Warsaw University of Technology

FACULTY OF BUILDING SERVICES,
HYDRO AND ENVIRONMENTAL ENGINEERING

Advanced R!

Surface Water Protection

Supervisor: Paweł Gilewski

05.01.2020
Zhiqi Wang
student ID: 309319

1. Abstract

In this report, I tried the interpolation in R! instead of data processing in GIS. I load a set of data about precipitation and run it in R to create maps, then I used several methods to modify and interpolated the map, such as Thiessen polygons and IDW, incase to make my predictions for Texas rainfall more accurate step by step. Moreover, I will also modify the details of the map I obtained and gave the analyzation.

2. Introduction

In case to predicted the precipitation in Texas, I runs a set of data about the precipitation in Texas and got the sampled precipitation, then I tried several methods to interpolated it, to predict the precipitation in Texas and compare the different predicted map to analyze which one is better.

3. Method

3.1 sampled precipitation

The sampled precipitation map shows in bellow, it can be seen from the figure that there are a total of 21 monitoring points in the Texas area, with rainfall ranging from 10 to 60 inches, divided into 5 classes, which are indicated by different colors.

Rainfall is low in the west and high in the east, and it is increasing from left to right.

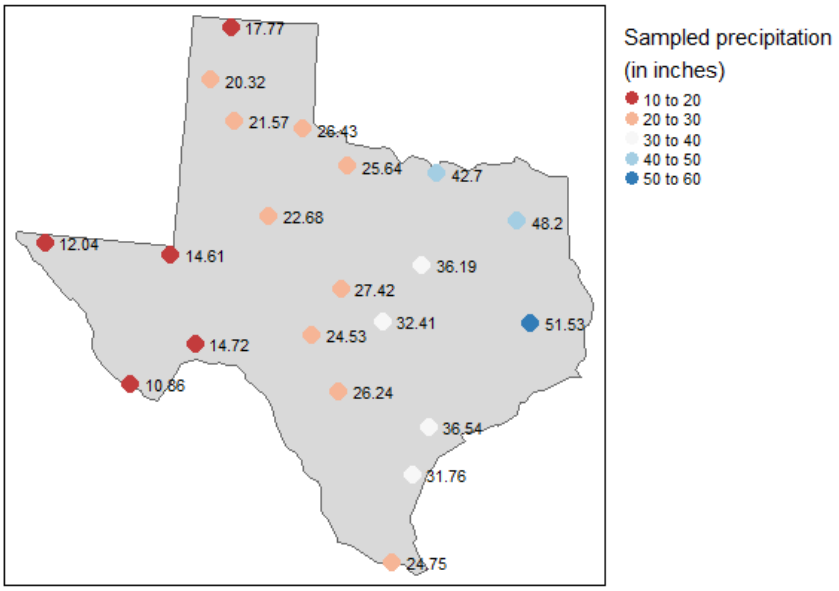

Fig 1 – The map of sampled precipitation in Texas

3.2 Thiessen polygons

Through the applied of the Thiessen polygons (or proximity interpolation) I used the Spatstat function, this method can divide an entire area into different small areas according to the different data.

In the figure 2, a) is the original image, and b) is the image after I changed the color to gradient blue **[1]**.

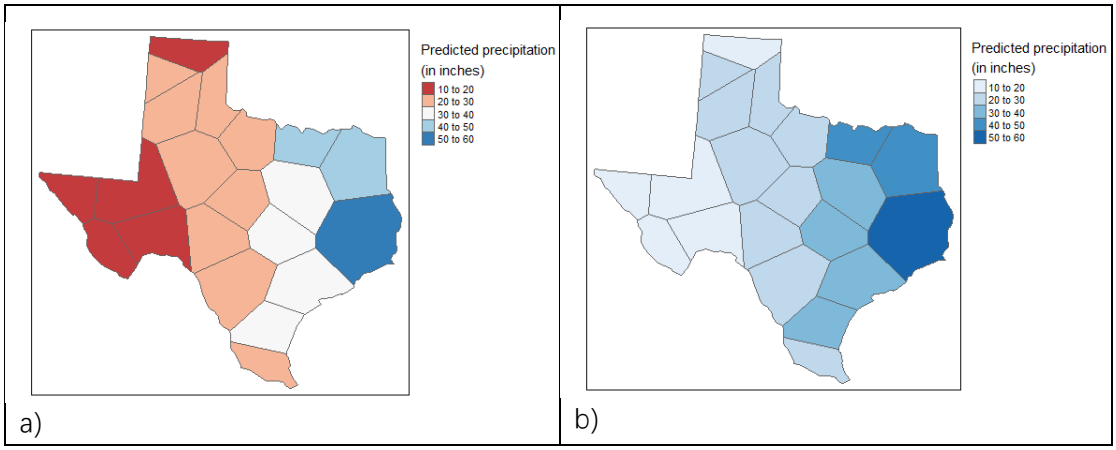

Fig 2 – The maps of predicted precipitation by Thiessen polygons

3.3 IDW

On the basis of Thiessen polygons, I applied the interpolation of IDW, normally the IDW displays the single monitoring point by color, in this phase, It blurs the boundaries of the polygon, the range of different colors clearly shows the stepwise geographical change of local precipitation. In addition, the rainfall classification on this map has also become finer, with as many as 9 classes.

In the figure 3, a) is the original image, and b) is the image after I changed the color to gradient blue.

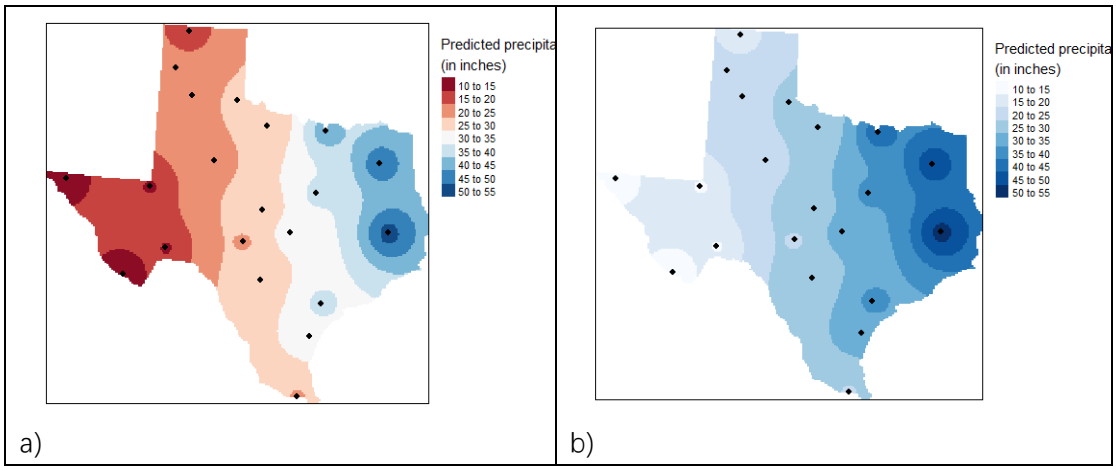

Fig 3 – The maps of predicted precipitation by IDW

3.3.1 Fine-tuning the interpolation

The data at this step of the map is still too subjective, so I use Fine-tuning the interpolation to reduce the error.

The red dotted line in the figure below is the observed value, and the black solid line is the predicted value after interpolation. When I change the idp, the predicted value changed, and the root mean square error RMSE **[2]** also changed,

So I can find the smallest deviation between the observed value and the simulated value, which is when idp = 2, the RMSE = 3.804498

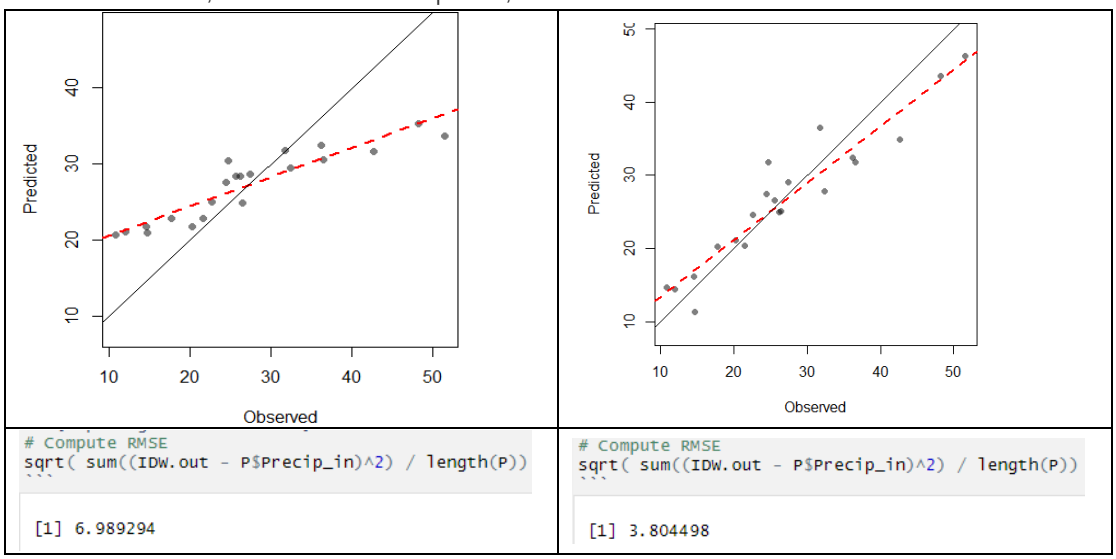

Fig 4 – the predicted curve after fine-tuning

3.3.2 Cross-validation

Next, I create a 95% confidence interval map of the interpolation model. It can be seen from the figure below that the leftmost points are dark in color, the middle point is very light, and the rightmost point is a little darker. Because the color depth represents its confidence, so I can draw the conclusion: The middle monitoring point has higher reliability.

a) is the original image, and b) is the image after I changed the idp from 2 to 11, It can be seen that the levels of b) are relatively clear and distinct.

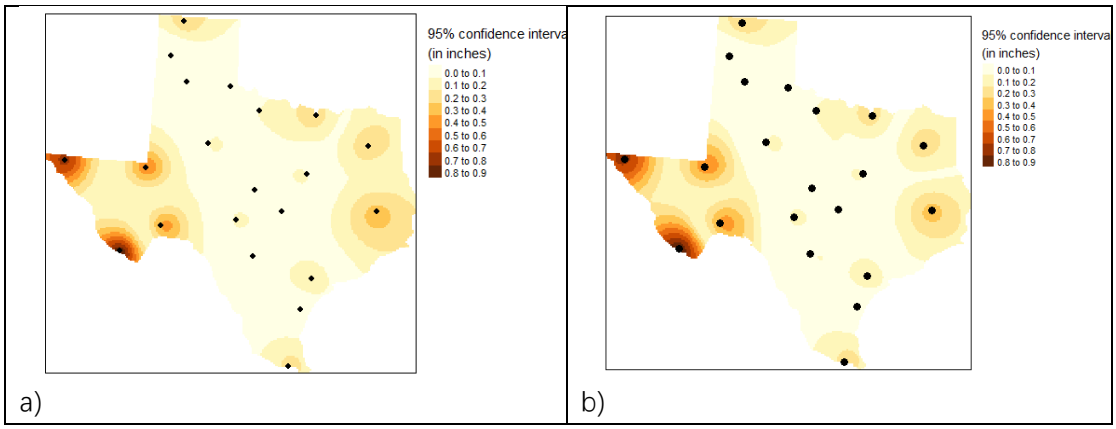

Fig 5 – the cross-validation

3.4 1st order polynomial fit

Next I fit a first-order polynomial to the interpolation model then draw the map, the data should fit the form of **precip=intercept+aX+bY**, the following map shows the results, It can be seen that the map is divided into multiple rectangular areas with diagonal lines, and the color of the map is stepped.

In the figure 6, a) is the original image, and b) is the image after I changed the color to "Spectral".

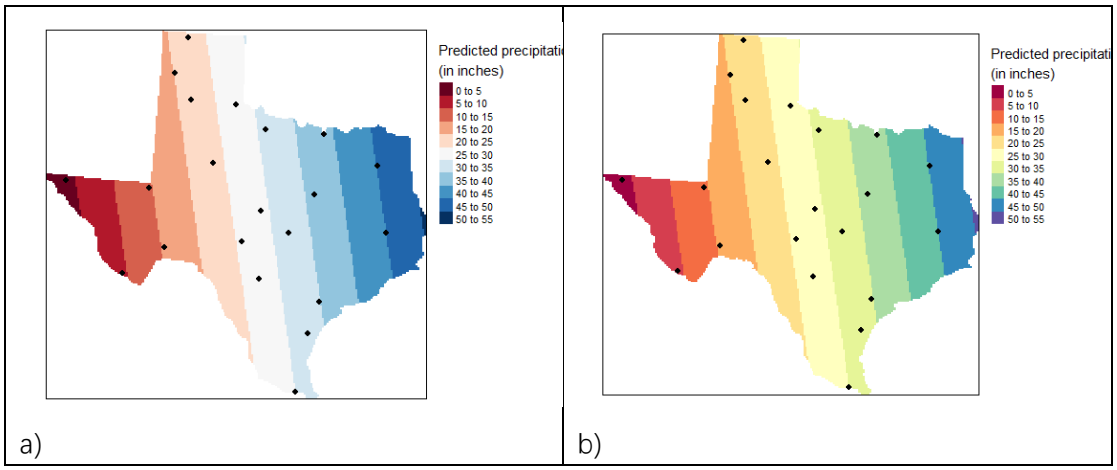

Fig 6 – the first-order polynomial

3.5 2nd order polynomial

Continue to modify the precipitation function, using the following second-order equation: **precip=intercept+aX+bY+dX2+eY2+fXY**, this second-order polynomial divide precipitation map with curves rather than straight line, thus, is more accurate and intuitive than the first-order fit.

In the figure 7, a) is the original image, and b) is the image after I changed the color to "YlGn" and I also changed the layers n from 12 to 7.

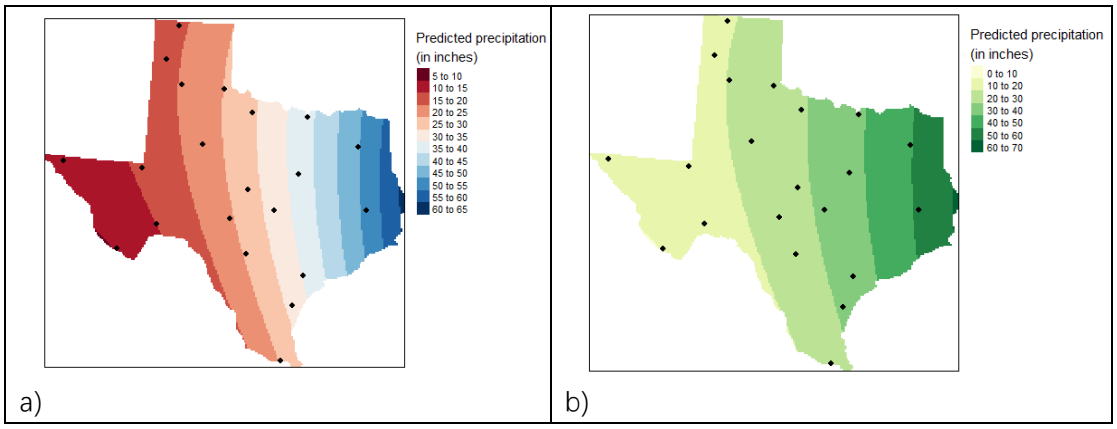

Fig 7 – the second-order polynomial

3.6 Kriging

From this part, I will find the regularity from the existing data, and then simulate the overall rainfall situation in this area, rather than starting from a single point to extrapolate.

3.6.1 Fit the variogram model [3]

In this part, I created a variogram model which need computed on the de-trended data.
The variogram is defined as the variance of the difference between two spatial points in the space.

As can be seen from the figure below, after the distance between the two points is taken to a certain large, its semivariance tends to stabilize.

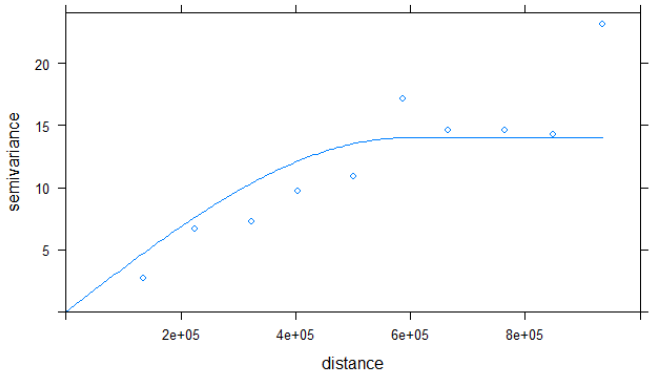

Fig 8 – the function of distance to semivariance

3.6.2 Generate Kriged surface

Next, use the variogram model to generate a kriged interpolated surface. It allows us to include a trend model and thus estimate unknown points on the surface.

In the figure 9, a) is the original image, and b) is the image after I changed the color to "Blues" and I also changed the form of legend “tm_legend(legend.outside=FALSE)” **[4]**.

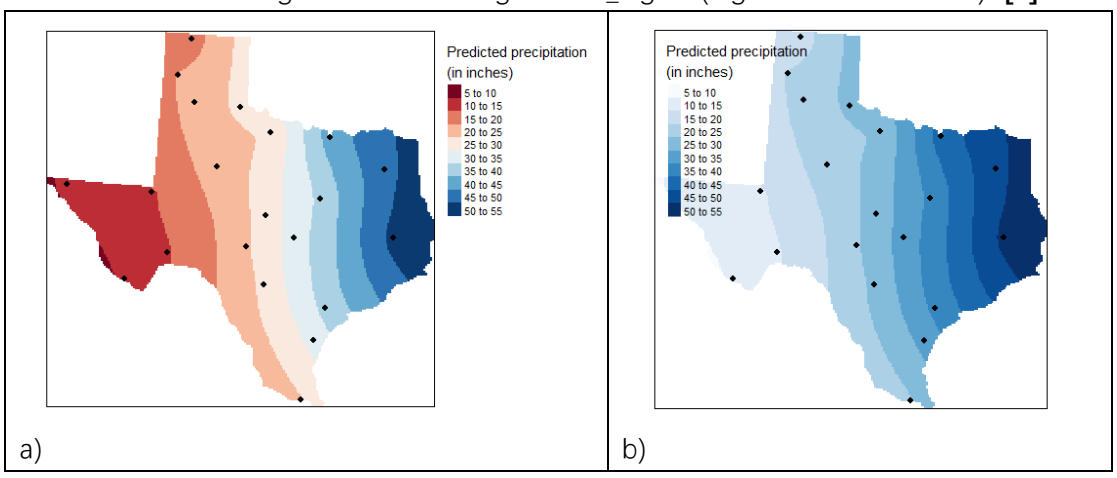

Fig 9 – generate Kriged surface

3.6.3 Generate the variance and confidence interval maps

At the end, I can also obtained variance and confidence interval (CI) maps, which proves the predictions for Texas precipitation from the side.

In the figure 10, a) is the original image, and b) is the image after I changed the color to “grays”.

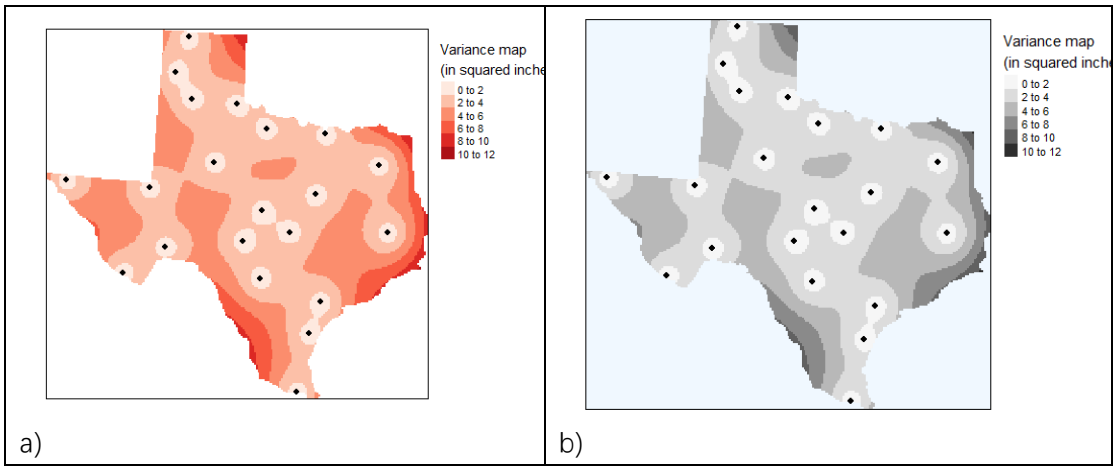

Fig 10 – the variance map

The 95% confidence interval map is more credible, this map can be simply understood as: the smaller the value, the lighter the place, the higher the reliability of the data.

In the figure 11, a) is the original image, and b) is the image after I changed the color to "YlOrRd", and the layers n changed from 7 to 12.

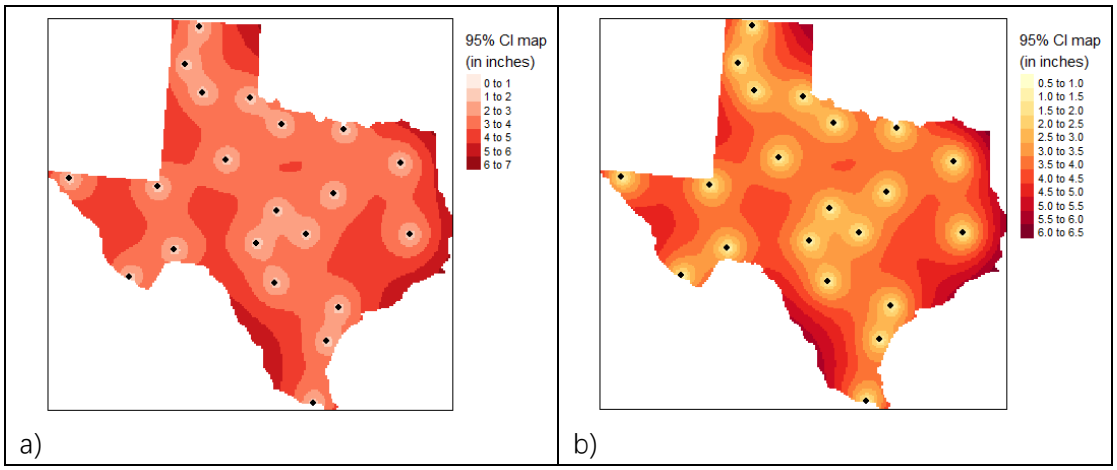

Fig 11 – the confidence interval (CI) map

4. conclusion

This report can be seen as two parts, the first part starts from Thiessen polygons, and finally gets the predicted precipitation after a series of interpolations. The second part starts from the first-order polynomial fitting, and finally gets the predicted precipitation by the Kriging method.

First part:

For predicted the precipitation in Texas in R!, I runs a load of data about the precipitation in Texas and got the sampled precipitation, then I used Thiessen polygons to polygon the map.

Next, applied the interpolation of IDW to blurs the boundaries of the polygon, clearly shows the stepwise geographical change of precipitation.

Next, I use Fine-tuning the interpolation to reduce the error, and found out that the idp can influence the root mean square error RMSE, when the idp = 11, the RMSE have smallest value = 3.804498, besides, I create a 95% confidence interval map to further accurate the data.

Second part:

In this part, my aim is to simulate the overall rainfall situation in this area, rather than starting from a single point to extrapolate.

After load the precipitation data, I applied first-order and second-order polynomial fitting to the map, make the map look more intuitive and change more regularly.

Next I applied Kriging method, I created a variogram model, then use the variogram model to generate a kriged interpolated surface.

Besides, I also got variance and confidence interval (CI) maps to helps to proves the predictions for Texas precipitation.

Compare:

Compare the map got from these two parts (IDW VS Kriging), I prefer to choose the second map because the first map has too little data to represent the overall trend, and the second one can predict the area as a whole through the calculation and fitting of the data.

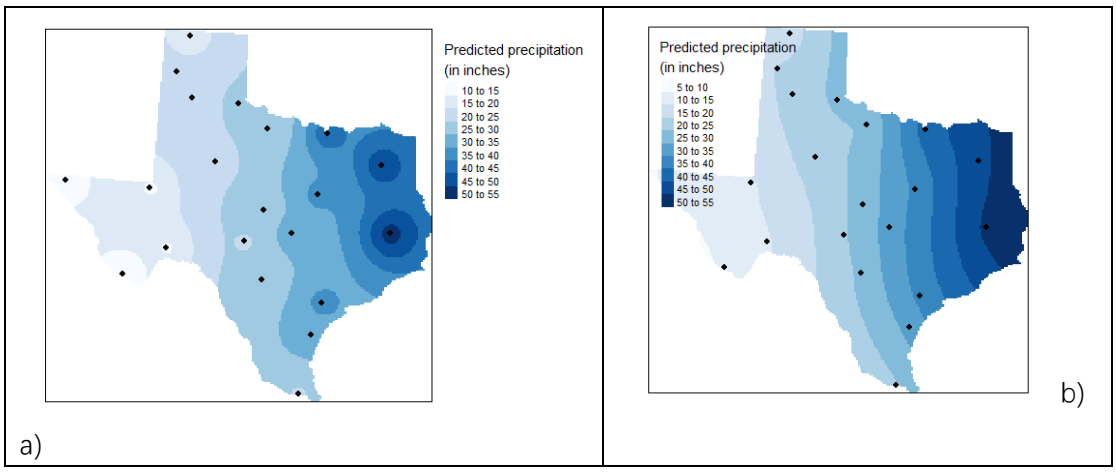

Fig 12– the compare of the predicted precipitation map between IDW and Kriging

5. Reference

[1] http://iccm.cc/colors-and-palettes-in-r-language/

[2] https://www.zhihu.com/question/24454827

[3]https://baike.baidu.com/item/%E5%8F%98%E5%BC%82%E5%87%BD%E6%95%B0/3737509?fromtitle=%E5%8F%98%E5%B7%AE%E5%87%BD%E6%95%B0&fromid=9070007

[4] https://blog.csdn.net/glodon_mr_chen/article/details/79496403