Precipitation prediction based on machine learning method
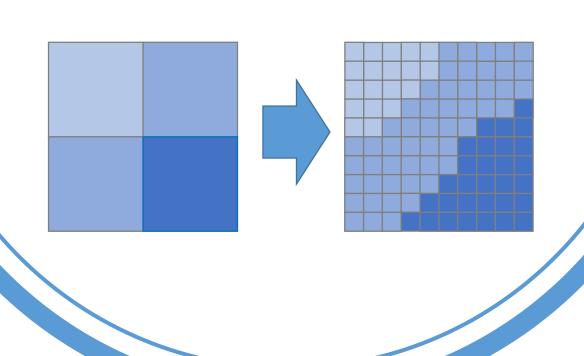
Jingwei Li (✉), Klug Hermann, Mittlböck Manfred

E-mail: s1061395@stud.sbg.ac.at

UNIVERSITÄT SALZBURG

Z_GIS

Downscaling Method

The resolution of original precipitation data is 25km. Based on ANN, the resolution of final result is 5km.
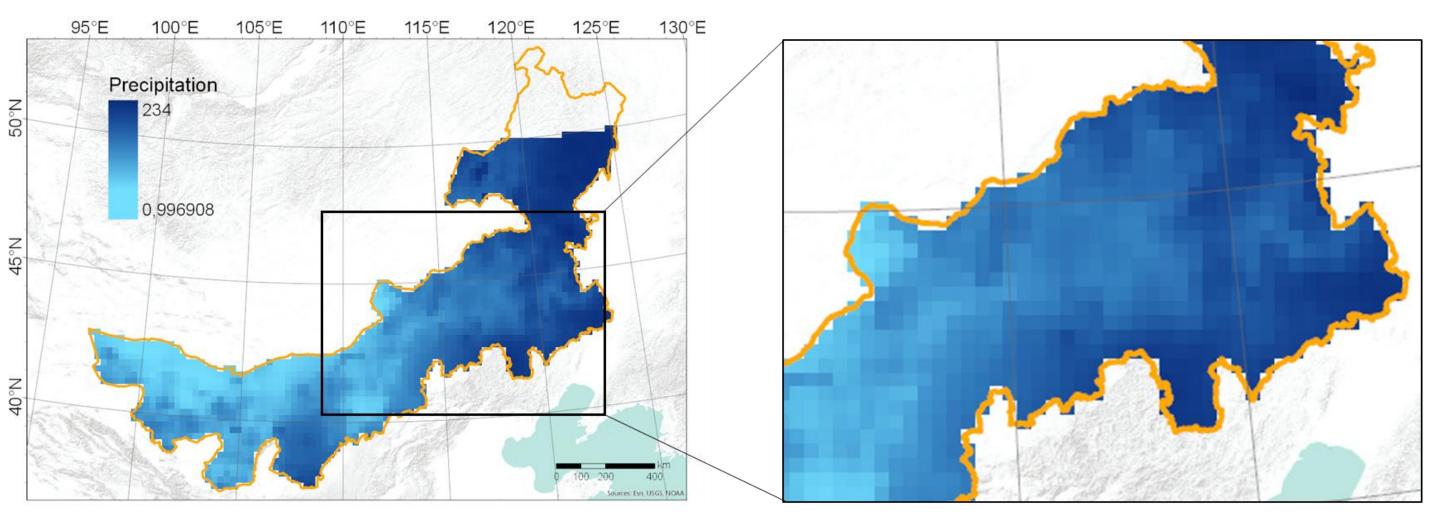
Precipitation observation based on satellite cannot cover all the place on the earth and the resolution is not enough for scientific researches. Artificial Neural Network (ANN) was used in the study. we predicted the distribution of precipitation based on normalized difference vegetation index (NDVI) and digital elevation model (DEM). Tropical Rainfall Measuring Mission (TRMM) data of June, 2018 was used for training and validation. We selected Inner Mongolia province, China as case study area, 70% data was used for training and 30 % for validation. The results was also compared with linear model. The results show that the method can efficiently predict the spatial distribution of precipitation.

Problems

Nowadays, Tropical Rainfall Measuring Mission (TRMM) is most common used precipitation data based on satellite observation. However, TRMM data cannot cover all the place on the earth and the resolution is not enough for scientific researches.

TRMM Monthly Data. The data only cover the area from 50°S to 50°N. The resolution of the data is 25km
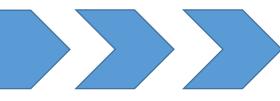
To solve these problems, the common used method is to establish a model based on precipitation data and some related factors. Traditional model is multilinear regression model (MLR) :

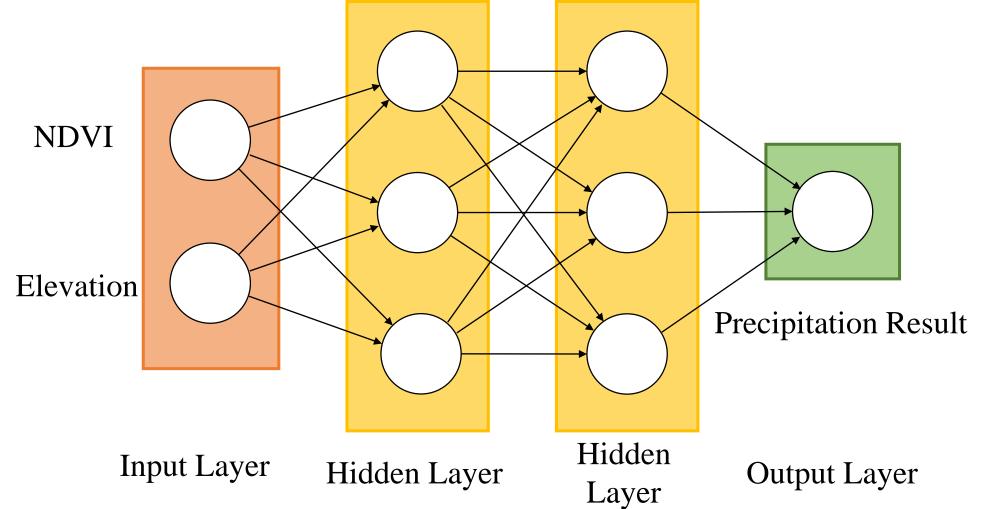$$\text{MLR Model: } y_i = a_0 + \sum_{i=1}^{n} a_i x_i$$

However MLR would lead to **overfitting problem** (means that MLR work well with training data but it work bad for predicting data).
And precipitation is considered as a **non-linear** procedure. A linear model can not predict precipitation well.
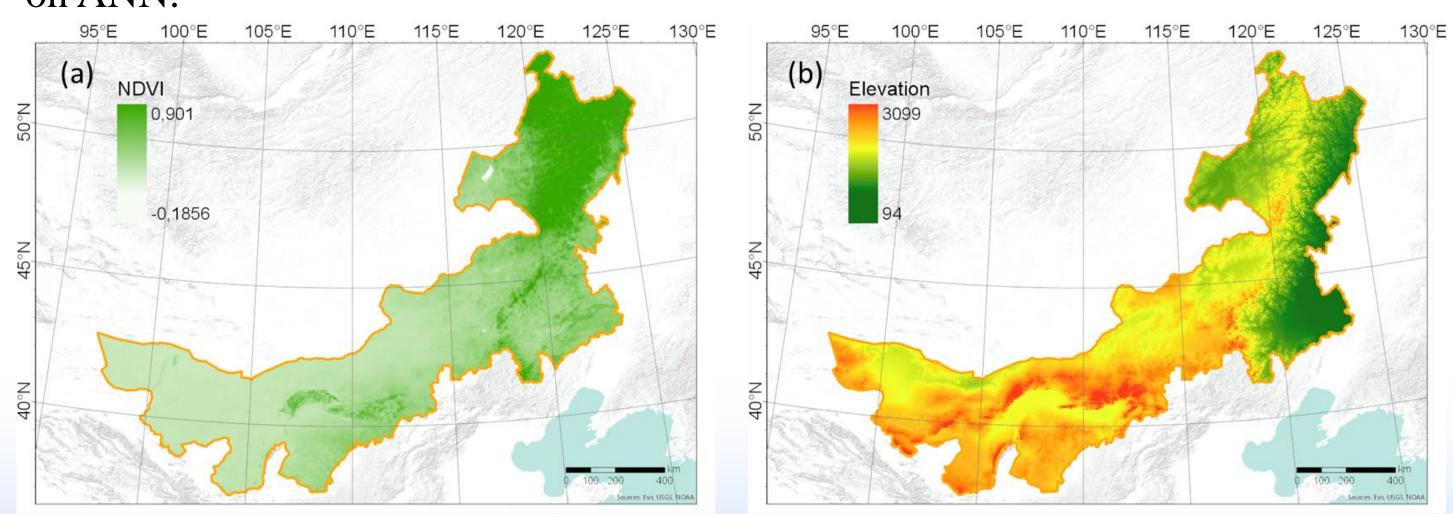
Method

Artificial Neural Network (ANN) is one of most popular machine learning method for predicting. It overcome overfitting problem. And it also a non-linear model used for prediction. ANN is also a non-linear method for predicting and it is more accurate than linear model.

Artificial Neural Network for predicting precipitation. Every circle in the figure is a neural. NDVI and elevation are regarded as important influence factors for precipitation. In this study, we use NDVI and elevation data for predicting precipitation based on ANN.
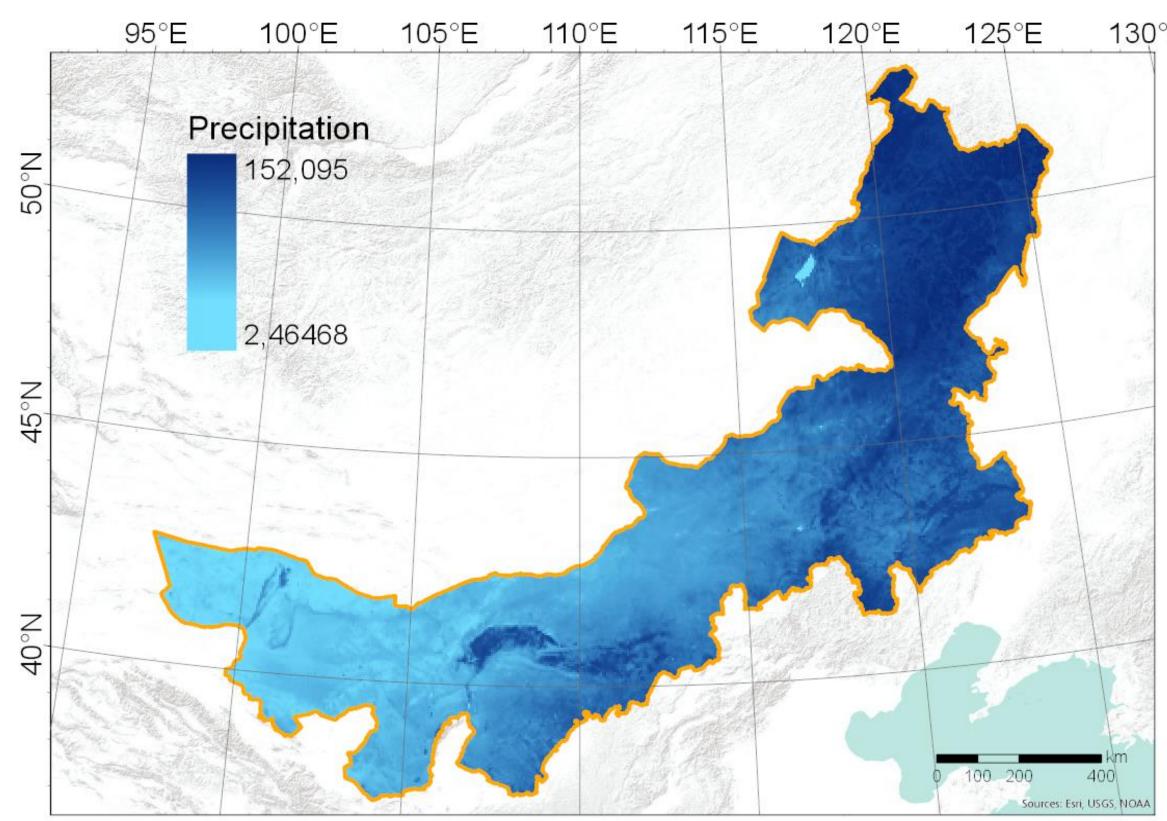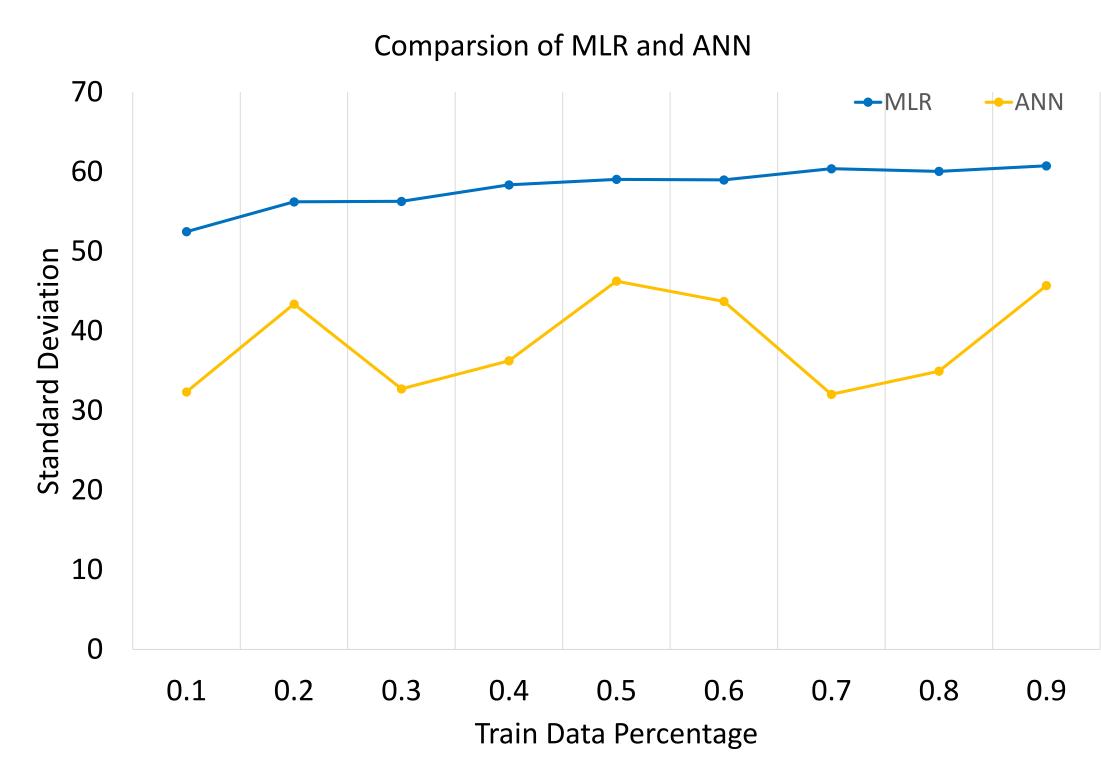
(a)NDVI Data (b)Elevation Data

Results

The results were obtained based on ANN method. 70% data (1173 sample points) were used for training and 30% (503 sample points) for validation.
The resolutions of elevation (5km) and NDVI (5km) is higher than TRMM data (25km). And the method not only can fill the uncovered part of TRMM data, but also can be a **downscaling** method (from 25 km to 5 km) to predict precipitation.

The predicted result based on ANN

The result was compared with MLR model. We used different percentage of training data to train the model of MLR and ANN. We found that the ANN method gave better result than MLR method.

Comparsion of MLR and ANN

The result was compared with MLR model. We used different percentage of training data to train the model of MLR and ANN. We found that the ANN method gave better result than MLR method.

Conclusion

Artificial Neural Network (ANN) is a good method for predicting precipitation. In the project, NDVI and elevation data were used as influence factors.

- We successfully predict the distribution of precipitation in study area. The uncovered part of TRMM data has been filled through the method.
- The method also can be used as a downscaling method to predict precipitation in the area.
- ANN work better than the traditional method (MLR).