

1990–2023 年青藏高原“一江两河”地区 30 米分辨率 耕地空间分布数据产品集

陈慧玲^{1,2}, 何国金^{1,2*}, 彭雪丽^{1,2}, 王桂周¹, 尹然宇¹

1. 中国科学院空天信息创新研究院, 北京 100094

2. 中国科学院大学, 北京 100049

摘要: 本研究基于 Landsat 5 和 Landsat 8 的 Level 2、Collection 2、Tier 1 数据集提取了 1990 年至 2023 年间的青藏高原“一江两河”地区耕地空间分布数据产品集。借助国产高分卫星影像, 本研究进行 2023 年的全区域耕地样本标注以及 2015 和 2020 年的验证点集生成, 应用 ResUNet 模型进行 Landsat 影像的耕地信息提取, 同时在模型跨传感器应用过程中进行了模型针对性微调。最终得到 1990、1995、2000、2005、2010、2015、2020 和 2023 年“一江两河”地区八期耕地分布产品。利用目视解译得到的点样本数据对 2015 年和 2020 年耕地分布产品进行精度验证, 得到了比较好的结果, 其中精确率 (Precision) 分别为 97.72%、96.84%, 召回率 (Recall) 分别为 91.99%、92.52%, F1 分别达到了 0.9477、0.9463。本数据产品集可为青藏高原“一江两河”地区的农业可持续发展和土地管理提供科学依据。

关键词: 耕地分布; 一江两河; Landsat; 长时序

数据库 (集) 基本信息简介

数据库 (集) 名称	1990–2023 年青藏高原“一江两河”地区 30 米分辨率耕地空间分布数据产品集
数据通信作者	何国金 (hegj@aircas.ac.cn)
数据作者	陈慧玲、何国金、彭雪丽、王桂周、尹然宇
数据时间范围	1990–2023 年
地理区域	地理范围为 28°17'–30°28'N, 87°04'–92°37'E。地理区域范围涉及西藏自治区的拉萨市、山南地区和日喀则地区的 18 个县 (市、区)。
空间分辨率	30 m
数据量	1.15 GB
数据格式	*.tif (GeoTIFF, byte)
数据服务系统网址	https://doi.org/10.57760/sciencedb.j00001.01185
基金项目	第二次青藏高原综合科学考察研究 (2019QZKK0307)
数据库 (集) 组成	本数据产品集包括 1990–2023 年“一江两河”地区 30 米分辨率耕地空间分布产品, 其中 1990–2020 年每 5 年为一期, 并包含 2023 年的产品, 共 8 期数据, 包含 8 个 *.tif 文件。



文献 CSTR:

32001.14.11-6035.csd.2024.0088.zh



文献 DOI:

10.11922/11-6035.csd.2024.0088.zh

数据 DOI:

10.57760/sciencedb.j00001.01185

文献分类: 地球科学

收稿日期: 2024-04-26

开放同评: 2024-05-06

录用日期: 2024-08-16

发表日期: 2024-10-24

* 论文通信作者

何国金: hegj@aircas.ac.cn

引言

在全球人口持续增长与气候变化日益严峻的双重挑战下，耕地的有效保护与合理利用已然成为维护粮食安全和生态平衡的重中之重^[1]。耕地不仅承载着农业生产的基本使命，更是维护生物多样性、推动地方社区经济发展的关键要素^[2-3]。其面积与质量的变化，深刻反映着自然规律与人类活动的交织影响^[4]。

耕地监测作为土地覆盖与土地利用研究中的核心环节，其重要性不言而喻。尽管全球范围内已涌现出众多土地覆盖与土地利用的共享产品，如 30 米分辨率的 GlobeLand30^[5]、GLC_FCS30^[6]、CLCD^[7]等，以及 10 米分辨率的 AIEC^[8]、CRLC^[9]、ESA WorldCover^[10]、Esri_LandCover^[11]等，这些产品为全球至区域尺度的土地覆盖研究提供了丰富的数据支撑。然而，对于青藏高原地区，这些共享产品存在较大的不确定性。

一方面，这些产品大多聚焦于土地利用的总体情况，精细化的耕地提取并非其主要目标，这在一定程度上影响了耕地数据的准确性。另一方面，由于青藏高原特殊的自然地理环境和独特的生态条件，这些共享产品的耕地分布空间一致性难以达到理想水平^[12]。这不仅阻碍了对耕地变化趋势的精准把握，也限制了对农业生产和生态环境变化的深入认识。此外，现有的共享产品还未更新到 2023 年，难以满足相关应用的时效性需求。

“一江两河”地区是西藏自治区的重要粮食产区，素有“高原粮仓”之称。但该区域面临复杂地形和脆弱生态环境带来的挑战，如土地退化、水土流失及气候变化的影响^[13]。因此，本研究“一江两河”为研究对象，使用 Landsat 5 和 Landsat 8 卫星数据，获取 1990–2023 年该区域耕地空间分布数据产品集。本数据产品集记录了该区域近三十年的耕地空间分布及变化情况。通过深入分析这些数据产品，能够更加准确地把握“一江两河”地区耕地变化的趋势和特征，这对于深化耕地变化的理解、指导农业可持续发展以及维护生态平衡具有重大意义。

1 数据采集和处理方法

根据国土资源部组织修订的国家标准《土地利用现状分类》（GB/T 21010-2017），本研究中的耕地指种植农作物的土地，包括熟地，新开发、复垦、整理地，休闲地（含轮歇地、休耕地），以种植农作物（含蔬菜）为主，间有零星果树、桑树或其他树木的土地。需要特别指出的是，本研究的耕地定义排除了以下两种情况：（1）耕地面积小于 0.09 公顷（最小宽度为 30 米）的小地块；（2）设施农业，如温室大棚等，因为这些地物在遥感影像中的光谱信息与其他类型的耕地明显不同。

本研究基于 Landsat 5 和 Landsat 8 数据集，结合地形和时相的复杂性，采用 ResUNet 模型，确保从遥感数据中有效提取耕地信息。为确保分析的精确性，特别对数据进行了云量控制及多时相分析处理，增强了耕地信息提取的有效性。研究流程及关键技术步骤见图 1。

1.1 数据来源

本研究选用了 USGS 提供的 Landsat 5 和 Landsat 8 卫星的 Level 2、Collection 2、Tier 1 数据集作为数据源。Landsat 5 数据集涵盖 1990 年至 2010 年，而 Landsat 8 数据集则覆盖 2015 年至 2023 年。选定的数据轨道号包括 137/039-040、138/039-040、139/039-040、140/039-040。为了提高研究数据的质量并减少云层干扰，本研究筛选了云量小于 30% 的影像，并利用 Landsat Collection 2 所附

带的质量评估 QA 波段进行去云处理。

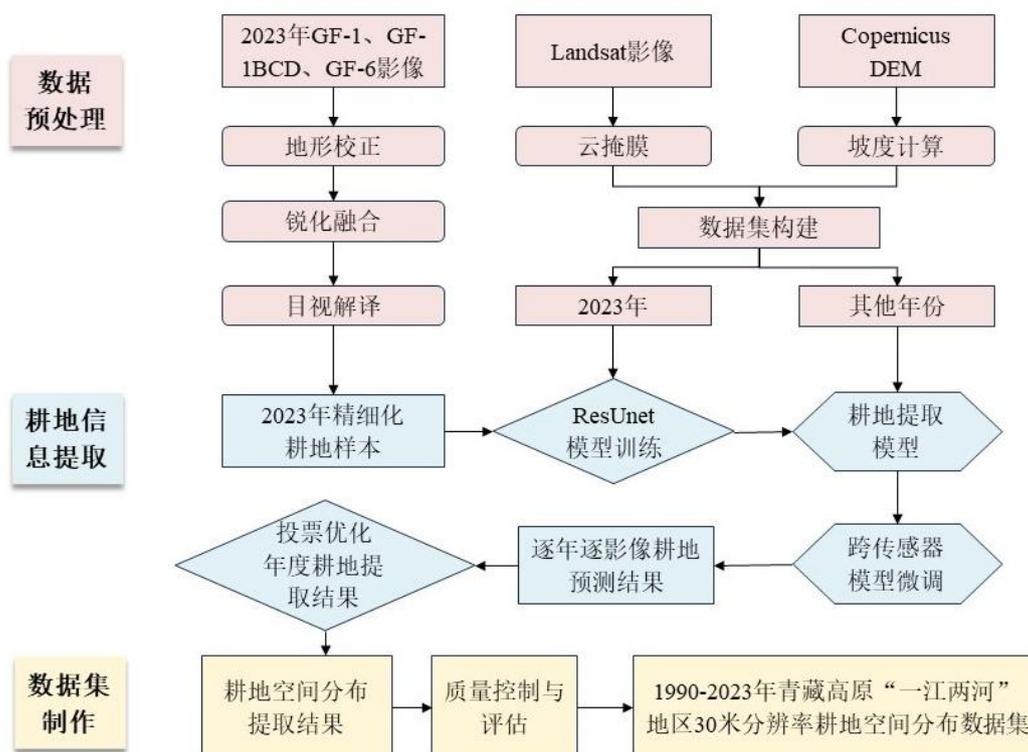


图 1 主要流程图

Figure 1 Main flow diagram

本研究从 Landsat 5 数据集中挑选了 6 个波段：蓝波段（SR_B1）、绿波段（SR_B2）、红波段（SR_B3）、近红外波段（SR_B4），以及 2 个短波红外波段（SR_B5 和 SR_B7），从 Landsat 8 数据集选择了对应的蓝波段（SR_B2）、绿波段（SR_B3）、红波段（SR_B4）、近红外波段（SR_B5）和两个短波红外波段（SR_B6 和 SR_B7）。此外，本研究还使用基于 Copernicus DEM^[14] 计算得到的坡度数据，作为额外的分类特征，并到影像数据中，构建数据集。耕地样本标注过程参考了 2023 年经过正射校正和锐化融合等预处理操作的国产高分卫星 2 米影像，标注了研究区内所有的耕地，并将这些样本转为 30 米分辨率的栅格影像，作为模型训练和预测样本数据集的基础。

1.2 耕地信息提取

本研究使用 2023 年的 Landsat 8 影像数据进行模型训练和验证，并采用 ResUNet^[15] 作为耕地信息提取模型。具体操作中，从 2023 年的影像中选取并处理了 6161 张切片，每张切片裁剪至 256 × 256 像素的尺寸。这些切片中，70%（4312 张）被随机选取作为训练集，而剩余 30%（1849 张）用作验证集，以测试模型的稳定性和泛化能力。关于训练参数，设置批量大小（batch size）为 16，总训练周期（epoch）为 100。初始学习率定为 0.0001，并采用 Adam 优化器来优化学习率。经过 100 个训练周期后，模型在训练集上 F1 得分达到了 0.8759，而在验证集上 F1 得分为 0.8636，显示出模型具有良好的性能和一致性。

由于 Landsat 5 和 Landsat 8 的波段设置存在些许差异，本研究对模型进行了针对性微调，以提升跨传感器的预测精度。具体来说，首先使用基于 Landsat 8 数据训练的模型对 2010 年的 Landsat 5

影像进行预测。通过目视比对预测结果与实际影像，选择较好部分的预测结果进行目视判别和校正，以生成 2010 年的真实耕地标签。基于这些校正后的影像，我们获得了 2394 张 256×256 像素的切片，其中 70%（1676 张）被用作训练集，而剩余 30%（718 张）用于验证集。在微调过程中，所有层的参数均保持可训练状态，批量大小仍为 16，训练周期设为 50。此次微调使模型在新的训练集和验证集上的 F1 得分分别达到 0.8682 和 0.8406，证明了模型在处理 Landsat 5 数据时的有效性。这样，我们得到了专门针对 Landsat 5 和 Landsat 8 数据的耕地信息提取模型。

为了减少云层干扰、提升预测效率及产品精度，在正式应用训练好的模型进行逐年逐影像的预测前，我们设计了一套影像筛选流程。具体筛选规则如下：首先，从每个月的影像中筛选出一张云量最少且低于 30% 的影像作为该月的代表影像。接着，在生长季（4 月至 10 月）选择两张云量最少的代表影像；非生长季再选择一张云量最少的代表影像。若生长季内合适的影像数量不足两张，从非生长季的影像中选取额外的影像以补充。采用这样的筛选策略是因为生长季的影像对于提高精度尤为关键，但生长季也常伴随较多雨水和云遮挡，因此额外选取一张非生长季的影像以补充信息，确保整体预测的准确性和稳定性。

最后，我们使用训练好的模型进行逐年、逐影像的预测。对于每一年的预测结果，我们将三幅筛选出的影像的预测结果进行集成，采用投票方式决定最终分类，即如果超过半数的影像预测为耕地，则该区域被判定为耕地。

2 数据样本描述

数据产品集采用 GeoTiff 格式存储，并设置为 Byte 数据类型。所有数据均以 Albers 投影格式保存，保证了空间分辨率统一为 30 米。整个数据产品集包含 8 期产品，分别对应 1990 年、1995 年、2000 年、2005 年、2010 年、2015 年、2020 年和 2023 年。每期产品对应一个 tif 文件。文件的命名遵循简洁的规则：“年份.tif”，方便用户快速识别和访问。每个产品文件包含的是二值图像，其中像素值为 1 代表耕地，0 则表示非耕地。图 2 展示了 2020 年“一江两河”地区耕地空间分布情况。图 3 为 1990 至 2023 年期间“一江两河”地区耕地空间分布的局部视图。

每期产品的耕地面积如图 4 所示。该地区耕地面积在不同年份略有波动，但总体保持相对稳定。耕地面积变化大致可分为两个阶段。第一阶段是 1990 年到 2010 年，这段时间内耕地面积相对稳定，1990 年的耕地面积为 138,519 公顷，到 2010 年略微增加到 140,723 公顷，增长约 1.59%，中间有所波动，但幅度不大。第二阶段是 2010 年到 2023 年，耕地面积呈小幅下降趋势，由 2010 年的 140,723 公顷下降到 2023 年的 136,449 公顷，下降约 3.03%。

3 数据质量控制和评估

为了评估本研究生成数据产品集的质量和可靠性，本研究构建了一个由 2508 个样本点组成的验证集。该验证点集基于本团队研制的 2015 年和 2020 年全国 2 m 分辨率镶嵌产品进行目视解译而成，其耕地与非耕地的点数量比例约为 1:1。通过使用此验证集，本研究计算了耕地类别的精确率（Precision）、召回率（Recall）、F1 得分（F1-Score）及耕地和非耕地的整体精度（Overall Accuracy）。此外还对 2020 年度本产品与其他几种共享产品（AIEC、ESA、Esiri 和 GLC-30）进行了比较，结果见表 1。

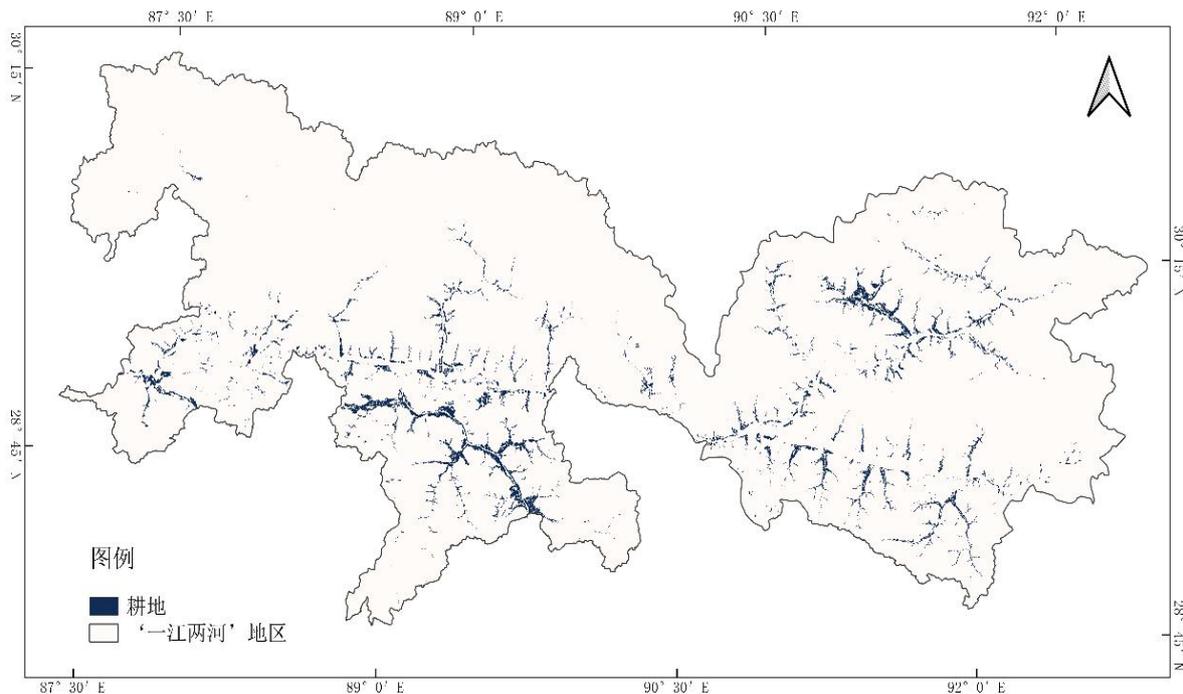
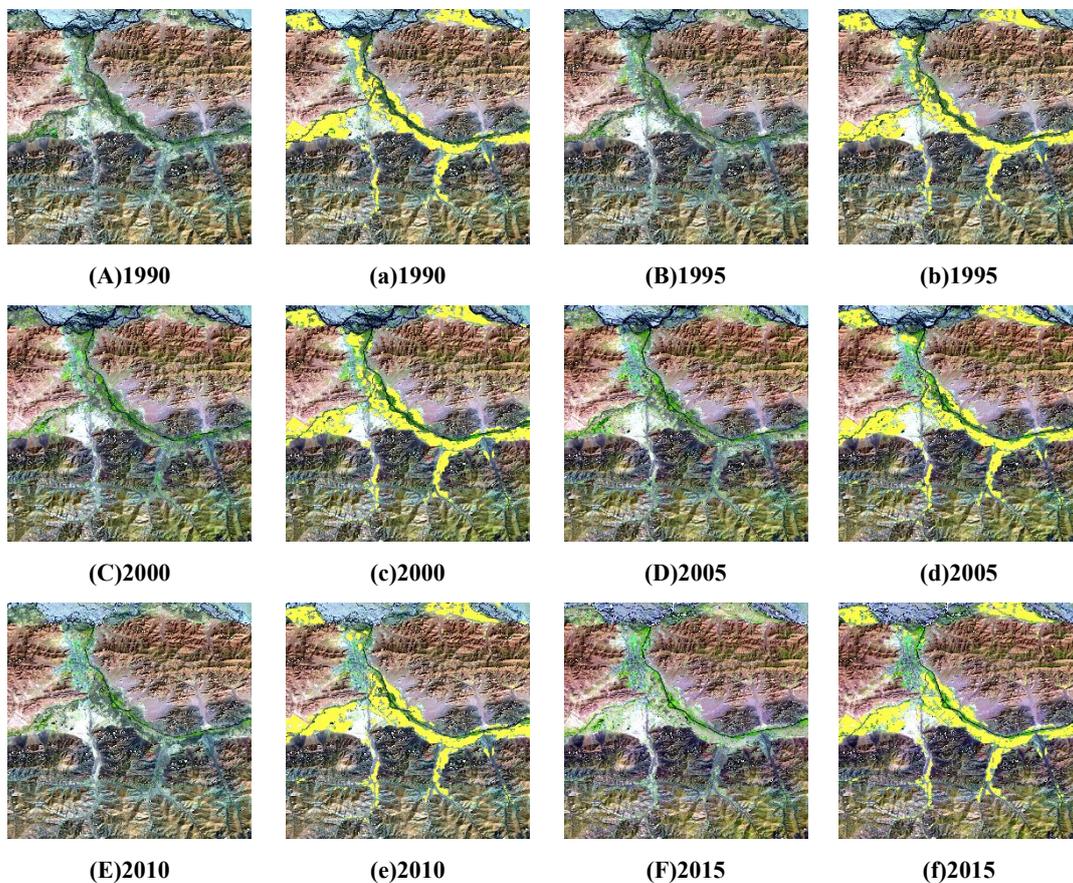
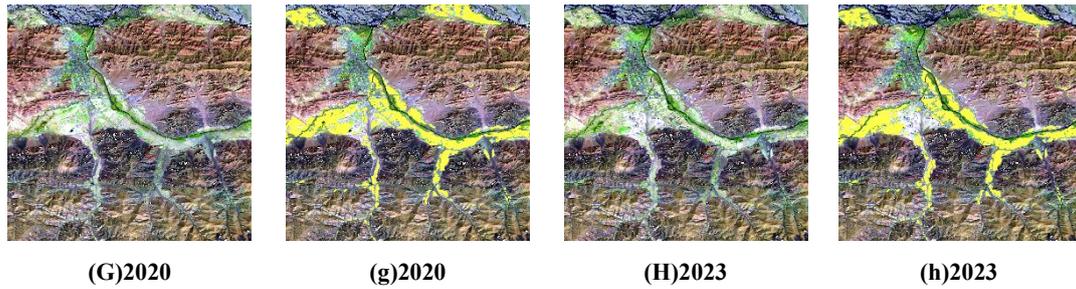


图 2 2020 年“一江两河”地区耕地空间分布图

Figure 2 Spatial distribution of cropland in the YarlungZangbo-Lhasa-Nyangqy River region in 2020





(A) – (H) 是 1990–2023 年的 Landsat 假彩色图像。其中，(A) – (E) 图像来自 Landsat 5，其红绿蓝通道分别对应短波红外 (SR_B5)、近红外 (SR_B4) 和红波段 (SR_B3)；(F) – (H) 图像来自 Landsat 8，其红绿蓝通道则分别使用短波红外 (SR_B6)、近红外 (SR_B5) 和红波段 (SR_B4)。(a) – (h) 是对应年份的耕地分布结果，黄色部分为耕地。

(A)–(H) are Landsat false-color images for 1990–2023. Specifically, (A)–(E) are from Landsat 5, with the red, green, and blue channels corresponding to Short-Wave Infrared (SR_B5), Near Infrared (SR_B4), and Red Band (SR_B3), respectively; (F)–(H) are from Landsat 8, with the red, green, and blue channels corresponding to Short-Wave Infrared (SR_B6), Near Infrared (SR_B5), and Red Band (SR_B4). (a)–(h) are the results of cropland distribution for the corresponding years, with cropland in yellow.

图 3 1990–2023 年“一江两河”地区耕地空间分布局部图

Figure 3 Spatial distribution of cropland in a segment of the YarlungZangbo-Lhasa-Nyangqy River region, 1990–2023

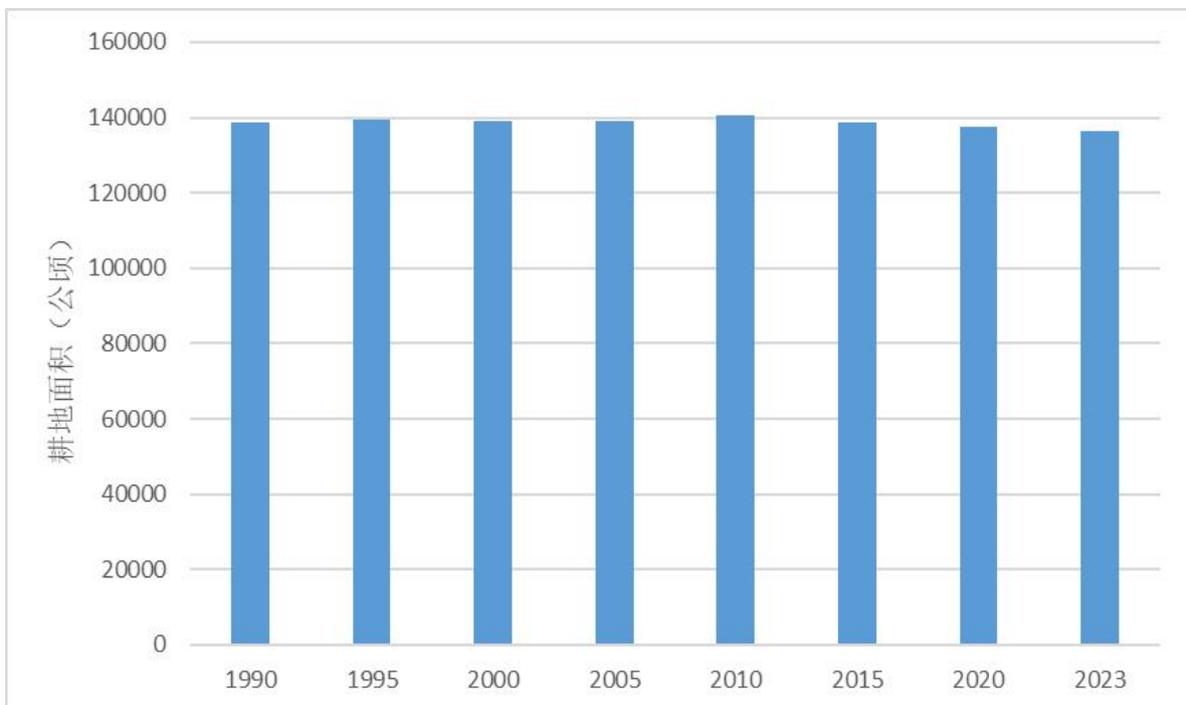


图 4 “一江两河”地区 1990–2023 年年度耕地面积

Figure 4 Annual cropland area in the YarlungZangbo-Lhasa-Nyangqy River region, 1990–2023

表 1 各产品耕地类的精度验证结果

Table 1 Validation results of farmland in each product

产品	空间分辨率 (米)	精确率	召回率	F1 得分	整体精确度
本产品 (2015 年)	30	97.72%	91.99%	0.9477	94.34%
本产品 (2020 年)	30	96.84%	92.52%	0.9463	94.46%
AIEC 产品 (2020 年)	10	95.11%	54.42%	0.6923	74.48%
ESA 产品 (2020 年)	10	95.04%	44.90%	0.6099	69.70%
Esiri 产品 (2020 年)	10	92.50%	40.06%	0.5591	66.67%
GLC-30 产品 (2020 年)	30	74.80%	86.17%	0.8008	77.39%

多数产品在精确率上表现相近且水平较高，但我们的产品在召回率方面显著优于其他产品，这表明我们的产品在耕地识别的完整性上具有明显的优势，能够更有效地减少漏分现象。与 GLC-30 产品相比，尽管 GLC-30 在召回率上表现较好（表明其能够较好地识别出耕地区域），但它在避免非耕地误判为耕地方面的表现不佳，导致较低的精确率。与之相对，本产品在确保较高召回率的同时，还能保持高精度率，显示出其在识别耕地时错分更少。因此本产品获得了优于其他产品的 F1-Score，2015 年及 2020 年的 F1 得分分别达到了 0.9477 及 0.9496。进一步来看，本产品在耕地及非耕地的整体分类精度上也表现出色，2015 年和 2020 年的整体精度分别达到了 94.34% 及 94.46%。由于样本数据的问题，未能对其他年份的数据进行精度验证。但考虑到 Landsat 系列数据具有较好的几何定位和辐射精度，且采用的耕地提取方法相同，可以认为其他年份的耕地精度与 2015 和 2020 接近。

另外本研究利用了西藏自治区及部分县域统计年鉴数据^[16-17]以及西藏自治区政府信息公开信息及年度报告，辅助评估“一江两河”地区耕地分布的精确度。考虑到该区域覆盖西藏自治区 18 个县（市、区），而统计数据来源不同且覆盖不全，本研究选择了数据公开的县（市、区）的统计年鉴中“耕地面积”数据作为辅助验证。图 5 展示了本产品耕地面积和统计年鉴耕地面积的散点图。从图 5 可以看出，散点图中的点大多分布在趋势线附近，这表明统计年鉴数据和产品数据之间存在较高的一致性。特别是尼木县、桑日县、曲水县等县的耕地面积在产品数据和统计年鉴数据中呈现出相似的趋势。整体来看，趋势线的斜率接近 1，说明两组数据间基本呈线性关系，且截距较小，进一步说明了产品耕地面积与统计年鉴耕地面积的高度相关性。

4 数据价值

本数据产品集提供了 1990–2023 年青藏高原“一江两河”地区的 30 米分辨率耕地空间分布情况。与现有的共享产品相比，本数据产品集在精确性和时效性方面具有明显优势。常见的大尺度土地覆盖产品，如 GlobeLand30 或 ESA WorldCover 等，通常聚焦于全国乃至全球尺度的土地利用和土地覆盖情况，这些产品在青藏高原的空间分布一致性仍存在较大的不确定性，而本数据产品专注于“一江两河”地区，使其在地区尺度上更为精确。此外，与多数数据产品相比，本数据产品集更新至 2023 年，提供了最近的耕地空间分布信息。

本数据产品集具有广泛的应用前景，不仅适用于农业生产管理和土地利用规划，还能支持生态

环境监测与气候变化研究。借助本数据产品集，可以深入分析人类活动和自然环境变化对耕地动态的具体影响，从而更全面地理解耕地变化背后的生态与社会经济因素。

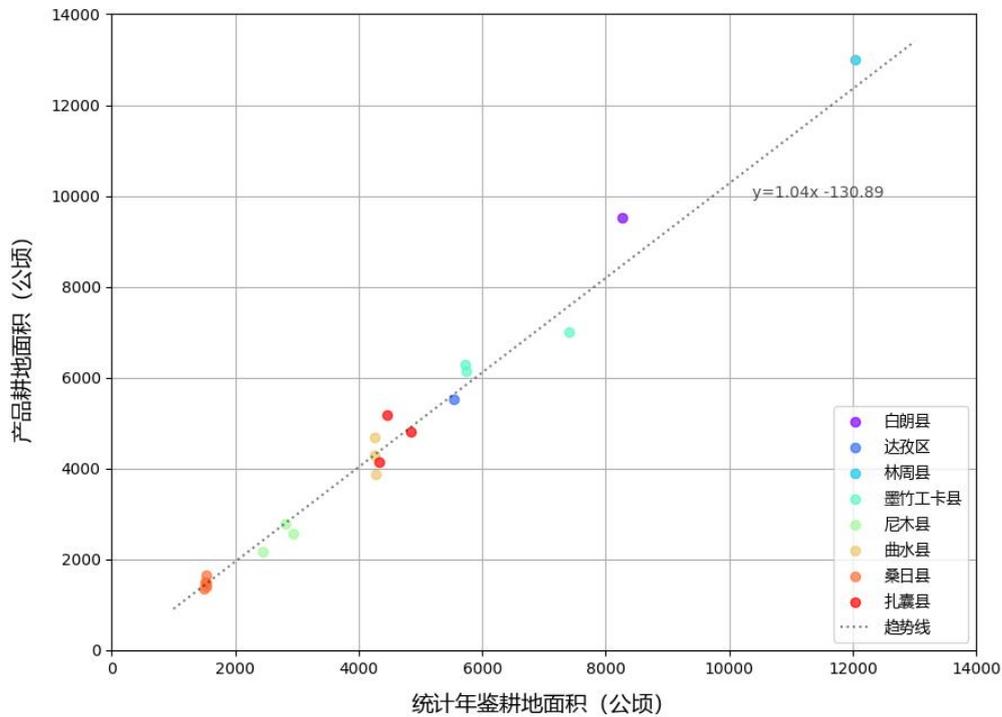


图 5 产品数据与统计年鉴的耕地面积对比散点图

Figure 5 Scatter plot of cropland area: comparison between product data and official statistical data

数据作者分工职责

陈慧玲（1999—），女，浙江省金华市人，学士，在读硕士，研究方向为遥感数据智能处理与信息挖掘。主要承担工作：实验设计，数据集构建，算法实现，图件制作，手稿撰写。

何国金（1968—），男，福建省龙岩市人，博士，研究员，研究方向为遥感数据智能处理与信息挖掘。主要承担工作：总体思路与实验方案设计，手稿修改、校对。

彭雪丽（1994—），女，江苏省徐州市人，硕士，在读博士，研究方向为遥感数据智能处理与信息挖掘。主要承担工作：模型训练，精度验证。

王桂周（1984—），男，山东省济宁市人，博士，高级工程师，研究方向为遥感图像智能处理。主要承担工作：高分数据标准化处理，数据集构建。

尹然宇（1995—），男，山东省临沂市人，博士，助理研究员，研究方向为遥感图像处理与信息智能挖掘。主要承担工作：高分数据整理。

参考文献

[1] GIRI C P. Remote sensing of land use and land cover principles and applications[M]. Boca Raton: CRC Press, 2012.

- [2] BADRELDIN N, ABU HATAB A, LAGERKVIST C J. Spatiotemporal dynamics of urbanization and cropland in the Nile Delta of Egypt using machine learning and satellite big data: implications for sustainable development[J]. *Environmental Monitoring and Assessment*, 2019, 191(12): 767. DOI: 10.1007/s10661-019-7934-x.
- [3] CORGNE S, HUBERT-MOY L, BETBEDER J. Monitoring of agricultural landscapes using remote sensing data[M]. *Land Surface Remote Sensing in Agriculture and Forest*. Amsterdam: Elsevier, 2016: 221–247. DOI: 10.1016/b978-1-78548-103-1.50006-6.
- [4] MOHAMMED HAMUD A, MOBARAK PRINCE H, ZULHAIDI SHAFRI H. Landuse/Landcover mapping and monitoring using Remote sensing and GIS with environmental integration[J]. *IOP Conference Series: Earth and Environmental Science*, 2019, 357(1): 012038. DOI: 10.1088/1755-1315/357/1/012038.
- [5] CHEN J, ZHANG J, ZHANG W W, et al. Continuous updating and refinement of land cover data product[J]. *National Remote Sensing Bulletin*, 2016, 20(5): 991–1001. DOI: 10.11834/jrs.20166250.
- [6] ZHANG X, LIU L Y, CHEN X D, et al. GLC_FCS30: global land-cover product with fine classification system at 30 m using time-series Landsat imagery[J]. *Earth System Science Data*, 2021, 13(6): 2753–2776. DOI: 10.5194/essd-13-2753-2021.
- [7] YANG J, HUANG X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019[J]. *Earth System Science Data*, 2021, 13(8): 3907–3925. DOI: 10.5194/essd-13-3907-2021.
- [8] LIU S, WANG H D, HU Y, et al. Land use and land cover mapping in China using multimodal fine-grained dual network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, 61: 4405219. DOI: 10.1109/TGRS.2023.3285912.
- [9] LIU Y H, ZHONG Y F, MA A L, et al. Cross-resolution national-scale land-cover mapping based on noisy label learning: a case study of China[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2023, 118: 103265. DOI: 10.1016/j.jag.2023.103265.
- [10] VAN DE KERCHOVE R, ZANAGA D, KEERSMAECKER W, et al. ESA WorldCover: Global land cover mapping at 10 m resolution for 2020 based on Sentinel-1 and 2 data[C]. *AGU Fall Meeting 2021*, New Orleans, USA, 2021.
- [11] KARRA K, KONTGIS C, STATMAN-WEIL Z, et al. Global land use/land cover with Sentinel 2 and deep learning[C]. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. Brussels, Belgium. IEEE, 2021: 4704–4707. DOI: 10.1109/IGARSS47720.2021.9553499.
- [12] ZHENG K Y, HE G J, YIN R Y, et al. A comparison of seven medium resolution impervious surface products on the Qinghai–Tibet Plateau, China from a user’s perspective[J]. *Remote Sensing*, 2023, 15(9): 2366. DOI: 10.3390/rs15092366.
- [13] 李丹, 田沛佩, 罗红英, 等. 西藏“一江两河”耕地生态安全时空格局与障碍诊断[J]. *农业机械学报*, 2020, 51(10): 213–222. DOI: 10.6041/j.issn.1000-1298.2020.10.024. [LI D, TIAN P P, LUO H Y, et al. Spatio-temporal characteristics and obstacle diagnosis of cultivated land ecological security in “one river and two tributaries” region in Tibet[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(10): 213–222. DOI: 10.6041/j.issn.1000-1298.2020.10.024.]
- [14] European Space Agency. Copernicus DEM-Global and European Digital Elevation Model (CO P-DEM) [EB/OL]. (2021-07) [2024-03-10]. <https://spacedata.copernicus.eu/collections/copernicus-digital>

l-elevation-model.

[15] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016.

[16] 国家统计局农村社会经济调查司. 中国县域统计年鉴（1990–2020）（县市卷）[M]. 北京: 中国统计出版社, 2021 [Rural Social and Economic Survey Division, National Bureau of Statistics. China Statistical Yearbook (County-level) (1990–2020) [M]. Beijing: China Statistics Press, 2021.]

[17] 西藏自治区统计局. 西藏统计年鉴（1990–2020）[M]. 北京: 中国统计出版社, 2020 [Bureau of statistics of Tibet Autonomous Region, National Bureau of Statistics. Tibet statistical yearbook(1990–2020) [M]. Beijing: China Statistics Press, 2020.]

论文引用格式

陈慧玲, 何国金, 彭雪丽, 等. 1990–2023 年青藏高原“一江两河”地区 30 米分辨率耕地空间分布数据产品集[J/OL]. 中国科学数据, 2025, 10(2). (2024-10-24). DOI: 10.11922/11-6035.csd.2024.0088.zh.

数据引用格式

陈慧玲, 何国金, 彭雪丽, 等. 1990–2023 年青藏高原“一江两河”地区 30 米分辨率耕地空间分布数据产品集[DS/OL]. V2. Science Data Bank, 2024. (2024-05-22). DOI: 10.57760/sciencedb.j00001.01185.

30-meter resolution cropland distribution dataset of the YarlungZangbo-Lhasa-Nyangqy River region, Tibet Autonomous Region, 1990–2023

CHEN Huiling^{1,2}, HE Guojin^{1,2*}, PENG Xueli^{1,2}, WANG Guizhou¹, YIN Ranyu¹

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, P.R. China

2. University of Chinese Academy of Sciences, Beijing 100049, P.R. China

*Email: hegj@aircas.ac.cn

Abstract: In this paper, the spatial distribution dataset of cropland in the Yarlung-Zangbo-Lhasa-Nyangqy River region from 1990 to 2023 was extracted based on Landsat 5 Level 2, Collection 2, Tier 1 and Landsat 8 Level 2, Collection 2, Tier 1 data. High-resolution images were used for the cropland sample labeling in 2023, and for generating validation point sets in 2015 and 2020. The ResUNet model was applied to extract cropland information from Landsat data. The model underwent fine-tuning to adapt to the characteristics of Landsat 5 data, ensuring consistency in accuracy across different sensors and years. Eight year's products were finally obtained, corresponding to the distribution of cropland in the region in 1990, 1995, 2000, 2005, 2010, 2015, 2020 and 2023. Among them, the precision of the 2015 and 2020 cropland distribution dataset is 97.72% and 96.84%, the recall is 91.99% and 92.52%, and the F1 reaches 0.9477 and 0.9463. This

dataset can provide scientific basis for the sustainable development of agriculture and land management in this area.

Keywords: cropland distribution; YarlungZangbo-Lhasa-Nyangqv River region; Landsat; long time series

Dataset Profile

Title	30-meter resolution cropland distribution dataset of the YarlungZangbo-Lhasa-Nyangqv River region, Tibet Autonomous Region, 1990–2023
Data corresponding author	HE Guojin (hegj@aircas.ac.cn)
Data authors	CHEN Huiling, HE Guojin, PENG Xueli, WANG Guizhou, YIN Ranyu
Time range	1990–2023
Geographical scope	The geographic range is 28°17'–30°28'N, 87°04'–92°37'E. The geographic area covers 18 counties (cities and districts) in Lhasa, Shannan and Rikaze regions of the Tibet Autonomous Region.
Spatial resolution	30 m
Data volume	1.15 GB
Data format	*.tif (GeoTIFF, byte)
Data service system	https://doi.org/10.57760/sciencedb.j00001.01185
Source of funding	The Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0307)
Dataset composition	This dataset includes the spatial distribution of cropland at 30-meter resolution in the YarlungZangbo-Lhasa-Nyangqv River region from 1990 to 2023, of which every five years from 1990 to 2020 is a phase, and 2023 is a phase, with a total of eight phases of data, including eight *.tif files.