



文献 CSTR:
32001.14.11-6035.csd.2024.0193.zh



文献 DOI:
10.11922/11-6035.csd.2024.0193.zh
数据 DOI:
10.57760/sciencedb.17175

文献分类: 地球科学

收稿日期: 2024-11-20
录用日期: 2025-06-24
发表日期: 2025-09-15

基于第二次全国土壤普查的云南土种志数据集

高美荣^{1,2*}, 朱波^{1,2}, 周明华^{1,2}, 王小国^{1,2}

1. 中国科学院、水利部成都山地灾害与环境研究所, 成都 610299
2. 四川盐亭农田生态系统国家野外科学观测研究站, 成都 610299

摘要: 本数据集是对1994年出版的《云南土种志》进行梳理、归纳加工而成, 该专著系统阐述了云南284个土种的分类、分布、主要理化性状、典型剖面特征、生产性能以及利用改良和培肥措施等。基于土壤科学信息化的表达特征, 考虑包含性、专业性和易检索等功能以及未来便于开发线上应用服务, 对专著中的文字信息和数据进行了结构化的采集和提取, 集成以土种ID和土种名称为关键索引字段的8个主体数据表, 分别为土壤分类系统、土种基本信息、土种典型剖面立地调查信息、土种典型剖面发生层调查信息及理化性状、土种剖面统计理化性状。因在第二次全国土壤普查中机械组成存在两种制式的换算, 为使理化性状表不显得冗余并尽量保持数据的完整性, 补充提取了典型剖面发生层土壤机械组成(卡庆斯基制)和剖面统计机械组成(国际制和卡庆斯基制)数据形成2个表, 也整理了土壤发生层、质地符号及名称2个附表方便读者使用数据。数据提取过程中, 采用智能识别技术并复核其准确性和完整性, 同时依据法定计量单位规范和最新行政区划对数据单位和地(州)、县、市名等进行了修订。本数据集不仅可服务于云南省第三次全国土壤普查, 还可用于区域土地质量评估、生态环境恢复和土壤科学研究等, 为合理利用土壤资源提供极具意义的科学支撑。

关键词: 云南; 土种; 第二次全国; 土壤普查

数据库(集)基本信息简介

数据库(集)名称	基于第二次全国土壤普查的云南土种志数据集
数据作者	高美荣, 朱波, 周明华, 王小国
数据通信作者	高美荣 (gmr@imde.ac.cn)
数据时间范围	1994 年以前
地理区域	云南省
数据量	1.30 MB
数据格式	*.xls
数据服务系统网址	https://doi.org/10.57760/sciencedb.17175
基金项目	国家重点研发计划(2023YFD1901202); 中国生态系统研究网络(CERN)-盐亭紫色土农业生态系统定位研究站运行项目; 国家科技基础条件平台-四川盐亭农田生态系统国家野外科学观测研究站运行服务项目
数据库(集)组成	数据集为一个 Excel 文件, 主要包含 8 个主体数据表和 2 个附表:

* 论文通信作者
高美荣: gmr@imde.ac.cn

数据库（集）组成

（1）土壤分类系统，列出云南土种所属的土纲、亚纲、土类、亚类和土属，以及除土属和土种以外的国标（2009）对应的代码；（2）土种基本信息，包含土种的地理分布与面积、土壤母质、主要性状和生产性能综述等信息；（3）土种典型剖面立地条件调查信息，包含有土种典型剖面采集地的地理位置、地形地貌、植被和气候以及剖面构型和母质类型等立地条件信息；（4）土种典型剖面发生层信息，主要是典型剖面各发生层的野外调查信息；（5）土种典型剖面理化性状；（6）土种剖面统计理化性状；（7）2个补充表是典型剖面各层次的机械组成（卡庆斯基制）数据和剖面机械组成（国际制和卡庆斯基制）数据；（8）2个附表分别是土壤发生层和质地的符号及名称。

引言

第二次全国土壤普查（简称“二普”）系统查明了我国土壤资源的类型、数量、分布、基本性状及生产力特征等，构建了我国土壤分类系统体系，量化评估了中低产田的空间格局与障碍因子，为农业综合开发、耕地整治、科学施肥等国家战略制订提供了关键数据支撑^[1]。20世纪90年代形成的系列普查成果虽已出版，但随着时间的推移，以纸质介质保存的原始数据面临载体老化、信息丢失等风险，在数字技术深度渗透的现代研究范式下，其数据获取效率、共享应用价值及长期保存能力均受到严峻挑战。历经40余年的耕作制度演变与农业现代化进程，我国土地利用模式发生结构性变革，土壤资源特征亦可能存在显著变化。在此背景下，国务院于2022年启动第三次全国土壤普查（简称“三普”），旨在通过四年时间“三普”的实施，全面查清农用地土壤质量家底，实现对土壤资源数量、质量和健康变化状况的全面了解和实时把控，总结土壤资源管理和利用的成功经验，打造全国“数字土壤”一张图的土壤数据库，提出面向绿色发展的国家层面土壤质量和健康问题的系统解决方案^[2]。

因此，面向区域“三普”的土壤基层分类及名称校准、土壤质量动态评估、土壤资源利用提升和学科研究的需求，第二次全国土壤普查数据的信息数字化和结构化重建，实现多源数据资源融合分析，可应用于云南省土壤“三普”工作的历史数据对比与质量演变分析，赋能区域土地质量评估、生态修复规划等实践；也可为土壤发生演化和资源可持续利用等科学研究提供科学数据支持，对保障国家粮食安全，促进社会全面协调可持续发展具有重要意义。

土种具有一定的空间分布和明显的区域性、生产性特点，也是地区性土壤基层分类的基本单元，在土壤科学研究和合理利用土壤资源等方面均具有重要的实践和指导意义^[3]。《云南土种志》是由云南省土壤普查办公室编著，云南科技出版社1994年出版的云南第二次全国土壤普查成果专著之一。该书系统阐述了云南284个土种的分类、分布、主要理化性状、典型剖面特征、生产性能以及利用改良和培肥措施等^[4]。基于土壤信息学的表达范式与服务导向，构建了多层次结构化的数据集，以土种ID与名称为核心索引，集成土壤分类系统、土种基本信息、典型剖面立地调查信息、土种典型剖面发生层调查信息及理化性状、土种剖面统计理化性状等8个主体数据表。其中，针对历史数据中土壤机械组成存在的双制式（卡庆斯基制与国际制）并存现象，为使理化性状表不显得冗余并尽量保持数据的完整性，补充提取了典型剖面发生层土壤机械组成（卡庆斯基制）和剖面统计机械组成（国际制和卡庆斯基制）数据形成2个表。另外也整理了土壤发生层和质地符号及名称附表，辅以对数据的理解和使用。

1 数据采集和处理方法

1.1 数据采集方法

云南土种志数据采集方法为：（1）获取正式出版的《云南土种志》纸质文献；（2）通过专业扫描设备完成全书数字化处理，采用图像增强技术提升扫描文档质量，运用 OCR 智能识别提取图文数据；（3）基于土壤学科知识信息特征及逻辑关系，深度解析专著内容，设计可包含且能完整表达土种属性和信息等核心内容的字段名和数据表结构，并确定数据表间关联主索引字段为土种 ID 和土种名称；（4）将识别并校对好的信息录入相应表中，形成有主索引字段的数据表集合；（5）对构建的数据集进行人工核验和专家审核，同时编制标准化元数据文档。

1.2 数据提取与集成

土种描述内容主要包括土种分类与归属、分布、主要性状、生产性能、典型剖面发生层调查信息及理化性状、土种剖面统计理化性状等内容。基于书中内容，遵循土壤科学数据元数据标准（GB/T 32739-2016）要求^[5]，在设计表字段和表结构时，需要充分考虑包含性、专业性和易检索等功能。部分描述性文字也是土种数据的重要表达，设置相应字段以文本格式一并集成到表中。该数据集共整理集成了 8 个主体数据表，也整理了剖面发生层和土壤质地符号及名称附表。以下为各数据表的提取指标或信息的简述。

1.2.1 云南-土壤分类系统

列出了第二次全国土壤普查中云南的土壤分类系统以及各土种的划分及归属，并参照中国土壤分类与代码（GB 17296-2009）^[6]对土类、亚类和土属名称及代码进行了规范化处理。其中亚纲、土类和亚类分别列出了在标准（GB 17296-2009）和“二普”中的名称作为对照，两者有不同的用红色字体给以标识。云南土壤分类原则、依据和级别等可参考《云南土壤》中第四章的详细阐述^[7]。

1.2.2 云南-土壤分类系统

该表主要提取土种的分布区域及面积、分布地形地貌、分布总面积、土壤母质等基础地理信息，还描述了该土种的剖面构型、主要性状和生产性能。

1.2.3 云南-土种典型剖面立地调查信息

基于各地、市（州）《土种志》开展云南省土壤普查资料汇总，选其符合土种划分概念、个体单元清晰、资料信息比较齐全者的剖面作为某一土种的典型剖面，该表提取并整合了典型剖面采集地的立地条件野外调查信息，包括采集地点的各级行政区域名称、小地形地貌、高程、土壤母质、自然植被、土地利用和气候条件等信息。其中，根据最新行政区划经查询和比对，在可能已经不存在了的原地名后标注了新地名。

1.2.4 云南-土种典型剖面立地调查信息

该表提取了典型剖面发生层各类野外调查基本信息，包括剖面发生层次的厚度、颜色、质地、结构、紧实度和根系多少等。其中所列土壤发生层名称是依据土壤发生层代码名称附表整理的，仅供参考。

1.2.5 云南-土种典型剖面理化性状

该表集成的是土种典型剖面发生层土壤理化性状的实验室分析数据，主要理化性质指标有土壤机械组成、质地、有机质含量、养分全量、C/N、速效养分、pH 值、有效微量元素、阳离子交换量和少量水解或代换酸等。其中，（1）仅剖面表层有有效微量元素数据；（2）土壤机械组成数据为国际制分级数据，若原著中还给出卡庆斯基制数据的，单独收录于云南-土种典型剖面机械组成（卡庆斯基制）表中；（3）因部分典型剖面无国际制土壤机械组成数据的或仅有卡庆斯基制数据时，原著中土壤质地只有描述性文字，作者依据原著中给出质地名称与符号的对应关系，匹配出部分可以用符号表达的列入表中，一些对应不上的仍旧用原著的文字描述；（4）另矿物组成、盐基含量等少量数据放在相对应的备注栏。

1.2.6 云南-土种剖面统计理化性状

该表整合的是同一土种的数个剖面理化性状的统计数据，提取的指标有剖面层次厚度、有机质、养分全量、有效养分、pH 值、阳离子交换量、盐基饱和度以及有效微量元素的平均值和统计样本数。

1.2.7 云南-土种典型剖面机械组成（卡庆斯基制）和云南-土种剖面统计机械组成

在土种典型剖面理化性状表中提取的土壤机械组成数据是国际制的，有极少部分的机械组成只有卡庆斯基制的数据，为了使该数据完整形成单独一张表；在土种剖面统计理化性状表中未设计土壤机械组成字段，补充形成土种剖面统计机械组成表，含卡庆斯基制和国际制数据。

1.2.8 附表 云南土壤发生层、质地符号及名称

在《云南土种志》专著中，典型剖面 and 统计剖面发生层均只给出代码，没有明确给出发生层名称，为了读者更好地理解和使用数据，提取了《云南土种志》编写说明中的剖面发生土层和土壤质地符号及其对应的名称，生成了 2 个附表。

1.3 土壤分类及命名的几点说明^[7]

1.3.1 云南土壤分类的原则和依据

（1）主要依据土壤发生学理论，将自然土壤与耕作土壤纳入一个系统中。

（2）土壤分类把成土条件、成土过程和土壤属性结合起来，有利于如实反映土壤类别之间的联系和区别。另土体构型是分类的重要指标。

（3）把科学性、生产性、群众性结合起来，仅仅围绕土壤普查的目的，在不违背分类原则的前提下，要考虑便于农、林、牧各业对土壤资源的应用。

（4）根据土壤类型在省内所占的面积比重和在利用上的重要程度，在分类层次上掌握宜粗则粗、宜细则细的原则。

1.3.2 分类级别

根据上述原则和依据，云南土壤分类分为土纲、亚纲、土类、亚类、土属、土种、变种 7 个级别，土类是基本分类单元，土种是主要基层分类单元，能更确切地反映地区特点。在本数据集中没

有出现变种级。

土种是主要基层分类单元，也是农民划分土壤的主要单元。同一土属中，由于土壤发育程度和土体构型间有量的差异加以划分。耕作土壤要根据熟化程度、肥力水平、质地结构反映在耕性上的差异，剖面性态、土体构型和生产性能以及主要障碍因素的情况加以划分。

1.3.3 群众命名土种的经验

总结群众对耕作土壤的分类命名经验是：（1）一般分三级，第一级反映土壤的耕性肥力和特殊问题；第二级以颜色分，反映肥瘦、水热情况以及其他条件对土壤属性在数量上的差异；第三级以夹杂物和底土情况加以区分。具有一定的科学性、系统性和生产性，便于应用。（2）土壤命名简短生动，突出反映土性的特点和问题。如油沙土，反映熟化度较高，耕性较好的轻质土壤；寡红土主要在寡字和母土是红壤；冷浸田既浸又冷；紧口砂既沙又板等。群众命名最重视耕性好坏，其次是养分含量、颜色、水热条件和特殊问题，综合起来，即重视熟化程度、生产性能和肥力水平。命名来源于土性，但是，由于受科学文化水平和地域范围等条件的限制，不可避免地会出现大量的同土异名或同名的情况。

2 数据样本描述

以下分别对该数据集的 8 个主体数据表和 2 个附表的结构和样本进行说明（表 1-10），以帮助数据使用者对数据的各字段名内涵有更好的理解。

2.1 数据表样本描述

8 个主体数据表其内容及结构，分别见表 1-8。

表 1 云南土壤分类系统表字段含义

Table 1 Field descriptions of the soil classification system in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN160
土种名称	文本型		土种名称	麻砂泥土
土纲代码-2009	文本型		土种所属国标土纲代码	A
土纲名称-2009	文本型		土种所属国标土纲名称	铁铝土
亚纲代码-2009	文本型		土种所属国标亚纲代码	A2
亚纲名称-2009	文本型		土种所属国标亚纲名称	温暖铁铝土
亚纲名称（二普）	文本型		土种所属亚纲名称	湿温铁铝土
土类代码-2009	文本型		土种所属国标土类代码	A21
土类名称-2009	文本型		土种所属国标土类名称	黄壤
土类名称（二普）	文本型		土种所属土类名称	黄壤
亚类名称-2009	文本型		土种所属国标亚类名称	典型黄壤
亚类代码-2009	文本型		土种所属国标亚类代码	A211

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
亚类名称（二普）	文本型		土种所属亚类名称	黄壤
土属名称（二普）	文本型		土种所属土属名称	麻黄土

表 2 云南土种基本信息表字段含义

Table 2 Field descriptions of basic information of soil species in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN001
土种名称	文本型		土种名称	红鸡粪土田
地理分布及地 形部位	文本型		土种的地理分布以及分 布的地形部位	大理（3.1 万亩）、红河（1.7 万亩）2 州地势 平缓，排灌方便的村镇附近，东川、楚雄、玉 溪、西双版纳等地州亦有零星分布。面积 6.3 万亩，占全省水稻土面积的 0.31%。
土种分布面积	浮点型,2	万亩	土种的分布面积	6.30
土壤母质	文本型		土种形成的土壤母质	红壤性母土水稻耕熟化发育而来
剖面构型	文本型		土种的剖面分层结构	A-P-C 或 A-P-B-C
主要性状	文本型		土种主要性状的综合描 述	耕层厚 16-29cm，有 10-17cm 的犁底层，质地 多为壤质粘土，通体呈微酸性反应，是本亚类 中熟化度较高的土种。据 6 个剖面分析统计， 表土有机质平均 3.14%，全氮 0.200%，全磷 0.120%，……
生产性能综述	文本型		土种的生产性能综合描 述	红鸡粪土田土体深厚，熟化度较高，宜耕性 好，宜种性较广，供肥性能平稳，……，发棵 好，属中上等肥力稻田。多为一年两熟，宜种 水稻、烤烟、玉米、小麦等粮经作物。水稻亩 产 350-500 公斤，小麦亩产 150-200 公斤。种 烤烟可获优质高产，……。应注意增磷钾肥， 推广配方施肥，……。

表 3 云南土种典型剖面立地调查信息表字段含义

Table 3 Field descriptions of site survey information of typical soil profiles in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN065
土种名称	文本型		土种名称	厚湿红土
典型剖面	文本型		典型剖面采集地点位置的综合描述	曲靖马龙区旧县街道

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
采集地点（新）				团结村旱耕地
市（州）	文本型		典型剖面采集地市（区）名称	曲靖
县（市、区）	文本型		典型剖面采集地县名称	马龙县（马龙区）
乡（镇）	文本型		典型剖面采集地乡镇名称	旧县乡（旧县街道）
村	文本型		典型剖面采集地村组及其他地理信息描述	旧县社区
组或其他	文本型			团结村
典型剖面点地形地貌和部位	文本型		典型剖面所在点的地形地貌和部位	旱耕地
典型剖面点高程	整数型	m	典型剖面所在的海拔高程	1980
典型剖面土壤母质	文本型		典型剖面土壤母质	残存较薄的古红土
典型剖面点年均温	浮点型,1	℃	典型剖面采样地所在区域的年均温	13.6
典型剖面点年降水	浮点型,1	mm	典型剖面采样地所在区域的年降水量	1024.0
典型剖面点≥10℃积温	浮点型,1	℃	典型剖面采样地所在区域的≥10℃积温	3900.0
典型剖面点无霜期	整数型	天	典型剖面采样地所在区域的无霜期	246
典型剖面点自然植被	文本型		典型剖面采样地的植被覆盖情况	
典型剖面点土地利用	文本型		典型剖面采样地的土地利用情况	农地
备注	文本型		典型剖面采样地所在区域的其他信息	采样时间 1983 年 3 月

表 4 云南土种典型剖面发生层信息表字段含义

Table 4 Field descriptions of horizon information of typical soil profiles in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN008
土种名称	文本型		土种名称	紫泥田
发生层代码	文本型		典型剖面发生层代码	A
发生层序号	整数型		典型剖面发生层序号	1
发生层厚度	整数型	cm	相应发生层厚度	16
发生层最上深度	整数型	cm	相应发生层最上深度	0
发生层最下深度	整数型	cm	相应发生层最下深度	16
发生层颜色	文本型		相应发生层土壤颜色	紫色（湿，5YR 6/3）
发生层质地	文本型		相应发生层土壤质地	粉砂质粘壤土
发生层结构	文本型		相应发生层土壤结构	块状
发生层松紧度	文本型		相应发生层土壤松紧度	稍紧实
发生层根系和其他	文本型		相应发生层土壤中根系多少等其他信息	有棕黄色锈纹，稻根多

表 5 云南土种典型剖面理化性状表字段含义

Table 5 Field definitions of physical and chemical properties of typical soil profiles in Yunnan

字段名称	数据类型及小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN212
土种名称	文本型		土种名称	厚灰汤泥
发生层代码	文本型		典型剖面发生层名称	A
发生层序号	整数型		典型剖面发生层序号	1
层次相对厚度	整数型	cm	相应发生层厚度	24
发生层最上深度	整数型	cm	相应发生层最上深度	0
发生层最下深度	整数型	cm	相应发生层最下深度	24
机械组成 2-0.02mm	浮点型,2	%	2-0.02mm (砂粒) 土壤颗粒含量占比	11.80
机械组成 2-0.2mm	浮点型,2	%	2-0.2mm 土壤颗粒含量占比	0.25
机械组成 0.2-0.02mm	浮点型,2	%	0.2-0.02mm 土壤颗粒含量占比	11.55
机械组成 0.02-0.002mm	浮点型,2	%	0.02-0.002mm (粉粒) 土壤颗粒含量占比	25.46
机械组成小于 0.002mm	浮点型,2	%	小于 0.002mm (粘粒) 土壤颗粒含量占比	62.74
质地	文本型		发生层土壤质地	C
有机质	浮点型,1	g/kg	发生层土壤中有机质含量	19.20
全氮	浮点型,2	g/kg	发生层土壤中全氮含量	1.04
全磷	浮点型,2	g/kg	发生层土壤中全磷含量	3.49
全钾	浮点型,1	g/kg	发生层土壤中全钾含量	8.60
C/N	浮点型,1		有机质中有机碳和有机氮的重量比	10.7
碱解氮	浮点型,1	mg/kg	用碱提取法所测得的土壤氮素	184.0
有效磷	整数型	mg/kg	土壤中能被植物利用的效磷含量	5.6
速效钾	整数型	mg/kg	土壤中能被植物利用的钾含量	65.0
水提 pH 值	浮点型,1		土壤中水提 pH 值	5.6
有效态锌	浮点型,2	mg/kg	易被植物利用的锌	0.60
水溶性硼	浮点型,2	mg/kg	被水溶解析出的硼	0.51
有效态钼	浮点型,2	mg/kg	易被植物利用的钼	0.11
有效态锰	浮点型,2	mg/kg	易被植物利用的锰	19.0
有效态铜	浮点型,2	mg/kg	易被植物利用的铜	2.4
盐基饱和度	浮点型,1	%	土壤中吸附的交换性盐基离子占交换性阳离子总量的百分数	22.9
水解酸	浮点型,1	cmol/kg(+)	在水解过程中产生的酸性物质	14.5

字段名称	数据类型及小 数位数	量纲	字段说明	数据样例
代换酸	浮点型,1	cmol/kg(+)	中性盐溶液中的阳离子所置换出的酸	5.9
CEC	浮点型,2	cmol/kg(+)	土壤阳离子交换量	18.8
备注	文本型		未在以上字段中体现的特征值	SiO ₂ 34.56%,.....

表 6 云南土种剖面统计理化性状表字段含义

Table 6 Field descriptions of statistical physical and chemical properties of soil profiles in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN003
土种名称	文本型		土种名称	红泥田
发生层代码	文本型		典型剖面发生层代码	A
发生层厚度统计样本数	整数型	个	层次相对厚度统计样品样本数	1
发生层厚度平均值	整数型	cm	相应发生层厚度平均值	13.00
有机质统计样本数	整数型	个	有机质含量统计样品样本数	8
有机质平均值	浮点型,1	g/kg	发生层土壤中有有机质平均值	27.1
全氮统计样本数	整数型	个	全氮统计样品样本数	9
全氮平均值	浮点型,2	g/kg	发生层土壤中全氮平均值	1.70
全磷统计样本数	整数型	个	全磷统计样品样本数	9
全磷平均值	浮点型,2	g/kg	发生层土壤中全磷平均值	1.20
全钾统计样本数	整数型	个	全钾统计样品样本数	9
全钾平均值	浮点型,1	g/kg	发生层土壤中全钾平均值	14.50
C/N	浮点型,2		有机质中有有机碳和有机氮的重量比	9.20
碱解氮统计样本数	整数型	个	碱解氮统计样品样本数	9
碱解氮平均值	浮点型,1	mg/kg	发生层土壤中碱解氮平均值	124.5
速效磷统计样本数	整数型	个	速效磷统计样品样本数	9
速效磷平均值	浮点型,1	mg/kg	发生层土壤中速效磷平均值	11.7
速效钾统计样本数	整数型	个	速效钾统计样品样本数	9
速效钾平均值	浮点型,1	mg/kg	发生层土壤中速效钾平均值	81.6
pH 值统计样本数	整数型	个	水提 pH 值统计样品样本数	9
pH 平均值或趋势值	浮点型,1		发生层土壤水提 pH 值或范围	6.5
阳离子交换量统计样本数	整数型	个	阳离子交换量统计样品样本数	5
阳离子交换量平均值	浮点型,1	cmol/kg(+)	发生层土壤阳离子交换量平均值	16.90
盐基饱和度统计样本数	整数型	个	发生层盐基饱和度统计样本数	3
盐基饱和度均值	浮点型,1	%	发生层盐基饱和度平均值	51.5

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
有效态锌统计样本数	整数型	个	发生层有效态锌统计样本数	37
有效态锌平均值	浮点型,2	mg/kg	发生层有效态锌平均值	2.68
水溶性硼统计样本数	整数型	个	发生层水溶性硼统计样本数	37
水溶性硼平均值	浮点型,2	mg/kg	发生层水溶性硼平均值	0.19
有效态钼统计样本数	整数型	个	发生层有效态钼统计样本数	37
有效态钼统计平均值	浮点型,2	mg/kg	发生层有效态钼统计平均值	0.27
有效态锰统计样本数	整数型	个	发生层有效态锰统计样本数	37
有效态锰平均值	浮点型,2	mg/kg	发生层有效态锰平均值	17.59
有效态铜统计样本数	整数型	个	发生层有效态铜统计样本数	37
有效态铜平均值	浮点型,2	mg/kg	发生层有效态铜平均值	7.58

表 7 云南土种典型剖面机械组成（卡庆斯基制）表字段含义

Table 7 Field descriptions of the mechanical composition (Kaqingsky system) of typical soil profiles in Yunnan

字段名称	数据类型及小 数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN057
土种名称	文本型		土种名称	棕油红土
发生层代码	文本型		典型剖面发生层代码	B
1-0.05mm	浮点型,2	%	1-0.05mm 土壤颗粒占比	7.98
0.05-0.01mm	浮点型,2	%	0.05-0.01mm 土壤颗粒占比	19.15
0.01-0.005mm	浮点型,2	%	0.01-0.005mm 土壤颗粒占比	12.57
0.005-0.001mm	浮点型,2	%	0.005-0.001mm 土壤颗粒占比	18.75
<0.001mm	浮点型,2	%	<0.001mm 土壤颗粒占比	33.19
土壤质地（卡庆斯基制）	文本型		卡庆斯基制土壤质地名称	轻粘土

表 8 云南土种剖面统计机械组成表字段含义

Table 8 Field descriptions of statistical mechanical composition of soil profiles in Yunnan

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
土种 ID	文本型		土种的代码	YN071
土种编码	文本型		土种代码	
土种名称	文本型		土种名称	紫泥田
发生层代码	文本型		剖面发生层代码	A
统计样本数	整数型	个	用于统计的样品个数	4
机械组成 2-0.2mm	浮点型,2	%	2-0.2mm 土壤颗粒含量占比	1.23
机械组成 0.2-0.02mm	浮点型,2	%	0.2-0.02mm 土壤颗粒含量占比	18.85

字段名称	数据类型及 小数位数	量纲	字段说明	数据样例
机械组成 0.02-0.002mm	浮点型,2	%	0.02-0.002mm (粉粒) 土壤颗粒含量占比	35.63
机械组成小于 0.002mm	浮点型,2	%	小于 0.002mm (粘粒) 土壤颗粒含量占比	44.29
土壤质地 (国际制)	文本型		国际制土壤质地名称或符号	LC
1-0.05mm	浮点型,2	%	1-0.05mm 土壤颗粒占比	31.06
0.05-0.01mm	浮点型,2	%	0.05-0.01mm 土壤颗粒占比	20.73
0.01-0.005mm	浮点型,2	%	0.01-0.005mm 土壤颗粒占比	8.48
0.005-0.001mm	浮点型,2	%	0.005-0.001mm 土壤颗粒占比	12.10
<0.001mm	浮点型,2	%	<0.001mm 土壤颗粒占比	27.78
土壤质地 (卡庆斯基制)	文本型		卡庆斯基制土壤质地名称	中壤土

2.2 云南土壤发生层、质地符号及名称附表

附表内容是其他数据集主表中土壤发生层符号及对应名称，土壤质地符号及对应名称。方便数据使用者查阅。

3 数据质量控制和评估

3.1 数据质量保证

对《云南土种志》的系统性研读基础上，对其内容和数据进行深度解析，依据土壤科学数据的信息化表达范式和逻辑，设计专业主体表、表结构以及字段名，采用智能识别技术对纸质版资料进行处理，在数据提取过程中，不断复核并完善各字段名称和数值的准确性和完整性；依据法定计量单位规范对数据单位进行了校正；根据最新行政区划的各级地名对原地（州）、县、市名等进行了修订；对纸质版中出现的个别错误信息在数据表中也进行了修正。集成专业数据集后，进行了三次仔细校对与核验，并开展专家抽查审核，同时编写标准的数据集元数据文档。

3.2 数据质量控制措施

3.2.1 法定计量单位转换

按照我国法定计量单位规定将专著中的一些指标的单位进行了修订：有机质、全氮、全磷和全钾原始数据的单位为%，均修订为 g/kg；有效态养分和有效微量元素的单位 ppm 修订为 mg/kg；阳离子交换量、水解酸和代换酸单位修订为 cmol/kg(+); 其他描述文本中单位未做修订，保持原有描述。为方便读者使用，以下列出之间的换算关系：1.0 %=10.0 g/kg，1.00 ppm=1.00 mg/kg，1.0 me/100g 土=1.0 cmol/kg(+)^[8-9]。

3.2.2 采样的深度表示

在专著的原始数据表中，发生层深度用相对厚度表示，使用者不易区分其绝对深度或深度范围。在数据表结构设计上，参照 ISO 28258^[10]，增添发生层最上深度和发生层最下深度两个字段，分别计算录入。

3.2.3 土壤分类系统的规范化

依照国家标准 GB/T 17296-2009^[6]，对《云南土种志》中的土壤分类系统进行了详细梳理，列出了各分类级别的名称及对应的代码和新旧名称的对照。

4 数据价值

《云南土种志》是云南第二次全国土壤普查资料汇总的一项重要成果，它是在各地州市编写的《土种志》基础上，选其符合土种划分概念、个体单元清晰、资料信息比较齐全者加以重新整理出 284 个土种。该成果实现了土种划分逐步向系统化、标准化、量化的方向发展，在当时填补了云南土壤分类系统的空白^[4]。

面向区域“三普”的土壤基层分类及名称校准、土壤质量动态评估、土壤资源利用提升和学科研究的需求，基于《云南土种志》重建的实现多源数据资源融合分析的该数据集，可应用于云南省土壤“三普”工作的历史数据对比与质量演变分析，赋能区域土地质量评估、生态修复规划等实践；也可为土壤发生演化和资源可持续利用等科学研究提供科学数据支持，对保障国家粮食安全，促进社会全面协调可持续发展具有重要意义。

数据作者分工职责

高美荣（1969—），女，陕西省佳县人，硕士，高级工程师，从事土壤学、生态学及其信息管理研究工作。主要承担数据采集与提取、数据表结构设计、数据质量控制和数据论文撰写。

朱 波（1966—），男，四川省仁寿县人，博士，研究员，从事水土界面碳氮迁移过程及氮循环研究。主要负责数据质量控制。

周明华（1983—），男，博士，四川省罗江县人，研究员，从事地球化学生物循环研究。主要负责数据应用与核验。

王小国（1977—），男，博士，湖北省红安县人，研究员，从事土壤碳循环研究。主要承担数据校对与核验。

参考文献

- [1] 全国土壤普查办公室. 中国土壤[M]. 北京: 中国农业出版社, 1998. [The Office National Soil Survey of China .The Soils of China[M]. Beijing: China Agriculture Press, 1998.
- [2] 国务院第三次全国土壤普查领导小组办公室.第三次全国土壤普查工作方案(农建发〔2022〕1号)[EB/OL]. (2022-02-24) [2024-11-13]. http://www.moa.gov.cn/govpublic/ntjsgls/202202/t20220224_6389491.htm. [The Office of the Leading Group for the Third National Soil Survey of the State Council.Work Plan for the Third National Soil Survey (Agricultural Construction Development Document [2022] No. 1)[EB/OL]. (2022-02-24) [2024-11-13]. http://www.moa.gov.cn/govpublic/ntjsgls/202202/t20220224_6389491.htm.]
- [3] 中国土壤学会. 中国土种、土属分类研究[M]. 南京:江苏科技出版社,1989. [Soil Science Society of China. A study on the classification of soil species and genera in China [M] Nanjing: Jiangsu Science

and Technology Press, 1989.]

[4] 云南省土壤普查办公室. 云南土种志[M]. 云南昆明: 云南科技出版社, 1994.[Yunnan Soil Survey Office, The Soil Species of Yunnan [M]. Kunming: Yunnan Science and Technology Press, 1994.]

[5] 国家质量监督检验检疫总局 中国国家标准化管理委员会. 土壤科学数据元数据: GB/T 32739—2016[S]. 北京: 中国标准出版社, 2017. [Standardization Administration of the People's Republic of China. Metadata for soil science data: GB/T 32739—2016[S]. Beijing: Standards Press of China, 2017.]

[6] 国家质量监督检验检疫总局 中国国家标准化管理委员会. 中国土壤分类与代码: GB/T 17296—2009[S]. 北京: 中国标准出版社, 2009. [Standardization Administration of the People's Republic of China. Classification and codes for Chinese soil: GB/T 17296—2009[S]. Beijing: Standards Press of China, 2009.]

[7] 云南省土壤肥料工作站, 云南省土壤普查办公室. 云南土壤[M]. 云南昆明: 云南科技出版社, 1996. [Yunnan Soil Fertilizer Workstation, Yunnan Soil Survey Office, Yunnan Soil.[M]. Kunming: Yunnan Science and Technology Press, 1996.]

[8] 施建平, 宋歌. 基于 Web 的中国土种数据库 [J]. 土壤, 2016, 48(6): 1246 - 1252. DOI:10.13758/j.cnki.tr.2016.06.026. [SHI J P, SONG G. Web based soil type database of China[J]. Soils, 2016, 48(6): 1246 - 1252. DOI:10.13758/j.cnki.tr.2016.06.026.]

[9] 高美荣, 张耀南, 康建芳. 基于全国第二次土壤普查的甘肃土种志数据集[J/OL]. 中国科学数据, 2022, 7(4): 172 - 182. DOI:10.11922/11-6035.ncdc.2021.0025.zh. [GAO M R, ZHANG Y N, KANG J F. A dataset of Gansu soil species based on the Second National Soil Survey[J/OL]. China Scientific Data, 2022, 7(4): 172 - 182. DOI:10.11922/11-6035.ncdc.2021.0025.zh.]

[10] Soil quality - Digital exchange of soil-related data: ISO 28258: 2013[S]. International Organization for Standardization, 2013.

论文引用格式

高美荣, 朱波, 周明华, 等. 基于第二次全国土壤普查的云南土种志数据集[J/OL]. 中国科学数据, 2025, 10(3). (2025-09-15). DOI:10.11922/11-6035.csd.2024.0193.zh.

数据引用格式

高美荣, 朱波, 周明华, 等. 基于第二次全国土壤普查的云南土种志数据集[DS/OL]. V3. Science Data Bank, 2025. (2025-06-24). DOI:10.57760/sciencedb.17175.

A dataset of Yunnan soil species based on the Second National Soil Survey

GAO Meirong^{1,2*}, ZHU Bo^{1,2}, ZHOU Minghua^{1,2}, WANG Xiaoguo^{1,2}

1. Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610299, P.R. China

2. Sichuan Yanting Agro-ecosystem Research Station in Chinese National Ecosystem Research Network, Chengdu 610299, P.R. China

Abstract: This dataset is compiled and processed from the *Soil Species of Yunnan* published in 1994, a monograph that systematically describes the classification, distribution, main physical and chemical properties, typical profile characteristics, production performance, and measures for improvement and fertilization of 284 soil species in Yunnan. Based on the expressive characteristics of soil science information and considering inclusiveness, professionalism, retrievability, and the potential for future online applications, we have structured and extracted the textual information and data from the monograph, and integrated eight core data tables with soil taxon ID and soil taxon name as the key index fields, namely Soil Classification System, Basic Information of Soil species, Typical Profiles Terrain Survey Information, Typical Profiles Horizon Survey Information and physicochemical properties, and statistical physicochemical properties of soil profiles. Because of the conversion of mechanical composition between two systems (the Kachinsky system and the International system) in the Second National Soil Survey, to avoid redundancy in the physicochemical property table while maintaining completeness of the data as much as possible, we additionally extracted data on the mechanical composition of horizons from typical profiles (Kaczynski system) and statistical profile mechanical composition (both International and Kaczynski systems), resulting in two supplementary tables. Further two appendices on soil horizons and texture symbols/names were organized to facilitate data use. In the process of data extraction, intelligent identification techniques were used, followed by verification of accuracy and completeness, while data units were standardized according to legal metrology specifications, and place names (prefectures, counties, and cities), were updated in line with the latest administrative divisions. This dataset can not only serve the Third National Soil Survey of Yunnan Province, but also be used in regional land quality assessment, ecological environment restoration and soil science research, providing meaningful scientific support for the rational use of soil resources.

Keywords: Yunnan; soil species; the second national; soil survey

Dataset profile

Title	A dataset of Yunnan soil species based on the Second National Soil Survey
Data authors	GAO Meirong, ZHU Bo, ZHOU Minghua, WANG Xiaoguo
Data corresponding author	Gao Meirong (gmr@imde.ac.cn)
Time range	Before 1994
Geographical scope	Yunnan Province
Data Volume	1.30 MB
Data formate	*.xls
Data service system	https://doi.org/10.57760/sciencedb.17175
Source of funding	National Key Research and Development Program (2023YFD1901202); Operation Project of Yanting Purple Soil Agro-ecological Experimental Station, Chinese Ecosystem Research Network; Operation Service Project of Sichuan Yanting Agro-

	ecological National Ecosystem Research Station, National Platform for Basic Conditions of Science and Technology.
Dataset composition	The dataset is provided as an Excel file and consists of 8 main data tables and 2 supplementary tables: (1) soil classification system, listing the soil orders, groups, subgroups and genus to which soil taxa Yunnan belong, along with the corresponding codes in the National Standard (2009), except for soil genera and soil taxa; (2) basic information on soil taxa, containing information on geographic distribution and area, parent material, main characteristics, and a summary of production performance; (3) Site survey information of typical soil profiles, recording site conditions of typical profile locations, including geographical coordinates, topography and geomorphology, vegetation, climate, profile configuration, and parent material type. (4) Occurrence horizon information of typical soil profiles, providing field survey data on each horizon of the typical profiles; (5) physicochemical properties of typical soil profiles; (6) Statistical physicochemical properties of soil profiles; (7) 2 supplementary tables– (a) mechanical composition data of horizons from typical profiles (Kaczynski system), and statistical profile mechanical composition data (International and Kaczynski systems); (8) 2 appendices, namely symbols and names of soil horizons and soil texture classes.