

# 降水量预测

预测15-20天内的海上降水量

# 第七次讨论会

## ➤总结

- 1、完善方法，使方法与结果都准确无误。
- 2、预处理好数据，提取出来6-9月份的数据。
- 3、得出详细的结果，做出图标。

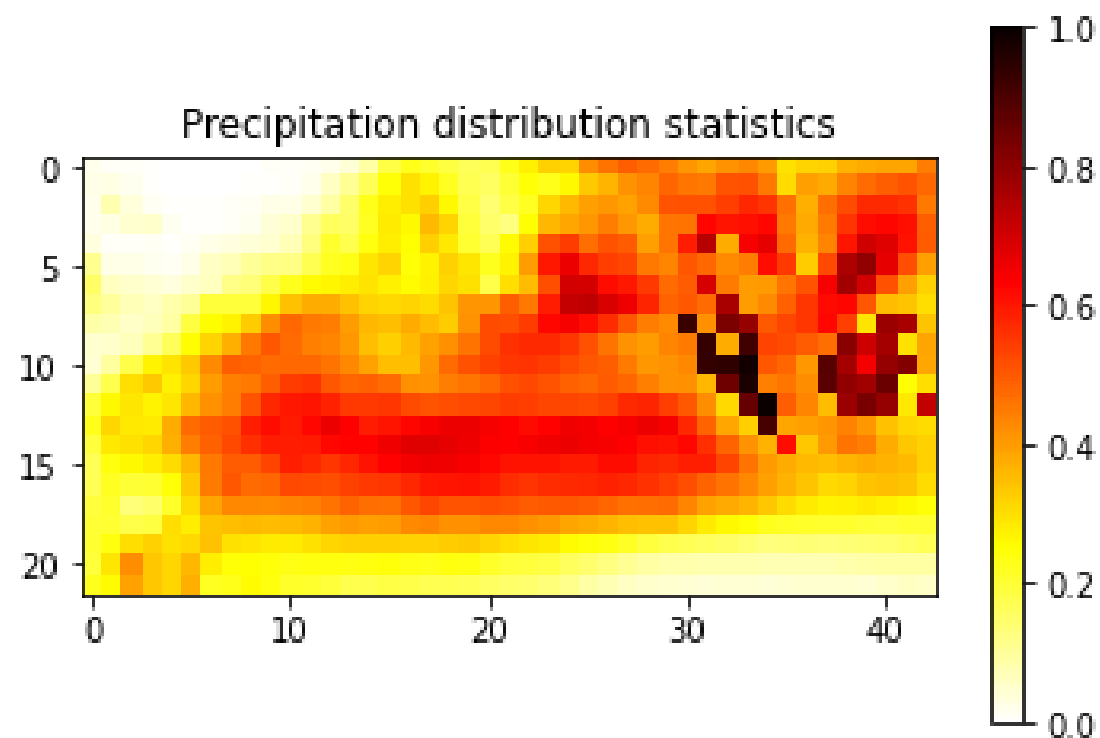
# 第八次讨论会

## ➤本周工作

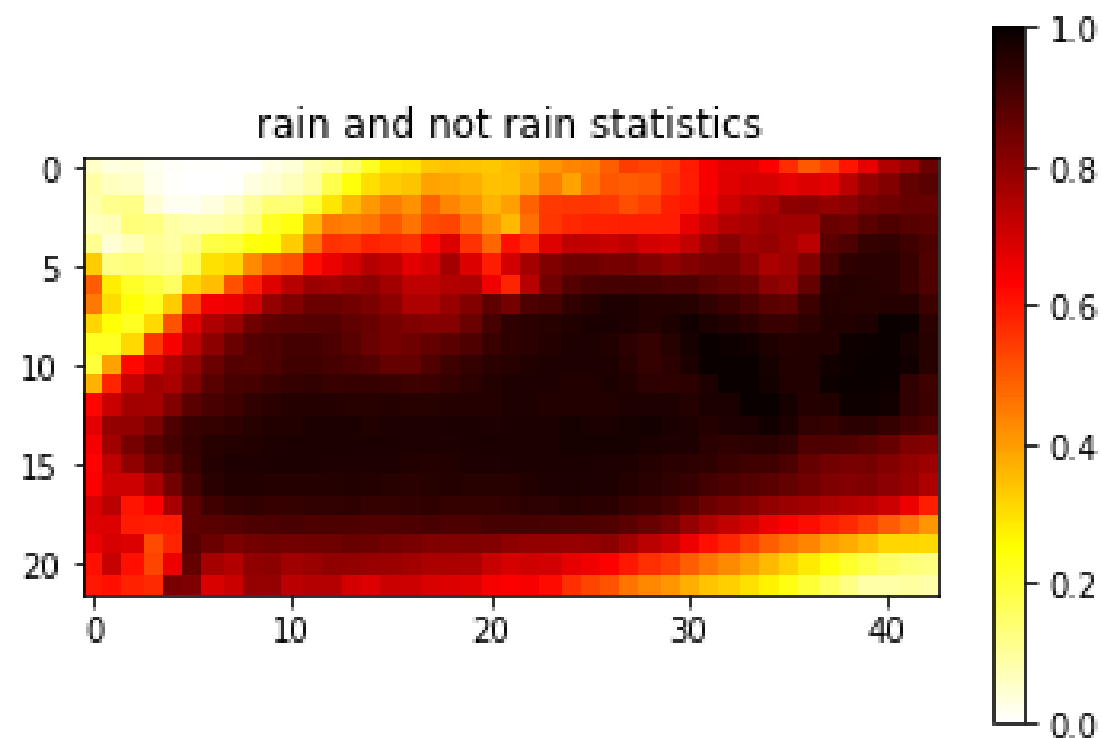
- 1、预处理数据，将数据整理出6-9月份的。
- 2、校准优化程序。
- 3、画出数据原始的分布图。
- 4、运行出所有点的结果，并画出分布图。

# 第八次讨论会

## ➤原始降水量分布



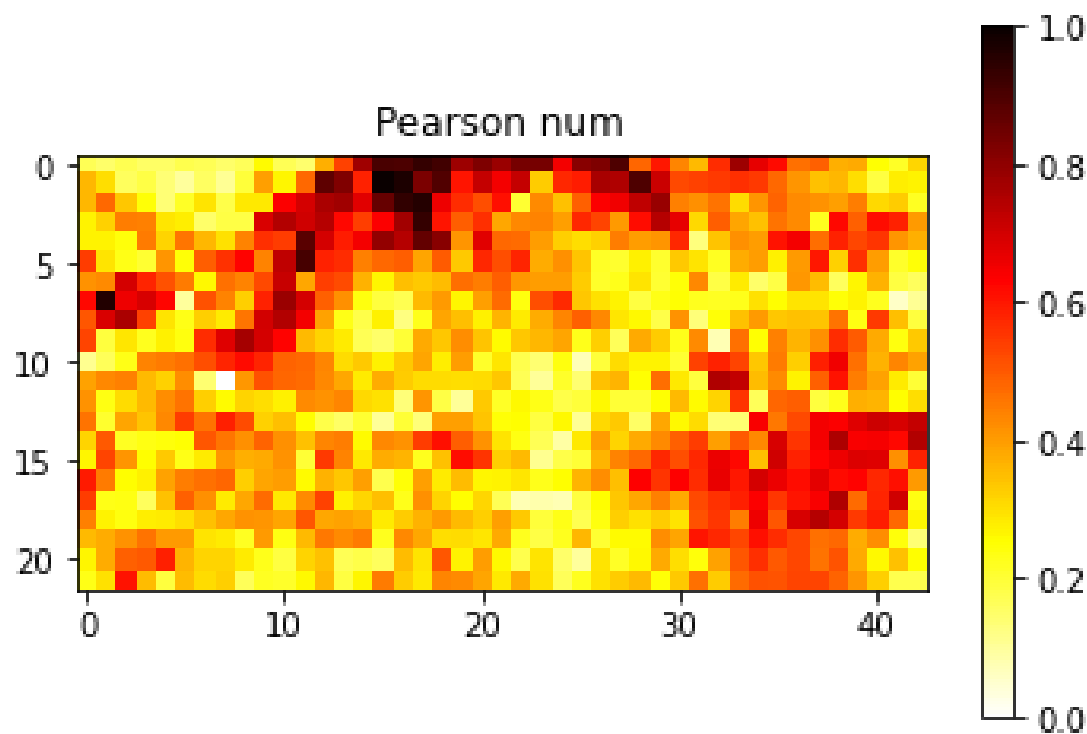
区域降水量总和分布图



是否降水，统计天数分布

# 第八次讨论会

## ➤各点的皮尔森系数分布情况



结果的另外一些数据统计分析:

空间总点数: 946

预测皮尔森系数大于0.6点数: 1 0.99%

预测皮尔森系数大于0.5点数: 18 1.8%

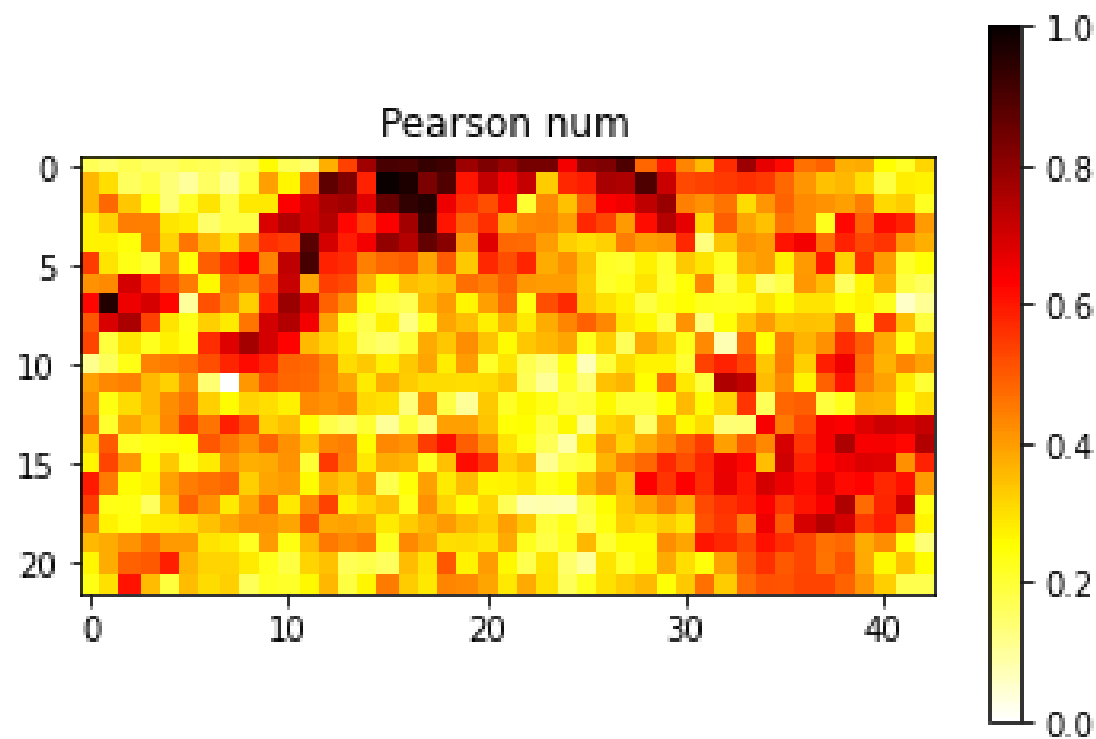
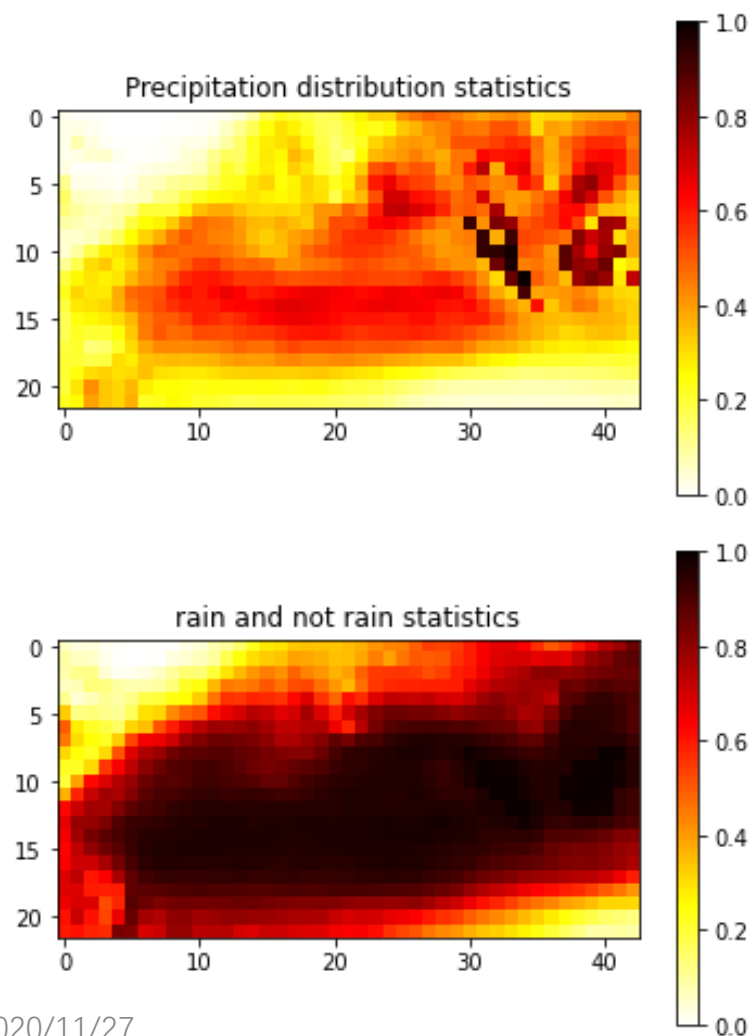
预测皮尔森系数大于0.4点数: 59 5.8%

预测皮尔森系数大于0.3点数: 174 17.2%

预测皮尔森系数的最大值: 0.6092

预测皮尔森系数的最小值: -0.1099

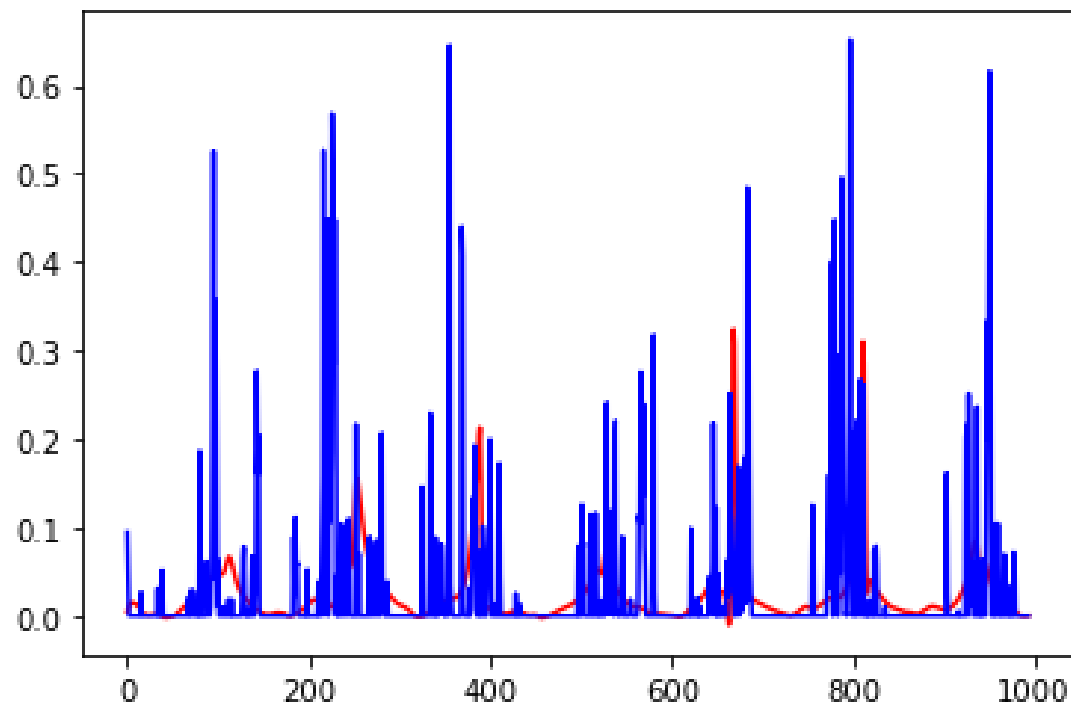
# 第八次讨论会



# 第八次讨论会

## ➤实验结果及分析:

- 1、模型不能对降水量进行很精确的预测，皮尔森系数普遍不高
- 2、模型对数据有很强的依赖性，即皮尔森系数高的值总是聚集在一起。说明，模型只适合部分数据的预测。
- 3、分析预测值会发现，模型可以预测出降水量大体的趋势，但有时计算出的皮尔森系数不是很高。



0.13384086399242973

# 第八次讨论会

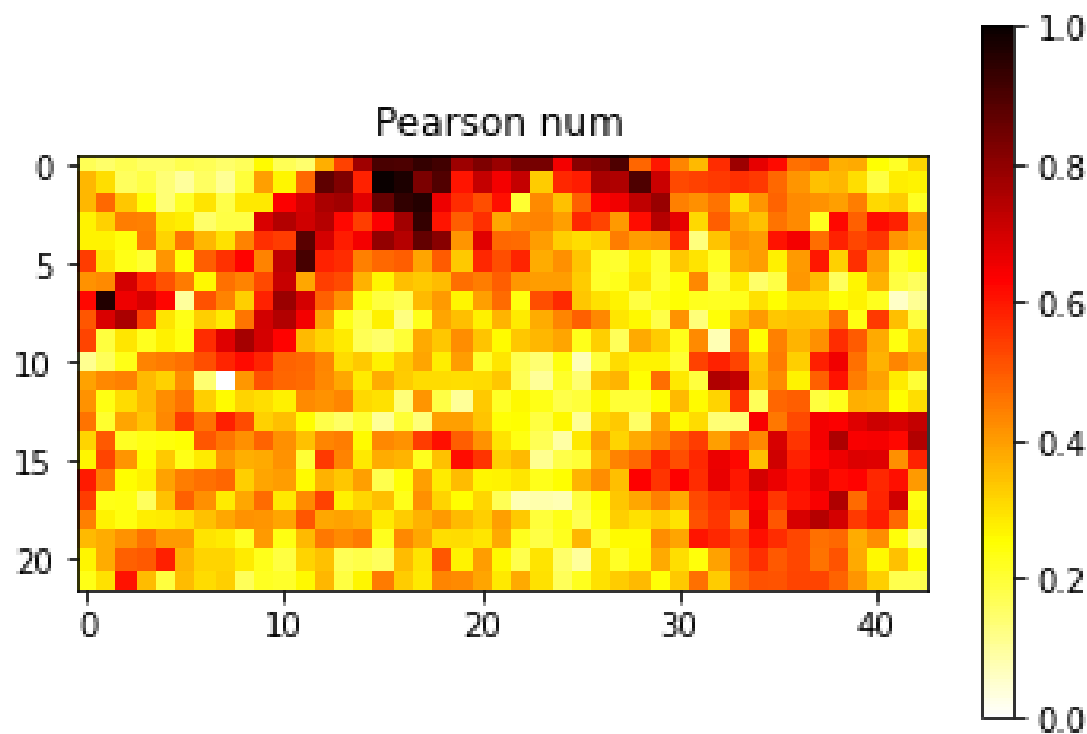
- 总结

输入数据中加入历史降水量数据



# 第九次讨论会

## ➤各点的皮尔森系数分布情况



结果的另外一些数据统计分析:

空间总点数: 946

预测皮尔森系数大于0.6点数: 3 0.3%

预测皮尔森系数大于0.5点数: 13 1.3%

预测皮尔森系数大于0.4点数: 60 5.9%

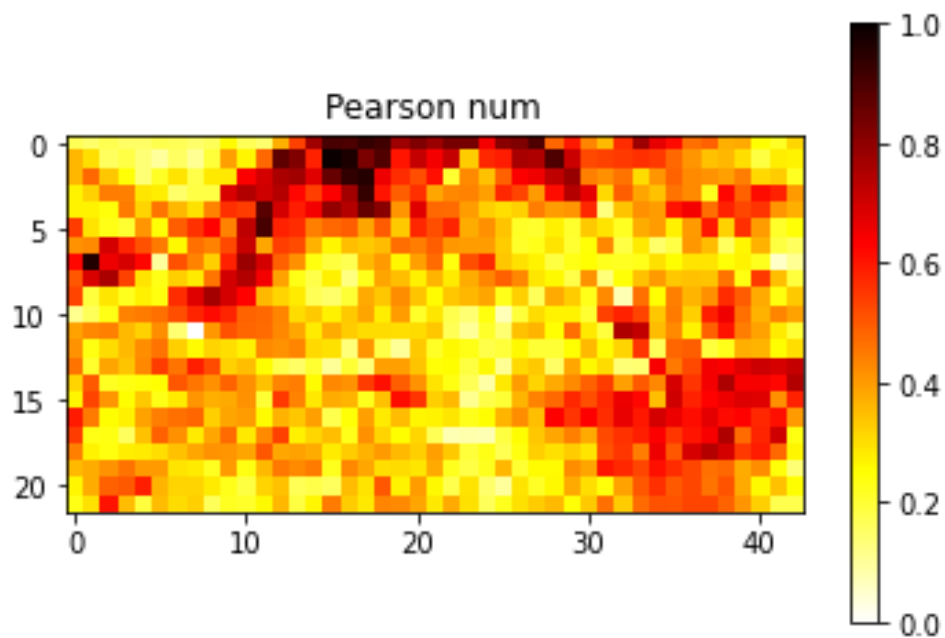
预测皮尔森系数大于0.3点数: 168 16%

预测皮尔森系数的最大值: 0.60926.

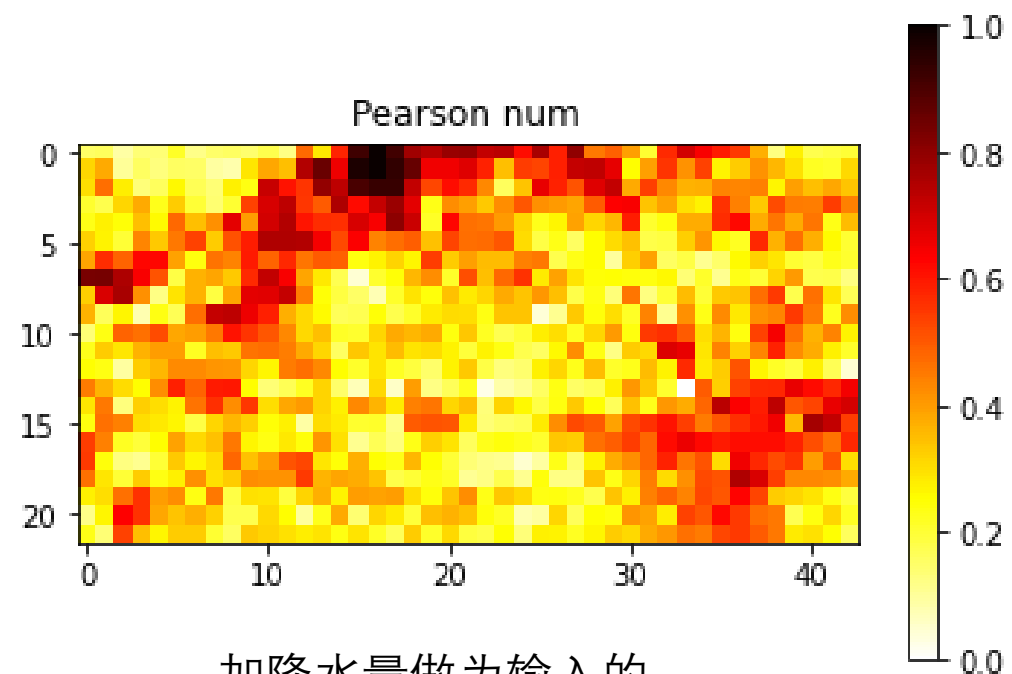
预测皮尔森系数的最小值: -0.1099

# 第九次讨论会

归一化，绘制的图像结果



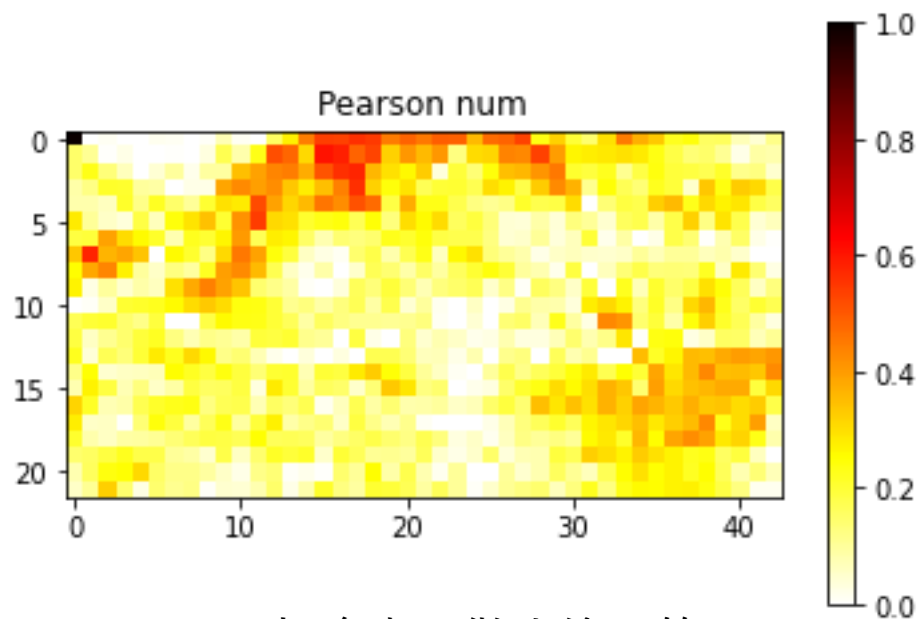
不加降水量做为输入的



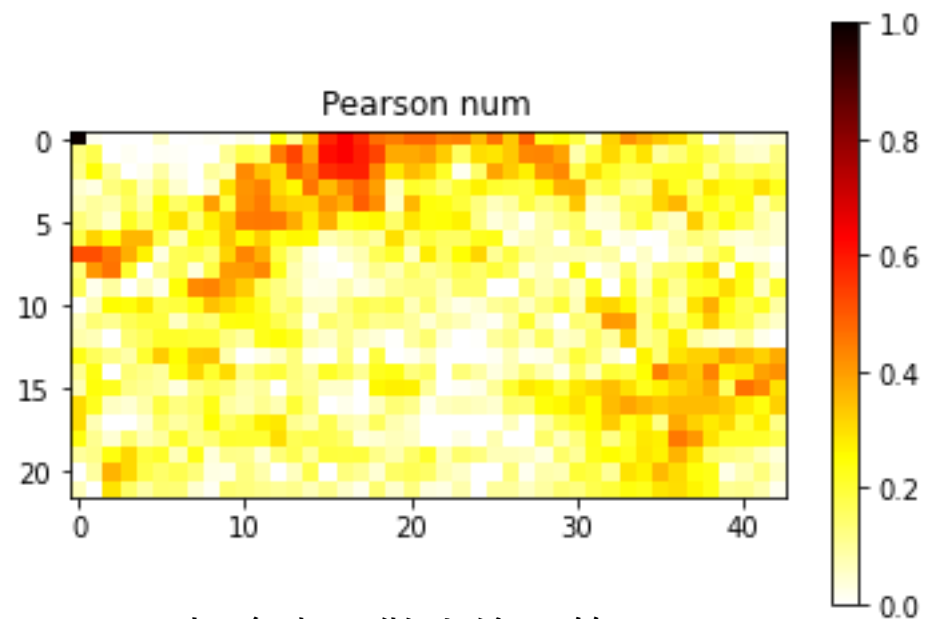
加降水量做为输入的

# 第九次讨论会

## 原数据绘制图像结果



不加降水量做为输入的



加降水量做为输入的

# 第九次讨论会

- 实验分析:

实验最后的效果不是很好，预测的超过0.6的，上升了2个，但是预测的超过0.5的，减少了5个，总体来说，有提升，但是提升效果特别微小。

- 进一步实验思路

- 1、随机生成20个数据，更改不同的hidden\_size
- 2、计算皮尔森系数时，只计算10天的

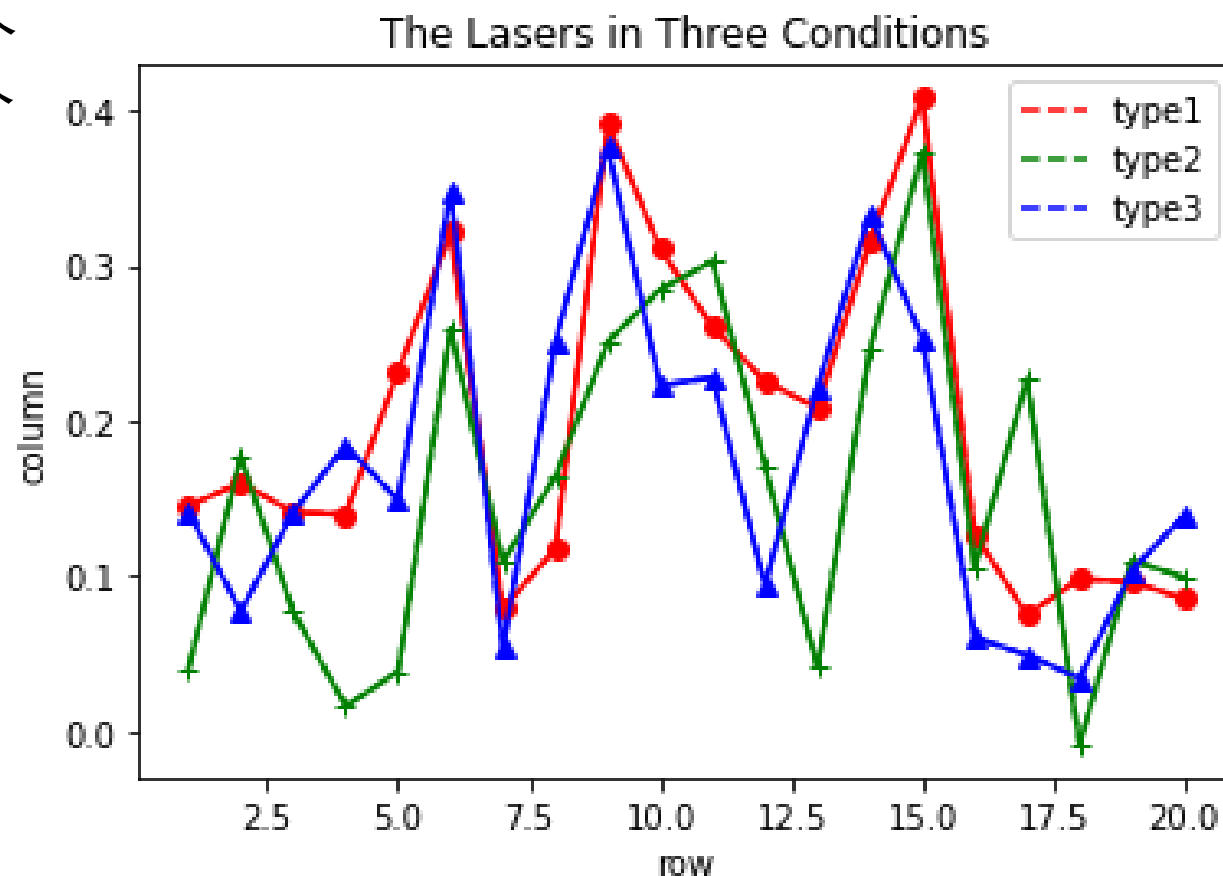
# 第九次讨论会

- 分别进行三次实验，随机选择20个点，以下三个实验都是在这二十个点上进行的。

Type1: hidden size=60&learn time=500

Type2: hidden size=60&learn time=750

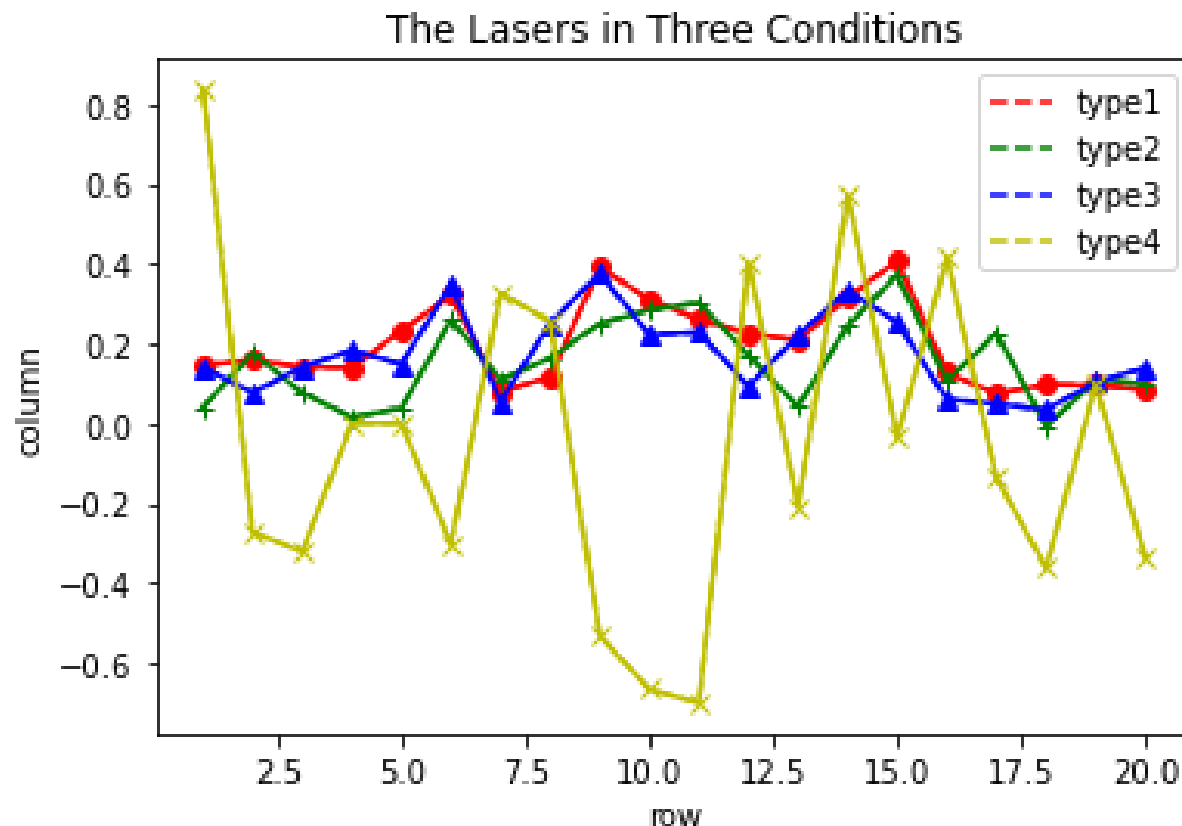
Type3: hidden size=90&learn time=500



# 第九次讨论会

在前面实验的基础上，做了新的实验。

前面的实验，是对后面预测的所有数据计算皮尔森系数，点太多可能影响计算的结果，于是，令Type4: hidden size=60&learn time=500，计算十天的皮尔森系数进行对比。



# 第九次讨论会

## • 实验结果分析

- 基于AUC思想，因为时间间隔一样，只计算各皮尔森系数的累加和即可，由此得到
- Type1: 3.9361343399950908
- Type2: 3.082052572559422
- Type3: 3.4592690555272188
- Type4: -0.9713070898523788

根据这个以上结果，和前面的图像，可以看出，增加训练次数，不会增加预测的效果，增加hidden的数量，也不会提高预测的准确性。

使用十天的数据计算皮尔森系数，会在某些点提高数据预测的准确性，但其中有偶然性的因素。

# 第九次讨论会

使用十天的数据，计算了所有空间点上的皮尔森系数  
得到了如下的结果：

max:0.9858583690991279

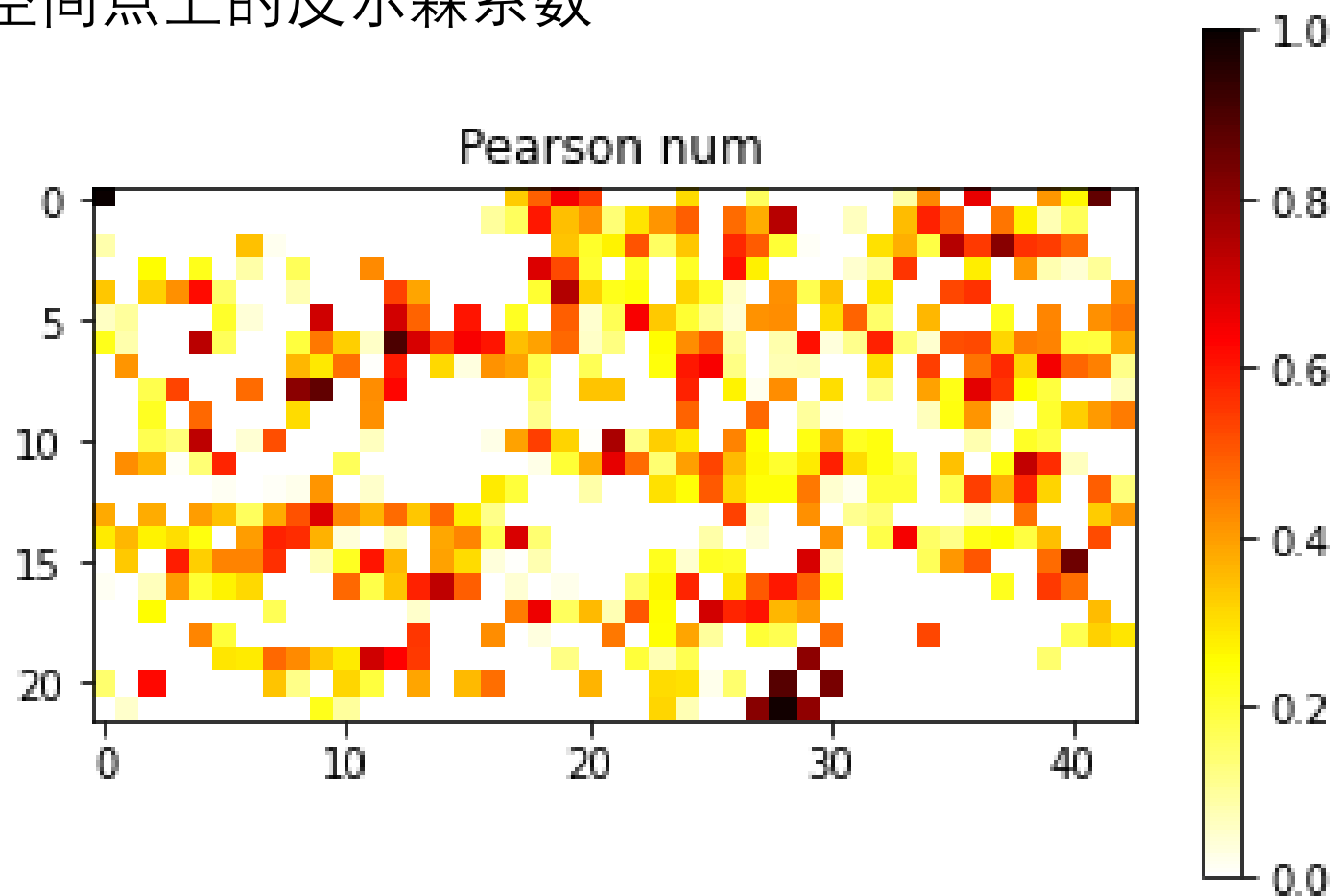
min:-0.9594048399046654

>0.7 sum=24 占比： 2.37%

>0.6 sum=49 占比： 4.84%

>0.5 sum=101 占比： 9.98%

>0.4 sum=179 占比： 17.69%





# 第九次讨论会

- 总结

$$f(x, y, t) = F_1(x, y) \times G_1(t) + F_2(x, y) \times G_2(t) + \dots$$

将预测的皮尔森系数画在世界地图上

# 第九次讨论会

2020年12月17日需要做的工作总结

- 1、画出带bar的世界地图上的热力图……………第三、画出带bar的热力图
- 2、运行20天之后的数据……………第一、先跑数据
- 3、画出之后利用5维cio数据跑的结果……………第二、画出五维的cio数据
- 4、整理ppt, 做好明天的演讲……………第四、预先准备下ppt
- 5、之后整理好数据, 井井有条, 好用好找。……之后工作, 整理数据

Ps: 明天带笔记本电脑

# 第十次讨论会

预测20天之后的降水量

皮尔森系数 $>0.3$  82 8.7%

皮尔森系数 $>0.4$  32 3.4%

皮尔森系数 $>0.5$  11 1.2%

皮尔森系数 $>0.6$  3 0.3%

# 第十次讨论会——(2020/12/18 和于老师、谭老师、周老师一起开会讨论。)

- 总结

- 1、降水量问题，要求是在感兴趣的区域中，有效果。
- 2、最好，优化效果在季节内。
- 3、计算一年内的，季节内的相关系数即可。
- 4、降水量可以进行带通滤波。
- 5、cio数据尽量使用一维。

# 第十次讨论会

## 下一步工作:

- 1、调整预测的点阵位置 $[-10,30][60,140]$ (ps:不是很确定)
- 2、使用一维的CIO数据, 一定要特别探究, 1Dcio数据与5Dcio数据对结果的影响, 如果影响给不大, 就使用一维的CIO数据。
- 3、进一步优化模型, 提高程序的运行效率。

## 疑问:

- 1、下一步优化模型怎么搞?
- 2、可不可以用到给出的别的参数。

# 第十一次讨论会

下一步工作：

- 1、将数据进行带通滤波，得到合适的结果。
- 2、使用原来的CIO和带通滤波后的降水量数据，对降水量情况进行预测。
- 3、使用project\_CIO和带通滤波后的降水量数据进行预测。

疑问：

- 1、下一步优化模型怎么设置？

# 第十一次讨论会

- 工作内容
- 1、将降水量信息进行带通滤波。
- 2、用原来的C10数据，和带通滤波后的降水量数据，预测降水量信息，得到结果，并画出图。
- 3、将project数据进行合适的分离，得到直接可用的数据，然后再使用带通滤波后的降水量数据，进行预测。

# 第十一次讨论会

## • 试验一 ——CIO数据与带通滤波后的降水量数据作为输入

最大值——0.7431793660837418

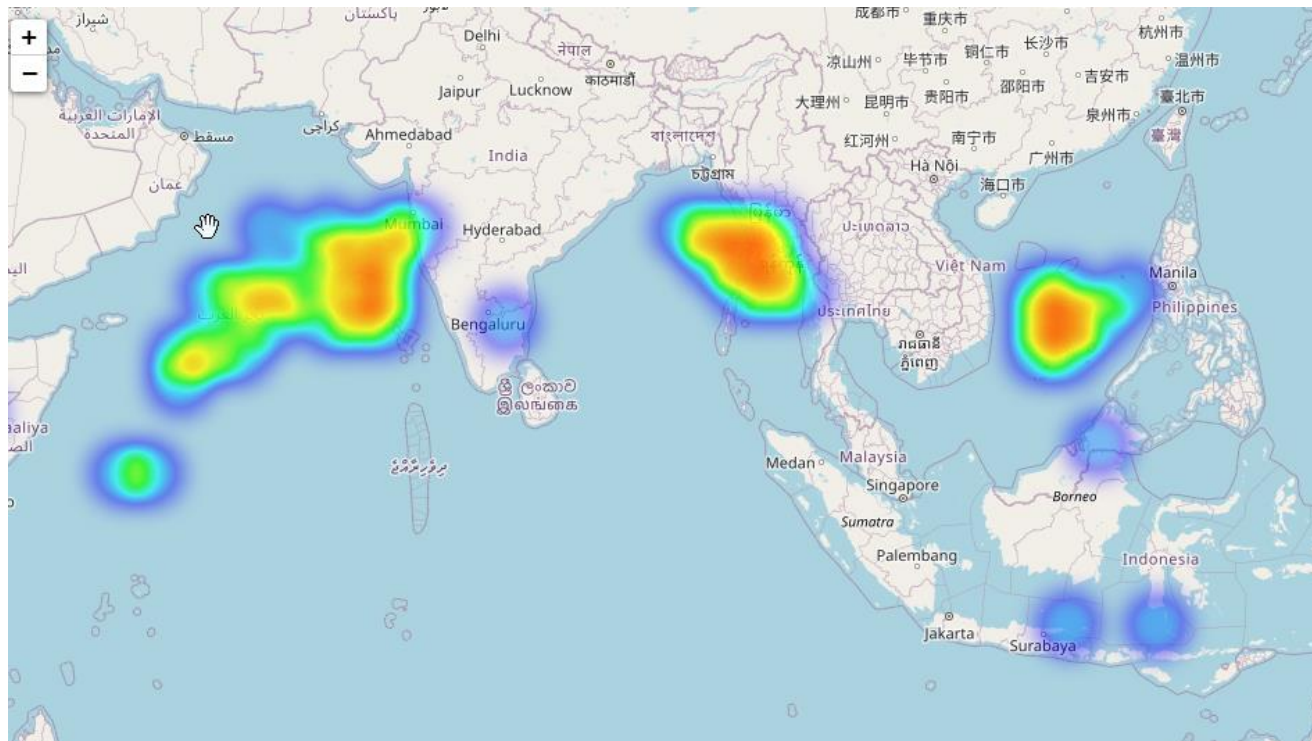
最小值—— -0.3256293158308732

皮尔森系数大于0.6    14    1.4%

皮尔森系数大于0.5    48    4.7%

皮尔森系数大于0.4    103    10.2%

皮尔森系数大于0.3    203    20.1%





# 第十一次讨论会

## • 试验二 ——proj\_CIO数据与带通滤波后的降水量数据作为输入

最大值——0.7431793660837418

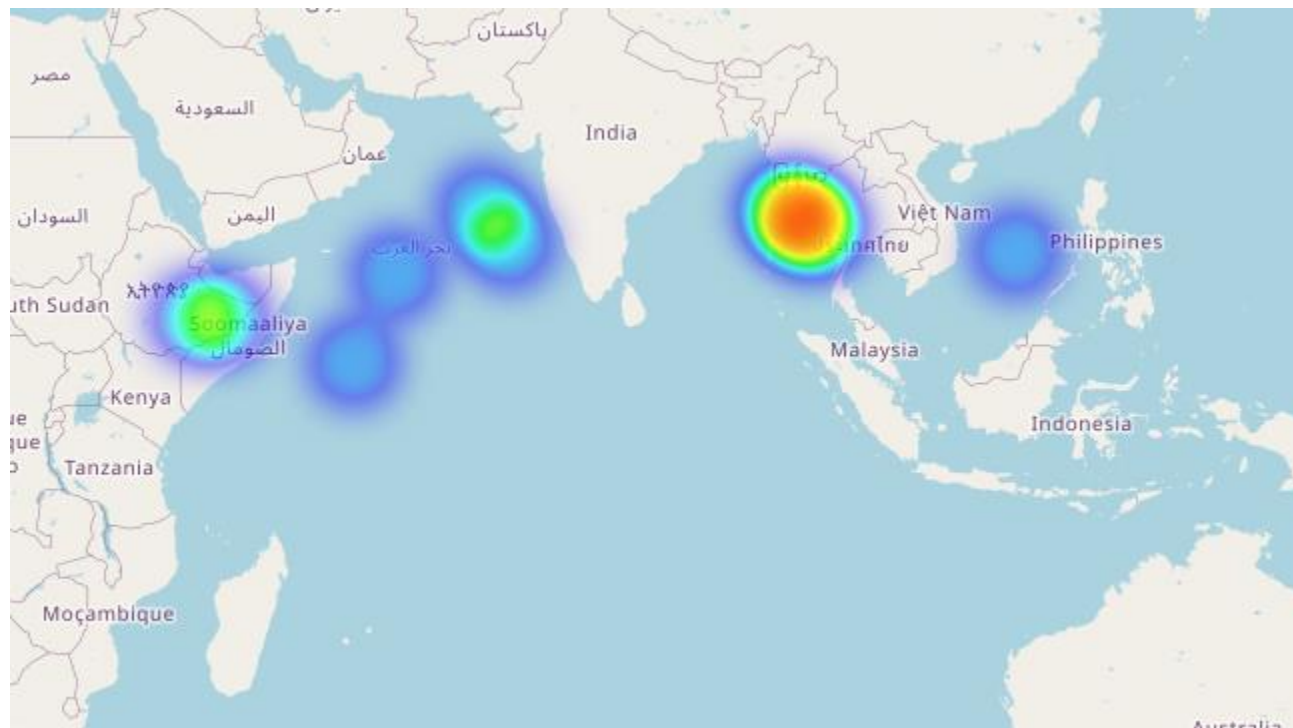
最小值—— -0.3256293158308732

皮尔森系数大于0.6    5        0.5%

皮尔森系数大于0.5    14       1.4%

皮尔森系数大于0.4    62       6.6%

皮尔森系数大于0.3    171      18.1%



# 第十一次讨论会

- 试验二 ——proj\_CIO数据与带通滤波后的降水量数据作为输入

更改后的模型

最大值: 0.6977332853612711

最小值: -0.22286163874439935

皮尔森系数大于0.6的: 17

皮尔森系数大于0.5的: 57

皮尔森系数大于0.4的: 134

皮尔森系数大于0.3的: 276

# 第十二次讨论会——线上联合讨论会

- 总结
  - 1、实验效果很好，但是是一些地方需要改进，进行标准化
  - 2、讲解LSTM
  - 3、写文稿
  - 4、模型创新（好像不用做）
- 工作
  - 1、画图（区域E40-E160,S10-N30，时间20天后的，皮尔森系数使用1年的来计算）。
  - 2、写文稿（这个不太懂）

# 第十二次讨论会——线上联合讨论会

- 不足
  - 1、区域需要调整
  - 2、周老师还是比较关心20天的情况
  - 3、计算皮尔森系数调成一年的数据