Long short-term memory networks for county-level corn yield estimation*

Haifeng Li
School of Geosciences and Info Physics, Central South University, Changsha, China
lihaifeng@csu.edu.cn

Yudi Wang
School of Geosciences and Info Physics, Central South University, Changsha, China

Renhai Zhong
College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China

Hao Jiang
College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China

Tao Lin
College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China
lintao1@zju.edu.cn

Abstract

Quantifying the response of crop yield to climate factors is important in agricultural systems. Many studies have worked on yield prediction through process-based simulation models and statistical models. Given the spatiotemporal explicit features of crop production, there exists a need to better understand the cumulative temporal effects of climate factors on crop production. To fill this gap, we build a Long Short-Term Memory (LSTM) model for weather-impacted corn yield prediction. The results show that LSTM model has a better performance (RMSE = 0.61 Mg ha^{-1}) in yield prediction than two other models: Lasso (RMSEP = 1.07 Mg ha^{-1}) and RF (RMSE = 0.64 Mg ha^{-1}) on the same test set. The results illustrate the potential of LSTM in crop yield prediction by considering the cumulative temporal impact of weather factors on crop yield.

2012 ACM Subject Classification Spatio-temporal modeling in machine learning applications

Keywords and phrases LSTM, Corn yield prediction, climate factors, cumulative temporal effect

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.xx

1 Introduction

In 2015, FAO estimated that 795 million people live without an adequate food supply [3]. With the increased global population, maintaining sustainable food supply becomes an international issue. How to improve crop productivity is critical to address food security. Accurate in-season yield prediction can support farmers to improve management and reduce yield loss caused by unfavorable weather conditions.

Process-based biophysical modeling and statistical modeling [4] are two popular approaches to quantify corn yield based on climate factors. Biophysical modeling is more suitable for site-specific yield analysis, whereas statistical modeling is often adopted in large-scale spatial analysis. Some researches further look into the integration of process-based and statistical

* corresponding author: Tao Lin, lintao1@zju.edu.cn.

Spatial big data and machine learning in GIScience, Workshop at GIScience 2018.
Editors: Martin Raubal, Shaowen Wang, Mengyu Guo, David Jonietz and Peter Kiefer;

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

modeling [6]. Under the rapid development of computing capabilities in recent years, artificial intelligence methods, such as Artificial Neural Network (ANN) [2] and Bayesian Network (BN) [5], have gradually been applied for agricultural yield prediction. These studies, however, often simply the temporal variations of yield-weather relationship and the cumulative effects of weather factors. In practice, not only the crop itself is growing over time, the impact of weather on the crop might also vary and accumulate throughout the growing season. Especially, the damage of extreme weather would possibly impose a long sequential impact on crop growth throughout the season. No remediations afterward can be applied to resolve the damage. There is a critical need, therefore, to integrate the cumulative effects of climate factors on crop production to better understand the interactions between crop and environmental factors.

Long Short-Term Memory (LSTM) model is a deep neural network that has been successful in learning sequence and tree structures [7]. It facilitates time-series analysis and handles complexity and nonlinearity functions by its unique structure. LSTM was developed to deal with the gradient vanishing and exploding problems. Previous studies have demonstrated that LSTM has a good performance in dealing with long sequential data in natural language modelling [7] and human trajectory prediction [1]. We would like to evaluate the performance of LSTM modeling in capturing dynamic temporal yield-weather relationships and yield prediction.

The objectives of this study are to: (i) develop a LSTM model to predict corn yield by cumulative climate factors; (ii) compare the prediction accuracy among LSTM, Random Forests (RF), and Lasso regression methods.

2 Methodology

2.1 Study Area and datasets

This study focuses on rainfed corn yield in the central and northern 11 states in the U.S. from 1970 to 2016. These 11 states are: Illinois, Indiana, Iowa, Kansas, Kentucky, Michigan, Minnesota, Missouri, Nebraska, Ohio and Wisconsin. The county-level non-irrigated corn yield data is from USDA's National Agriculture Statistics Service [8]. To capture the impact of climate factors on corn yield, we calculate the detrend yield (yield influenced by climate factors) by linear regression of corn yield versus year as the predictable variable in the models. The daily county-level climate data is obtained from Applied Climate Information System (ACIS) Web Service [9]. The climate data used in this study include: maximum daily temperature (Tmax), minimum daily temperature (Tmin), and daily precipitation (PRCP). During the corn growth period (from week 20 to 39), weekly Tmax, Tmin and PRCP is calculated and transferred into sequential vectors as the input of LSTM model. In addition, Min-Max scaling is used to scale the input into a range from zero to one.

2.2 Models

The structure of LSTM model in this study includes four layers: input layer, hidden layers and output layer (Figure 1). The input is a time sequence $X = \{x_1, x_2, \cdots, x_T\}$, x_t is a vector which includes climate factors.$x_t = [\mathrm{Tmax}_t, \mathrm{Tmin}_t, \mathrm{PRCP}_t]$, T is 20, the length of time sequence and t is the time, represent the week in corn growth period, from week 20 to 39. The hidden layers are two LSTM layers composed of LSTM cells, in which information is selectively transported and stored. The output is detrend yield y calculated out by all input vectors in one time sequence of corn growth period, from week 20 to 39.

To make a comparison between LSTM and other models, we build a Lasso regression model ($\lambda = 0.003$) and a RF model as baselines. All three models are trained based on the same training set, which is randomly selected 80% of the total sample set. The remaining 20% of the total sample set is used as the test set, where RMSE is used as the performance indicator of yield prediction accuracy.

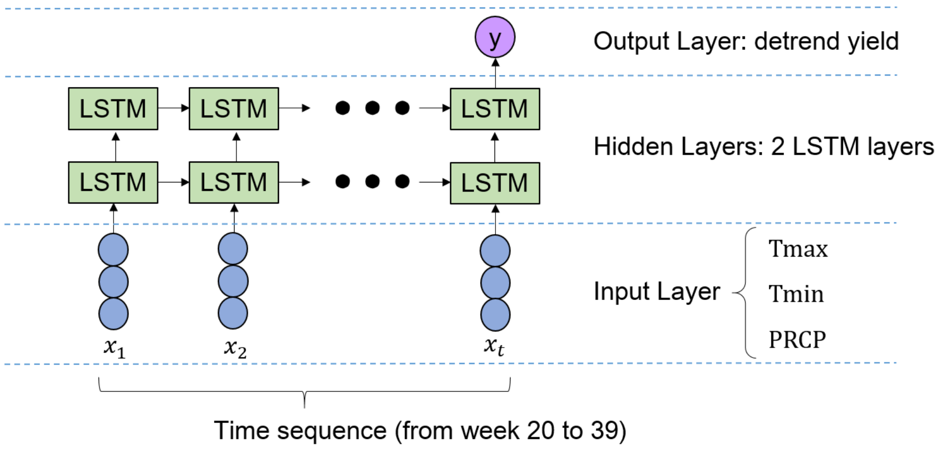

Figure 1 The structure of LSTM.

2.3 Results and Discussion

We compare the accuracy of LSTM model with two typical models: RF and Lasso regression (Table 1). The results show that machine learning models, LSTM (RMSE = 0.61 Mg ha^{-1}) and RF (RMSE = 0.64 Mg ha^{-1}), outperform traditional linear regression model, Lasso (RMSE = 1.04 Mg ha^{-1}) in yield prediction on the same test set. Compared to RF, LSTM has a less degree of overfitting and a slight improvement on prediction accuracy. The improved accuracy by the LSTM model is possibly due to its structure designed for capturing not only the direct impact at each time period but accumulated effect of weather on crop yield throughout the entire growing season. The degree of impact by weather may vary temporally as the requirement of water and nutrients by crop varies at different growing stages. In addition to the accumulated temporal impact, LSTM is more suitable to capture nonlinear impact of weather on crop, especially considering the extreme events such as heating and drought during the growth period.

Model	RMSE Training Set (34868 obs)	RMSE Test Set (8717 obs)
LSTM	0.53	0.61
RF	0.31	0.64
Lasso	1.07	1.04

Table 1 The RMSE of LSTM, RF and Lasso models.

To better understand the accumulated and nonlinear impact of climate factors on corn yield, some future work are proposed: (i) evaluate the performance of in-season prediction; (ii) quantify the sensitivity of yield to climate factors in different corn growing stages by adding an Attention mechanism into the LSTM model; (iii) incorporate the spatial correlation features of the crop yield-weather relationship in the LSTM model.

3 Conclusions

We build a Long Short-Term Memory model to predict the county-level rainfed corn yield in response to climate factors for 11 states in the U.S. from 1970 to 2016. The results show that the LSTM model has a better prediction accuracy than traditional Lasso regression model. Compared to the RF, LSTM has a slightly better prediction accuracy and less degree of overfitting. The results demonstrate the potential of LSTM in crop yield prediction and temporal analysis of yield-weather interactions. The mechanism of how the LSTM model captures the cumulative temporal effect of climate factors on corn yield needs to be further studied.

References

1 Al. Alahi, K. Goel, V. Ramanathan, et al. Social lstm: Human trajectory prediction in crowded spaces. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 961–971, 2016.

2 R. Alvarez. Predicting average regional yield and production of wheat in the argentine pampas by an artificial neural network approach. *European Journal of Agronomy*, 30(2):70–77, 2009.

3 FAO. The state of food insecurity in the world. Meeting the 2015 international hunger targets: Taking stock of uneven progress, 2015.

4 D.B. Lobell and M.B. Burke. On the use of statistical models to predict crop yield responses to climate change. *Agricultural and Forest Meteorology*, 150(11):1443–1452, 2010.

5 N.K. Newlands and L. Townley-Smith. Predicting energy crop yield using bayesian networks. In *The Fifth IASTED International Conference*, volume 711, pages 014–106, 2010.

6 M.J. Roberts, N.O. Braun, T.R. Sinclair, et al. Comparing and combining process-based crop models and statistical models with some implications for climate change. *Environmental Research Letters*, 12(9):095010, 2017.

7 M. Sundermeyer, R. Schlüter, and H. Ney. LSTM neural networks for language modeling. In *13th Annual Conference of the International Speech Communication Association*, 2012.

8 USDA-NASS. Quick Stats 2.0. SDA-NASS, Washington, DC. `http://www.nass.usda.gov/quickstats/`, 2017. [Online; accessed 20 August].

9 RCC-ACIS web service. `http://www.rcc-acis.org/`, 2017. [Online; accessed 21 August].