

## 2000–2020 年中国陆地生态系统年总初级生产力 数据集

ISSN 2096-2223

CN 11-6035/N



文献 CSTR:

32001.14.11-6035.csd.2023.0037.zh



文献 DOI:

10.11922/11-6035.csd.2023.0037.zh

数据 DOI:

10.57760/sciencedb.o00119.00077

文献分类: 地球科学

收稿日期: 2023-01-13

开放同评: 2023-02-03

录用日期: 2023-04-26

发表日期: 2023-05-12

范仁雪<sup>1</sup>, 朱先进<sup>1,29\*</sup>, 陈智<sup>2,3\*</sup>, 于贵瑞<sup>2,3</sup>, 张维康<sup>2</sup>, 韩朗<sup>4</sup>,  
王秋凤<sup>2,3</sup>, 陈世莘<sup>5</sup>, 刘绍民<sup>6</sup>, 王辉民<sup>2</sup>, 闫俊华<sup>7</sup>, 谭俊磊<sup>8</sup>,  
张法伟<sup>9</sup>, 赵风华<sup>2</sup>, 李英年<sup>9</sup>, 张一平<sup>10</sup>, 石培礼<sup>2</sup>, 朱教君<sup>11</sup>,  
吴家兵<sup>11</sup>, 赵仲辉<sup>12</sup>, 郝彦宾<sup>13</sup>, 沙丽清<sup>10</sup>, 张玉翠<sup>14</sup>, 姜世成<sup>15</sup>,  
顾峰雪<sup>16</sup>, 吴志祥<sup>17</sup>, 张扬建<sup>2,3</sup>, 周莉<sup>18</sup>, 唐亚坤<sup>19</sup>, 贾丙瑞<sup>5</sup>,  
李玉强<sup>8</sup>, 宋清海<sup>10</sup>, 董刚<sup>20</sup>, 高艳红<sup>8</sup>, 蒋正德<sup>11</sup>, 孙聃<sup>7</sup>,  
王建林<sup>21</sup>, 何其华<sup>22</sup>, 李新虎<sup>23</sup>, 王飞<sup>24</sup>, 魏文学<sup>25</sup>, 邓正苗<sup>25</sup>,  
郝翔翔<sup>26</sup>, 李彦<sup>23</sup>, 刘晓利<sup>27</sup>, 章熙锋<sup>28</sup>, 朱治林<sup>2,3</sup>

1. 沈阳农业大学, 农学院, 沈阳 110866
2. 中国科学院地理科学与自然资源研究所, 生态系统网络观测与建模重点实验室, 中国生态系统研究网络综合研究中心, 北京 100101
3. 中国科学院大学, 资源与环境学院, 北京 100049
4. 天津大学, 地球系统科学学院, 表层地球系统科学研究院, 天津 300072
5. 中国科学院植物研究所, 植被与环境变化国家重点实验室, 北京 100093
6. 北京师范大学, 地理科学学院, 地表过程与资源生态国家重点实验室, 北京 100875
7. 中国科学院华南植物园, 广州 510650
8. 中国科学院西北生态环境与资源研究院, 兰州 730000
9. 中国科学院西北高原生物研究所, 西宁 810008
10. 中国科学院西双版纳热带植物园, 云南勐腊 666303
11. 中国科学院沈阳应用生态研究所, 沈阳 110016
12. 中南林业科技大学, 长沙 410004
13. 中国科学院大学, 生命科学学院, 北京 100049
14. 中国科学院遗传与发育生物学研究所, 农业资源研究中心, 石家庄 050021
15. 东北师范大学, 长春 130024
16. 中国农业科学院农业环境与可持续发展研究所, 北京 100081
17. 中国热带农业科学院橡胶研究所, 海口 570100
18. 中国气象科学院, 北京 100081

\* 论文通信作者

朱先进: xianjin1985@163.com

陈智: chenz@igsnr.ac.cn

19. 西北农林科技大学, 陕西杨凌 712100
20. 山西大学, 太原 030006
21. 青岛农业大学, 山东青岛 266109
22. 中国科学院成都生物研究所, 成都 610041
23. 中国科学院新疆生态与地理研究所, 乌鲁木齐 830011
24. 内蒙古农业大学, 呼和浩特 010018
25. 中国科学院亚热带农业生态研究所, 长沙 410125
26. 中国科学院东北地理与农业生态研究所, 长春 130102
27. 中国科学院南京土壤研究所, 南京 210008
28. 中国科学院水利部成都山地灾害与环境研究所, 成都 610041
29. 辽宁盘锦湿地生态系统国家观测研究站, 辽宁盘锦 124000

**摘要:** 陆地生态系统年总初级生产力 (AGPP) 是粮食生产和陆地生态系统固碳的基础。分析 AGPP 的时空变化可以为保障区域粮食安全和减缓气候变化趋势提供理论依据, 但需以准确评估区域 AGPP 为基础。基于 ChinaFLUX 的长期联网观测数据和公开的数据集, 本研究构建了中国陆地生态系统 AGPP 数据集, 并结合生物、气候和土壤因素, 利用随机森林回归树通过模拟单位叶面积的 AGPP 构建了中国 AGPP 评估模型, 生成了 2000–2020 年中国陆地生态系统 AGPP 数据, 空间分辨率为 30 arcsecond, 数据格式为 tiff。本数据可以为模型模拟提供验证数据, 也可以为区域生产力、生态质量、陆地碳汇评估及管理提供数据支撑。

**关键词:** 中国; 陆地生态系统总初级生产力; 随机森林回归树模型; 时空格局

### 数据库 (集) 基本信息简介

数据库 (集) 名称	2000-2020 年中国陆地生态系统年总初级生产力数据集
数据通信作者	朱先进 (xianjin1985@163.com); 陈智 (chenz@igsnr.ac.cn)
数据作者	范仁雪, 朱先进, 陈智, 于贵瑞, 张维康, 韩朗, 王秋风, 陈世莘, 刘绍民, 王辉民, 闫俊华, 谭俊磊, 张法伟, 赵风华, 李英年, 张一平, 石培礼, 朱教君, 吴家兵, 赵仲辉, 郝彦宾, 沙丽清, 张玉翠, 姜世成, 顾峰雪, 吴志祥, 张扬建, 周莉, 唐亚坤, 贾丙瑞, 李玉强, 宋清海, 董刚, 高艳红, 蒋正德, 孙聘, 王建林, 何其华, 李新虎, 王飞, 魏文学, 邓正苗, 郝翔翔, 李彦, 刘晓利, 章熙锋, 朱治林
数据时间范围	2000–2020年
地理区域	中国
空间分辨率	30 arcsecond
数据量	4.30 GB
数据格式	.xlsx, *.tiff
数据服务系统网址	<a href="https://doi.org/10.57760/sciencedb.o00119.00077">https://doi.org/10.57760/sciencedb.o00119.00077</a>
基金项目	国家自然科学基金 (32071585, 32071586)
数据库 (集) 组成	数据集包括站点观测AGPP数据和区域AGPP数据, 共22个文件, 站点观测AGPP数据为站点观测结果, 包括872行20列, 166个站点, 872个站点年, 5种生态系统类型等 (具体信息见表1); 区域AGPP为21个数据文件, 为逐年AGPP空间分布结果。

## 引言

年总初级生产力 (AGPP) 是一年内通过植被光合作用固定的有机物总量, 是形成粮食产量和陆地生态系统碳汇的基础<sup>[1]</sup>。分析 AGPP 的时空变化可以明确粮食产量及陆地碳汇的时空分布, 对于保障粮食安全和缓解气候变化起到至关重要的作用<sup>[2]</sup>, 而对 AGPP 时空变化的分析需要以准确评估区域 AGPP 为基础。

现有区域 AGPP 评估结果存在不确定性需要拓展新的区域评估方法。已有大量研究基于过程模型<sup>[3-4]</sup>或者机器学习语言<sup>[5-7]</sup>模拟短时间 (半小时、日尺度、8 天尺度及月尺度) 的总初级生产力累加得到区域 AGPP 结果, 实现区域 AGPP 评估, 但该方法存在参数异质性<sup>[8]</sup>、GPP 环境响应异质性<sup>[9]</sup>等局限。基于经典统计学, Zhu 等人<sup>[10]</sup>利用年均气温 (MAT) 和年总降水量 (MAP) 将站点 AGPP 直接扩展至区域水平, 但复杂的 AGPP 空间变异规律很难采用有限的环境变量通过多元回归关系来完全模拟<sup>[10]</sup>。因而, 需要以站点观测 AGPP 为基础, 通过站点 AGPP 向区域的扩展以评估区域 AGPP, 降低 AGPP 数据的不确定性。

本数据集基于 ChinaFLUX 长期联网观测数据及公开数据获得 AGPP 观测数据集, 结合气候土壤生物栅格数据获得各生态系统的对应辅助数据, 构建得到 AGPP 原始观测数据集。同时, 基于随机森林回归树, 整合气候土壤生物因素, 并通过模拟单位叶面积的 AGPP (PAGPP) 实现站点 AGPP 向区域的扩展, 生成 2000–2020 年中国陆地生态系统 AGPP 数据集。随机森林回归树是机器学习算法的一种, 通过随机抽取样本和特征, 建立多棵相互不关联的决策树, 通过并行的方式获得预测结果。每棵决策树都能通过抽取的样本和特征得出一个预测结果, 通过综合所有树的结果取平均值, 得到整个森林的回归预测结果。本数据集涵盖最新时间跨度, 与各类产品在空间以及时间变化上的一致性较高为分析中国陆地生态系统 AGPP 时空变化、评估陆地碳收支提供数据支撑。

## 1 数据采集和处理方法

本数据集涵盖站点观测数据的收集及区域 AGPP 的评估两个过程。站点观测数据的收集涉及观测 AGPP 的收集和相关气候土壤生物要素的获取。基于观测 AGPP, 通过评估不同扩展手段、变量组合及扩展途径的差异, 扩展站点 AGPP 至区域水平, 实现 AGPP 的区域评估, 主要流程见图 1。

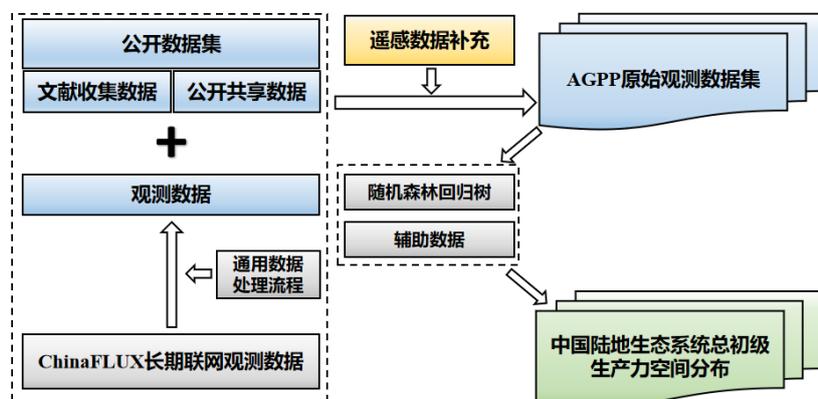


图 1 中国陆地生态系统总初级生产力时空格局生成主要流程示意图

Figure 1 Schematic diagram of the main process of generating spatio-temporal pattern of annual gross primary productivity of terrestrial ecosystems in China

## 1.1 站点观测数据收集

站点观测 AGPP 源于 ChinaFLUX 观测数据及公开数据。

自 2002 年以来, ChinaFLUX 已在 62 个生态系统陆续开展涡度相关观测, 积累了 332 个站点年的原始观测结果。基于 ChinaFLUX 原始观测数据, 利用 ChinaFLUX 通用数据处理流程对数据质量进行控制和插补<sup>[1]</sup>, 获得各站点年 ChinaFLUX 观测数据的 AGPP 值。

公开数据包括两部分: 文献数据和通量网络公开共享数据。以“涡度相关”或“eddy covariance”为关键词, 本研究自中国知网 (CNKI) 和 Web of Science 检索中国区域涡度相关观测的中英文文献。对于搜索到的条目, 本数据集逐篇筛选包含 AGPP 观测数值的文献进行提取, 具体筛选标准包括: (1) 文献报道的结果为涡度相关直接观测的 AGPP 数值, 排除部分文献使用模型模拟报道的结果; (2) 具有一年以上 (含一年) 连续观测数据; (3) 有总初级生产力年累积量或所有月份月总量的数值报道<sup>[2]</sup>。经过筛选, 本数据集逐篇提取各文献报道的站点信息 (包括经纬度、海拔、气候要素、观测年份), 并以年为单位记录各 AGPP 的观测数值。除了文献数据, 本数据集同步查询了已有通量观测研究网络如 FLUXNET、ChinaFLUX 共享的公开数据, 进一步补充观测数据集。公开数据共包括 128 个生态系统, 540 个站点年。

同时, 为了避免同一站点同一观测年份出现多个数据源的重复数据 (如 ChinaFLUX 数据共享与文献数据均有某一站点某一年的观测), 如果某一站点拥有多个数据源的数值, 本研究选用 ChinaFLUX 使用统一数据处理流程所得结果。基于文献搜索和数据整编, 共获得 166 个站点, 641 个站点年, 跨越 2002–2020 年的 AGPP 数据。

辅助数据是基于观测 AGPP 进行尺度扩展的重要工具, 但文献数据及通量观测网络公开数据提供辅助数据的一致性也存在差异, 如文献收集的 MAT 和 MAP 与站点观测值的相关系数分别为 0.95 和 0.88, 而海拔可以达到 0.99。因此, 本研究基于观测数据的地理信息提取相应气候土壤生物辅助数据。气候要素包括 6 个变量, 分别为: 年均气温 (MAT)、年总降水量 (MAP)、年总光合有效辐射 (PAR)、年总潜在蒸散 (PET)、年均饱和水汽压差 (VPD)、年均 CO<sub>2</sub> 密度 ( $\rho_{c,yr}$ )。土壤要素包括 3 个变量: 土壤年均湿度 (SM)、土壤有机碳含量 (SOC)、土壤总氮含量 (STN)。生物要素包括 2 个变量: 年均叶面积指数 (LAI) 和年最大叶面积指数 (MLAI)。

各站点年的 MAT、MAP 及饱和水汽压差 (VPD) 数据源自英国气候研究组织 (CRU) 的长时间序列结果<sup>[3]</sup>。基于 CRU 的 0.5° 空间分辨率的逐月气温、降水量、水气压数据, 结合世界气象组织 (www.worldclim.org) 的各月长期平均气温、降水量和水气压结果, 本研究利用 Delta 空间降尺度方法生成中国区域 1 km 空间分辨率的数据产品<sup>[4]</sup>。

基于气候研究组织 (Climate Research Unit, CRU) 的长时间序列逐月潜在蒸散 (PET) 数据, 利用通过样条插值法将 0.5° 空间分辨率的逐月 PET 至 30 arcsecond (约 1 km), 进而求得各年 PET。

年总光合有效辐射 (PAR) 数据采用全球陆地表面卫星数据产品 (Global Land Surface Satellite, GLASS) 的数据 (0.05°)。基于 GLASS 的逐日 PAR 数据<sup>[15-16]</sup>累加得到年 PAR 值, 并使用样条插值方法插值到 30 arcsecond (约 1 km)。

本研究利用全球陆地表面卫星 (global land surface satellite, GLASS) 产品提取 8 天尺度的吸收光合有效辐射所占比例 (FAPAR) 和逐日的光合有效辐射, 并基于 8 天尺度的 FAPAR 及逐日 PAR 计算得到该年的 APAR 数据, 并使用样条插值方法插值到 30 arcsecond (约 1 km)。

年均 CO<sub>2</sub> 密度 ( $\rho_{c,yr}$ ) 则基于大气 CO<sub>2</sub> 浓度 ( $b_c$ )、CO<sub>2</sub> 摩尔质量 (44 g mol<sup>-1</sup>) 和当前状态的气

气体摩尔体积 ( $V$ ) 计算而得到, 其中大气  $\text{CO}_2$  浓度采用 Mauna Roa 的观测值来替代, 当前状态的气体摩尔体积采用理想气体状态方程结合当地气压、MAT 计算而得到。当地气压则利用年均气温和当地海拔采用压高公式计算而得到<sup>[17]</sup>。

本研究利用空间分辨率为  $0.05^\circ$ 、时间分辨率为年的微波遥感反演的 2002–2018 年中国土壤水分降尺度数据集<sup>[18]</sup>, 表征观测年份的土壤湿度, 并使用样条插值方法插值到 30 arcsecond (约 1 km)。

基于全球土壤数据库 (Global Soil Dataset for use in Earth System Models, GSDE) 的表层 5 cm 深的土壤有机碳和土壤总氮密度及土壤容重数据, 计算得到各站点的表层土壤有机碳 ( $\text{gC m}^{-2}$ ) 和全氮含量 ( $\text{gN m}^{-2}$ )<sup>[19]</sup>。

年均叶面积指数 (LAI) 和最大叶面积指数 (MLAI) 基于空间分辨率为 1 km 的 MODIS 数据计算而获取。本研究利用经过验证的 8 天尺度 MODIS 数据<sup>[20]</sup> (<http://globalchange.bnu.edu.cn/research/lai>) 计算各像元的年均 LAI 及年内 LAI 最大值 (MLAI)。

## 1.2 区域 AGPP 评估

区域 AGPP 评估涉及站点 AGPP 向区域水平的扩展, 涉及扩展工具、扩展变量及扩展途径的选择。比较 6 种扩展工具 (多元线性回归、偏最小二乘回归、BP 神经网络、支持向量机、随机森林回归树和增强回归树) 的模拟效果, 已有研究发现随机森林回归树是最佳扩展工具<sup>[21]</sup>; 比较 4 种变量组合 (光温水、气候变量、气候+土壤变量、气候+土壤+生物变量) 的模拟效果, 已有研究发现气候+土壤+生物变量是实现站点 AGPP 扩展至区域水平的最优变量组合<sup>[21]</sup>; 比较 4 种扩展途径 (直接扩展 AGPP、扩展单位叶面积的 AGPP、扩展光能利用率、扩展单位叶面积的光能利用率) 的模拟效果, 已有研究发现扩展单位叶面积的 AGPP 是实现站点 AGPP 扩展至区域水平的最优扩展途径。

本研究将 641 个站点年随机划分为训练集 (70%) 和测试集 (30%), 分别用于训练和测试映射方案。所有的制图工具都显著地反映了观测到的 AGPP 变化 ( $p < 0.01$ ), 但它们的预测性能存在明显差异, 这在训练集和测试集中均有发现。机器学习优于经典统计方法, 其中随机森林回归树的预测性能最好。在训练集中经典统计方法仅解释了观测到的 AGPP 变化的 60%; 机器学习算法预测 AGPP 解释的百分比则增加到 78% 以上; 其中 BP 神经网络预测 AGPP 的表现最差, 占观测 AGPP 变化的 78%, 支持向量机和增强回归树预测的 AGPP 分别占观测 AGPP 变化的 96% 和 89%, 随机森林回归树在绘制 AGPP 方面表现最佳,  $R^2$  为 97%, RMSE 为  $116.18 \text{ gC m}^{-2} \text{ yr}^{-1}$ 。在测试集中经典统计方法仅解释了 71% 观测到的 AGPP 变化; 机器学习算法优于经典统计方法, 表现为解释力的增强; BP 神经网络解释了观测 AGPP 变化的 82%, 而支持向量机和增强回归树预测 AGPP 的解释率分别提高到 85% 和 79%; 随机森林回归树在绘制 AGPP 方面表现最佳,  $R^2$  为 85%, RMSE 为  $272.91 \text{ gC m}^{-2} \text{ yr}^{-1}$ 。因此, 基于站点观测数据集, 本研究利用随机森林回归树结合气候、土壤和生物数据构建站点向区域扩展的单位叶面积 AGPP (PAGPP) 模型, 结合 LAI 实现站点向区域水平的 AGPP 扩展, 并用前述 (1.1 站点观测数据收集) 区域气候土壤生物数据生成本数据集中的区域 AGPP, 完成区域 AGPP 评估。

## 2 数据样本描述

### 2.1 命名格式

本数据集包括 22 个文件, 即 1 个.xlsx 电子表格文件和 21 个.tiff 栅格文件。电子表格文件存储

站点观测的 AGPP 数据及提取的气候土壤生物辅助数据，命名为中国陆地生态系统站点观测年总初级生产力数据集.xlsx。21 个.tiff 栅格文件是 2000–2020 年逐年 AGPP 空间分布数据，命名规则为 AGPP\_20XX.tiff，其中 XX 代表不同的年份。例如：AGPP\_2000.tiff 是 2000 年中国陆地生态系统的总初级生产力空间格局数据。

## 2.2 数据样本

本数据集包括 1 个站点观测的 AGPP 数据集及 21 个区域 AGPP 数据。站点观测的 AGPP 数据集共有 166 个站点 872 个站点年观测结果，由 20 列组成，包括站点名称、生态系统类型、观测年份、AGPP 数值、气候土壤生物变量、参考文献等。各列的具体含义及实例参见表 1。

表 1 站点观测 AGPP 数据集的组成

Table 1 The composition of site observation AGPP datasets

序号	列名	单位	示例
1	站点顺序	—	1
2	站点名称	—	尖峰岭森林
3	生态系统类型	—	森林
4	纬度	°N	18.73
5	经度	°E	108.89
6	观测年份	年	2006
7	AGPP	gC m <sup>-2</sup> yr <sup>-1</sup>	1919
8	MAT	°C	20.84
9	MAP	mm yr <sup>-1</sup>	1183.66
10	PAR	MJ m <sup>-2</sup> yr <sup>-1</sup>	2842.79
11	PET	mm yr <sup>-1</sup>	1256.89
12	VPD	kPa	0.53
13	$\rho_{c, yr}$	mgCO <sub>2</sub> m <sup>-3</sup>	715.04
14	SM	m <sup>3</sup> m <sup>-3</sup>	0.56
15	SOC	gC m <sup>-2</sup>	1841.67
16	STN	gN m <sup>-2</sup>	133.38
17	LAI	m <sup>2</sup> m <sup>-2</sup>	4.04
18	MLAI	m <sup>2</sup> m <sup>-2</sup>	6.20
19	重复标志(1 为重复)	—	0
20	参考文献	—	CHEN H, LU W, YAN G, et al. Typhoons exert significant but differential impacts on net ecosystem carbon exchange of subtropical mangrove forests in China. Biogeosciences, 2014, 11(19), 5323-5333.

重复标志：为了避免同一站点同一观测年份拥有多个数据源的 AGPP 数据，参照以下标准进行筛选：1) 如果同时有 ChinaFLUX 处理数据和文献数据，使用 ChinaFLUX 处理数据作为观测 AGPP；2) 如果有多个文献数据源，使用最新文献数据作为观测 AGPP。与最终应用 AGPP 相同站点年的其

他数据源数据，标记为重复数据，用标记符号 1 来表示。数据使用中如遇其他问题，可详细参考本课题组先前发表的论文：ZHU X J, YU G R, CHEN Z, et al. Mapping Chinese annual gross primary productivity with eddy covariance measurements and machine learning[J]. *Science of the Total Environment*, 2023, 857: 159390. DOI: 10.1016/j.scitotenv.2022.159390.

21 个区域 AGPP 数据表征 2000–2020 年中国陆地生态系统 AGPP 数据，每一年为一个文件，文件采用 WGS-84 坐标系，共 4320 行 7560 列，纬度范围为 18°N~54°N，经度范围为 73°E~136°E，空间分辨率为 30 arcsecond（约 1 km）。

### 3 数据质量控制和评估

本数据集数据质量控制和评估体现在站点观测数据集和区域 AGPP 数据两个方面。

站点观测数据集涵盖 ChinaFLUX 长期观测数据和公开数据，其数据质量控制方法也有所差异。ChinaFLUX 长期观测数据是依托各台站长期观测结果使用 ChinaFLUX 通用数据处理流程进行数据质量控制与插补，而 ChinaFLUX 通用数据处理流程已经得到国内外一致认可，故 ChinaFLUX 长期观测数据具有较高的数据可信度。公开数据主要包括文献发表数据及各通量网络公开共享数据。文献发表结果已经经过同行评议，其数据质量也有较高可信度。通量网络公开共享数据则采用公认的数据处理方法对数据质量进行控制，故也有较高可信度。

区域 AGPP 数据是基于站点观测数据集利用随机森林回归树所生成，其质量控制与评估体现在模型构建与生成数据评估两个方面。基于站点观测结果，我们前期研究发现，基于随机森林回归树模拟 PAGPP 可以解释训练集 97% 的观测 AGPP 时空变异，对测试集观测 AGPP 时空变异的解释比例也达到 85%，体现出所构建的模型对观测数据具有较高的重现度。在生成数据评估方面，我们比较了本数据与已有 AGPP 数据产品（GLASS 的光能利用率结果、数据驱动的 FLUXCOM 结果及过程模型的北美碳计划结果）在空间变化、时间变化及总量上的一致性。结果发现，在空间分布方面，本数据结果与各类产品的空间一致性均较高，与各产品的回归斜率近似于 1，但与数据驱动结果（FLUXCOM）的一致性最高， $R^2$  高于 0.95；在时间变化方面，本数据结果与 GLASS 及北美碳计划结果具有较大差异，但与数据驱动结果具有较高一致性，回归斜率近似于 1， $R^2$  达到 0.6；在总量方面，本数据结果生成的 AGPP 总量约为  $7.03 \pm 0.45 \text{ PgC yr}^{-1}$ ，与数据驱动结果及 GLASS 结果具有较高一致性，也落于各模型结果的范围内，同时略高于部分文献结果，但主要与本项目所得结果涵盖最新的时间跨度且 AGPP 年际呈增加趋势有关<sup>[21]</sup>。因此，本研究中区域 AGPP 数据具有较高可信度。

### 4 数据使用方法和建议

本数据集收录了站点观测的 AGPP 数据及基于站点观测结果所生成的区域 AGPP 数据，站点观测数据可以为模型模拟提供验证数据，并为分析 AGPP 时空变化规律提供数据基础。区域 AGPP 数据可以为区域生产力及生态质量、陆地碳汇等评估提供数据支撑。但本研究中的关键变量在获取方法上均存在些许不足，因此本数据集仍存在一定的不确定性，主要包括以下三部分：

(1) 站点观测数据的气候土壤生物辅助数据可能存在一定的不确定性。本数据集中，站点辅助数据是基于站点经纬度信息自区域栅格数据中提取所得，目的是使不同站点的辅助数据具有一致的来源，但区域栅格数据本身具有一定的不确定性，故导致站点观测数据的气候土壤生物辅助数据存

在一定的不确定性。

(2) 区域 AGPP 数据适用于中国陆地生态系统 AGPP 时空格局的评估, 尽管该数据可以解释测试集 85% 的 AGPP 时空变化, 但其对于单一站点可能会存在一定误差。

(3) 区域 AGPP 数据的年际变化可能存在一定的不确定性。本数据使用所有年份观测结果生成随机森林回归树模型, 故对区域 AGPP 年际变异的重现性相对较弱, 这也在本数据与已有数据产品在时间变化一致性评估时有所体现<sup>[21]</sup>。因此, 在使用本数据分析 AGPP 年际变化时, 本数据的年际变化趋势可能与站点观测结果有一定的偏差。

数据使用中如遇其他问题, 可详细参考本研究组先前发表的论文: ZHU X J, YU G R, CHEN Z, et al. Mapping Chinese annual gross primary productivity with eddy covariance measurements and machine learning[J]. Science of the Total Environment, 2023, 857: 159390. DOI: 10.1016/j.scitotenv.2022.159390.

本数据集由国家生态科学数据存储库 (<https://ecodb.scidb.cn/>) 提供数据服务。

## 致 谢

感谢 ChinaFLUX 的长期联网观测联盟和公开的数据集发表作者以及各位为碳通量研究做出贡献的科研人员, 感谢大家的辛勤付出以及为科学进步做出的贡献。

## 数据作者分工职责

范仁雪 (1994—), 女, 在读硕士, 全球变化生态学。主要承担工作: 数据汇编, 论文撰写。

朱先进 (1985—), 男, 博士, 副教授, 全球变化生态学。主要承担工作: 数据处理, 论文撰写, 方法指导。

陈智 (1986—), 女, 博士, 副研究员, 主要研究方向为全球变化与碳循环。主要承担工作: 数据调查。

于贵瑞 (1959—), 男, 博士, 研究员, 研究方向为生态系统生态学、全球变化与碳氮水循环。主要承担工作: 总体方案设计。

张维康 (1990—), 女, 博士, 主要研究方向为全球变化与碳循环。主要承担工作: 数据收集。

韩朗 (1994—), 女, 博士, 主要研究方向为全球变化与碳循环。主要承担工作: 数据收集。

王秋凤 (1973—), 女, 博士, 研究员, 研究方向为全球变化与碳氮水循环。主要承担工作: 方法指导。

陈世莘, 刘绍民, 王辉民, 闫俊华, 谭俊磊, 张法伟, 赵风华, 李英年, 张一平, 石培礼, 朱教君, 吴家兵, 赵仲辉, 郝彦宾, 沙丽清, 张玉翠, 姜世成, 顾峰雪, 吴志祥, 张扬建, 周莉, 唐亚坤, 贾丙瑞, 李玉强, 宋清海, 董刚, 高艳红, 蒋正德, 孙聘, 王建林, 何其华, 李新虎, 王飞, 魏文学, 邓正苗, 郝翔翔, 李彦, 刘晓利, 章熙锋, 朱治林: 数据生产者, 主要承担: 各台站数据的监测和质量控制。

## 参考文献

[1] STEFFEN W, NOBLE I, et al. The terrestrial carbon cycle: implications for the Kyoto protocol[J]. Science, 1998, 280(5368): 1393–1394. DOI: 10.1126/science.280.5368.1393.

- [2] STUART CHAPIN F III, MATSON P A, MOONEY H A. Principles of Terrestrial Ecosystem Ecology[M]. New York, NY: Springer New York, 2002. DOI: 10.1007/b97397.
- [3] LI Z Q, YU G R, XIAO X M, et al. Modeling gross primary production of alpine ecosystems in the Tibetan Plateau using MODIS images and climate data[J]. Remote Sensing of Environment, 2007, 107(3): 510–519. DOI: 10.1016/j.rse.2006.10.003.
- [4] ANAV A, FRIEDLINGSTEIN P, BEER C, et al. Spatiotemporal patterns of terrestrial gross primary production: a review[J]. Reviews of Geophysics, 2015, 53(3): 785–818. DOI: 10.1002/2015rg000483.
- [5] XIAO J F, ZHUANG Q L, LAW B E, et al. A continuous measure of gross primary production for the conterminous United States derived from MODIS and AmeriFlux data[J]. Remote Sensing of Environment, 2010, 114(3): 576–591. DOI: 10.1016/j.rse.2009.10.013.
- [6] YAO Y T, LI Z J, WANG T, et al. A new estimation of China’s net ecosystem productivity based on eddy covariance measurements and a model tree ensemble approach[J]. Agricultural and Forest Meteorology, 2018, 253/254: 84–93. DOI: 10.1016/j.agrformet.2018.02.007.
- [7] YAO Y T, WANG X H, LI Y, et al. Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years[J]. Global Change Biology, 2018, 24(1): 184–196. DOI: 10.1111/gcb.13830.
- [8] GARBULSKY M F, PEÑUELAS J, PAPALED, et al. Patterns and controls of the variability of radiation use efficiency and primary productivity across terrestrial ecosystems[J]. Global Ecology and Biogeography, 2010, 19(2): 253–267. DOI: 10.1111/j.1466-8238.2009.00504.x.
- [9] GUO Q, HU Z M, LI S G, et al. Contrasting responses of gross primary productivity to precipitation events in a water-limited and a temperature-limited grassland ecosystem[J]. Agricultural and Forest Meteorology, 2015, 214/215: 169–177. DOI: 10.1016/j.agrformet.2015.08.251.
- [10] ZHU X J, YU G R, HE H L, et al. Geographical statistical assessments of carbon fluxes in terrestrial ecosystems of China: results from upscaling network observations[J]. Global and Planetary Change, 2014, 118: 52–61. DOI: 10.1016/j.gloplacha.2014.04.003.
- [11] YU G R, WEN X F, SUN X M, et al. Overview of ChinaFLUX and evaluation of its eddy covariance measurement[J]. Agricultural and Forest Meteorology, 2006, 137(3/4): 125–137. DOI: 10.1016/j.agrformet.2006.02.011.
- [12] YU G R, ZHU X J, FU Y L, et al. Spatial patterns and climate drivers of carbon fluxes in terrestrial ecosystems of China[J]. Global Change Biology, 2013, 19(3): 798–810. DOI: 10.1111/gcb.12079.
- [13] HARRIS I, OSBORN T J, JONES P, et al. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset[J]. Scientific Data, 2020, 7(1): 109. DOI: 10.1038/s41597-020-0453-3.
- [14] PENG S Z, DING Y X, LIU W Z, et al. 1 km monthly temperature and precipitation dataset for China from 1901 to 2017[J]. Earth System Science Data, 2019, 11(4): 1931–1946. DOI: 10.5194/essd-11-1931-2019.
- [15] CHENG J, LIANG S L. Estimating the broadband longwave emissivity of global bare soil from the MODIS shortwave albedo product[J]. Journal of Geophysical Research: Atmospheres, 2014, 119(2): 614–634. DOI: 10.1002/2013jd020689.

- [16] CHENG J, LIANG S L, VERHOEF W, et al. Estimating the hemispherical broadband longwave emissivity of global vegetated surfaces using a radiative transfer model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(2): 905–917. DOI: 10.1109/TGRS.2015.2469535.
- [17] ZHU X J, YU G R, WANG Q F, et al. Approaches of climate factors affecting the spatial variation of annual gross primary productivity among terrestrial ecosystems in China[J]. *Ecological Indicators*, 2016, 62: 174–181. DOI: 10.1016/j.ecolind.2015.11.028.
- [18] MENG X J, MAO K B, MENG F, et al. A fine-resolution soil moisture dataset for China in 2002–2018[J]. *Earth System Science Data*, 2021, 13(7): 3239–3261. DOI: 10.5194/essd-13-3239-2021.
- [19] SHANGGUAN W, DAI Y J, DUAN Q Y, et al. A global soil data set for earth system modeling[J]. *Journal of Advances in Modeling Earth Systems*, 2014, 6(1): 249–263. DOI: 10.1002/2013ms000293.
- [20] YUAN H, DAI Y J, XIAO Z Q, et al. Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling[J]. *Remote Sensing of Environment*, 2011, 115(5): 1171–1187. DOI: 10.1016/j.rse.2011.01.001.
- [21] ZHU X J, YU G R, CHEN Z, et al. Mapping Chinese annual gross primary productivity with eddy covariance measurements and machine learning[J]. *Science of the Total Environment*, 2023, 857: 159390. DOI: 10.1016/j.scitotenv.2022.159390.

## 论文引用格式

范仁雪, 朱先进, 陈智, 等. 2000–2020 年中国陆地生态系统年总初级生产力数据集[J/OL]. *中国科学数据*, 2023, 8(2). (2023-05-12). DOI: 10.11922/11-6035.csd.2023.0037.zh.

## 数据引用格式

范仁雪, 朱先进, 陈智, 等. 2000–2020 年中国陆地生态系统年总初级生产力数据集[DS/OL]. *Science Data Bank*, 2023. (2023-02-22). DOI: 10.57760/sciencedb.o00119.00077.

## A dataset of annual gross primary productivity in China's terrestrial ecosystems during 2000–2020

FAN Renxue<sup>1</sup>, ZHU Xianjin<sup>1,29\*</sup>, CHEN Zhi<sup>2,3\*</sup>, YU Gurui<sup>2,3</sup>, ZHANG Weikang<sup>2</sup>, HAN Lang<sup>4</sup>,  
WANG Qiufeng<sup>2,3</sup>, CHEN Shiping<sup>5</sup>, LIU Shaomin<sup>6</sup>, WANG Huimin<sup>2</sup>, YAN Junhua<sup>7</sup>, TAN Junlei<sup>8</sup>,  
ZHANG Fawei<sup>9</sup>, ZHAO Fenghua<sup>2</sup>, LI Yingnian<sup>9</sup>, ZHANG Yiping<sup>10</sup>, SHI Peili<sup>2</sup>, ZHU Jiaojun<sup>11</sup>, WU  
Jiabing<sup>11</sup>, ZHAO Zhonghui<sup>12</sup>, HAO Yanbin<sup>13</sup>, SHA Liqing<sup>10</sup>, ZHANG Yucui<sup>14</sup>, JIANG Shicheng<sup>15</sup>,  
GU Fengxue<sup>16</sup>, WU Zhixiang<sup>17</sup>, ZHANG Yangjian<sup>2,3</sup>, ZHOU Li<sup>18</sup>, TANG Yakun<sup>19</sup>, JIA Bingrui<sup>5</sup>,  
LI Yuqiang<sup>8</sup>, SONG Qinghai<sup>10</sup>, DONG Gang<sup>20</sup>, GAO Yanhong<sup>8</sup>, JIANG Zhengde<sup>11</sup>, SUN Dan<sup>7</sup>,  
WANG Jianlin<sup>21</sup>, HE Qihua<sup>22</sup>, LI Xinhui<sup>23</sup>, WANG Fei<sup>24</sup>, WEI Wenxue<sup>25</sup>, DENG Zhengmiao<sup>25</sup>,  
HAO Xiangxiang<sup>26</sup>, LI Yan<sup>23</sup>, LIU Xiaoli<sup>27</sup>, ZHANG Xifeng<sup>28</sup>, ZHU Zhilin<sup>2,3</sup>

1. College of Agronomy, Shenyang Agricultural University, Shenyang 110866, P. R. China

2. Synthesis Research Center of Chinese Ecosystem Research Network, Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, P. R. China
3. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, P. R. China
4. Institute of Surface-Earth System Science, School of Earth System Science, Tianjin University, Tianjin 300072, P. R. China
5. State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, P. R. China
6. State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, P. R. China
7. South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, P. R. China
8. Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, P. R. China
9. Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, P. R. China
10. Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, P. R. China
11. Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, P. R. China
12. Central South University of Forestry and Technology, Changsha 410004, P. R. China
13. University of the Chinese Academy of Sciences, College of Life Sciences, Beijing 100049, P. R. China
14. Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang 050021, P. R. China
15. Northeast Normal University, Changchun 130024, P. R. China
16. Institute of Environmental and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences, Beijing 100081, P. R. China
17. Rubber Research Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 570100, P. R. China
18. Chinese Academy of Meteorological Sciences, China Meteorological Administration, Beijing 100081, P. R. China
19. Northwest A&F University, Yangling 712100, P. R. China
20. Shanxi University, Taiyuan 030006, P. R. China
21. Qingdao Agricultural University, Qingdao 266109, P. R. China
22. Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, P. R. China
23. Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, P. R. China
24. Inner Mongolia Agricultural University, Hohhot 010018, P. R. China
25. Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, P. R. China
26. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, P. R. China
27. Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, P. R. China

28. Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, P. R. China

29. Liaoning Panjin Wetland Ecosystem National Observation and Research Station, Panjin 124000, P. R. China

\* Email: xianjin1985@163.com (ZHU Xianjin); chenz@igsnr.ac.cn (CHEN Zhi)

**Abstract:** The annual gross primary productivity (AGPP) is the basis of food production and carbon sequestration in terrestrial ecosystems. An accurate assessment of regional AGPP can provide a theoretical basis for analyzing the spatiotemporal variation of AGPP and ensuring regional food security and mitigating climate change trends. Based on Chinese Flux Observation and Research Network (ChinaFLUX) measurements and public datasets, we produced a dataset of annual gross primary productivity over China's terrestrial ecosystems was constructed. In combination with biological, climatic, and soil factors, we used the random forest regression tree to construct the assessment model of China AGPP by simulating the AGPP of unit leaf area. The dataset of annual gross primary productivity over China's terrestrial ecosystems during 2000-2020 was generated with a spatial resolution of 30arcsecond and a data format of tiff. The dataset can provide validation data for model simulation, as well as data support for regional productivity, ecological quality, and assessment and management of terrestrial carbon sinks.

**Keywords:** China; annual gross primary productivity of terrestrial ecosystem; random forest regression tree model; spatiotemporal variation

### Dataset Profile

<b>Title</b>	A dataset of annual gross primary productivity over China's terrestrial ecosystems during 2000-2020
<b>Data corresponding author</b>	ZHU Xianjin(xianjin1985@163.com), CHEN Zhi(chenz@igsnr.ac.cn)
<b>Data author(s)</b>	FAN Renxue, ZHU Xianjin, CHEN Zhi, YU Gurui, ZHANG Weikang, HAN Lang, WANG Qiufeng, CHEN Shiping, LIU Shaomin, WANG Huimin, YAN Junhua, TAN Junlei, ZHANG Fawei, ZHAO Fenghua, LI Yingnian, ZHANG Yiping, SHI Peili, ZHU Jiaojun, WU Jiabing, ZHAO Zhonghui, HAO Yanbin, SHA Liqing, ZHANG Yucui, JIANG Shicheng, GU Fengxue, WU Zhixiang, ZHANG Yangjian, ZHOU Li, TANG Yakun, JIA Bingrui, LI Yuqiang, SONG Qinghai, Dong Gang, GAO Yanhong, JIANG Zhengde, SUN Dan, WANG Jianlin, HE Qihua, LI Xinqu, WANG Fei, WEI Wenxue, DENG Zhengmiao, HAO Xiangxiang, LI Yan, LIU Xiaoli, ZHANG Xifeng, ZHU Zhilin
<b>Time range</b>	2000-2020
<b>Geographical scope</b>	China
<b>Spatial resolution</b>	30 arcsecond
<b>Data volume</b>	4.30 GB
<b>Data format</b>	.xlsx, *.tiff
<b>Data service system</b>	<a href="https://doi.org/10.57760/sciencedb.o00119.00077">https://doi.org/10.57760/sciencedb.o00119.00077</a>
<b>Source(s) of funding</b>	National Natural Science Foundation of China (32071585, 32071586)

<b>Dataset composition</b>	The dataset consists of 22 subsets in total, including site observation AGPP data and regional AGPP data. The site observation AGPP is the observation result of the site, including 872 rows and 20 columns, 166 sites, 872 site years, 5 ecosystem types, etc. (See Table 1 for the details). The regional AGPP contains 21 data files, which is the spatial distribution result of AGPP year by year.
----------------------------	--