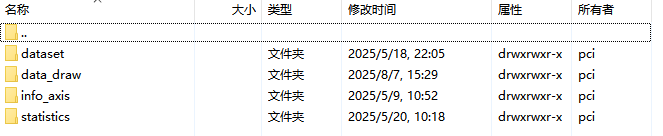
项目结构：



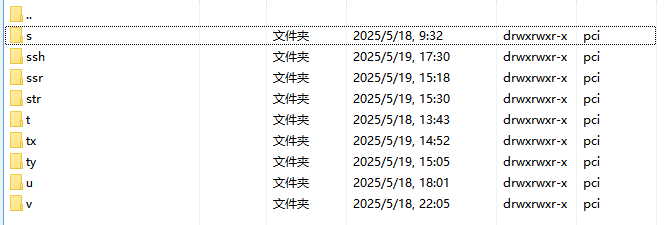
详细说明：

1. data文件夹中存放 数据集加载所需的全部数据，包含训练和测试数据，其文件夹结构如下：

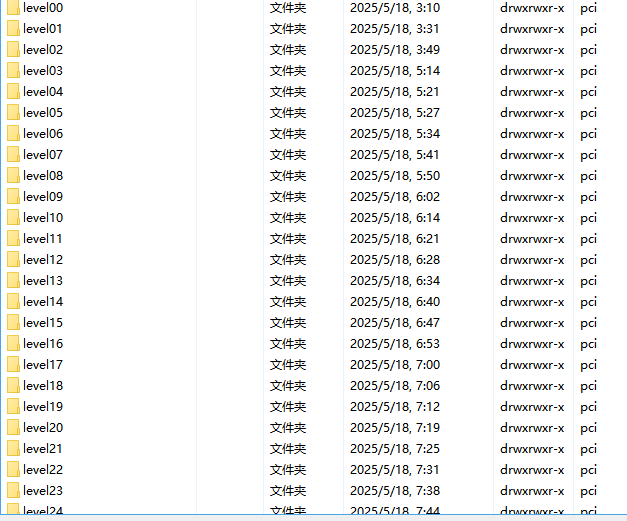


* 1. dataset文件夹中 包含data 和 mask文件夹，其中data文件夹用于存储多个 二维空间物理场单元文件，mask文件夹用于存储 二维空间物理场单元文件 相应的掩码（地形信息，陆地为True，海洋为False）
     1. data文件夹详细结构

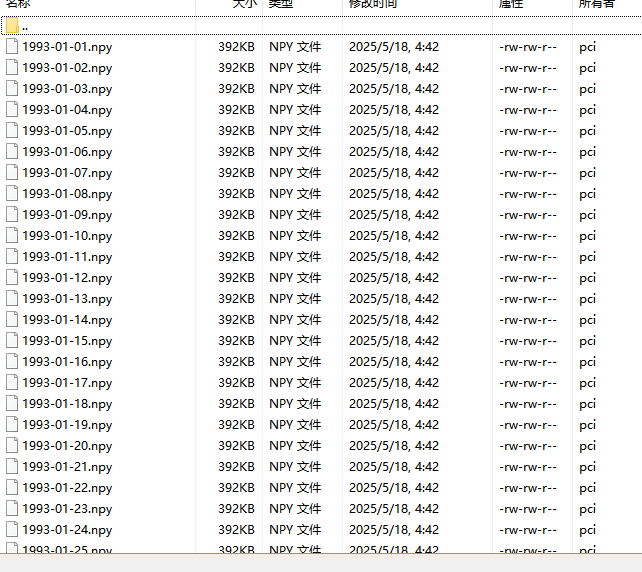
data文件夹中（变量目录）>>>



data - s 文件夹中（水层目录）>>>



data - s - level00文件夹中（时间目录）>>>



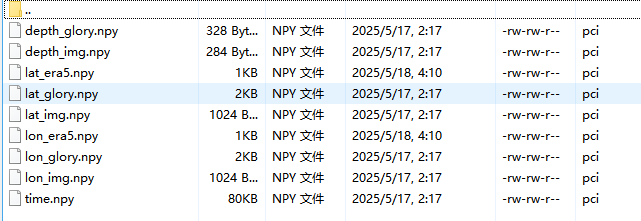
备注：上述以盐度s为例展示了文件结构，其他变量的存放方式是相同的。

* + 1. mask文件夹详细结构

与data文件夹的结构一致

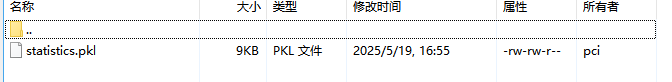
* 1. data\_draw文件夹忽略即可，存放用于图像信息
  2. info\_axis文件夹存放高维矩阵的轴信息（类似nc数据的经纬度、深度、时间说明信息）

info\_axis文件夹截图如下：



* 1. statistics文件夹存放所有数据的统计信息，比如各变量的最大值、最小值、均值和标准差，用于后续的数据归一化操作

该文件夹截图如下



1. data\_manage文件夹存放数据处理程序，步骤1中的data文件夹中的数据均由该文件夹下的程序处理生成

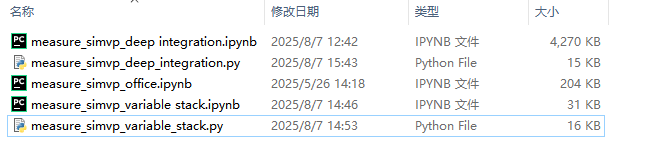
data\_manage文件夹结构如下：



* 1. 01 Data divide glory2d.ipynb程序用于把 形状为(Time, 1，Lat，Lon)的高维矩阵分割为Time × 1个(Lat，Lon)的单元数据（文件）,适用于Glory再分析数据集中的二维变量，如ssh；
  2. 01 Data divide glory3d.ipynb程序用于把 形状为(Time, Depth，Lat，Lon)的高维矩阵分割为Time × Depth个(Lat，Lon)的单元数据（文件）,适用于Glory再分析数据集中的三维变量，如uo、vo等；
  3. 02 Data Check.ipynb程序用于检验生成数据的完整性；
  4. 03 ERA5 Data Segmentation.ipynb程序用于把 形状为(Time, Lat，Lon)的高维矩阵分割为Time 个(Lat，Lon)的单元数据（文件）,适用于ERA5再分析数据集变量；
  5. 04 Calculate Wind Stress.ipynb程序为风应力计算程序
  6. 05 Statistics.ipynb程序为统计各变量各层级的全局信息，如最大最小值、均值和标准差；

1. measure文件夹下存放了测试（推理）程序，用于评估模型性能或生成预报产品

其文件夹结构如下：



* 1. measure\_simvp\_deep integration.ipynb程序是考虑垂直结构的测试或推理程序
  2. measure\_simvp\_variable stack.ipynb程序是不考虑垂直结构的测试或推理程序
  3. 其他两个.py文件忽略即可

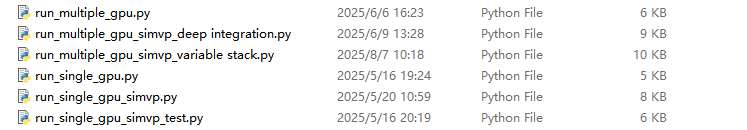
1. package文件夹下存放 所有自定义库 和 第三方库，自定义库包括数据加载器、常用工具等程序，第三方库包括 GitHub下载的模型架构，由于逻辑介绍起来过于复杂，此处不作介绍，直接调用即可
2. result文件夹下存放所有的预训练模型、训练日志、模型配置信息

该文件夹结构如下所示：



* 1. checkpoint文件夹下存储 预训练模型以及用于训练热启动的相关参数
  2. config文件夹存储 模型配置信息
  3. logs文件夹下存储 训练日志，基于tensorboard生成的监控文件

1. 训练代码



上述代码，均为训练代码，后缀multiple\_gpu表示多gpu训练，后缀single\_gpu表示单gpu训练，后缀deep integration表示考虑垂直结构的训练程序（用于生成 考虑垂直结构的预训练模型），后缀variable stack表示不考虑垂直结构的训练程序（用于生成 不考虑垂直结构的预训练模型）