



地球系统数值模拟装置科学数据 共享服务平台用户培训

汇报人：肖霞

2023年6月7日





提纲

1

简介

2

科学数据共享门户

3

公共数据池



1

简介



一、简介



数据直接调用

成果共享/推广



- 想要实时预报数据为为短期气候预测、天气预报和空气质量预报提供实时的初始场和驱动场 ✓
- 想要历史预报数据做算法训练 ✓
- 想要长期的、不同种类的分析/再分析数据进行分析对比、为数值模拟提供背景场、评估模式结果 ✓
- 想要直接使用前沿的科学数据成果 ✓
- 想要定制数据能直接使用 ✓



- × 数据获取不及时
- × 数据存储空间不够
- × 数据下载网络不畅
- × 数据倒来倒去耗费时间、增加存储
- × 原始数据不是想要的的数据，还要花大量时间进行再处理



- ✓ 同步官网获取数据
- ✓ PB级空间支持数据存储和处理
- ✓ 专线国际带宽保障数据下载
- ✓ 装置内公共数据和平台自主知识产权数据直接调用
- ✓ 专业数据团队提供数据预处理和后处理服务



一、简介

公共数据池

将研究人员在科研过程中使用频率高的模式数据、分析/再分析数据、同化数据、遥感数据等国内国际数据通过不同的渠道获得并**建立公共数据池**，由装置统一管理（整理、更新和分发），用于装置内模式运行、研究和评估过程中所需数据的支持服务，同时向所内科研人员开放分发

- 🔗 **用户人群**：有大装置账户的用户、所内科研人员
- 🔗 **使用权限**：大装置系统、所内IP
- 🔗 **访问方式（所内）**：<http://sharedata.iap.ac.cn>
- 🔗 **调用方式（装置内）**：</sharedata/dataset>
- 🔗 **数据详情**：访问网址或路径，获取数据详单

- ✔ **用户人群**：全网
- ✔ **数据类别**：科学数据实体、科学数据描述信息、科学数据辅助工具软件
- ✔ **用户分类**：数据汇交用户，数据使用用户
- ✔ **使用权限**：全网、所内、装置内、用户组
- ✔ **共享方式**：元数据共享、实体数据共享、保护期共享
- ✔ **访问方式**：<http://earthlab-data.iap.ac.cn/>

集数据**汇交、发布、共享、推广、下载和管理**于一体的综合性门户平台，主要用于**具有自有知识产权成果数据的全网发布**

科学数据共享门户

2

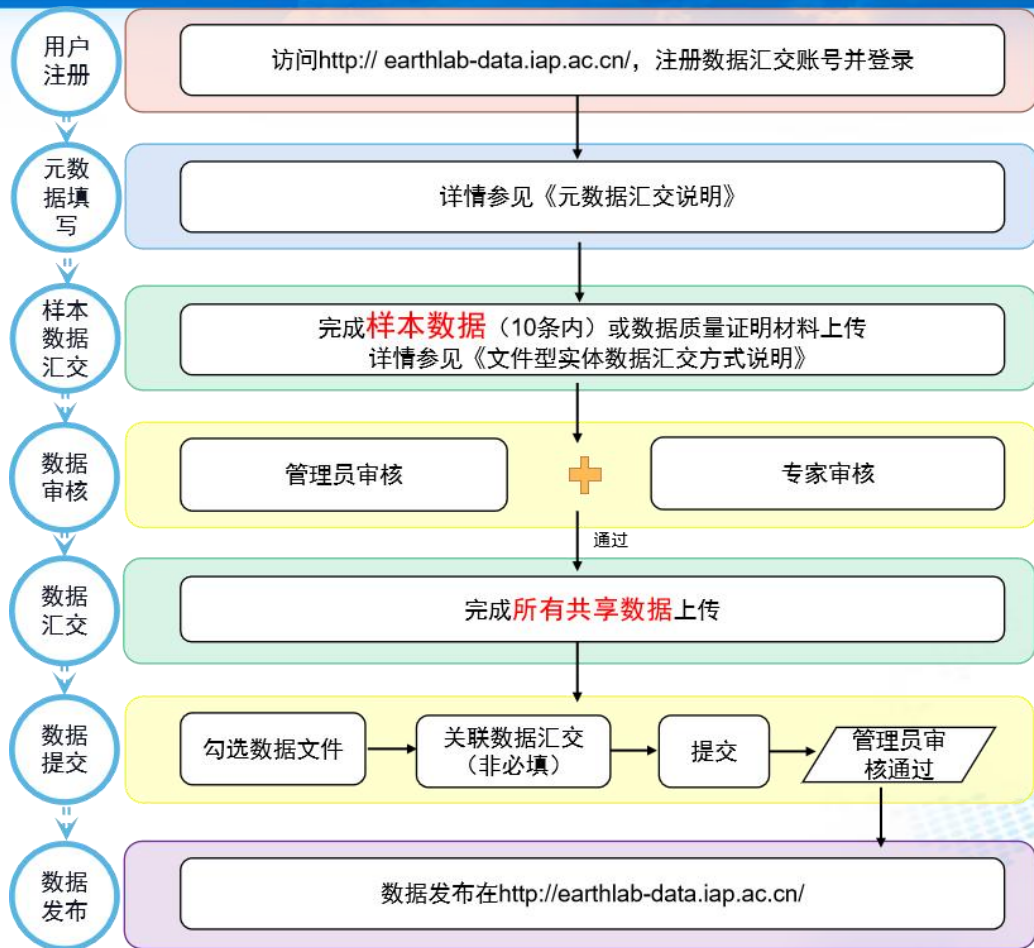
科学数据共享门户





二、科学数据共享门户——数据汇交

汇交流程





二、科学数据共享门户——数据汇交



注册及登录方式





二、科学数据共享门户——数据汇交

*选择角色: 汇交人员

*用户账号: 请输入用户账号, 区分大小写, 长度在6-30位

注册密码: |(8-20位, 字母(至少两位)、数字和特殊字符!@#\$\$%^&()?)等

密码确认: (8-20位, 字母(至少两位)、数字和特殊字符!@#\$\$%^&()?)等

*真实姓名: 请输入真实姓名

*性别: 男

*联系电话: 请输入联系电话, 如是座机为: 区号 (-) 电话号码 格式

*电子邮件: 请输入电子邮件

*从事职业: 科研院所研究人员

*单位名称: 请输入单位名称

*申请理由: 请输入申请理由

*验证码: 请输入验证码  换一张

我已阅读并承诺《地球系统数值模拟大科学装置数据中心服务条款》

[用户注册](#)

地球系统数值模拟装置

 请输入用户账号

 请输入密码

 请输入验证码 E u vq 换一张

[用户注册](#) [忘记密码?](#)

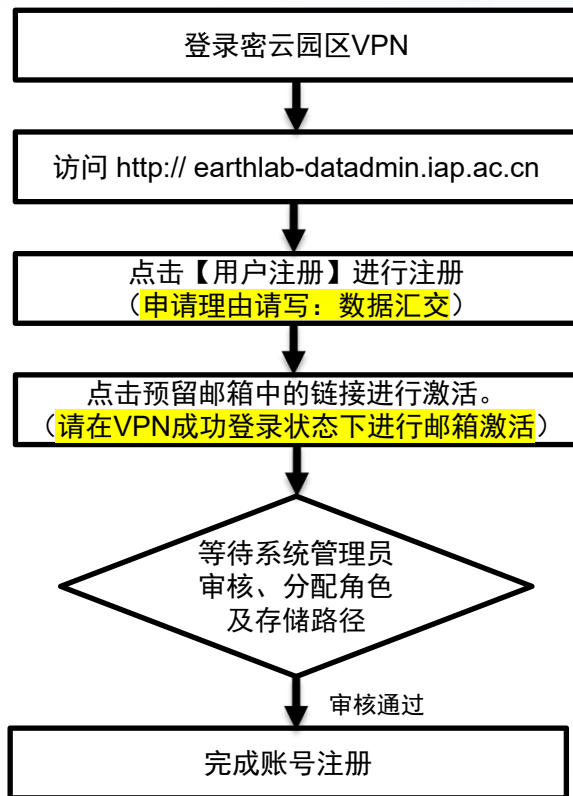
[登录](#)





二、科学数据共享门户——数据汇交

P.S





二、科学数据共享门户——数据汇交

地球模拟海量数据管理

首页



汇交录入员

肖霞 退出

关系型实体数据管理

文件型实体数据管理

数据汇交管理

数据汇交统计

订单管理

数据共享查看

欢迎您！汇交录入员

作为整个汇交系统的数据录入用户，您将负责对被分配到体系级别（如课题等）进行数据的汇交与管理，将数据组织成数据集形式进行提交发布。

您可以通过以下方式进行不同类型数据汇交：

1. 关系型实体数据管理

通过“创建新表”进行关系数据表的数据汇交。具体提供3种方式：

- ①对于本地存放的结构完整的Excel文件，可以直接“导入式建表”，数据内容将汇交为关系库表。
- ②“在线建表”方便用户直接从零开始构建一张新的关系库表。
- ③“关联创建新表”方便高级用户直接通过SQL语句或根据现有库表内容关联构建新的数据表。“数据字段配置”用于配置数据字段在门户展示的字段名称及类型，如可配置为文件型、枚举型等。“数据记录管理”用于对数据的增删改查。

2. 文件型实体数据管理

该功能提供文件的多种上传方式，如本地上传、FTP上传、文件夹构建与管理。

3. 数据汇交管理

进行数据的组织和管理，包括数据集、论文等不同资源类型数据的汇交。通过填写元数据、选择相应的实体文件或库表，构建关联关系，完成一条数据集的汇交。为方便快捷汇交构建数据集等资源，此模块还提供了批量元数据的导入与导出。

4. 数据汇交统计

展现本人已汇交数据的统计，已汇交发布数据的访问统计。

5. 共享数据查看

如果体系构建时，允许主节点下数据共享，则此模块可以查看到同级的其他汇交人员已汇交成功的数据。

详细使用说明可以参考用户手册。





二、科学数据共享门户——数据汇交

地球模拟海量数据管理

文件实体汇交管理



汇交录入员

肖霞 退出

关系型实体数据管理

文件型实体数据管理

数据汇交管理

数据汇交统计

订单管理

数据共享查看

FTP上传 本地上传 新建文件夹

返回上级 | file

- test
- dataFile
- HCEO
- GFS

选择文件并上传

一次最多能上传100条

拖拽文件到这里 ...
支持多文件同时上传

选择 ...

退出

修改日期

| |
|---------------------|
| 2023-03-06 14:23:28 |
| 2023-06-06 21:33:57 |
| 2023-01-12 11:34:26 |
| 2022-06-27 10:14:44 |

实体数据上传





二、科学数据共享门户——数据汇交

地球模拟海量数据管理

文件实体汇交管理



汇交录入员

肖霞 退出

关系型实体数据管理

文件型实体数据管理

数据汇交管理

数据汇交统计

订单管理

数据共享查看

FTP上传

本地上传

新建文件夹

基于文件名搜索

检索

返回上级 | file

| 文件名 | 大小 | 修改日期 |
|------|----|---------------------|
| test | - | 2023-03-06 14:23:28 |
| | | 1:33:57 |
| | | 1:34:26 |
| | | 0:14:44 |

FTP上传

FTP服务器地址: ftp://10.64.202.10

FTP账号:

FTP密码:

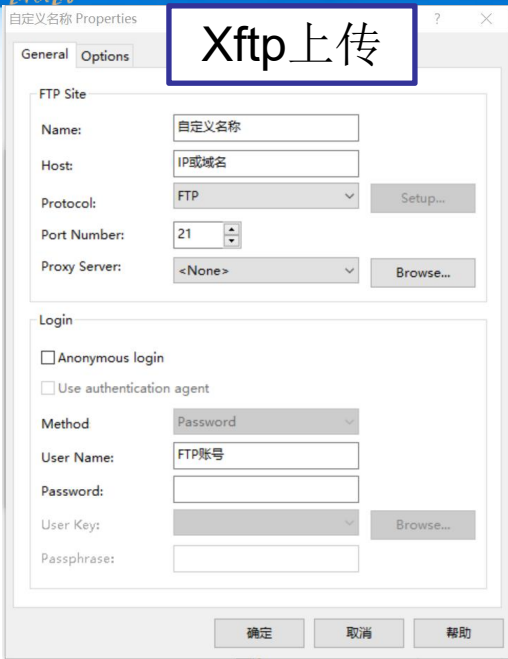
实际路径: /home/vftpuser/dataExchange/dataAuditT/nodeThreeT/nodeThreeT/file/

取消





二、科学数据共享门户——数据汇交





二、科学数据共享门户——数据汇交

```
[root@wings disk2]# ftp ftpUserRemmitter-XX@10.64.202.10
Password:
ftp ftpUserRemmitter-XX@10.64.202.10:~> ls
drwxrwxrwx  4 1000    0          4096 Jul 15 11:30 file
ftp ftpUserRemmitter-XX@10.64.202.10:~/> cd file
ftp ftpUserRemmitter-XX@10.64.202.10:/file> ls
drwxr-x---   3 0      0          4096 Jun 27 02:14 GFS
drwxr-xr-x   9 1007   0          4096 Jul 15 09:19 dataFile
ftp ftpUserRemmitter-XX@10.64.202.10:/file> mirror -R 本地目录名
```

(1) 所内网传输（北郊）

`lftp ftp-username@ FTP 服务器 IP`

Password:

(2) 外网传输

`lftp ftp-username@earthlab-dataftp2.iap.ac.cn`

Password:

(3) 装置内传输

为了不让其他人通过ps命令看到你的密码，应在 .netrc 文件中配置 FTP 服务器账号登录信息，避免在命令行上键入密码：

① vim ~/.netrc

添加文件内容如下（x.x.x.x 是 FTP 服务器 IP）：

machine x.x.x.x login ftpuser password ftpuser_password

② chmod 0600 ~/.netrc

之后再执行 lftp 命令时，只提供登录名即可，lftp 命令会自动去 ~/.netrc 去找密码，实现自动登录。

`lftp ftp-username@ FTP 服务器 IP`

数据在服务器



二、科学数据共享门户——数据汇交

元数据信息填写

- 根据提示，按要求完成元数据信息填写

基本的元数据信息包括：数据集中文名、数据集英文名、数据简介、封面图片、关键词、数据版本、数据变量、时空范围、时空分辨率、数据来源、数据格式、数据类别、标识码

创建和发布信息包括：研发团队（中英文）、研发单位（中英文）、联系人、联系电话及电子邮箱

The screenshot displays the 'Basic Information' (基本信息) section of a data submission form. It includes a world map showing dust emission and deposition rates. The form fields are as follows:

- 数据集名称(中文):** 全球沙尘气溶胶循环数据集
- 数据集名称(英文):** Global dust cycle simulated by CAS-ESM2.0
- 数据简介:** CAS-ESM2.0在观测和气候模式下的沙尘模拟结果, 包括1960年10月至2010年10月10日平均和12年平均, 介绍相关格式和数据的文章已发表于《Journal》
- DOI标识符:** 10.12392/IAP-Earthlab.2022.62a824c4001de1841a7a57a
- CSTR标识符:** 36001.IAP-Earthlab.2022.62a824c4001de1841a7a57a
- PID标识符:** 21.86101.3.IAP-Earthlab.2022.62a824c4001de1841a7a57a
- 关键词:** Dust emission, Dust deposition, Dust concentration

Below the form, there is a 'Usage' (引用方式) section with detailed citation instructions in Chinese and English, and a 'Creation and Release Information' (创建和发布信息) section with fields for research team, unit, contact person, phone, and email.

服务热线

010-6488-9273

IP地址: 10.10.10.10

联系我们: 010-6488-9273

地址: 北京市海淀区中关村大街

邮编: 100080

技术支持: 010-6488-9273



二、科学数据共享门户——数据汇交

- 1、**元数据字段的定制和扩展：**可在基本元数据字段的基础上，根据数据集特征，扩展元数据字段，使数据描述更加清晰
- 2、**免费为具有自主知识产权的数据或数据集提供唯一标识码，**包括doi, cstr, pid

本平台的标识码格式：

DOI: 10.12392/IAP-Earthlab.年份+序号

CSTR: 36001.IAP-Earthlab.年份+序号

PID: 21.86101.3/IAP-Earthlab.年份+序号

DOI标识符: 10.12392/IAP-Earthlab.2022.62c824c400f1de1841a7a57a

CSTR标识符: 36001.IAP-Earthlab.2022.62c824c400f1de1841a7a57a

PID标识符: 21.86101.3/IAP-Earthlab.2022.62c824c400f1de1841a7a57a



二、科学数据共享门户——数据汇交

3、可根据需求**自主选择共享方式**，可以对**实体数据进行共享**，也可以**只共享元数据信息**，还可以**自主设定数据保护期**，选择实体数据共享的时间，进行数据推广

全网共享：所有用户均可对数据进行查看和下载

所内共享：仅所内用户可对数据进行查看和下载。所内用户为具有所内邮箱的用户。

装置内共享：装置用户可对数据信息进行查询和调用。装置内用户为具有装置账号的用户，可以查询、下载和调用所有数据集，装置账号失效则数据中心用户权限降为普通用户。

用户组共享：汇交人员可选择有数据查看和下载权限的用户组进行数据集共享。如有指定的用户组，可以由项目组向管理员线下申请新建用户组，然后在数据汇交时进行选择，详情见《大装置数据中心用户组申请表》。

数据申请：可在数据详情页进行数据集权限申请，汇交人员可在已有用户权限的基础上添加用户及用户组。



二、科学数据共享门户——数据汇交

4. 自动创建数据引用信息

数据引用格式：研发团队主要人员（前三位）+数据集名称+版本号
或版本时间+研发单位+发布平台（earthlab）+doi+cstr+pid

数据引用(中文): 气溶胶和大气化学模式团队,全球沙尘气溶胶循环数据集,V2020,中国科学院大气物理研究所,earthlab,10.12392/IAP-Earthlab.2022.62c824c400f1de1841a7a57a,36001.IAP-Earthlab.2022.62c824c400f1de1841a7a57a,21.86101.3/IAP-Earthlab.2022.62c824c400f1de1841a7a57a

数据引用(英文): Team for Aerosol and Atmospheric Chemistry Model,Global dust cycle simulated by CAS-ESM2.0,V2020,Institute of Atmospheric Physics, Chinese Academy of Sciences,earthlab,10.12392/IAP-Earthlab.2022.62c824c400f1de1841a7a57a,36001.IAP-Earthlab.2022.62c824c400f1de1841a7a57a,21.86101.3/IAP-Earthlab.2022.62c824c400f1de1841a7a57a

文献引用(中文): 暂无

文献引用(英文): Wu, C., Lin, Z., Liu, X., Ji, D., Zhang, H., Li, C., & Lin, G. (2021). Description of dust emission parameterization in CAS-ESM2 and its simulation of global dust cycle and East Asian dust events. *Journal of Advances in Modeling Earth Systems*, 13, e2020MS002456. <https://doi.org/10.1029/2020MS002456>



二、科学数据共享门户——数据汇交

5. 自动生成数据批量下载脚本

数据文件 | 附件 | 变更说明

返回上级 | 全球沙尘气溶胶循环数据集 > mont

Python脚本下载

| 文件名称 | 大小 | 时间 | 操作 |
|---|-----------|---------------------|----|
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-08.nc | | | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-07.nc | | | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-06.nc | 592.76 MB | 2022-07-05 22:03:56 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-05.nc | 592.76 MB | 2022-07-05 22:03:07 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-04.nc | 592.76 MB | 2022-07-05 21:58:29 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-03.nc | 592.76 MB | 2022-07-05 21:57:26 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-02.nc | 592.76 MB | 2022-07-05 21:53:11 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2010-01.nc | 592.76 MB | 2022-07-05 21:52:10 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2009-12.nc | 592.76 MB | 2022-07-05 21:47:10 | 下载 |
| <input checked="" type="checkbox"/> AMIPC5X_dust_tune_landfrac.cam2.h0.2009-11.nc | 592.76 MB | 2022-07-05 21:46:23 | 下载 |

新建下载任务

网址: blob:http://earthlab-data.iap.ac.cn/ffa6d6k

名称: download20230113.py 13.98 KB

下载到: 可视化分系统双周报 剩: 704.00 GB 浏览

[使用迅雷下载](#) [直接打开](#) [下载](#) [取消](#)



二、科学数据共享门户——数据汇交

6. 提供API服务

| 用户评价 | API | 点击查看api使用详情 | |
|-----------|---|-----------------------------|------|
| API | URL | 格式 | HTTP |
| 获取总门户统计信息 | http://earthlab-data.iap.ac.cn:80/api/TotalPortalStatisticsInf | JSON | GET |
| 获取元数据 | http://earthlab-data.iap.ac.cn:80/api/sdoMetadata/resourceId | JSON | GET |
| 文件批量下载 | http://earthlab-data.iap.ac.cn:80/api/downloadFilesScript?resourceId=resourceId&filePath=filePath&listName=fileName1,fileName2 | 文件 | GET |



二、科学数据共享门户——数据汇交

7. **提供数据统计服务**：包括数据汇交统计、数据访问及下载统计，协助掌握数据使用情况



| 序号 | 名称 | 汇交存储量 | 汇交文件数 |
|----|-------------------------|----------|-------|
| 1 | 美国国家环境预报中心全球预报系统数据... | 45.22 TB | 91227 |
| 2 | 淮南气候环境综合观测试验站RPG-HAT... | 0.11 MB | 2 |
| 3 | 淮南气候环境综合观测试验站RPG-HAT... | 0.11 MB | 2 |
| 4 | 淮南气候环境综合观测试验站RPG-HAT... | 0.11 MB | 2 |
| 5 | 淮南气候环境综合观测试验站HMB-KST... | 0.03 MB | 1 |
| 6 | 淮南气候环境综合观测试验站HMB-KST... | 0.03 MB | 1 |
| 7 | 淮南气候环境综合观测试验站HMB-KST... | 0.03 MB | 1 |





二、科学数据共享门户——数据汇交

8. 提供定制化服务

可根据汇交用户需求：

- 1) 建立定制化数据分类体系
- 2) 定制化元数据信息发布
- 3) 申请用户组数据共享



二、科学数据共享门户——数据汇交

实测数据

分析/再分析数据
野外观测数据
等

模拟预报数据

大气/海洋/陆面/大气
化学/耦合模式等各种
模拟输出结果

卫星与遥感数据

联合体和遥感所数据
以及其他卫星产品数
据

科技项目数据

重大科技项目成果数
据

排序方式: 名称 访问量 下载量 更新时间

过滤条件

数据分类

- 实测数据
- 卫星遥感数据
- 模拟预报
- 分析/再分析数据
- 国际常用数据

- 野外观测数据
- 大气理化过程观测数据



淮南气候环境综合观测试验站RPG-HATPRO-G3型微波辐射计观测数据

RPG-HATPRO-G3型微波辐射计采用42通道14接收机并行和波导捷变频技术, 具有三种扫描方式, 天顶模式(固定垂直向上对天顶观测)、边界层多仰角扫描模式以及全天空2D扫描, 保证在整个对流层和边界层获取大气不同高度、不同方位的准确和高分辨率观测数据。该微波辐射计能在高时空分辨率下测定不同的大气参数, 是一款被动遥测垂直上空大气温度、湿度和液态水廓线数...

数据集 2023-01-12 01:19:25

访问量: 1 下载量: 0



淮南气候环境综合观测试验站

观测试验站HMB-KST型固态毫米波云雷达观测数据

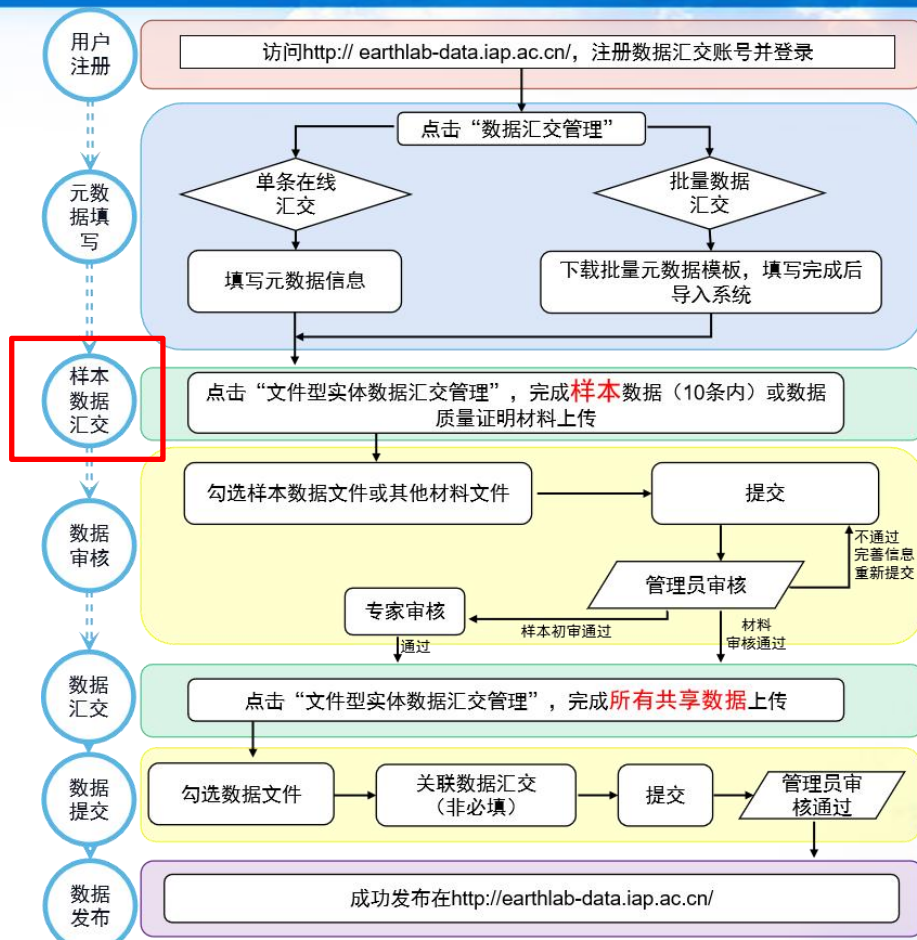
工作于Ka波段, 工作频率为33.44GHz±4Hz, 采用全固态体制, 利用云滴/雨滴对电磁波的散射作用和退极化双效应, 用于非降水云、弱降水云和弱降水的观测

数据集 2023-01-12 01:19:28

访问量: 1 下载量: 0



二、科学数据共享门户——数据汇交





二、科学数据共享门户——数据使用





二、科学数据共享门户——数据使用

(1) 不同用户权限设置

| 用户名称 | 用户权限 | 申请条件 |
|-------|-------------------------------------|----------|
| 普通用户 | 所有数据集元数据信息浏览，全网共享数据集实体数据查看及下载 | 无 |
| 所内用户 | 所有数据集元数据信息浏览，全网共享及所内共享数据集实体数据的查看及下载 | 所内邮箱 |
| 装置内用户 | 所有数据集元数据和实体数据的查看、下载及调用 | 有效的大装置账号 |



二、科学数据共享门户——数据使用

(2) 数据信息访问和查询

在首页可以看到最近更新、热门访问的数据集，查看数据统计信息、服务与公告，可访问所有数据集列表，可按关键词、标签、数据分类搜索筛选数据集；可查看数据集详情。

(3) 多种下载方式

- ✦ 直接下载；
- ✦ 提供数据批量下载脚本，勾选所需的文件或文件夹，点击“python脚本下载”，生成批量下载脚本，将脚本上传至服务器运行脚本完成数据下载；
- ✦ 同时也可根据API信息，自行编写脚本下载。



二、科学数据共享门户——数据使用

数据共享门户平台已汇交**自主研发**的数据集**10个**，包括LICOM海洋模式试验数据集、FGOALS模式试验数据集、全球沙尘气溶胶循环数据集、中国高分辨率大气污染再分析数据集等，数据量达**332TB**，数据记录数超过**100万条**。

<http://earthlab-data.iap.ac.cn/>



2

公共数据池





三、公共数据池

公共数据池中共享的公共数据指的是科研过程中使用频率高的预报数据、分析/再分析数据、遥感数据等国内国际下载获取的数据。公共数据池面向所内（所内IP直接访问和下载）和装置内（大装置用户直接调用）提供公共数据服务，目前已同步官网更新以下8套常用预报数据集和再分析数据集、3套收费的试用数据集：

- (1) **GFS** (00, 06, 12, 18时)
- (2) **GDAS** (00, 06, 12, 18时)
- (3) **CFS** (00, 12时)
- (4) **EC预报数据** (0.4° × 0.4°)
- (5) **ERA5再分析** (全球/全高度层/全变量)
(1940年-今, 历史数据逐步补齐中)
- (6) **ERA5-Land** (1950年-今)
- (7) **FNL** (1999年-今)
- (8) **JRA-55** (1958年-今)
- (9) **EC预报数据** (0.125° × 0.125°)
- (10) **AWS数据** (中国气象局2170个地面基本气象观测站数据)
- (11) **Sounding数据** (中国气象局87个探空站数据, 同步WMO的GTS国际交换数据)

公开近2周数据
公开近3天数据

以及不定期更新其余31套国际常用数据。数据量达PB级

➤ 所内: <http://sharedata.iap.ac.cn>

➤ 装置: </sharedata/dataset>

数据详情请访问上述地址中的readme

```

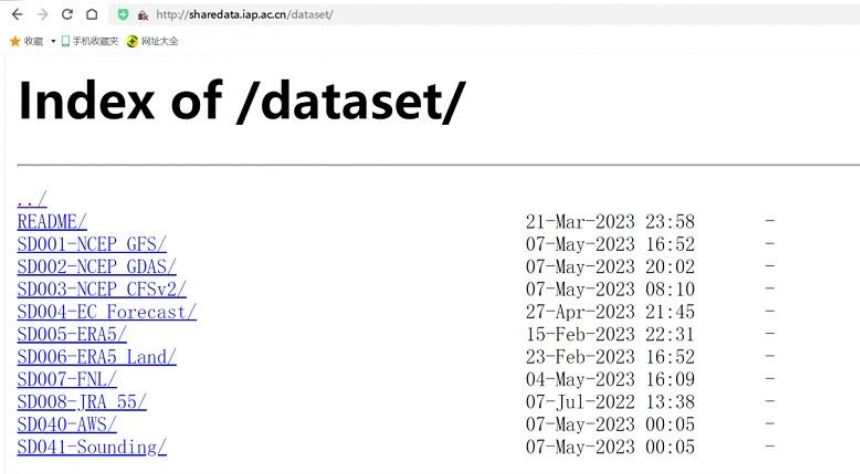
/sharedata/dataset/historical
[user-xx@server10 historical]$ ls
SD009-NCEP_Reanalysis      SD020-GPCP          SD031-0AFLUX
SD010-ERA_interim          SD021-GPCP          SD032-NOCS2
SD011-20thC_ReanV2c        SD022-CPC_PREC      SD033-NSIDC_MEASURES_Freeze_Thaw
SD012-HYCOM_NCODA_GLBa0.08 SD023-CPC_TEMP      SD034-NOAA_OISSTv2
SD013-NCEP_GODAS           SD024-HadIOD         SD035-SODA
SD014-SD014                SD025-HadISST       SD036-WOA
SD015-CUR_ts               SD026-HadSST3.1     SD037-GTSPP
SD016-GHCN_CAMS            SD027-HadSST4.6     SD038-OSCAR
SD017-PREC                  SD028-HadCRUT4.6    SD039-Surf0cean_CO2_Atlas
SD018-TRMM_3B42            SD029-EN4
SD019-CMAP                  SD030-NOAA_ERSST

```

```

/sharedata/dataset/realtime
[user-xx@server10 realtime]$ ls
README      SD004-EC_Forecast  SD008-JRA_55
SD001-NCEP_GFS  SD005-ERA5         SD040-AWS
SD002-NCEP_GDAS SD006-ERA5_Land    SD041-Sounding
SD003-NCEP_CFSv2 SD007-FNL

```





问题受理联系方式（包括新增数据需求、数据使用方面的问题或故障、成果发布等），请发送邮件至：earthlab-data@mail.iap.ac.cn



欢迎大家通过共享平台
进行科学数据成果推广

谢谢大家

