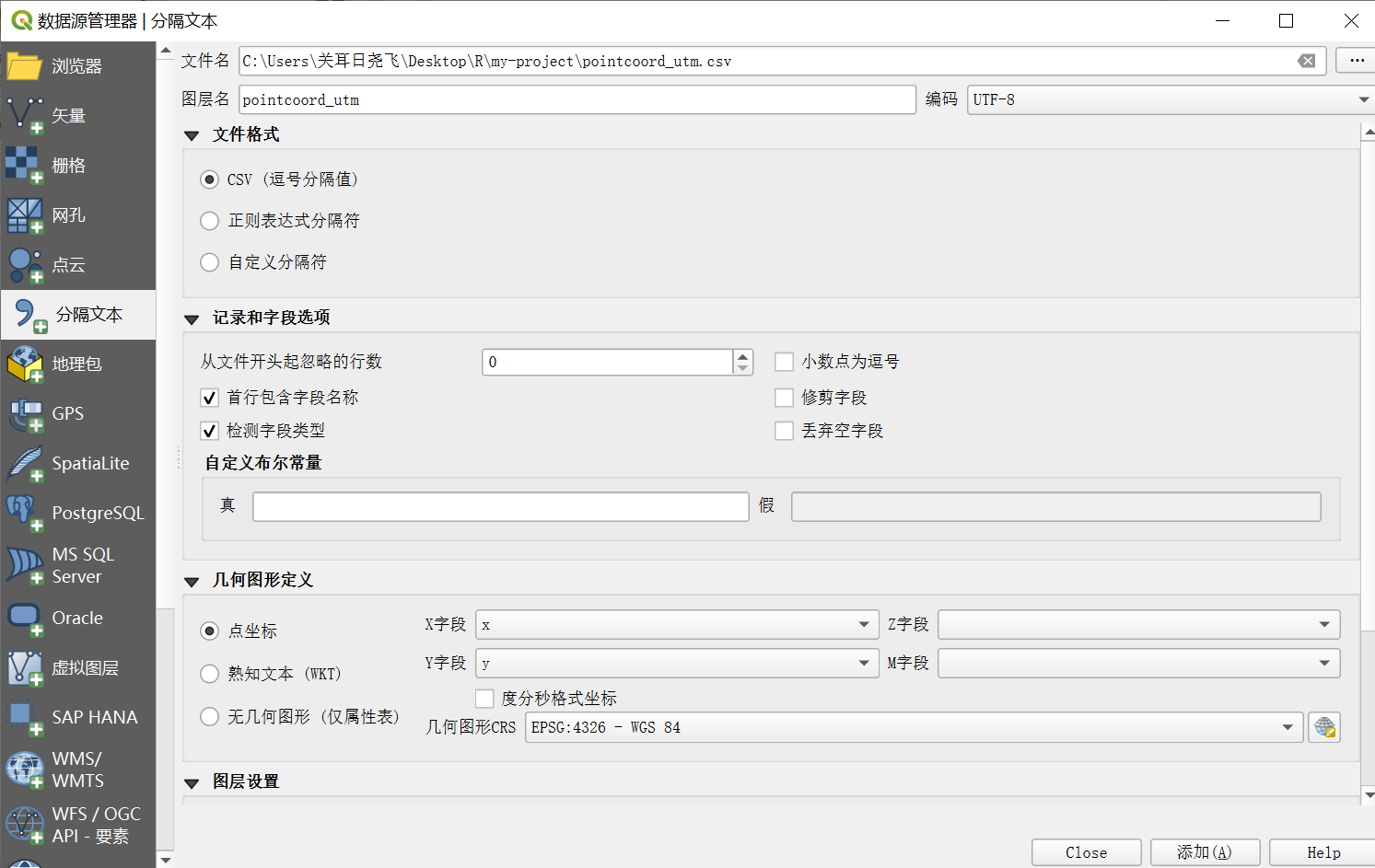
1. 对于Doubs数据集，采样点坐标为千米，但没有参考系信息，只有通过QGIS确定各点地理坐标。请简述QGIS操作过程并给出对应点的地理坐标。
2. 从GitHub库中下载**pointcoord\_utm.csv**文件（Doubs数据集中的UTM坐标）

保存该数据文件于"C:\Users\关耳日尧飞\Desktop\R\my-project"。

1. 打开QGIS，点击工具栏**图层 → 添加图层 → 添加分隔文本图层 → 导入Doubs数据集点信息成为图层中的点数据，命名图层名为pointcoord\_utm。**



1. 成功将Doubs数据集加载到QGIS中，**pointcoord\_utm图层**如下图所示：

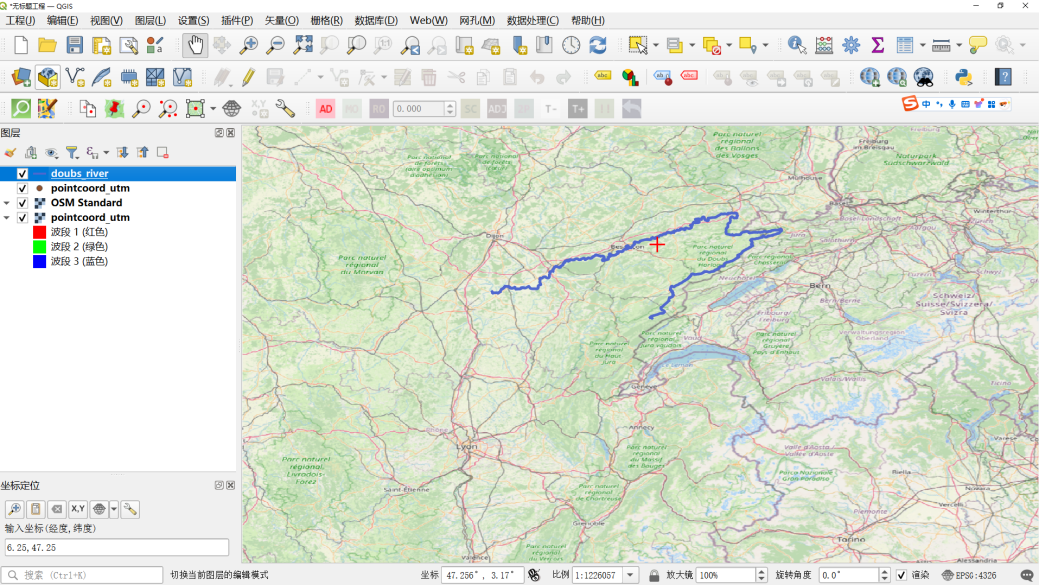


1. 检查**pointcoord\_utm**图层的属性，确保其中包含了采样点的**UTM**坐标。



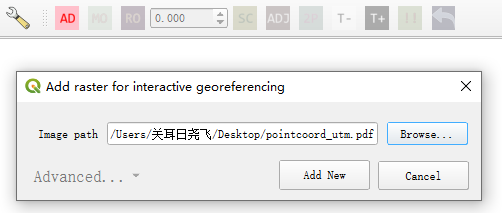
1. 点击**工程 → 导出 → 导出地图为PDF格式文件，命名为pointcoord\_utm.pdf用于后续配准操作。**

****（六）**将上次作业保存的**doubs\_river.geojson****文件导入到画布中：**图层 → 添加图层 → 添加矢量图层，并添加一个地图底图：Web → QuickMapServices → OSM → OSM Standard。**



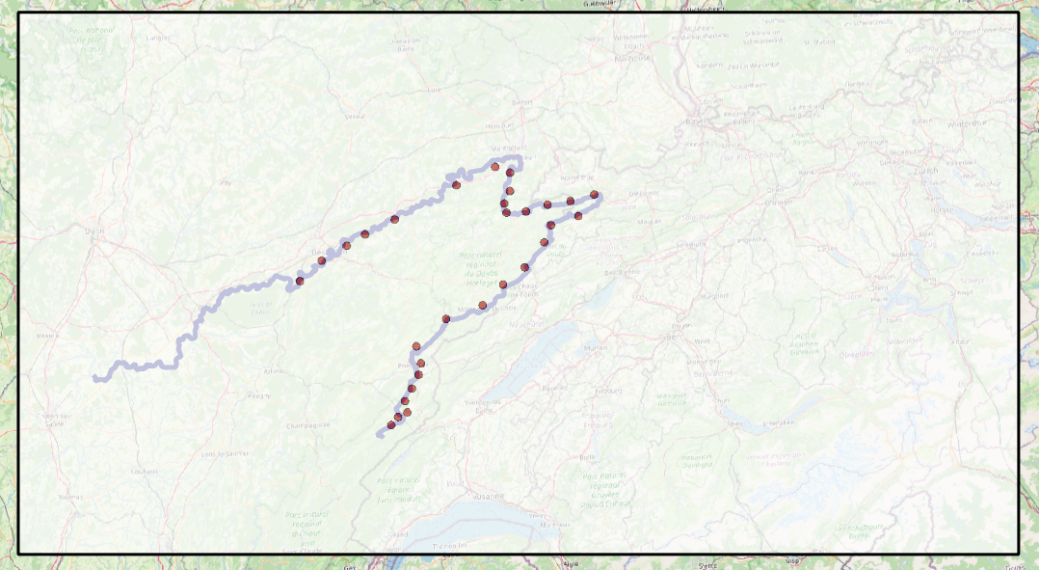
****（七）**为了更方便地进行接下来的配准操作，需要将搜索到的河流数据与导入的数据点图**（**pointcoord\_utm.pdf**）**进行重叠，这样可以更好地在配准的时候进行画布找点。**

**这里首先安装一个**插件**Freehand raster georeferencer，利用该插件的工具，再次添加pointcoord\_utm.pdf**文件导入到画布中；

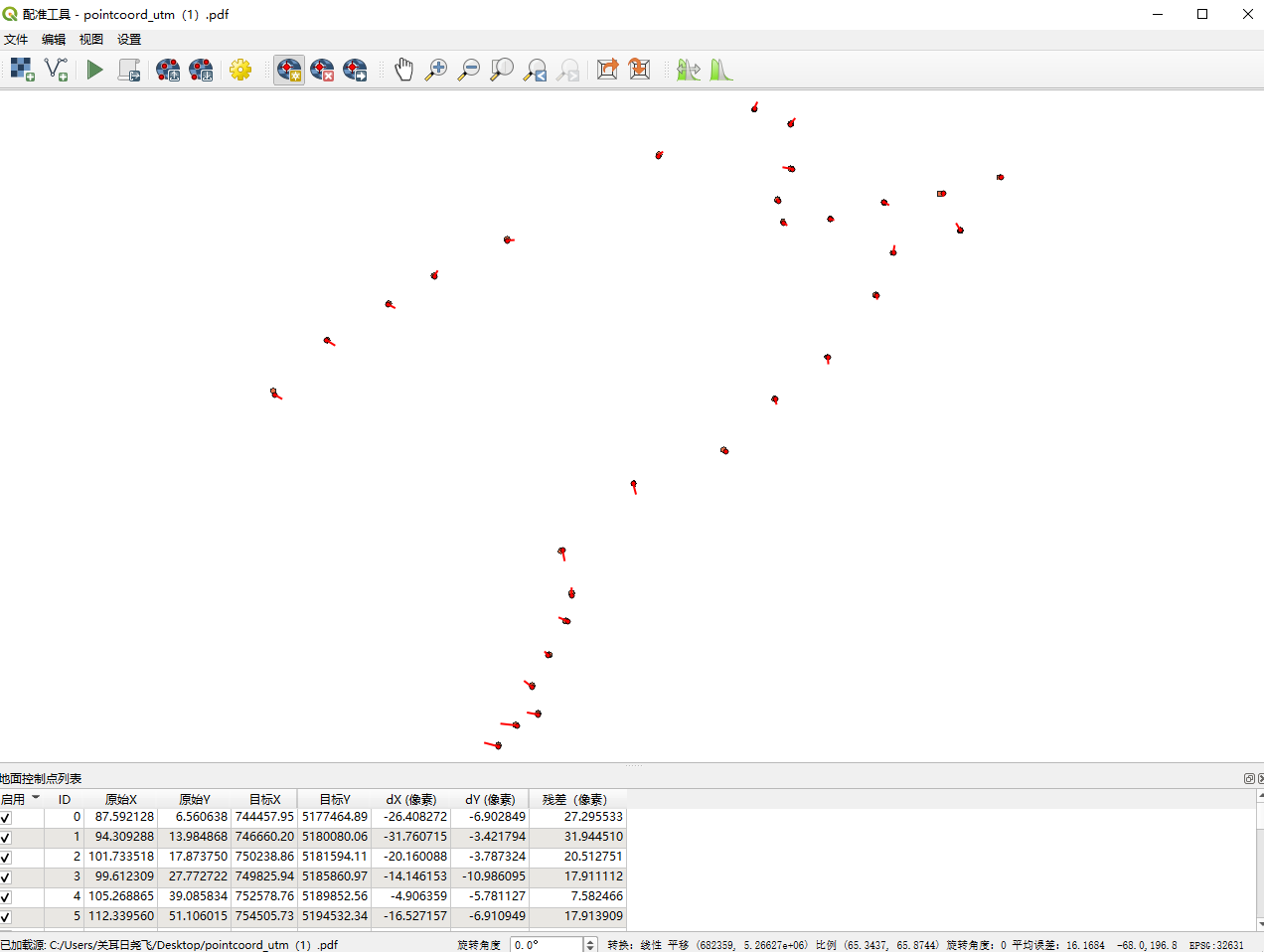


随后就可以**使用Freehand Raster Georeferencer的移动和缩放功能，将搜索到的河流数据与导入的数据点进行重叠。**

**（在进行两者重叠时，需要通过SC来进行等比例放大缩小，实现与河流线图进行基本重合，并且需要使用合适的坐标参考系统，EPSG：32631）**



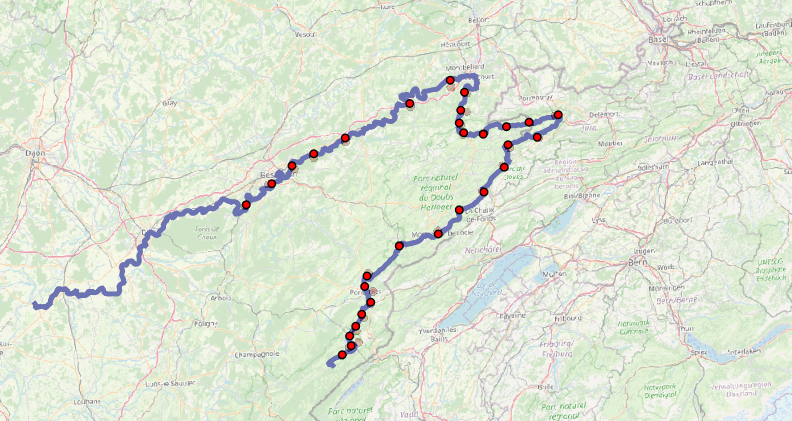
**（八）使用配准工具，根据**doubs\_river**河流数据，对前面导出的pointcoord\_utm.pdf**文件进行配准：**点击图层 **→ 配准工具 → 打开格栅添加配准文件 → 对PDF中的每一个点进行配准，尽量使得点落在河流线段上**。**



**（九）**完成配准操作后**点击开始校准**，获得**pointcoord\_utm已修改.tif**文件，此时该图层也已经加载到QGIS中，同时保存校准后的文件。



1. 校准完成后，可以在QGIS页面中看到校准好的数据点

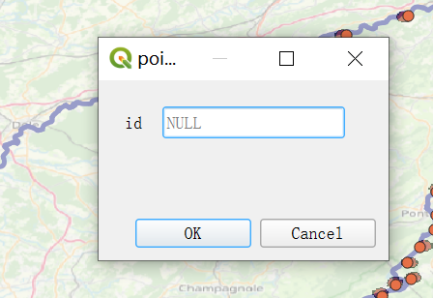
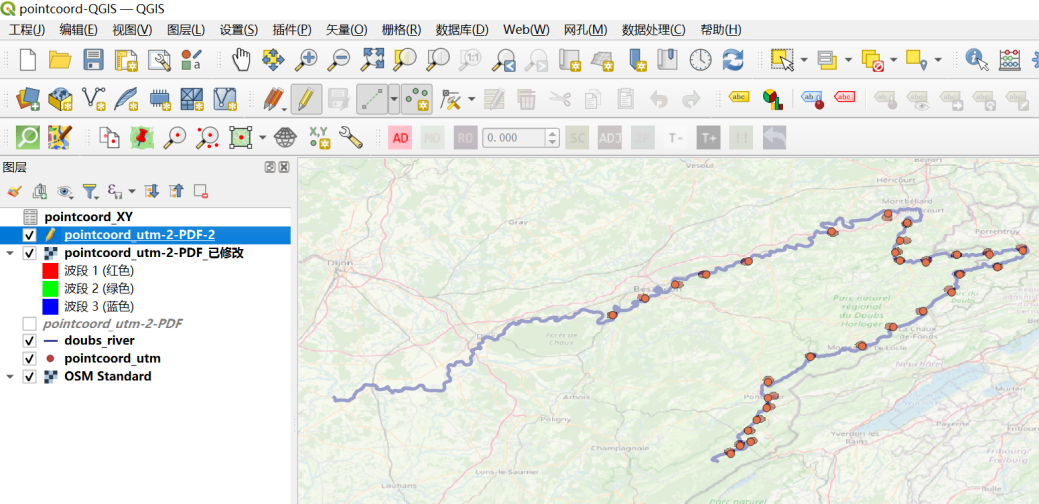


1. **为了**实现提取采样点的对应地理坐标，还需要新建一个**形状文件图层。**

（切记此时需要将CRS更改为EPSG：4326）



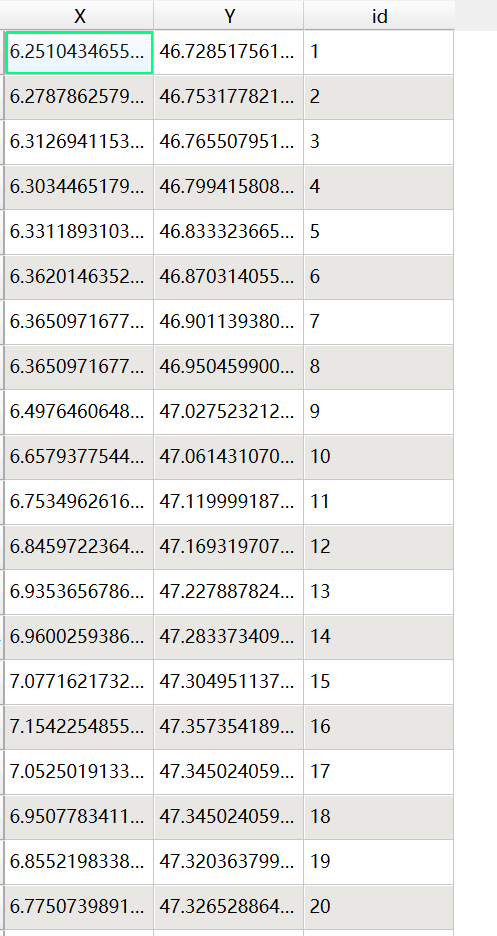
**（十二）**根据配准后的栅格图层，将采样点矢量化，并对应标注各点的id。



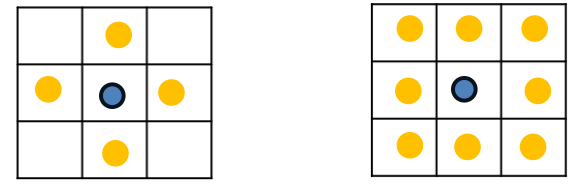
（十三）将矢量点数据的地理坐标保存为**csv格式**



1. **读取该文件的属性表，实现提取采样点的对应地理坐标。**



1. 为确定是否存在空间自相关，需要确定空间邻近，包括Rook和queen相邻，请分别指出如下两图中蓝色位置的空间邻近，并给出它们的空间权重矩阵。



首先将上述3×3的格子进行编号，以便后续回答问题

|  |  |  |
| --- | --- | --- |
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

第一张图表示的是Rook相邻，蓝色位置的空间临近为蓝色圈的上**（2）、下（8）、左（4）、右（6）四个**黄色圈。

空间权重矩阵为：（从上到下9行依次为1、2、3、4、5、6、7、8、9的权重矩阵）

0 1 0 1 0 0 0 0 0

1 0 1 0 1 0 0 0 0

0 1 0 0 0 1 0 0 0

1 0 0 0 1 0 1 0 0

0 1 0 1 0 1 0 1 0

0 0 1 0 1 0 0 0 1

0 0 0 1 0 0 0 1 0

0 0 0 0 1 0 1 0 1

0 0 0 0 0 1 0 1 0

第二张图表示的是Queen相邻，蓝色位置的空间临近为蓝色圈的左上**（1）、上（2）、右上（3）、左（4）、右（6）、左下（7）、下（8）、右下（9）八个**黄色圈。

空间权重矩阵为：（从上到下9行依次为1、2、3、4、5、6、7、8、9的权重矩阵）

0 1 0 1 1 0 0 0 0

1 0 1 1 1 1 0 0 0

0 1 0 0 1 1 0 0 0

1 1 0 0 1 0 1 1 0

1 1 1 1 0 1 1 1 1

0 1 1 0 1 0 0 1 1

0 0 0 1 1 0 0 1 0

0 0 0 1 1 1 1 0 1

0 0 0 0 1 1 0 1 0

3. 在空间数据探索性分析中，常见Moran‘s scatter plot和Lagged mean plot两种图，依据图判断是否存在空间自相关。请回答如下问题：

1）根据Moran’s scatter plot图，如何分辨空间自相关？

1. 两图中都有一直线，此直线分别表达的意思是什么？

（1）

|  |  |  |
| --- | --- | --- |
| **情况** | **图像特征** | **解释** |
| 正的空间自相关 | 散点呈**右上-左下方向斜线**，大致沿正斜率分布  （倾斜向上） | 说明**高值聚集在高值附近，低值聚集在低值附近** |
| 负的空间自相关 | 散点呈**右下-左上方向斜线**，大致沿负斜率分布  （倾斜向下） | 说明**高值邻近低值，低值邻近高值** |
| 无空间自相关 | 散点杂乱无序，没有明显线性趋势 | 各单元之间没有空间上的联系 |

（2）

|  |  |  |
| --- | --- | --- |
| **图类型** | **直线含义** | **备注** |
| **Moran’s scatter plot** | 直线是**回归拟合线**，斜率 = Moran’s I值 | - 如果斜率 > 0，说明有**正空间自相关** - 如果斜率 < 0，说明有**负空间自相关** |
| **Lagged Mean Plot** | 直线是**x = y 的45度线**，参考线 | - 主要用于观察单元自身值和邻居均值的一致性 - 点分布在45度线附近 → 自身值≈邻居均值，表明存在正空间自相关 |

## 4.通过探索性分析，结果表明doubs河流中鱼群的物种多样性存在一定程度的空间自相关，请简述如何将空间自相关纳入鱼群物种多样性-环境要素关系模型中。

解答：

以doubs河流中的鱼群物种多样性数据为例：

构建模型时，target为各采样点鱼类的shannon多样性指数（shan\_div字段），而features则为前面已筛选过、排除了高度共线性（pearson相关系数 > 0.8，或其他自定义阈值）且与待预测target密切相关的环境变量，最终保留carea（面积）、cslope（坡度）、pen、deb、pH、dur、dbo这几个字段；在此基础上，构建一定缓冲区范围内的空间格网，将鱼类丰度点数据分层后，计算各层上点到各格网的缓冲区距离（即已知点-预测点距离，buffer\_distance），然后将各层buffer\_distance数据作为新的features引入机器学习建模，评估并筛选较为重要的缓冲区距离变量以此实现空间自相关的纳入。