

RICE YIELD MODELING IN CHINA USING CLIMATE DATA WITH DEEP NEURAL NETWORK

S. B. Cho (1), Y. W. Lee (1) *

¹ Pukyong National University, 45 Yongsoro, Nam-gu, Busan 48513, South Korea

Email: chosb401@gmail.com; modconfi@pknu.ac.kr

KEY WORDS: Climate Change, Grain Yield, Rice Yield, Deep Learning

ABSTRACT: Climate change is an important factor in vegetation growth, and it is very significant to understand the relationship between climate change and rice yield. China is a food-importing country whose grain consumption is higher than grain production, and which relies on imports of rice, soybean, wheat and other grains. Therefore, in order to secure food security for 1.4 billion people in China, it is necessary to grasp the relationship between climate change and rice yield. In this study, 16 administrative districts in China were selected and designated as study area. This study used annual rice production from the USDA(United States Department of Agriculture) for each of China's major administrative regions from 1979 to 2009, as well as average climate data from July to August, which were meteorological observations collected from the CRU(Climatic Research Unit). Using this data, the rice crop was increased in 10 administrative regions in China and the reduction in rice harvest in 6 administrative areas was confirmed. The relationship between selected rice production and climate change was nonlinear and modelled using a deep neural network to train odd rows and verify even rows of data. Therefore, a more quantitative analysis of the relationship between climate change and rice yield changes has been made possible through our prediction model. This study is expected to contribute to better food self-sufficiency in China and forecast future grain yields.

1. INTRODUCTION

Climate change for agriculture and food security are effective world-wide, and they are very important for China. Whole changes of climate change are fairly uncertified. China's large land mass is extended to experience different warmings and probably other changes as well, not quite different from each other's supply and acquisition in 1996. Climate change is an important factor in plant growth. The growth of plants is closely related to food safety. This supply of food is also economically related. In China, demand for food consents when China's 1.6 billion people by the year 2030 begin importing granaries (Brown, 1995). Therefore, identifying the relationship between climate change and grain yield and predicting future grain yield has become an important part of food safety.

In this study, the model was built using the Multiple Linear Regression(MLR) technique for linear relationship analysis and the artificial neural network technique for nonlinear relationship analysis (Jian and Cotton, 2014) to predict the yield of rice in China, with 10 climate data as descriptive variables. The two models were selected for accuracy by analyzing and comparing the MBE, RMSE, MAE, and Corr.

2. DATA AND METHOD

2.1 Research Area and Period of Study

In this study, 16 out of 25 provinces were selected as research areas. The provinces having extremely low or high data of rice yield were excluded. The areas chosen as research areas are Anhui, Fujian, Gansu, Guangdong, Guangxi, Hebei, Heilongjiang, Hunan, Jiangsu, Jilin, Ningxia Hui, Shaanxi, Shandong, Shanxi, and Zhejiang. July and August were the study period from 1979

* Corresponding author

to 2009 (Figure 1).

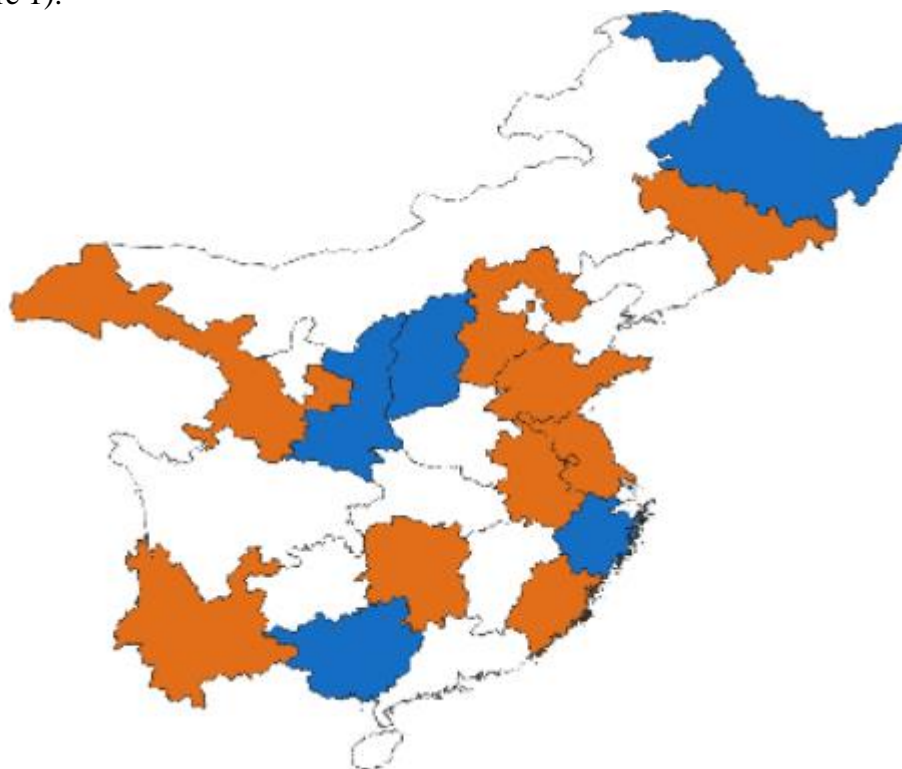


Figure1. Research area in China (rice yield increasing areas in orange and decreasing areas in blue, 1979-2009)

2.2 Data for Use

For this study, the rice production data and the rice cultivation data were USDA(United States Department of Agriculture) data, and the climate variable data were the Climate Research Unit(CRU) data.

2.2.1 USDA Rice Harvest Data

USDA provides data on production, supply and distribution of aggregate commerce for U.S. and key processing and counting counts. In this study, the rice yield was calculated using the annual rice production data provided by USDA and the rice cultivation area data as the rice yield per unit area(ton/ha) (Figure 2).

2.2.2 CRU Climate Variables Data

CRU produces a range of climate datasets, covering temperature, precipitation, pressure and circulation, both global regional. In this study, monthly data for cloud cover(CLD, %), diurnal temperature range(DTR, degree Celsius), ground frost frequency(FRS, days), potential evapotranspiration(mm/day), precipitation(PET, mm/month), near-surface temperature minimum(TMN, degree Celsius), near-surface temperature maximum(TMX, degree Celsius), vapor pressure(VAP, hPa), and wet day frequency(WET, days) were used.

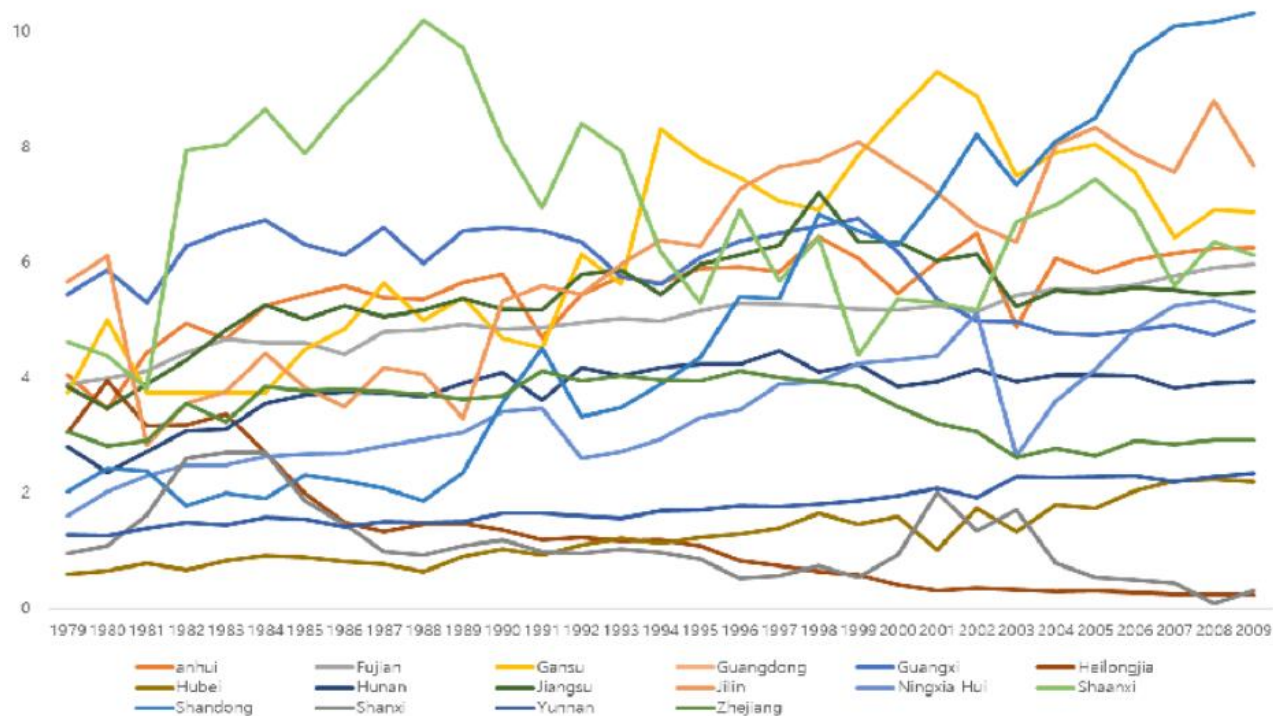


Figure 2 Rice yield variation of 16 provinces in China (ton/ha)

2.3 Research Method

In this study, the Multiple Linear Regression (MLR) and Deep Neural Network (DNN) techniques were used to compare and evaluate the accuracy of the yield of rice crops due to climate change.

2.3.1 MLR (Multiple Linear Regression)

The MLR collects the equation of relationships between multiple description variables and one response variable and utilizes it for tracking and forecasting. The MLR techniques used in this study for rice yield estimation by nine different climatic variables as explanatory variables.

2.3.2 DNN (Deep Neural Network)

DNN is a kind of artificial intelligence technique based on a multilayered neural network with input and output layers and a number of hidden layers. The classic neural network has a local minima problem in which an optimization process often stops at a locally, rather than globally, optimized state. In addition, generic machine learning methods sometimes have problems with overfitting, in which they cannot handle data with outliers due to excessive learning from the given dataset. Such problems can be resolved by DNNs through an intensive optimization process in a deep network structure. In this study, the input layer has nine different climatic variables, and the output layer was set to rice yield. Through a number of optimization experiments, the hidden layers consisted of 300-300-300, and the number of epochs was set to 3000 (Figure 3).

2.3.3 Training and Validation

MLR and DNN were used to train and validate the rice yield prediction models. We had two schemes for a total of 496 matchups: (1) training by odd rows and validation by even rows and (2) training by even rows and validation by odd rows. Finally, two sets of validation statistics were merged, and mean bias error (MBE), mean absolute error (MAE), root mean square error (RMSE), and correlation coefficient (Corr.)

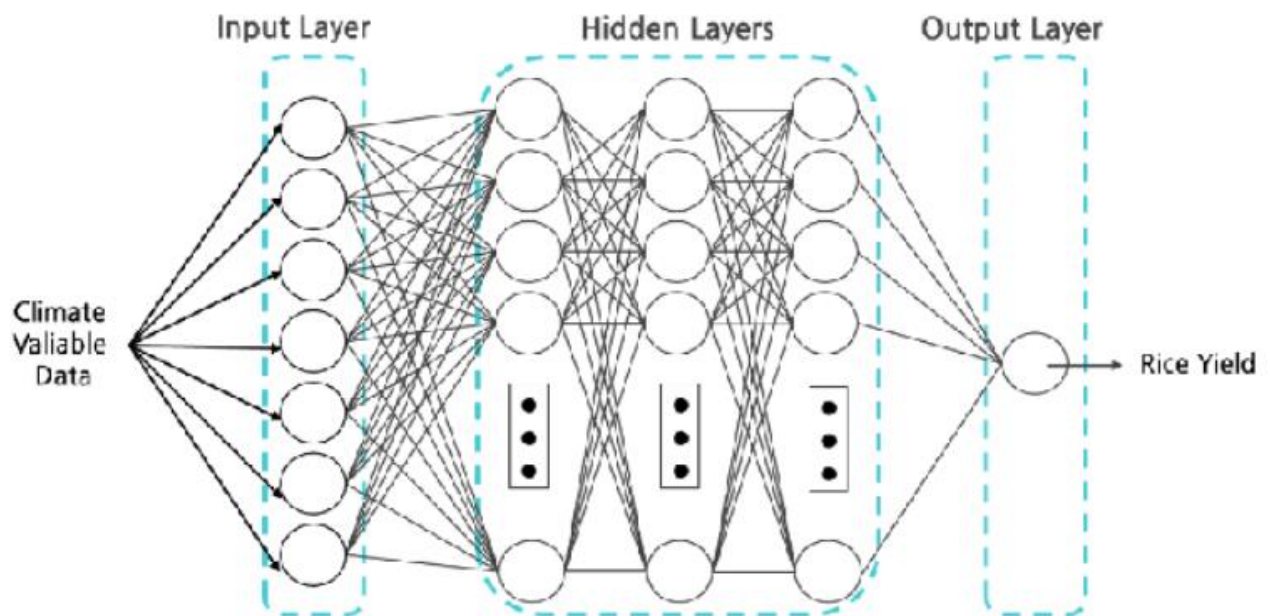


Figure 3. Structure of deep neural network

3. RESULTS

The training-validation for MLR and DNN models were carried out for odd and even rows, respectively, and the validation statistics were summarized in Table 1. The DNN showed better results than MLR. The MAE was 32% smaller $((1.404-1.067)/1.067)$ and the correlation coefficient was 33% higher than $((0.732-0.551)/0.551)$.

The values of the observation versus estimation data using MLR and DNN were expressed as a scatterplot (Figure 4 and 5). The scatterplot of DNN seemed closer to the one-to-one line, which indicates the estimation was more accurate and stable.

	MLR	DNN
MBE	-1.506e-12	-0.059
MAE	1.404	1.067
RMSE	1.813	1.500
Corr.	0.551	0.732

Table 1. Validation statistics

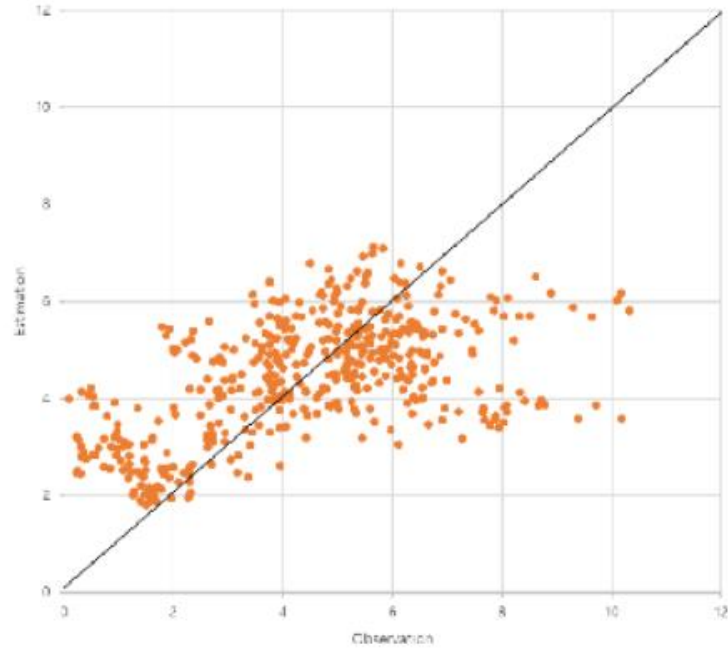


Figure 4. Scatterplot of test accuracy of MLR

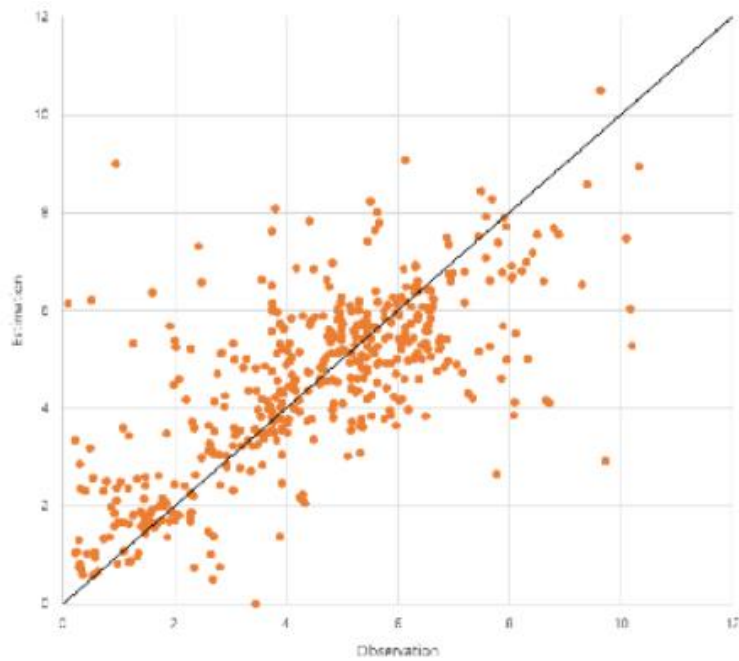


Figure 5. Scatterplot of test accuracy of DNN

4. CONCLUSIONS

In this study, we tried to make more accurate analyses and predictions in order to recognize and improve the problems of food shortages that could arise from climate change. In a training-validation experiment, the DNN model showed much higher accuracy than MLR model. It is very important for China, which has a large population, to keep its main food. The economic benefits of predicting rice yields in the future can also be greatly appreciated by taking advantage of various climate factors that change the cause of climate change.'

Future studies will find quantitative reasons for selecting variables, increasing the accuracy of the predictions.

REFERENCES

Barry, S., Cai, Y., 1996. Climate change and escalation in China. *Global Environmental Change*, 205-214.

Brown, R., 1995. Who Will Feed China? *Wake Up Call A Small Planet*, 163 pages.

Jiang, H., Cotton, W.R., 2004, Soil Moisture estimation use an alternative natural network: a flexible study, *Canadian Journal of Remote Sensing*, 827-839.