Predicting Global Wheat Crop Yield Using a Combination of CNN and LSTM

Sungjoon Park

May 7, 2024

Abstract

This research proposes a deep neural network model that predicts global wheat crop yield using time-series data of global weather, aiming to address the inefficiencies of current region-specific models. To this end, a combination of CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory) networks is implemented, where the CNN block processes encode input images and the LSTM block play a role in predicting wheat crop yield in an auto-regressive manner. This model is meaningful in that it predicts the global crop field and uses a spatiotemporal approach.

Compliance in Generative AI Assistance (GAIA) Policy

1. This report was initially drafted by me, Sungjoon Park, and was refined with the aid of Copilot (GPT-4 Turbo), ensuring full compliance with the regulations outlined in the course syllabus.
2. Generative AI: OpenAI GPT-4 Turbo
3. I rectified grammatical errors and adapted the text to conform to academic writing standards through the Generative AI.
4. I utilized Generative AI in order to enhance the report, ensuring the delivery of accurate information and fulfilling academic writing objectives.

Introduction

The stability of global grain prices is significantly influenced by the grain supply, which in turn, hinges on the production output of key grain-producing regions worldwide. Traditionally, models designed to forecast grain productivity have predominantly relied on weather data specific to these regions [1][2]. However, given the increasing prevalence of teleconnections, such as El Niño or La Niña phenomena, it becomes imperative to investigate the potential of enhancing prediction accuracy by utilizing global weather time-series data, rather than solely depending on regional weather data. Prior to validating this approach, the development of an appropriate model is a crucial prerequisite.

In response to this challenge, this research proposes a novel approach: the use of a deep neural network trained to predict global wheat crop yield using global weather information as input. This approach requires implementing a model that can deal with spatiotemporal data to predict a single number. To this end, image datasets, such as time-series data for global average temperature heat maps, are used as inputs to capture the spatiotemporal information. Along with these datasets, a combination of CNN (Convolutional Neural Network) and LSTM (Long Short-Term Memory) networks are applied. The performance of this model will be evaluated by mean squared error (MSE) metric. Meanwhile, this research focuses only on wheat crop yield rather than dealing with all grains.

Related Work

Klumpenburg et al. (2020) [2] conducted a comprehensive review of the existing literature to identify the major research questions pertaining to crop yield prediction. These questions can be distilled into three primary inquiries: Which machine learning algorithms have been employed in the literature for crop yield prediction? What features have been utilized in the literature for crop yield prediction using machine learning? And what evaluation parameters and approaches have been adopted in the literature for crop yield prediction?

In response to the first research question, Klumpenburg et al. enumerated 37 features that have been used in the reviewed papers. Among these, seven features—temperature, soil type, rainfall, crop information, soil maps, humidity, and pH-value—account for 50% of the usage. The study also highlighted the most frequently used machine learning algorithms: neural networks, linear regression, random forest, support vector machine, and gradient boosting tree. Neural networks were used 27 times, constituting 40.2% of the total cases, thereby indicating their recognition as the most accurate predictor by researchers. However, it is important to note that these neural networks were specifically used for predicting regional crop yield, not for global prediction.

Lastly, the study found that previous research primarily used evaluation parameters derived from the mean square error (MSE), namely the root mean square error (RMSE), R-squared, and MSE. These methods accounted for 73.6% of the total cases. In other instances, researchers employed the mean absolute error (MAE), which is robust to outliers, the mean absolute percentage error (MAPE), and the reduced simple average ensemble (RSAE), among others.

Approach

This research endeavor employs time-series images as input data, which is subsequently transformed into encoded features. The proposed neural network architecture initiates with a CNN block that processes the images to encode latent representations. These representations are then sequentially fed into an LSTM block. The culmination of this process is the LSTM network forecasting the global wheat crop yield for the forthcoming year, based on the features from the preceding 12 months derived from the CNN block. Although the window length could be a hyperparameter, it is fixed at 12 in this study.

To elaborate, the architecture utilizes the CNN block to process global weather images and extract relevant features. These features are then relayed to the LSTM block, which operates within a many-to-one feed-forward process, facilitating the generation of a unified prediction. In this project, ResNet15 serves as the CNN block without any fine-tuning, leaving the LSTM block as the sole component subject to training.

However, this structure presents two challenges. The first is its inability to reflect technological advancements. The target variable of this research, global wheat crop yield, exhibits an upward trend as depicted in Figure 1. Productivity enhancement is influenced by several factors, such as fertilizers, farm equipment, and institutional infrastructure. As the input data does not encompass these factors, an assumption is made to substitute them: the introduction of an auto-regressive conduit when predicting crop yield. The final layer of the LSTM block generates hidden states, which are subsequently fed into the last fully-connected layer tasked with predicting crop yield. Here, the previous crop yield value is concatenated to the hidden states to incorporate the auto-regressive path.

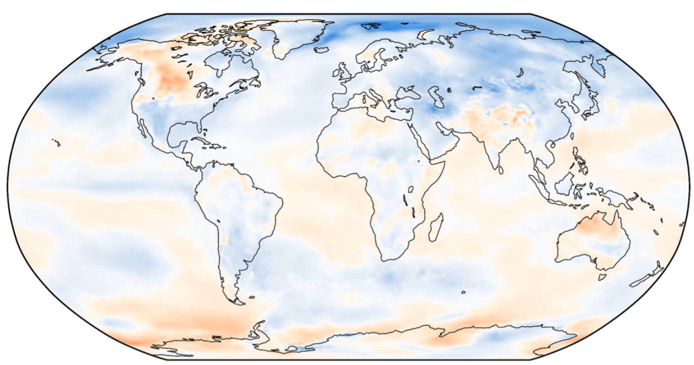

Figure 1: Upward Trend in Global Wheat Crop Yield

The second challenge is that this architecture is required to learn to predict the same target despite receiving different 12-month window ResNet15 features. For instance, the model is trained to predict the crop yield in 2020 if the input data comprises monthly features from February 2019 to January 2020. Similarly, when the input features for the LSTM block range from June 2019 to May 2020, this LSTM block will be trained to predict the same crop yield. To address this, positional encoding values are added to each feature generated by ResNet15, enabling the model to recognize the temporal location of each monthly data set.

Datasets

1. **Input data (Climate Information)**

   (a) **Near-Surface Air Temperature Monthly Data**
   - Source: C3S (Copernicus Climate Change Service)
   - Period: Jan. 2015 $\sim$ Dec. 2019
   - Frequency: hourly - Preprocessing: Converting into monthly images via averaging NumPy arrays

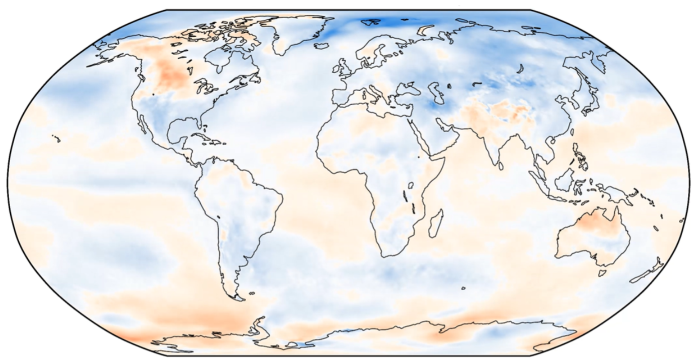

Figure 2: Near-Surface Air Temperature sample

   (b) **Global Monthly Precipitation**
   - Source: C3S (Copernicus Climate Change Service)
   - Period: Jan. 1979 $\sim$ Dec. 2019
   - Frequency: Monthly

2. **Output data (Wheat Crop Yield)**
   - Source: Our World in Data
   - Period: 1970 $\sim$ 2019
   - Frequency: Yearly

3. **LSTM inputs**
   - Input pairs: Features from ResNet15 (A) and a lagged crop yield value (B)

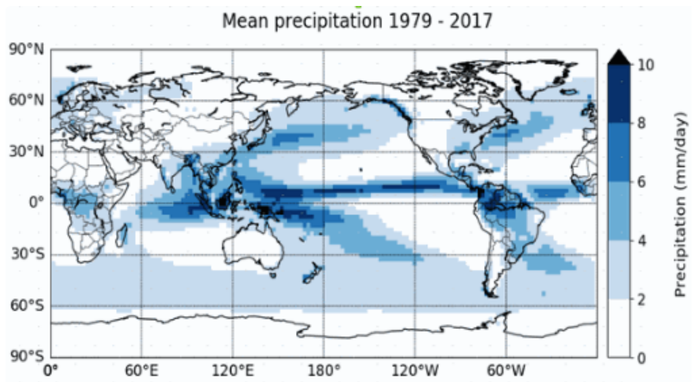

Figure 3: Global Precipitation sample

(a) Shape of A: (Batch size=1, The number of months=12, Feature size)
- Examples of A
    - The first tensor: [Batch 1, Jan. 1971:Dec. 1971, Features]
    - The second tensor: [Batch 2, Feb. 1971:Jan. 1972, Features]

(b) Shape of B: (Batch size=1, 1)
- Examples of B
    - The first tensor: [Batch 1, 1970 Crop yield]
    - The second tensor: [Batch 2, 1971 Crop yield]

- Target: (Batch size=1, 1)
    - The first target: [Batch 1, 1971 Crop yield]
    - The second target: [Batch 2, 1972 Crop yield]

Training

The training configuration is as follows:

1. The number of epochs: 100
2. Optimizer: Adam
3. Learning rate: 1e-3
4. Loss function: mean-squared-error

The sole hyper-parameter is the hidden size of the LSTM block. Training models with different hyper-parameters will be a future task.

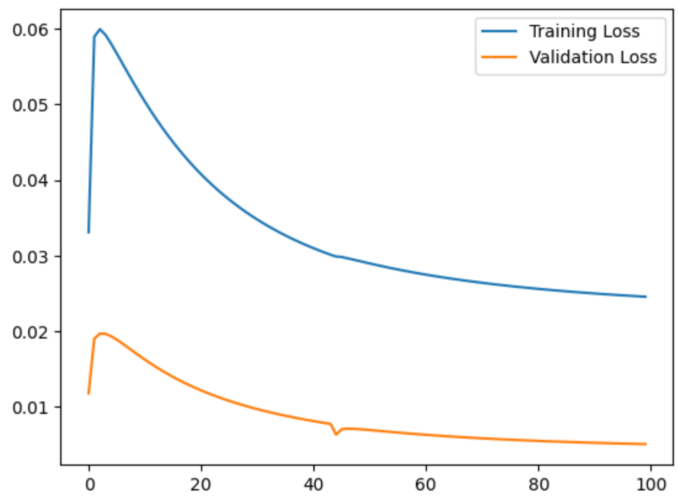

Figure 4: MSE Losses When a Model with Hidden Size of 64 Trained

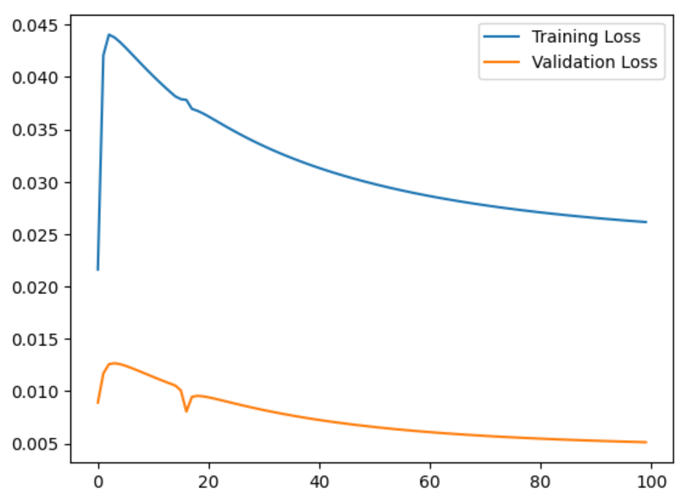

Figure 5: MSE Losses When a Model with Hidden Size of 128 Trained

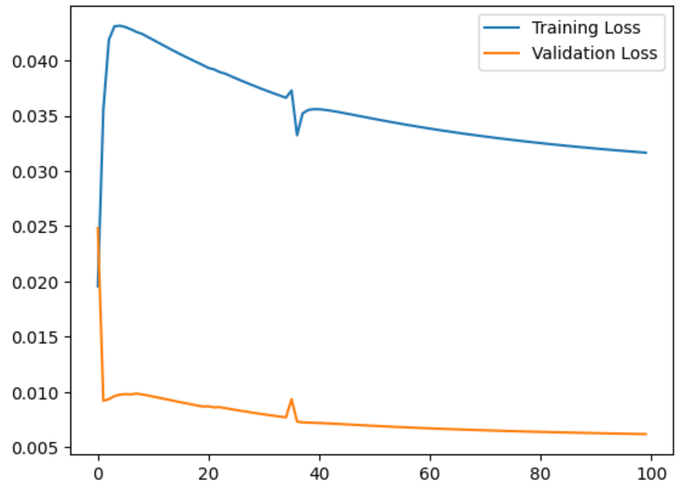

Figure 6: MSE Losses When a Model with Hidden Size of 256 Trained

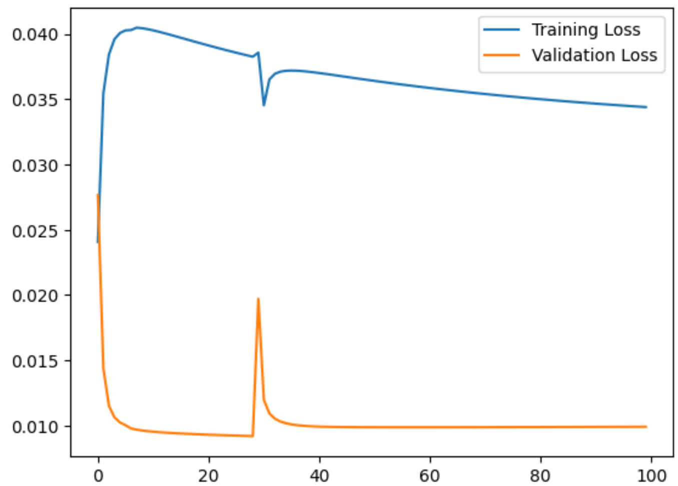

Figure 7: MSE Losses When a Model with Hidden Size of 512 Trained

Evaluation Results

The test mse results are 0.0057 (hidden size of 64), 0.005 (hidden size of 128), 0.0014 (hidden size of 256), 0.00059 (hidden size of 512). Therefore, when the hidden size of the LSTM block is the largest, the model performs best. Since the optimal hidden size needs to be checked by increasing it, further investigation to conduct hyperparameter-tuning would be the next task of this project.

Conclusion

This study explores the potential of combining CNN and LSTM network in predicting crop yields by utilizing temperature and precipitation heat maps. This approach allows for the integration of global image data into predictive models. Furthermore, the model is designed to simultaneously predict global wheat crop yields using information from around the world. Given the significant impact of agricultural products on inflation, a phenomenon recently termed as 'agflation', and the increasing variability in agricultural supply across different regions due to climate changes, the accurate and timely prediction of global agricultural product prices becomes crucial. In the context of global warming, this model serves as a valuable tool for economists, enabling them to effectively predict and respond to changes in the global agricultural market. In addition to it, the future research of this project will involve several enhancements. Firstly, incorporating additional inputs into the model, such as humidity data and soil type maps is required. This is because these factors could potentially provide more nuanced insights into the problem at hand. Secondly, a direct comparison of this model's performance with previous studies was not feasible due to differences in the scales of MSE. Hence, devising a method for making these comparisons possible will be crucial to validate the usefullness of this model. Lastly, there are additional hyperparameters such as the method of positional encoding and different ResNet pre-trained models, which can help optimization through hyperparameter-tuning.

References

[1] Ansarifar, J., Wang, L., & Archontoulis, S. V. (2021). An interaction regression model for crop yield prediction. *Scientific reports*, 11(1), 17754. https://doi.org/10.1038/s41598-021-97221-7

[2] Klompenburg, T.V., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric., 177*, 105709.