

基于 MODIS 的宁夏县级尺度枸杞估产数据集 (2010–2019)

ISSN 2096-2223

CN 11-6035/N

王涛¹, 刘立波¹, 张鹏^{1*}, 王晓丽^{2,3,4*}



文献 CSTR:

32001.14.11-6035.nasdc.2022.0001.zh

文献 DOI:

10.11922/11-6035.nasdc.2022.0001.zh

数据 DOI:

10.12205/E0290.20220110.36.cs.1925

文献分类: 农学

收稿日期: 2022-01-17

开放同评: 2022-03-03

录用日期: 2022-04-11

发表日期: 2022-09-26

1. 宁夏大学信息工程学院, 银川 750021
2. 中国农业科学院农业信息研究所, 北京 100081
3. 国家农业科学数据中心, 北京 100081
4. 中国农业科学院国家南繁研究院, 海南三亚 572024

摘要: 针对现有遥感估产数据集过度依赖于实测数据和人工经验的问题, 本文基于宁夏 16 县 MODIS 高光谱影像、年际枸杞统计产量数据和枸杞种植区域矢量图, 构建了一种多波段、多时相融合的县级尺度遥感影像估产数据集。首先, 采用重投影、重采样和时间序列补充 3 种方法对 MODIS 多产品影像数据进行预处理, 保证了遥感影像数据时空的一致性; 然后, 对上述处理结果进行波段融合和时间序列融合操作, 增强了波段信息及时间序列信息的丰富度; 最后, 运用直方图降维方法对各县融合影像数据进行降维, 克服了数据稀疏的问题。实验结果表明, 通过卷积神经网络自动提取本文数据集的影像特征, 估产模型在 MRE、RMSE 和 R^2 上分别达到了 14.52%、859.23 吨和 0.83, 验证了本数据集的准确性。本数据集可为宁夏县级尺度枸杞年际产量预测提供重要数据支持, 对地区农业可持续发展和科学研究具有重要意义。

关键词: 估产数据集; 多时相; 多波段; MODIS

数据库(集)基本信息简介

数据库(集)名称	基于MODIS的宁夏县级尺度枸杞估产数据集(2010–2019)
数据作者	刘立波
数据通信作者	张鹏(pengzhang123@foxmail.com), 王晓丽(wangxiaoli@caas.cn)
数据时间范围	2010 – 2019年
地理区域	宁夏境内(35° 14' –39° 23' N, 104° 17' –107° 39' E)
数据量	157 MB
数据格式	*.npz, *.tif, *.csv
数据服务系统网址	http://dx.doi.org/doi:10.12205/E0290.20220110.36.cs.1925
基金项目	宁夏重点研发计划项目(2020BFG02013)
数据库(集)组成	本数据集由2010–2019年枸杞年鉴统计产量数据和MODIS影像组成。其中枸杞年鉴统计产量数据为2010–2019年宁夏中宁、贺兰等16县年际产量数据, 为csv文件, 包含7个属性列, 分别为年份、省份名、省份编号、县名、县级编号、枸杞种植面积(单位: 公顷)和产量(单

* 论文通信作者

张鹏: pengzhang123@foxmail.com

王晓丽: wangxiaoli@caas.cn

数据库（集）组成	位：吨），总计160条产量数据；MODIS影像数据为160个大小是 $32 \times 26 \times 13$ 的直方图矩阵，包含三位维度，分别是像素区间个数、时相和波段数，以“hist”加年份和地名进行命名，为npz文件，并提供与其相对应的MODIS影像文件（未经直方图降维方法处理的波段、时相融合结果，以“img”加年份和地区名命名，为tif文件，也包含3个维度，分别是影像宽度、高度和波段数）。影像命名中的年份和地区名与产量数据中的“Year”和“County”列名相对应。
----------	---

引言

估产数据集是表征农作物长势和产量的重要信息，被广泛用于作物长势分析、产量预测等研究。然而，传统的估产数据集制作方法采用人工实地测量土壤墒情、作物干重、气象条件等参数，从而构建数据集^[1-3]。该方法速度慢、工作量大，且过度依赖于实测数据。

近年来，遥感影像因其低成本、覆盖面广等优势，常被用于提取归一化植被指数（Normalized Difference Vegetation Index, NDVI）和增强型植被指数（Enhanced Vegetation Index, EVI）。如 Quarmby^[4-6]等利用红外波段和近红外波段计算得到了 NDVI，Bolton 等^[7-8]利用红外、近红外和蓝光波段计算得到了 EVI；此外，满卫东等^[9]基于 AVHRR 影像制作了基于辽宁省 1982 至 2009 年 AVHRR 归一化植被指数数据集（<https://www.osgeo.cn/data/wef57>）。更进一步，Aerial Intelligence 公司 2017 年发布的美国冬小麦产量预测数据集（<https://aerialintel.blob.core.windows.net/recruiting/datasets/wheat-2013-supervised.csv> 和 <https://aerialintel.blob.core.windows.net/recruiting/datasets/wheat-2014-supervised.csv>）不仅提取遥感影像中的 NDVI 和 EVI，而且加入了气象、地理位置等特征，提升了估产特征的多样性。但在这类数据集中，NDVI、EVI 等手工特征仅用少许的波段计算得到，忽略了其余波段重要信息，且特征的选取过度依赖人工经验，具有一定局限性。

因此，针对以上问题，本文摒弃手工特征制作的方法，制作了一种基于 MODIS 高光谱遥感影像的、多波段和多时相融合的宁夏县级尺度枸杞估产数据集，用于卷积神经网络特征的自动提取，简化特征提取操作的同时，进一步增强了特征丰富度。

1 数据采集和处理方法

1.1 数据来源

本文采用的实验数据由宁夏回族自治区遥感影像、枸杞种植区域矢量图和年际枸杞产量 3 类数据组成。其中，遥感影像采用 MODIS 高光谱影像数据，来源于 EARTHDATA 网站（<https://search.earthdata.nasa.gov/>），行列号为 h26v04，时间范围为枸杞生长季内的每年第 97 天至第 297 天，包括 MOD09A1、MOD13A1、MYD11A2、MCD15A2H 4 种类型的 MODIS 产品数据（表 1）；枸杞种植区域矢量图由宁夏农林科学院研究人员实地记录枸杞种植区域经纬度制成，为 shp 文件；年际枸杞产量数据来源于宁夏回族自治区统计局，包括 2010–2019 年宁夏 16 县（县级市区）枸杞种植面积和实际产量。其中，2019 年产量数据如表 2 所示，枸杞种植总面积为 27960 公顷，总产量为 94843 吨，平均产量为 3.39 吨/公顷。其中，同心县、红寺堡区等地区枸杞产量均高于 9000 吨，属于高产地区；金凤区、平罗县等地区产量均在 1000 吨左右，属于中产地区；其余为低产地区。

表 1 影像及矢量数据表

Table 1 The table of remote sensing images and vector data

数据类别	数据名称	投影坐标系	空间分辨率/m	时间分辨率/天	影像个数/个	波段描述
MOD09A1	地表反射率	正弦曲线投影	500	8	1040	Band1(620–670 nm)
						Band2(841–876 nm)
						Band3(459–479 nm)
						Band4(545–565 nm)
						Band5(1230–1250 nm)
						Band6(1628–1652 nm)
						Band7(2105–2155 nm)
MOD13A1	植被指数	正弦曲线投影	500	16	1040	NDVI
						EVI
MYD11A2	地表温度	正弦曲线投影	1000	8	1040	LST of day
						LST of night
MCD15A2H	叶面积指数和光合有效辐射	正弦曲线投影	500	8	1040	LAI
						FPAR
矢量图	枸杞种植区域矢量图	西安 80 投影	1	-	1	-

表 2 2019 年产量数据表

Table 2 The table of yield data in 2019

地区	面积/公顷	产量/吨
同心县	6495	27610
中宁县	7986	26756
沙坡头区	5917	22148
红寺堡区	3104	9452
金凤区	798	1795
西夏区	646	1453
惠农区	386	1084
永宁县	344	1050
平罗县	360	980
贺兰县	195	686
原州区	856	620
利通区	539	493
青铜峡市	133	282
大武口区	89	267
灵武市	37	104
盐池县	75	63

1.2 数据处理

1.2.1 重投影与重采样

为了保证遥感影像数据空间位置的一致性, 首先, 利用 MODIS 影像重投影工具 (MODIS Reprojection Tool, MRT) 工具将 MODIS 遥感影像和枸杞种植区域掩膜数据重投影为基于 WGS-84 椭球体的 UTM 投影; 然后, 将 MYD11A2 影像和枸杞矢量数据重采样为 500 m, 使 MYD11A2 影像和枸杞矢量数据与其余 MODIS 数据产品的空间分辨率相互统一。

1.2.2 时间序列补充

由于 MOD13A1 的时间分辨率为 16 天, 其余 MODIS 数据产品为 8 天 (表 1), 为了保证时间序列的完整性, 根据式 (1) 采用上下影像求平均的方法对枸杞生长季内缺失影像进行补充, 如将 MOD13A1 第 97 天和第 113 天影像的均值作为第 105 天的影像数据。

$$I_i = \frac{I_{i-8} + I_{i+8}}{2} \quad (i = 105, 121, 137, \dots, 289) \quad (1)$$

式中, I_i 、 I_{i-8} 和 I_{i+8} 分别为 MOD13A1 第 i 天、第 $i-8$ 天和第 $i+8$ 天影像数据, i 的取值范围为 [105, 289], 时间间隔为 8 天。

1.2.3 波段融合和时间序列融合

为了提高波段信息的丰富度, 分别提取了同一景 MOD09A1、MOD13A1 等遥感影像中的 band1-band7、NDVI、EVI 等 13 个波段 (表 1), 并对其进行了融合。随后, 为了进一步融合影像不同时相上的枸杞生长信息, 对波段融合结果在时间维度上进行融合, 最终形成了 10 张波段数为 13, 时相为 26 的时间序列影像。

1.2.4 枸杞种植区域提取

遥感图像是基于像素点的, 包括了各种地面覆盖类型, 为了使各县年鉴统计产量与其遥感影像相对应, 利用了不同县域枸杞种植矢量图提取上述融合数据中的种植区域, 以中宁县枸杞种植区域提取为例。首先, 采用 ArcGIS 软件从宁夏行政区划矢量图提取出中宁县行政区划矢量图; 然后, 根据中宁县行政区划矢量图左上角和右下角经纬度裁剪得到中宁县融合影像 (图 1); 接着, 调用 GDAL (Geospatial Data Abstraction Library) 库中的 warp 函数以实现枸杞中宁枸杞种植区域矢量图 (图 2) 对图 1 的裁剪, 从而得到中宁市枸杞种植区域图 (图 3); 最后, 以此类推, 提取出其余县枸杞种植区域。

1.2.5 直方图降维

鉴于影像数据集的稀疏性, 无法采用端到端方式训练估产模型。因此, 本文将影像无差别划分为 32 个像素区间, 进而将影像中每个波段不同像素值映射至不同区间, 以达到直方图降维的目的。

2 数据样本描述

基于 MODIS 的宁夏县级尺度枸杞估产数据集 (2010–2019) 主要由遥感影像数据和年鉴统计产量数据组成, 具体数据样本描述如下。



图 1 中宁县高光谱影像

Figure 1 Hyperspectral image of Zhongning County



图 2 中宁县枸杞种植区域矢量图

Figure 2 Vector map of wolfberry planting area in Zhongning County

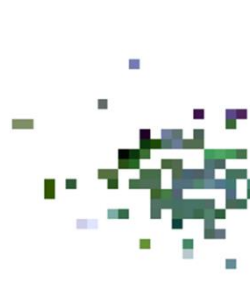


图 3 中宁县枸杞种植区域图

Figure 3 The area of interest for wolfberry planting in Zhongning County

2.1 遥感影像数据描述

遥感影像数据以 MODIS 高光谱影像为数据源，经上述方法处理后，共形成了 160 个大小为 $32 \times 26 \times 13$ （32 为像素区间数，26 为时相数，13 为波段数）的直方图矩阵，以“hist_年份_地名.npy”的方式进行命名，存放于“dataset/hist_data”文件夹中。其中，16 个县（县级市区）（表 2）中的每个地区各涵盖 10 年直方图矩阵数据，时间范围为 2010–2019 年，直方图矩阵数据示意如图 4 所示。

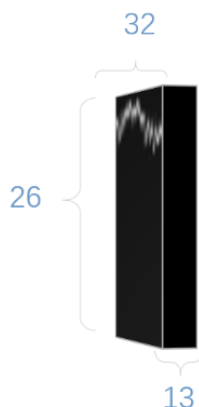


图 4 直方图矩阵数据示意图

Figure 4 Schematic diagram of histogram matrix data

2.2 统计产量数据描述

为了方便用户对数据的加载及使用，本文提供了与直方图矩阵年份和地名相对应一致的年鉴统计产量数据。该数据为表格数据，保存于“dataset/yield_data/yield_data.csv”文件中，共包含 7 个属性列，分别为年份、省份名、省份编号、县名、县级编号、枸杞种植面积和产量，总计 160 条产量数据，平均每个县（县级市区）包含 10 条 2010–2019 年的枸杞产量数据，部分产量数据如表 3 所示。

表 3 部分产量数据表

Table 3 The table of partial yield data

年份	省份名	省份编号	县名	县级编号	枸杞种植面积/公顷	产量/吨
2010	宁夏回族自治区	1	银川市	1	601	2038
2011	宁夏回族自治区	1	银川市	1	783	2657
2012	宁夏回族自治区	1	银川市	1	839	2846
2013	宁夏回族自治区	1	银川市	1	918	3113
2014	宁夏回族自治区	1	银川市	1	724	2455
2015	宁夏回族自治区	1	银川市	1	826	2668
2016	宁夏回族自治区	1	银川市	1	1164	3492
2017	宁夏回族自治区	1	银川市	1	1065	3398
2018	宁夏回族自治区	1	银川市	1	1322	4414
2019	宁夏回族自治区	1	银川市	1	1901	4274
2010	宁夏回族自治区	1	永宁县	2	194	659
2011	宁夏回族自治区	1	永宁县	2	253	859
2012	宁夏回族自治区	1	永宁县	2	271	920
2013	宁夏回族自治区	1	永宁县	2	297	1006
2014	宁夏回族自治区	1	永宁县	2	234	794
2015	宁夏回族自治区	1	永宁县	2	267	862
2016	宁夏回族自治区	1	永宁县	2	373	1176
2017	宁夏回族自治区	1	永宁县	2	371	1594
2018	宁夏回族自治区	1	永宁县	2	227	960
2019	宁夏回族自治区	1	永宁县	2	344	1050

3 数据质量控制和评估

为了验证本文数据集的可用性，本文分别从定性和定量两个角度评价本数据质量和准确性，具体如下：

3.1 定性评价

通过可视化对比高中低产量所对应的直方图降维结果（图 5）可见，在高产、中产和低产中波段 1、波段 7、NDVI 和 EVI 波段明显存在视觉差异，表明可以从本数据集中提取到有利的特征，用于表征枸杞的长势和产量。此外，在时间维度上，像素区间最大值基本集中于第 180 天左右，正值枸杞植株长势茂盛期，符合枸杞的生长规律，说明本文数据集质量较好。

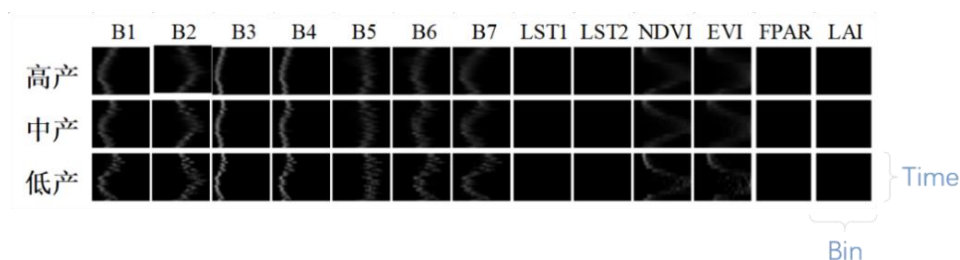


图 5 可视化结果图

Figure 5 Visualized result graph

3.2 定量评价

为了进一步说明该数据集的准确性, 本文以平均相对误差 (Mean Relative Error, MRE)、均方根误差 (Root Mean Square Error, RMSE) 和决定系数 (Coefficient of Determination, R^2) 为评价指标, 计算公式如式 (2)–式 (4) 所示, 分别基于本文数据集和 Aerial Intelligence 公司发布的美国冬小麦产量预测数据集, 设置了两个对比实验组, 分别是:

(1) 以本文数据集为数据源, 构建了卷积神经网络 (Convolutional Neural Networks, CNN), 用于枸杞长势特征的抽取, 进而采用全连接网络 (Fully Connected Network, FCN)、支持向量回归 (Support Vector Regression, SVR) 和岭回归 (Ridge Regression, RR) 方法预测枸杞年际产量。卷积神经网络结构如图 6 所示, 第一层为输入层, 为本文数据集, 训练集、验证集和测试集划分比率为 7:2:1; 第二层为特征提取层, 由 6 个卷积层构成, 卷积核数量分别是 128、256、256、512、512、512, 卷积核大小全为 3×3 ; 最后一层为全连接层, 将特征图映射至 2048 维 embedding (嵌入向量) 空间, 以表征枸杞的长势信息。

(2) 与上述方法相比, 对比实验基于冬小麦产量预测数据集, 仅用单一的 FCN、RR 和 SVR 预测冬小麦产量。其中, FCN 包含 5 个隐藏层, 维度分别为 1024、512、256、128 和 1, 实验结果如表 4 所示。

表 4 对比实验结果表

Table 4 The table of comparative experiment results

方法	MRE(%)	RMSE(吨)	R^2
SVR	18.82	1115.60	0.70
RR	15.38	911.68	0.77
FCN	18.11	1073.35	0.67
CNN+SVR(ours)	13.90	822.54	0.82
CNN+ RR (ours)	14.01	852.48	0.72
CNN+FCN(ours)	14.52	859.23	0.83

由表 4 可知, 在本文数据集上, MRE 和 RMSE 分别为 14.52%、859.23 吨, 且 R^2 达到了 0.83, 均优于对比数据集。与 2017 年冬小麦产量预测比赛数据集相比, 本文数据集采用波段融合和时间序列融合方法分别融合了不同波段和时相上的影像信息, 增强了特征丰富度, 所以基于本文数据集的 CNN 回归方法, 在 MRE 和 RMSE 上分别下降了 1.48% 和 89.14 吨, R^2 上升了 0.06, 说明增加了 0.06% 的特征可由回归方法解释, 验证了本文数据集的准确性。

$$\text{MRE} = \frac{1}{n} \sum_{i=1}^n \frac{T_i - P_i}{T_i} \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (T_i - P_i)^2}{n}} \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (T_i - P_i)^2}{\sum_{i=1}^n (T_i - \bar{T})^2} \quad (4)$$

其中, T_i 为第 i 个县枸杞统计产量, P_i 为第 i 个县产量预测值, $\bar{T}$ 为 16 县统计数据平均产量, 单位为吨。

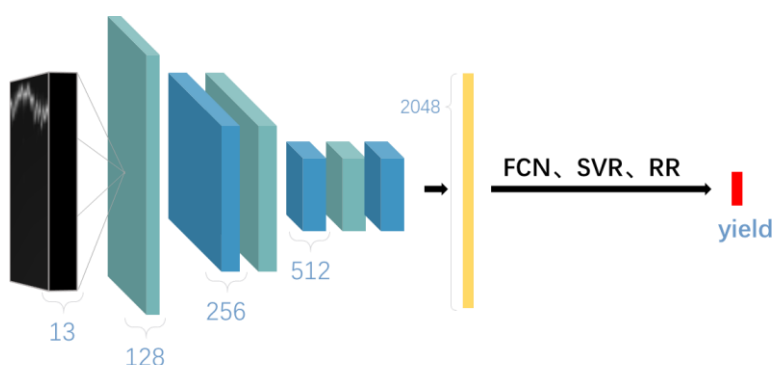


图 6 卷积神经网络结构图

Figure 6 Convolutional neural network structure diagram

4 数据价值

目前, 公开的遥感估产数据集非常有限, 且目视解译纹理特征等相关提法过度依赖于实测数据和人工经验。因此, 本文构建了一种基于 MODIS 的宁夏县级尺度枸杞估产数据集, 具有高光谱、多波段、多时相等特点, 可用于卷积神经网络特征自动提取农作物长势特征, 为作物产量预测研究提供数据支撑, 同时为遥感多时相估产影像数据集的制作提供了一定的参考价值。

数据作者分工职责

王涛 (1996—), 男, 四川人, 硕士, 研究方向为遥感影像处理。主要承担工作: 图像处理、论文撰写。

刘立波 (1974—), 女, 宁夏人, 博士, 教授, 中国计算机学会会员, 研究方向为智能信息处理。主要承担工作: 本数据集的规划与建立。

张鹏 (1975—), 男, 宁夏人, 博士, 副教授, 主要从事人工智能及农业信息化研究。主要承担工作: 影像数据与产量数据的收集。

王晓丽 (1982—), 女, 河北人, 博士, 助理研究员, 主要从事科学数据管理相关研究。主要承担工作: 论文撰写指导。

参考文献

- [1] MEDIN K. Crop yield estimation and crop insurance in Sweden[J]. Review of the International Statistical Institute, 1965, 33(3): 414. DOI:10.2307/1401703.
- [2] SENGUPTA J, JACOB P. A Technique for Yield-Estimation Surveys in Cinchona Plantations[J]. Sankhyā: The Indian Journal of Statistics, Series B (1960-2002), 1968, 30(1/2).
- [3] 李付琴, 田国良. 小麦单产的遥感—气象综合模式研究[J]. 环境遥感, 1993(3): 202–210. [LI F Q, TIAN G L. AfResearch on remote sensing-meteorological model for wheat yield estimation[J]. National Remote Sensing Bulletin, 1993(3): 202–210.]
- [4] QUARMBY N A, MILNES M, HINDLE T L, et al. The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction[J]. International Journal of Remote Sensing, 1993, 14(2): 199–210. DOI:10.1080/01431169308904332.
- [5] 徐勇, 黄雯婷, 窦世卿, 等. 2000—2020 年西南地区植被 NDVI 对气候变化和人类活动响应特征[J]. 环境科学, 2022, 43(6): 3230–3240. DOI:10.13227/j.hjlx.202108107. [XU Y, HUANG W T, DOU S Q, et al. Responding mechanism of vegetation cover to climate change and human activities in southwest China from 2000 to 2020[J]. Environmental Science, 2022, 43(6): 3230–3240. DOI:10.13227/j.hjlx.202108107.]
- [6] 湛稳, 孙立群, 李晴岚, 等. 一种基于图论重构 MODIS EVI 时间序列数据集的新方法[J]. 地球信息科学学报, 2022, 24(4): 738–749. [CHEN W, SUN L Q, LI Q L, et al. A new method to reconstruct MODIS EVI time series data set based on graph theory[J]. Journal of Geo-Information Science, 2022, 24(4): 738–749.]
- [7] BOLTON D K, FRIEDL M A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics[J]. Agricultural and Forest Meteorology, 2013, 173: 74–84. DOI:10.1016/j.agrformet.2013.01.007.
- [8] XIE F, FAN H. Deriving drought indices from MODIS vegetation indices (NDVI/EVI) and Land Surface Temperature (LST): is data reconstruction necessary? [J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 101: 102352. DOI:10.1016/j.jag.2021.102352.
- [9] 满卫东, 王宗明, 刘明月, 等. 1990—2013 年东北地区耕地时空变化遥感分析[J]. 农业工程学报, 2016, 32(7): 1–10. [MAN W D, WANG Z M, LIU M Y, et al. Spatio-temporal dynamics analysis of cropland in Northeast China during 1990-2013 based on remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(7): 1–10.]

论文引用格式

王涛, 刘立波, 张鹏, 等. 基于 MODIS 的宁夏县级尺度枸杞估产数据集 (2010–2019) [J/OL]. 中国科学数据, 2022, 7(3). (2022-09-26). DOI:10.11922/11-6035.nasdc.2022.0001.zh.

数据引用格式

刘立波. 基于 MODIS 的宁夏县级尺度枸杞估产数据集 (2010–2019) [DS/OL]. 宁夏大学信息工程学

院. 国家农业科学数据中心, 2022. (2022-01-11). DOI: 10.12205/E0290.20220110.36.cs.1925.

A dataset of MODIS-based county-level wolfberry yield estimation in Ningxia (2010–2019)

WANG Tao¹, LIU Libo¹, ZHANG Peng^{1*}, WANG XiaoLi^{2,3,4*}

1. School of Information Engineering, Ningxia University, Yinchuan 750021, P.R. China

2. Agricultural Information Institute of CAAS, Beijing 100081, P.R. China

3. National Agriculture Science Data Center, Beijing 100081, P.R. China

4. National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya 572024, P.R. China

*Email: pengzhang123@foxmail.com (ZHANG Peng); wangxiaoli@caas.cn (WANG Xiaoli)

Abstract: To solve the problem that the current remote sensing yield estimation dataset is overly relying on measurement data and manual experience, we constructed a dataset of multi-band and multi-temporal remote sensing image yield estimation fusion. First, to ensure the temporal and spatial consistency of remote sensing image data, we used reprojection, resampling, and time series supplementation to preprocess MODIS image data. Then, we performed band fusion and time series fusion for the above-mentioned processing results, enhancing the richness of band information and time series information. Finally, we used the histogram dimensionality reduction method to reduce the dimensionality of the fusion data of all counties, which overcame the problem of data sparseness. According to the experimental results, the image features of the dataset automatically extracted through the convolutional neural network and the production estimation model in this article have reached 14.52%, 859.23 tons and 0.83 on MRE, RMSE and R^2 respectively, which shows the accuracy of this dataset. The dataset provides important data support for the prediction of the annual yield of wolfberry at the county level in Ningxia, which is of great significance to the sustainable development of regional agriculture and scientific research.

Keywords: yield estimation dataset; multi-temporal; multi-band; MODIS

Dataset profile

Title	A dataset of MODIS-based county-level wolfberry yield estimation in Ningxia (2010–2019)
Data corresponding author	ZHANG Peng (pengzhang123@foxmail.com), WANG Xiaoli (wangxiaoli@caas.cn)
Data author	LIU Libo
Time range	2010–2019
Geographical scope	Ningxia (35°14′–39°23′N, 104°17′–107°39′E)
Data volume	157MB
Data format	*.npz, *.tif, *.csv
Data service system	< http://dx.doi.org/doi:10.12205/E0290.20220110.36.cs.1925 >

Source of funding	Ningxia Key R&D Project (2020BFG02013)
Dataset composition	This dataset is composed of the statistical yield data of the Chinese Ningxia wolfberry yearbook from 2010 to 2019 and MODIS images. The former is the interannual yield data of 16 counties including Zhongning and Helan in Ningxia from 2010 to 2019, which is a “csv” file containing 7 attribute columns, namely year, province name, province number, county name, and county level. Number, planting area and yield of wolfberry, a total of 160 rows of data; the latter contains 160 histogram matrices with a size of 32×26×13, including three dimensions, namely the number, time phase and band number of pixel intervals named after “hist” plus the year and place name. They are “npz” files, providing corresponding MODIS image files. In addition, the band and time phase fusion results not processed by the histogram dimensionality reduction method are named “img” plus the year and the region name. They are “tif” files, also containing three dimensions, namely the image width, height and band number. The year and area names in the image naming correspond to the “Year” and “County” column names in the yield data.