

## Abstract

## Data Source

The amount of pollutants observed in each station is provided by Yang Yu. The shape of China and all provinces can be found on GADM<sup>1</sup>.

## Interpolation

There are many multivariate interpolation method for 2D interpolation, such as bilinear, bicubic methods (wiki). However, when applying these methods, there are always some empty parts in the map. It is because these methods are used to interpolate regular grid, but the position of stations does not form a regular grid. For this sort of scattered data, radial basis function(RBF) is used for interpolation in the paper (wiki). The function form is rather simple. For any point  $\mathbf{x}$ ,

$$y(\mathbf{x}) = \sum_{i=1}^N w_i \phi(\|\mathbf{x} - \mathbf{x}_i\|). \quad (1)$$

$\mathbf{x}_i$  are those observation points and  $\phi$  is a customized function. The estimation of  $w_i$  can be solved by linear least square.

Below several different functions are tested respectively.

In these images, those interpolated by “quadratic” methods have a lot of empty parts caused by negative value. However, this does not happen in “linear” methods. When comparing “linear” methods  $1/(1+r)$ ,  $e^{-r}$ ,  $r$ , the images in Xinjiang are quite different. In  $e^{-r}$ , concentration degrades so rapidly that only points close to observation station are not green. In contrast to  $e^{-r}$ ,  $r$ -interpolation degrades much slower and most of Xinjiang is red.  $1/(1+r)$  is the intermediate between  $e^{-r}$  and  $r$ . Due to lack of data, we cannot determine which function is the best. For simplicity,  $r$ -interpolation is used throughout the project.

One of the advantage of RBF is that  $y$  is linear. Therefore, when it needs to compute the average value interpolation, calculating interpolation in each hour and then taking average for each point are not necessary. We can directly compute the average value for each station and then calculate the interpolation, which saves much of the time.

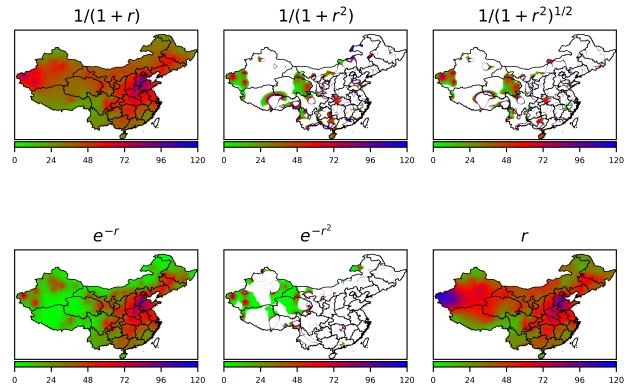


Figure 1: Interpolation of annually average concentration of PM2.5 using different  $\phi(r)$

## Space Distribution

The high accuracy distributions in China of PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, O<sub>3</sub>, and CO, which are the main air pollutants, are got by the data and interpolation method mentioned above. The regularities of space distributions of these air pollutants are summarized in this part.

By an overview of these figures, a general impression of the air quality over China can be got.

In the figures which represent PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, and CO, there are two areas which are in the south of Xinjiang and around Hebei with the most serious pollution, which are the blue areas in the figures. The distribution of the red areas, which are the areas with pollution not so serious, is radial from these two points. In the south of China and the northeast of China, the pollution is slight in comparison.

It is special for the figure which represent O<sub>3</sub>. It shows that the distribution is average and that the area with the maximum concentration of O<sub>3</sub> is around the northeast of Qinghai, the middle of Shaanxi and the west of Inner Mongolia.

In the sheet, the pollutants are considered in seven main

<sup>1</sup><https://gadm.org/index.html>

Region	PM2.5	PM10	SO2	NO2	O3	CO
Northeast	42.13	70.99	20.88	22.53	58.06	0.86
East	51.32	84.23	24.16	28.49	56.18	1.00
North	43.33	84.73	28.48	25.69	63.58	1.08
Central	61.08	97.41	24.57	27.23	54.96	1.16
South	36.28	55.87	15.45	20.07	52.06	0.95
Southwest	33.46	69.44	17.64	19.15	55.06	0.93
Northwest	54.36	131.96	24.10	24.96	59.49	1.19

Table 1: Annually average concentration of pollutants in seven regions

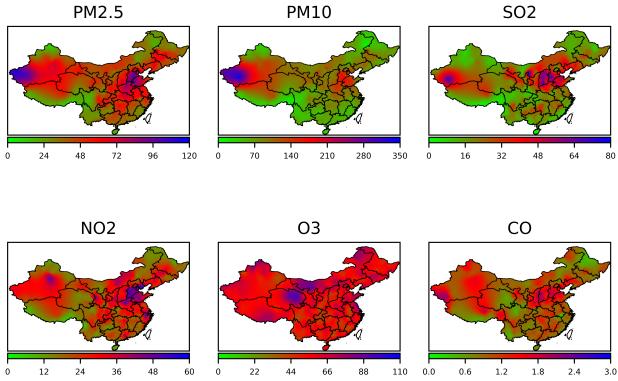


Figure 2: Distribution of annually average concentration of pollutants

regions of China, which are Central China, North China, South China, East China, Northeast, Northwest, and Southwest. The sheet shows the annual average concentrations of different pollutants in different regions, which is calculated by the result of interpolation method. In keeping with the overview, the annual average concentrations of all pollutants in Northwest, South China, and Southwest are obviously smaller than those in Central China, North China, East China, and Northwest.

As it is mentioned above, the two regions with the most serious pollution in China are in Xinjiang and around Hebei. However, the factors that make the high pollution are not the same for these two regions. The heavy industry with high emission in Hebei, such as iron and steel industry and cement industry, is one of the factors that make the poor air quality. However, the factors in Xinjiang are most natural factors but not man-made, such as climate and physiognomy. For example, the desert and sand storm will increase the concentration of particulate matter a lot.

There is something that should be explained. Because of the maldistribution of the observation points, the estimation by interpolation in some regions, such as in the west of China and Inner Mongolia, may be different with the fact. The whole Inner Mongolia is considered in the North China. As a result, it is not strange to see that the pollution in North

China is not so serious, although the pollution in Beijing-Tianjin-Hebei Region is the most serious as we know.

## Time Distribution

The data also shows some regularities of the time distribution of pollutants. In this part, the regularities will be analysed in one day, in a week, and in the whole year. That is because human activities, such as work and rest, changes in one day cycle, in one week cycle, and in one year cycle. In the whole year, the climatic conditions change with the seasons, which may be the factors that affect pollutants concentration. Thus, it is a reasonable guess that the concentration of pollutants will have the certain regularity in one day, in one week, and in the whole year.

X

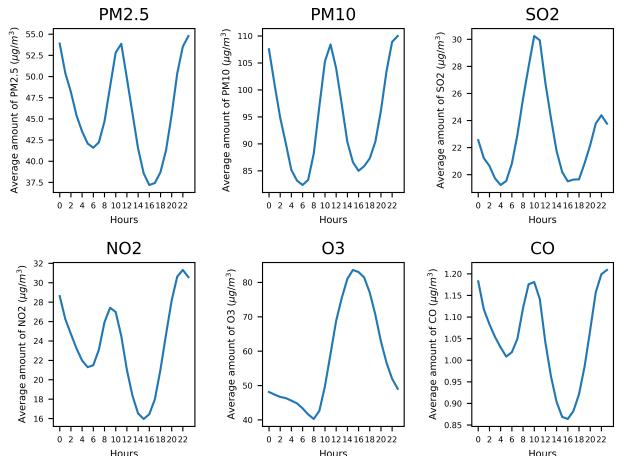


Figure 3: Daily average concentration of pollutants

The figures shows the change curves of average concentration over 24 hours for different pollutants. The curves of PM<sub>2.5</sub>, PM<sub>10</sub>, CO, and NO<sub>2</sub> are similar, which have two peaks with similar height. one peak appears around 10h, and the other peak appears around 22h. The curve of SO<sub>2</sub> also has two peaks, but the peak appears around 22h is much lower than the other. The curve of O<sub>3</sub> is special, which only has one peaks that is around 15h.

According to some previous researches (yuyang), human daily activity is a factor of the regularity. The peak around

10h, which is time for trips, appears because of the traffic. The concentration increases at afternoon because of human activities after work, such as traffic and cooking. The concentration reaches another peak around 22h because the industrial production at night. The curve of  $O_3$  seems strange, because being different with other pollutants, it is produced at high temperature from volatile organic compound and oxynitride. Thus its curve reaches the peak around 3h, when the temperature is highest.

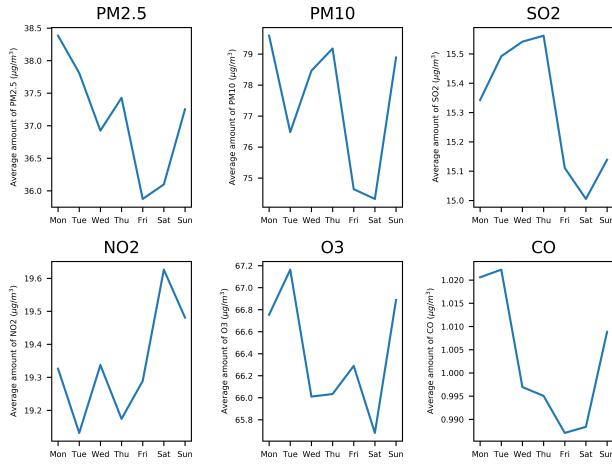


Figure 4: Weekly average concentration of pollutants

The figures shows the change curves of average daily concentration from Monday to Sunday. However, the curves seem zigzag and irregular, which can not reflect any regularities of weekly concentration changes. That means that the regularity of human weekly activities is not shown in the concentration changes.

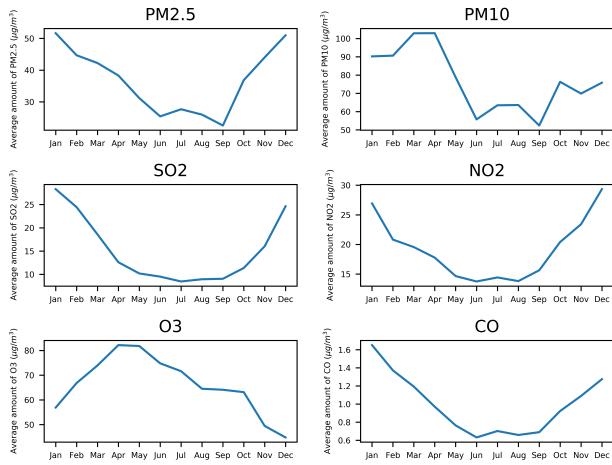


Figure 5: Monthly average concentration of pollutants

The figures shows the change curves of average monthly concentration over the whole year. As it shows, the shape of concentration curves of  $PM_{2.5}$ ,  $PM_{10}$ ,  $SO_2$ ,  $NO_2$ , and  $CO$

is like a letter U. The concentration decreases from January to June, keeps a low value from June to September, and increases from September to December. The curve of  $O_3$  is special again. It is an inverted-U with a peak around April and May.

## Reason (TODO)

### The Independence of Space Distribution and Time Distribution

The regularities of space distribution and time distribution of pollutants have been concluded above. Another valuable topic to discussed is the independence of space distribution and time distribution. In other words, that is the space universality of time distribution and the time universality of space distribution. To simplify the analyse, AQI is used to represented the integrative concentration of six main air pollutants.

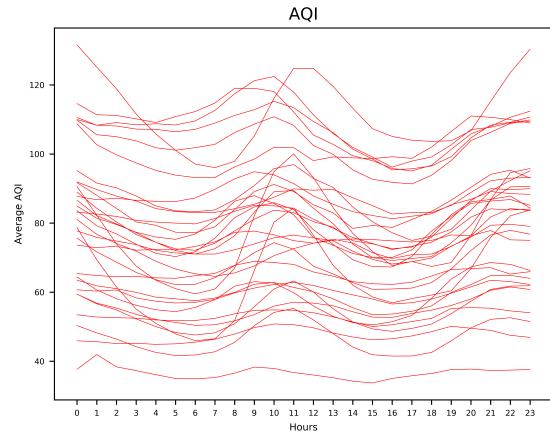


Figure 6: Daily average concentration of AQI for all provinces except Taiwan (lack of data)

The figure analyses the change curves of average AQI over 24 hours for every province in China. The shape of curves in the figure are all similar, which shows that the regularity of daily distribution is general in all province in China. It is interesting that some provinces with low pollution have curves with little fluctuation. For example, the curve that represents Hainan, which is at the bottom of the figure, is nearly a horizontal line. This educes a reasonable educt that the regions with good air quailty have less changes of AQI.

The figure shows the change curves of monthly average AQI over the whole year for every province in China. It can be seen that all curves are roughly U-shaped, which conforms with the regularity summarized above and shows that the regularity is spacial general. Similar to the daily distribution, the conclusion that lower AQI is with less changes is also applied in this figure.

It can be found that the curve of Xinjiang, which is at the top of the figure, is the most special. A possible reason

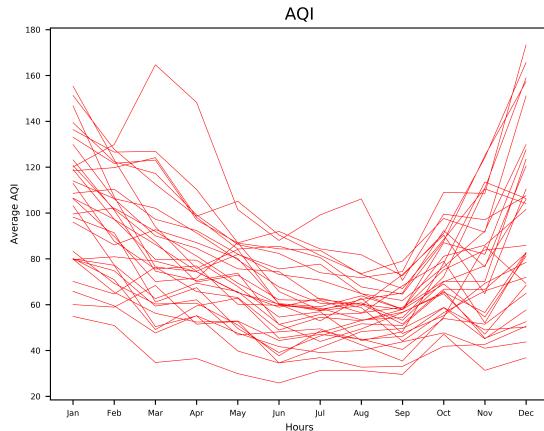


Figure 7: Monthly average concentration of AQI for all provinces except Taiwan (lack of data)

is the unique natural and geographical conditions in Xinjiang. Another interesting phenomenon is that the AQI increases quickly between September and October in many provinces, which means that this is a general phenomenon in China. Similar to the analyses for 2014 in former research, the reason may be some special pollution processes in October, which leads to the decline in air quality across China. However, it can not be proved by the data.

The figure is the space distribution of AQI in China in spring, summer, autumn, and winter. It can be found that the distribution does not change a lot in four seasons, which means that the space distribution is general in the whole year. The visible difference is the integral level of AQI, and the regularity has been discussed above. As the figure shows, the AQI in spring and winter is higher than that in summer and autumn, which matches the conclusion above.

According to the discussion above, the independence of space distribution and time distribution of air pollutants has been verified preliminary.

### **Comparison with Standard**