对照区形状，调整网格的位置，使采样覆盖范围最大化。

如果有多个对照区，对每个对照区分别采用系统布点法进行布点。

如果无法找到均匀的对照区，也可适当结合判断布点法进行对照区点位布设，确保点位代表性。

对照区土壤生物观测样点设置，参照 HJ 710.10 中的观测样地和样点的设置。

如布设的点位落在以下位置，可根据实际情况进行调整：

- 集中施用农药、肥料等投入品的位置；
- 土壤存在颜色、气味等异常情况的位置；
- 水土流失严重或表土被破坏的位置；
- 其它可能影响土壤特征污染指标、理化性质指标、生物指标的位置。

6.5.4 调查深度

垂直方向采样层位原则上应与评估区相同（确保土壤性质相同的前提下，尽可能采集相同深度样品），除非有证据表明不同层位土壤中相关指标没有明显变化。

6.6 样品采集、分析检测

历史数据获取和对照区调查应采用与评估区调查相同的样品采集、保存、流转和分析检测方法。

土壤钻探和土壤环境样品采集、保存要求参照 HJ 25.2。涉及钻探时，应为无浆液钻进，全程套管跟进。土壤样品的流转要求参照 HJ/T 166。涉及挥发性有机污染物时，应符合 HJ 1019 的相关规定。土壤生物样品采集、保存和处理要求参照 HJ 710.10。

针对 5.3 确定的调查指标，进行土壤环境样品分析检测和生物观测。土壤环境样品分析检测方法应符合 GB 36600 要求；涉及农用地时，应符合 GB 15618 要求。土壤生物样品观测方法应符合 HJ 710.10 要求，在对土壤动物计数时，仅计数最大直径大于 2 mm 的个体，主要为大型土壤动物。

6.7 质量保证与质量控制

为保证土壤生态环境基线调查数据具有代表性、准确性、精密性、可比性和完整性，质量控制与质量保证应涉及土壤生态环境基线调查的全过程。土壤环境样品采集、保存、流转、检测过程质量控制参照 HJ 25.2；特征指标涉及挥发性有机物时，质量控制还应遵循 HJ 1019 的相关规定。土壤生物样品采集、保存、流转、观测过程质量控制参照 HJ 710.10。

6.8 数据处理分析与基线确定

6.8.1 数据分布类型检验

检验样本是否符合正态分布的常见方法包括图形法（包括直方图、P-P 图、Q-Q 图等）、偏度峰度检验法、Shapiro-Wilk 检验法（W 检验）、Kolmogorov-Smirnov 检验法（K-S 检验）和卡方检验。不同方法的适用情形见表 2。

图形法、偏度峰度检验法、W 检验方法的使用参照 GB/T 4882，K-S 检验、卡方检验方法的使用参照附录 A。

表 2 不同数据分布类型检验方法的适用情形

检测方法	样本数量	适用情形
图形法	/	适用于对是否服从正态分布进行定性判断
偏度峰度检验法	/	适用于有特定信息表明真实分布与正态分布的可能差别的情形
W 检验	$n \leq 50$	适用于没有特定信息表明真实分布与正态分布的可能差别的情形， n 较大时可综合图形法进行正态分布判定 ^a
K-S 检验	$n > 50$	适用于没有特定信息表明真实分布与正态分布的可能差别、且无法使用其他方法进行检验的情形， n 较大时可综合图形法进行正态分布判定 ^a
卡方检验	$n > 100$	适用于分组数据

^a表示样本量很大的情况下，即使检验结果 $P < 0.05$ ，数据来自的总体也可能是服从正态分布的，需结合直方图、P-P、Q-Q 的图示法进行综合判断。

6.8.2 异常值判别与处理

如数据存在明显异常或数据的变异系数 C_v 偏大（砂质土壤为 0.5，较细土壤为 0.75），则应启动异常值判别。

判别样本异常值的常用方法包括 Nair 检验法、格拉布斯（Grubbs）检验法、狄克逊（Dixon）检验法、偏度峰度检验法、箱线图法。不同方法的适用情形见表 3。

排除可疑的异常值后，进行正态分布检验。排除异常值后如果服从正态分布且已知总体标准差，采

用 Nair 检验法进行异常值判别；如果总体标准差未知且限定异常值不超过 1 个，根据样本数量，从 Grubbs、Dixon 检验法中选择适用的方法进行异常值判别；如果总体标准差未知且限定异常值个数超过 1 个，选择偏度峰度检验法或 Dixon 检验法进行异常值判别。如果不服从正态分布，根据样本数量，选择 Walsh 检验法或箱线图法进行异常值判别。经异常值判别，如果属于异常值，且满足剔除条件，则进行剔除，如果不满足剔除条件，则予以保留。

Nair、Grubbs、Dixon、偏度峰度检验法的使用参照 GB/T 4883，Walsh 检验法的使用参照附录 B，箱线图法的使用参照 HJ 1185。

表 3 不同异常值判别方法的适用情形

检测方法	样本数量	是否需要满足正态分布	是否可以检测多个异常值
Nair 检验法	$n \leq 100$	是	是
Grubbs 检验法	/	是	否
Dixon 检验法	$n \leq 100$	是	是
偏度峰度检验法	/	是	是
Walsh 检验法	$n > 60$	否	是
箱线图法	$n \geq 6$	否	是

对于所判别的异常值，按照以下方式进行处理：

a) 检查原始记录，如果是样品采集、保存流转、分析检测、数据输入等过程的过失或错误等导致的异常数据，应予以更正或剔除；

b) 查阅区域土壤环境背景值相关数据，初步判断异常值是否在背景范围内；如明显高于背景值，进一步参照 GB/T 39792.1 中的同源性分析，根据目标元素或化合物含量特征、组合特征、同位素特征等与污染源的相似程度，判断是否来源于污染；或参照 HJ 1185，采用富集系数等方法判断异常原因。若异常值来源于污染，则剔除；若来源于高背景，应予以保留；

c) 若判别出的异常值不止一个，按异常值数字从大到小的顺序逐个判断，逐个处理。

记录被更正或剔除的异常值及其理由。处理后样本数量不能满足统计要求的，应补充样本数据。

6.8.3 基线确定

未进行异常值剔除的，按照剔除前的正态分布检验结果选择数据统计分析方法；进行了异常值剔除的，按照剔除后的正态分布检验结果选择数据统计分析方法。

分别统计每个层位土壤的基线水平。

对于服从正态分布的数据，当污染或破坏导致评价指标升高时，采用历史或调查数据的 90% 参考值上限（算术平均数 + 1.65 标准差）作为基线；当污染或破坏导致评价指标降低时，采用历史或调查数据的 90% 参考值下限（算术平均数 - 1.65 标准差）作为基线。对于不服从正态分布的数据，当污染或破坏导致评价指标升高时，采用历史或调查数据的第 90 百分位数作为基线；当污染或破坏导致评价指标降低时，采用历史或调查数据的第 10 百分位数作为基线。

6.9 补充调查

当存在以下两种情况时，应开展补充调查：

a) 基于前期调查获取的有效数据，计算实际变异系数 C_V 和相对偏差 m ，若 C_V 大于计算基础点位数量时的设定值且 m 不可接受（一般可接受范围为 20%~30%）时，需重新计算基础点位数量，若计算出的基础点位数量大于实际点位数量，应开展补充调查；

b) 若 $C_V > 1$ ，表明存在较大空间变异性，应开展补充调查，补充点位数量 ≥ 5 个。

7 基于其他方法确定土壤特征污染物生态环境基线

7.1 参考相关环境质量标准

当调查指标为土壤特征污染物，且历史数据或对照区调查不可行时，可根据评估区土地利用方式，参考适用的土壤环境质量国家标准（如 GB 15618、GB 36600 等）、当地地方标准确定基线，主要有以下情形：

a) 评估区位于工业园区或其它工业聚集区，周边存在一种或多种可能影响土壤特征污染物含量的因素；

b) 评估区存在复杂的土地利用历史，且前期土地使用者造成的土壤污染情况不明；

c) 其它无法选择合适对照区的情形。

当缺乏适用的标准时，可参考国外政府部门或国际组织发布的相关标准。当同时存在基于保护人体健康、保护地下水和保护生物受体等不同类型的质量标准时，应从严选择。

7.2 开展专项研究

当无法获取历史数据和对照区数据，且无可用的土壤环境质量标准时，参照 GB/T 39792.1，开展专项研究，以确定土壤生态环境基线，具体包括：

a) 对于土壤环境质量标准中不涵盖的指标，可以基于健康风险评估或生态风险评估反推可接受风险水平下的污染物浓度值，作为基线水平；

b) 对于土壤环境质量标准中不涵盖但地下水质量标准中涵盖的指标，且土壤中污染物可能迁移到地下水，导致地下水中污染物超过相关标准，则基于包气带和含水层污染物的迁移转化模型推算保护地下水的土壤污染物浓度限值，作为基线水平；

c) 其它用于确定基线水平的专项研究。

附 录 A
(资料性附录)
正态分布检验方法

A.1 卡方检验

H0: 样本的总体分布服从某特定分布

H1: 样本的分布不服从某特定分布

计算样本的均值和标准差作为总体样本均值 (μ) 和标准差 (σ) 的估计;

根据实际情况将样本数据划分为 k 个区间, 每个区间的期望频数相等, 为 $1/k$;

查标准正态分布表, 得到标准正态分布下划分每个区间的临界值 (z), 再计算每个临界值对应的区间界限 (x), 计算方法见公式 (A.1):

$$x = \mu + z \times \sigma \quad (\text{A.1})$$

式中: x ——每个临界值对应的区间界限;

μ ——总体样本均值;

z ——标准正态分布下划分每个区间的临界值;

σ ——总体样本标准差。

根据计算的区间界限 (x), 将样本数据的每一个数据分配到对应的区间, 然后统计实际频数 f_i 。最后由实际频数 f_i 与期望频数 F_i 的差值计算卡方统计量 x^2 , 计算方法见公式 (A.2):

$$x^2 = \sum_{i=1}^n \left[\frac{(f_i - F_i)^2}{F_i} \right] \quad (\text{A.2})$$

式中: x^2 ——卡方统计量;

f_i ——实际频数;

F_i ——期望频数。

按照公式 (A.3) 计算自由度 ν :

$$\nu = k - r - 1 \quad (\text{A.3})$$

式中: ν ——自由度;

k ——区间数;

r ——估计的分布参数数量。

根据 α (通常取 0.05) 和自由度, 查卡方分布表, 得到卡方值。比较卡方统计量 x^2 和卡方值, $x^2 <$ 卡方值, 接受 H0, 数据服从正态分布, 否则拒绝 H0。

A.2 K-S 检验

K-S 检验是比较一个频率分布与理论分布或者两个观测值分布的检验方法。该方法以样本数据的累计频数分布与特定的理论分布比较 (比如正态分布), 如果两者差距小, 则推论样本分布取自某特定分布。

假设检验问题:

H0: 样本的总体分布服从正态分布

H1: 样本的分布不服从正态分布

$F_n(x)$: 样本的累计分布函数

$F_0(x)$: 理论分布函数

D : $F_0(x)$ 与 $F_n(x)$ 差值的绝对值最大值

根据 n 和 α (通常取 0.05), 查表得到 $D(n, \alpha)$, n 为样本数量;

将 D 与 $D(n, \alpha)$ 相比较, 如果 $D < D(n, \alpha)$, 则接受 H_0 , 数据服从正态分布, 否则拒绝 H_0 。

附 录 B
(资料性附录)
Walsh 检验法

假设 X_1, X_2, \dots, X_n 表示从小到大排序的数据, 采用 Walsh 检验法进行异常值判别的步骤如下:

第一步: 确定被怀疑的异常值数量 $r, r \geq 1$;

第二步: 计算 c, k, b, a , 计算方法分别见公式 (B.1)、公式 (B.2)、公式 (B.3)、公式 (B.4):

$$c = [\sqrt{2n}] \quad (\text{B.1})$$

式中: $[\]$ ——将值向上取整数 (如 3.24 表示为 4);

n ——样本数量。

$$k = r + c \quad (\text{B.2})$$

式中: r ——被怀疑的异常值数量;

$$b^2 = 1/\alpha \quad (\text{B.3})$$

式中: α ——显著性水平;

$$a = \frac{1 + b\sqrt{(c-b^2)/(c-1)}}{c-b^2-1} \quad (\text{B.4})$$

如果 $60 < n \leq 220$, α 取 0.10; 如果 $n > 220$, α 取 0.05;

第三步: 如果 $X_{r-(1+a)X_{r+1}+aX_k} < 0$, r 对应的最小值是异常值;

第四步: 如果 $X_{n+1-r-(1+a)X_{n-r}+aX_{n+1-k}} > 0$, r 对应的最大值是异常值。