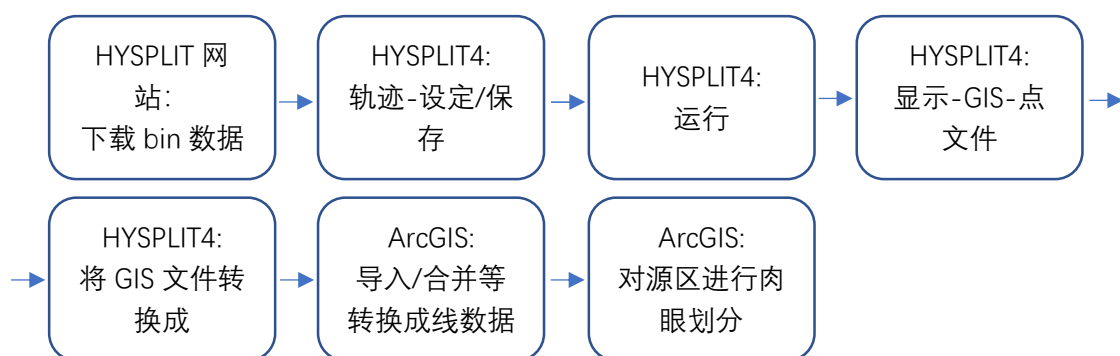
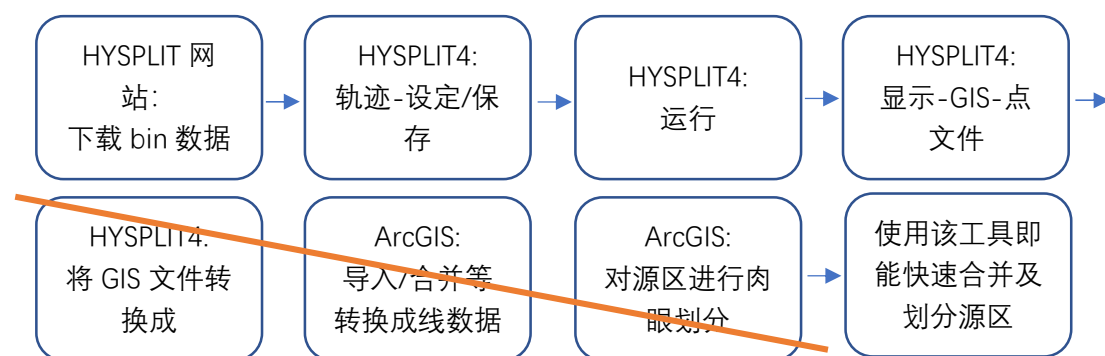


HYSPLIT 结果批量合并和对源区分类工具说明书

这是一个非常简易使用的小工具，本工具意在对 HYSPLIT4 的中途的获得结果（TXT 文件）进行合并与对源区进行分类，主要起简化重复步骤的作用。传统地，对于使用 HYSPLIT4 离线工具的生成轨迹（步骤 2-3），转换成 TXT 文件（步骤 4），再转换成 ArcGIS 文件（步骤 5），导入到 ArcGIS 中并对源区进行划分（步骤 6-7），需要进行如下步骤：



当需要使用量较大的数据时，有些步骤重复并且会耗费大量精力，本工具主要对后面 3 步进行简化：



通过对这些步骤进行简化，能省去大量时间，并且可以减少因为肉眼分类而产生的错误。

1 运行原理

主要对使用 HYSPLIT4 工具的显示-GIS 点文件为基础的 TXT

文件（如何才能得到该结果，请看第 3 部分）。因为每条轨迹会生成 1 个 TXT 文件，我们需要将其归纳起来，比如在例子中，将其以 年月\TXT 文件即如：

199803\GIS_traj_ps_01.txt

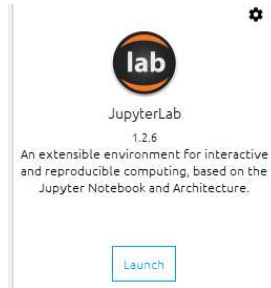
进行保存。本例中每日生成 1 条轨迹，并以月进行归纳（还能以其他尺度的文件夹及文件名保存，请看第 2 部分）。打开 TXT 可以看到，每个 TXT 文件实际上就是 ArcGIS 的点数据。

对于合并功能：其主要是对所有文件夹下所有 TXT 文件进行合并，每个 TXT 文件会归纳到同 1 个特征值中，这样，在后续 ArcGIS 中导入点数据文件后，能使用“点转换为线”工具进行统一转换。

对于源区分类功能：可以看到每个 TXT 文件下的点都有坐标，该功能可以对所有文件夹下所有 TXT 文件，把最后第 N 个（默认为最后 1 个，即轨迹在时间上最开始的点。这个 N 值可以修改。），按照预设好的多个方形区域进行分区。

2 使用方式

因为脚本所涉及模块占用大量空间，封装成 EXE 文件会占用大量空间，因此不进行封装。为了打开 ipynb 文件，首先需要下载 Anaconda 才能使用（<https://www.anaconda.com/>）（主要需要 Python 3 和 Pandas 模块）。安装后，使用其中的 JupyterLab



，该功能会在您浏览器里打开 JupyterLab（某些国内浏览器可能无法支持，请使用 Firefox 或者 Chrome 等）。在左侧打开到 MassHYSP.ipynb 所在位置，双击打开，然后可以看到界面：

```
[1]: # -*- coding: utf-8 -*-
import numpy as np
import pandas as pd
import sys
import os

if os.access('Classification.xlsx', os.F_OK):
    classi = pd.read_excel('Classification.xlsx', 'Region', index_col=None, na_values=['NA'])
    lat1 = classi['Upper-Left Lat.']
    lat2 = classi['Lower-Right Lat.']
    logi1 = classi['Upper-Left Logi.']
    logi2 = classi['Lower-Right Logi.']
    Num = classi['Number']
else:
    print('Error: Can not find Classification.xlsx - Sheet\'Region\', the function of Classification is invalid.')

def numberjudge(x):
    for charc in str(x):
        L = [x for x in range(10)]
        L = list(map(str, L))
        if charc in L:
            pass
        else:
            return False
    return True

def zero(x):
    x = str(x)
    if len(x) == 1:
        x = '0' + x
    return x

def classify(lat, logi):
    for z in range(0, len(classi)):
        if lat2[z] < lat < lat1[z] and logi1[z] < logi < logi2[z]:
            return Num[z]

frq, prd, fstartdate, pref, finalrow = 'M', 12, '1998/03', 'GIS_traj_ps_', 1
if os.access('COMETG' os.F_OK):
```

接下来，以本样例作为说明，说明其如何工作：

先解压 Example 文件，直接解压到文件内，使得：



名称	修改日期	类型	大小
199803	2020/8/6 18:05	文件夹	
199804	2020/8/6 18:05	文件夹	
199805	2020/8/6 18:05	文件夹	
199903	2020/8/6 18:06	文件夹	
199904	2020/8/6 18:06	文件夹	
Classification.xlsx	2020/8/7 16:43	Microsoft Excel 工作表	9 KB
Example.rar	2020/8/7 18:10	WinRAR 压缩文件	329 KB
MassHYSP.ipynb	2020/8/7 17:24	IPYNB 文件	12 KB
说明书.pdf	2020/8/7 17:33	Adobe Acrobat 文档	710 KB

首先，由于本软件会先推算所有潜在的文件夹、文件，再对存在的文件进行处理。所以要对数据基础命名方式进行了解，打开任意一个样例的文件夹（如 199803），可以发现里面有许多 TXT，这里，以 1998 年 3 月每日做一次模拟得到的结果，分别做了 1998 年 3-5 月，1999 年 3-4 月，可以发现其命名方式是：

前缀名对应指令 pref, 默认为 GIS_traj_ps_

\MassHYSP\199803\GIS_traj_ps_01.txt

文件名, 对应指令 fstartdate, frq 及 prd


每个文件夹从 01 开始, 最大为 99

fstartdate 指的是文件夹名的最初的日期

frq 每个文件夹名间隔的频率, 如 Y、M、D 各代表年、月、日 (可另外指定)

prd 指的是由上述 2 个参数指定的开始日期和间隔的频率后, 向后推算的文件夹的日期的个数。

上述图中，框内英文代表参数设置的指令。在点击图中的运行

后 ，可以输入这些指令，这些指令分别有：

freq: 每个文件夹名间隔的频率，在窗口内输入“freq”（回车键）后，再填入“Y”、“M”、“D”（回车键），分别可指定为年、月、日（其由 pandas 产生，对其熟悉可自行另外指定小时、分钟、秒等）。由于正确填入 fstartdate 后，会对 freq 进行自动纠正（只限于年、月、日），所以在使用时可以不用理会本参数。

fstartdate: 文件夹名的最初的日期。在文件夹中命名时，应当是连续的“年月日”，比如“1998”（对应 freq 为“Y”）、199803（对应 freq 为“M”）、19980301（对应 freq 为“D”）等。在程序中，先输入“fstartdate”（回车键），然后对应地输入：“1998”、“1998/03”、“1998/03/01”（回车键），注意程序中输入时，必须增加斜杠分割年、月、日，否则会出现错误。

prd: 上述 2 个参数指定的开始日期和间隔的频率后，向后推算的文件夹的日期的个数。比如本样例中，fstartdate 设置为“1998/03”（程序内），如果 prd 指定为 12，那么，它运行会遍历从 1998 年 03 月（即 199803 的文件夹）到 1999 年 02 月（即 199902 的文件夹），这总共有 12 个文件夹。同样，如果在

fstartdate 设置为“1998”，prd 指定的是 10，那么它运行会从 1998 年（即 1998 的文件夹）运行到 2007 年（即 2007 的文件夹）

pref: 前缀名，默认为 GIS_traj_ps_。可以根据需要修改前缀名。

您需要做的，是在设定完这些参数之后，让程序完全覆盖您全部的文件夹以及当中的轨迹的 TXT 文件。注意的是，只支持下一级一级

的文件夹，并且每个文件夹内参数的编号（上图蓝色标识）范围只能从 01 到 99。另外，当中的文件夹或轨迹的中间的编号有缺失并不会影响程序的运行，您只需要把范围覆盖完全即可。

在设定完参数之后，通过另外的指令（输入指令并按回车键）进行运行：

end: 关闭脚本。想多次使用时，或做错时，请务必输入 end 关闭。

parascan: 查看现在的参数。

save: 保存这些参数。在下次运行脚本时, 会自动读入。

源区分类:

finalrow: 指的是分类依据的点, 默认是轨迹最开始的点, 即 1。如果有需要, 可以进行修改。

classify: 运行后, 对源区进行划分。源区的设置, 可以打开 Classification.xlsx 进行设置。第 A 列为源区的名字 (建议设定为数字)。以源区为一个矩形看待, 之后第 B、C 列分别是各源区的左上角的点的纬度、经度, 而 D、E 列则是右下角的点的纬度、经度。可以增加无数多个源区, 注意, 请保持源区下方的单元格为空值, 否则可能会大大增加运算量。

轨迹合并:

merge: 运行后, 把全部数据由最开始日期的文件夹的编号为 01 的轨迹开始, 把所有存在的轨迹的点进行合并。这会比较花时间, 另外可能占用较多的内存。

现在我们尝试运行一下, 在运行代码之后, 由于初次运行, 会默认设置参数。可以看到 (随后可以输入 parascan 来查看):

```
frq=M, prd=12, fstartdate=1998/03, pref=GIS_traj_ps_, finalrow=1
```

由于源区划分原本也已指定, 可以先通过输入 classify 指令试运行, 会文件夹下生成一个 Classi_result_() 的 excel 表, 打

开其后发现，其只做了 1998 年 3 月到 1998 年 5 月的轨迹分类，这是因为 1998/03 (fstartdate) 往后推 12 个单位 (prd)，只能覆盖到 1999/02，即没有覆盖 1999 年 3-4 月。现在，我们输入 prd，回车键，再输入 24，回车键。由于这时已经推了 24 个单位的日期，能覆盖到 2002 年 02 月，所以再打开新的 excel 表可以发现 1998 年 3-5 月，1999 年 3-4 月全部都进行了分类，其顺序是按时间顺序的。类似地，我们也能做 merge，会生成 Merge_result_() 的 excel 表，merge 的结果是以每个轨迹单独划分一条线数，然后合并了所有点的 excel 表，之后我们可以在 ArcGIS 中进行合并，将在第 3 节中讲解。至此，样例已经全部讲解完毕。您可以修改其他参数进行尝试。

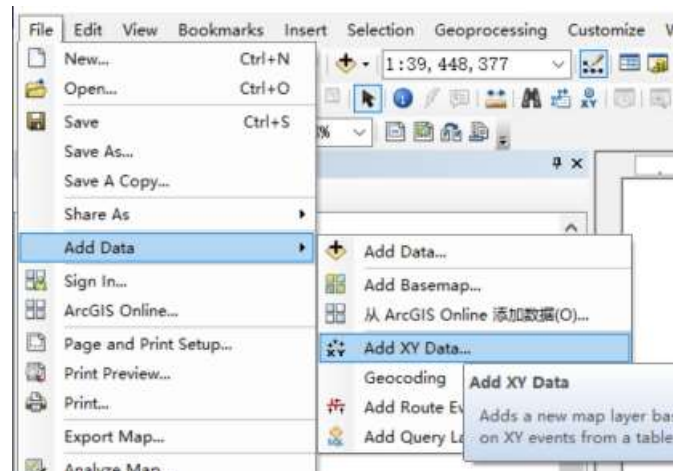
3. 后续及之前的流程

3.1 后续流程

源区分类的后续：打开源区分类的表格后，可以发现其 B 列为轨迹的名、C 列是分类依据及提取的点、D、E 列分别是该点的纬度、经度，F 是所在源区。可以直接将源区与相应精度的降水量等数据进行联系，这样就可以获得各源区来的降水量了。请务必注意每条轨迹应当和器测数据相对应。

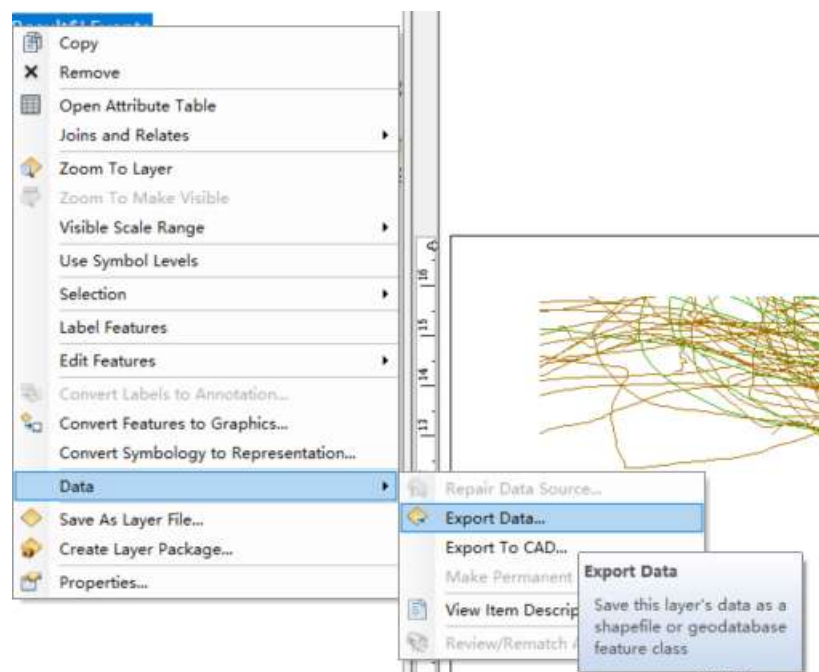
合并的后续：由于 ArcGIS 无法直接导入该表，首先打开该 excel 表，然后另存为 xls 文件。然后在 ArcMap 中以点的方式

导入该表：

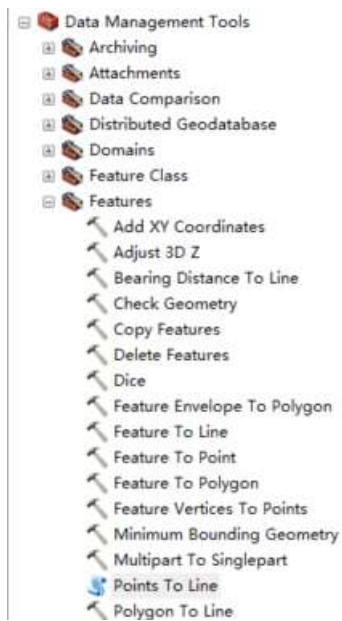


接下来其 XY 值会自动匹配，按确认即可。

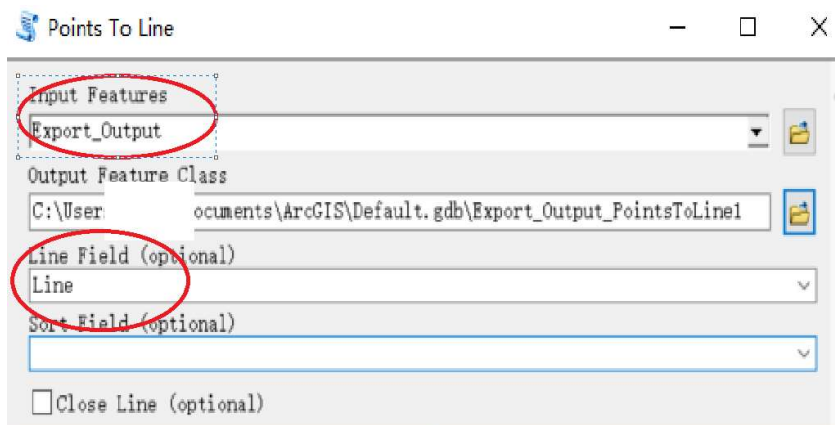
导入后，右键点数据，输出为 Data：



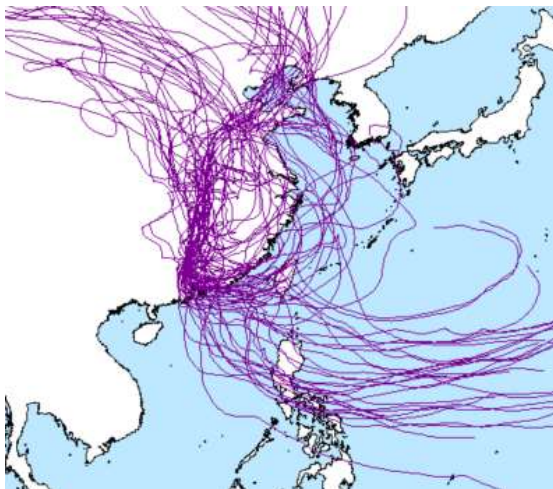
，保存 shp 等您熟悉的位置，因为点比较多，会花比较长时间。接下来提示按确认即可导入到本文件中，您也可以自动导入。接下来使用 points to line 工具，



接下来注意，输入的是刚刚输出的 shp 文件，然后 Line Field 指定为 Line 列



然后按确认，之后就能转化为线文件了：



该结果也能与之前的分类结果相匹配。

3.2 之前的流程

如何获得 TXT 文件？首先，您需要在 HYSPLIT 上下载 offline 版本的 HYSPLIT (https://ready.arl.noaa.gov/HYSPLIT_hytrial.php)。然后您需要安装 ActiveTcl 才能运行 HYSPLIT，关于这个问题，在 HYSPLIT 安装后请看其中的 readme。然后您还需要轨迹模拟的基础数据，可以在 (<https://www.ready.noaa.gov/ready2-bin/extract/extracta.pl>) 中下载。关于这些，都汇总在 HYSPLIT 的首页 (<https://ready.arl.noaa.gov/HYSPLIT.php>)。

进入 HYSPLIT4 后，选择 Trajectory - Setup Run，设置时间、纬度等，在 Add Meteorology Files 处，加入对应时间的空间的轨迹模拟的基础数据，最后选择 Save。再选择 Trajectory - Run Model，即可运行（若报错，请注意基础数据不能太多，同时应该覆盖上述的空间和时间，另外，参数输入也应注意避免出错）。这里其他保持默认状况下，只需要对 GIS Out 项选择 GIS-point 即可，然后上方 Output Postscript 就是输出 TXT 的文件名。以本工具为例，应当为 traj01 到 traj99。之后，在 C:\hysplit\working\下即能找到该 TXT。为了使用本工具，应当对其按照日期的文件夹进行归类。

4. 另外一些提示

事实上，HYSPLIT 提供了更多功能非常丰富而强大的后续处

理软件、代码，这些您可以在 HYSPLIT 首页最下方看到。比如也是基于 Python 的，在 <https://github.com/mscross/pysplit> 中也有更加强大的功能的代码。本工具只是个非常简陋的工具，只能实现一点简单的功能。感谢使用！

本工具只是对 HYSPLIT 结果进行整合，并无直接使用 HYSPLIT，而使用这些结果时，请注意**必须**引用 HYSPLIT 相关文章。对于本工具，需要咨询或者出现了些问题、建议等可以积极联系 sieyalixnet@163.com。若使用本工具进行研究，还希望能在文章中进行提及，非常感谢！

Sieyalixnet