

---

# A Comprehensive Multimodal Dataset for Climate-Conscious Prediction of Crop Yields

---

## Abstract

Accurate forecasting of crop yields is crucial for maintaining food security and promoting sustainable agricultural methods. While AI has shown significant promise in various scientific domains, the creation of deep learning models for crop yield prediction has been constrained by the absence of an expansive, publicly accessible, multimodal dataset that encompasses adequate information. To address this limitation, we introduce CropNet, the first terabyte-scale, publicly available, multimodal dataset designed for climate-aware crop yield predictions across the contiguous United States at the county level. The CropNet dataset integrates three types of data: Sentinel-2 Imagery, WRF-HRRR Computed Dataset, and USDA Crop Dataset, covering over 2200 U.S. counties over six years (2017-2022). This dataset is designed to help researchers develop versatile deep learning models for accurate and timely county-level crop yield predictions, considering both short-term weather variations during the growing season and long-term climate change impacts. Additionally, we offer the CropNet package, which includes three types of APIs to facilitate data downloading for specific times and regions of interest and to support the flexible development of deep learning models for precise crop yield predictions. Extensive experiments using various deep learning solutions on the CropNet dataset confirm its general applicability and effectiveness in climate-conscious crop yield predictions. The CropNet dataset is officially released on Hugging Face Datasets, and the CropNet package is available on the Python Package Index (PyPI).

## 1 Introduction

The accurate estimation of crop yields is vital for proactive agricultural planning, timely adjustments to management policies, informed financial decision-making, and ensuring national food security. Recent progress in deep neural networks (DNNs) has led to remarkable performance in various fields. Building on these advancements, numerous studies have utilized spatial-temporal DNNs to enhance the timeliness and accuracy of crop yield predictions. However, these studies often rely on individually curated and limited datasets, resulting in somewhat moderate prediction accuracy. There is a pressing need for new, extensive, and deep learning-ready datasets specifically designed for widespread use in crop yield forecasting.

Recent studies have introduced open and large-scale datasets based on satellite imagery or meteorological parameters, which are adaptable to agricultural tasks like crop type classification. However, these datasets have two primary limitations that prevent their direct application to general crop yield predictions. First, they lack the essential ground-truth crop yield data, making them unsuitable for predicting crop yields. Second, they offer only a single data modality, either satellite images or meteorological parameters. Accurate crop yield predictions often require the simultaneous monitoring of crop growth and the capture of meteorological variations that affect yields, necessitating multiple data modalities. To date, the creation of a large-scale, multimodal dataset specifically for county-level crop yield predictions remains an unresolved challenge.

In this research, we aim to develop such a dataset, named CropNet, which is the first terabyte-sized, publicly accessible dataset with multiple modalities, specifically designed for county-level crop yield predictions across the United States (U.S.) continent. The CropNet dataset comprises three data modalities: Sentinel-2 Imagery, WRF-HRRR Computed Dataset, and USDA Crop Dataset, covering 2291 U.S. counties from 2017 to 2022. Specifically, the Sentinel-2 Imagery from the Sentinel-2 mission provides two types of satellite images, agriculture imagery (AG) and normalized difference vegetation index (NDVI), for detailed monitoring of crop growth. The WRF-HRRR Computed Dataset, derived from the WRF-HRRR model, offers daily and monthly meteorological parameters, accounting for short-term weather variations and long-term climate change, respectively. The USDA Crop Dataset, sourced from the USDA Quick Stat website, contains annual crop yield information for four major crops (corn, cotton, soybean, and winter wheat) grown in the contiguous U.S., serving as the ground-truth label for crop yield prediction tasks.

## 2 Data Sources

The CropNet dataset is constructed from three distinct data sources, as detailed below:

**Sentinel-2 Mission:** Launched in 2015, the Sentinel-2 mission is a crucial Earth observation initiative. It offers multi-spectral satellite images with 13 spectral bands and a high revisit frequency of 5 days. These images are valuable for various applications, including climate change monitoring and agricultural oversight.

**WRF-HRRR Model:** The High-Resolution Rapid Refresh (HRRR) is a forecast modeling system based on the Weather Research & Forecasting Model (WRF). It provides hourly forecasts of weather parameters for the entire United States continent with a spatial resolution of 3 km. We use the HRRR assimilated results archived at the University of Utah, which include several parameters relevant to crop growth, such as temperature, precipitation, wind speed, relative humidity, and radiation, starting from July 2016.

**USDA:** The United States Department of Agriculture (USDA) offers annual crop information for major crops cultivated in the U.S. at the county level, including corn, cotton, soybeans, and wheat. The statistical data, dating back to 1850, includes planted areas, harvested areas, production, and yield for each crop type.

### 3 Our CropNet Dataset

#### 3.1 Motivation

Large-scale, multimodal data that include satellite images, numerical meteorological weather data, and crop yield statistics are essential for monitoring crop growth and correlating weather variations with crop yields. These data are crucial for making timely and precise crop yield predictions at the county level. Currently, there is no such open and extensive dataset available for county-level crop yield prediction. In this benchmark article, we introduce CropNet, an open and large-scale dataset with multiple modalities, including visual satellite images, numerical meteorological parameters, and crop yield statistics across the U.S. continent. It is important to note that not all U.S. counties are suitable for crop planting; therefore, our dataset includes data from 2291 out of 3143 counties. This multimodal dataset is invaluable for researchers and practitioners to design and test various deep learning models for crop yield predictions, considering both short-term growing season weather variations and long-term climate change impacts on crop yields.

#### 3.2 Overview of Our CropNet Dataset

The CropNet dataset consists of three data modalities: Sentinel-2 Imagery, WRF-HRRR Computed Dataset, and USDA Crop Dataset, spanning from 2017 to 2022 across 2291 U.S. counties. Given that crop planting is highly dependent on geography, the dataset includes the number of counties for each crop type in the USDA Crop Dataset. The four major crops included are corn, cotton, soybeans, and winter wheat, with satellite imagery and meteorological data covering all 2291 counties. An overview of the CropNet dataset is provided in Table 1. The total size of the dataset is 2362.6 GB, with 2326.7 GB of visual data for Sentinel-2 Imagery, 35.5 GB of numerical data for the WRF-HRRR Computed Dataset, and 2.3 MB of numerical data for the USDA Crop Dataset. Sentinel-2 Imagery contains two types of satellite images (AG and NDVI), both with a spatial resolution of approximately 40 meters (covering an area of 9x9 km with 224x224 pixels) and a revisit frequency of 14 days. The WRF-HRRR Computed Dataset provides daily or monthly meteorological parameters gridded at a spatial resolution of 9 km in one-day or one-month intervals. The USDA Dataset offers county-level crop information for four types of crops, with a temporal resolution of one year.

Table 1: Dataset comparison

Dataset	Size (GB)	Data Modality
SEVIR	970	Satellite Imagery
DENETHOR	254	Satellite Imagery
PASTIS	29	Satellite Imagery
WorldStrat	107	Satellite Imagery
RainNet	360	Satellite Imagery
ENS-10	3072	Meteorological Parameters
Our CropNet Dataset	2362	Satellite Imagery
		Meteorological Parameters Crop Information

#### 3.3 Data Collection and Preparation

**Sentinel-2 Imagery:** We acquire satellite images from the Sentinel-2 mission using the Sentinel Hub Processing API at a processing level of Sentinel-2 L1C. We set a maximum cloud coverage of 20%, with three spectral bands (B02, B08, and B11) for AG images and two bands (B04 and B08) for NDVI images. Satellite images are obtained every 14 days instead of the original 5 days to avoid a large number of duplicate images. Each county is partitioned into multiple grids with a resolution of 9x9 km, each corresponding to one satellite image. The downloaded satellite images for one U.S. state, spanning one season, are stored in one Hierarchical Data Format (HDF5) file. The HDF5 file format is chosen for its ability to save disk space, store data in multidimensional arrays, and store descriptive information for the satellite images.

**WRF-HRRR Computed Dataset:** The WRF-HRRR Computed Dataset is derived from the WRF-HRRR model, which produces hourly GRID files containing meteorological parameters across the contiguous U.S. at a spatial resolution of 3x3 km. Our CropNet dataset includes nine crop growth-relevant meteorological parameters: averaged temperature, precipitation, relative humidity, wind gust, wind speed, downward shortwave radiation flux, maximal temperature, minimal temperature, and vapor pressure deficit (VPD). VPD is calculated using the formula:

$$\begin{aligned} T_C &= T_K - 273.15, \\ e_{sat} &= \frac{610.7 \times 10^{(7.5 \times T_C)/(237.3 + T_C)}}{1000}, \\ e_{air} &= e_{sat} \times \frac{RH}{100}, \\ VPD &= e_{sat} - e_{air}. \end{aligned} \tag{1}$$

We align the resolution of the WRF-HRRR Computed Dataset with that of Sentinel-2 Imagery by using the latitude and longitude of the centric point in the 9x9 km grid to find the nearest 3x3 km grid in the WRF-HRRR model. Meteorological parameters from the 3x3 km grid and its surrounding eight grids represent a region gridded at 9x9 km. Daily meteorological parameters are computed from hourly data, and monthly parameters are derived from daily data. These parameters are stored in Comma Separated Values (CSV) files, which also include the FIPS code, latitude, and longitude of each grid.

**USDA Crop Dataset:** Data from the USDA Crop Dataset is retrieved from the USDA Quick Statistic website using a newly developed web crawler. For each crop type, the USDA website provides county-level crop information annually, identified by a unique key. Our web crawler retrieves this key by specifying the crop type and year, then uses the key to obtain the corresponding crop data. The downloaded data is stored in a CSV file, which includes additional information such as FIPS code, state name, and county name. The data format is unified to store production and yield information in separate columns for easy access by Python libraries like pandas.

Our CropNet dataset targets county-level crop yield predictions across the contiguous U.S. continent. We use the FIPS code to fetch data for each county, including HDF5 files for Sentinel-2 Imagery, CSV files for daily and monthly meteorological parameters, and a CSV file for the USDA Crop Dataset. Configurations are stored in a JSON file for enhanced accessibility.

## 4 Experiments and Results

We evaluated the general applicability of our CropNet dataset to various deep learning solutions through three scenarios of climate change-aware crop yield predictions: Crop Yield Predictions, One-Year Ahead Predictions, and Self-Supervised Pre-training.

### 4.1 Experimental Settings

**Approaches:** We employed ConvLSTM, CNN-RNN, GNN-RNN, and MMST-ViT models for crop yield predictions. Additionally, we considered two self-supervised learning (SSL) techniques: MAE and MM-SSL within the MMST-ViT, representing unimodal and multimodal SSL techniques, respectively. These methods were adapted to fit the CropNet data in our experiments.

**Metrics:** We used Root Mean Square Error (RMSE), R-squared (R<sup>2</sup>), and Pearson Correlation Coefficient (Corr) to assess the effectiveness of the CropNet dataset. Lower RMSE and higher R<sup>2</sup> or Corr values indicate better prediction performance.

### 4.2 Performance Evaluation for 2022 Crop Yield Predictions

Experiments were conducted on the CropNet dataset for 2022 crop yield predictions using satellite images, daily weather conditions during growing seasons, and monthly meteorological conditions from 2017 to 2021. The models used were ConvLSTM, CNN-RNN, GNN-RNN, and MMST-ViT. Table 2 presents the overall performance results for each crop. All models achieved excellent prediction performance with our CropNet data. For instance, ConvLSTM, CNN-RNN, GNN-RNN, and MMST-ViT showed low RMSE values for soybean yield predictions. These results validate that our CropNet dataset is well-suited for LSTM-based, CNN-based, GNN-based, and ViT-based models, demonstrating its general applicability. MMST-ViT achieved the best performance across all scenarios, with the lowest RMSE values and highest R<sup>2</sup> and Corr values for predicting corn, cotton, soybeans, and winter wheat yields. This superior performance is attributed to MMST-ViT’s novel attention mechanisms, which capture the effects of both growing season weather variations and climate change on crop growth. This experiment demonstrates that our CropNet dataset can provide timely and precise crop yield predictions, which are essential for making informed economic decisions and optimizing agricultural resource allocation.

### 4.3 Performance of One-Year Ahead Predictions

Predicting crop yields well in advance of the planting season is crucial for farmers to make early crop planting and management plans. We used the CropNet dataset one year before the planting season to predict the next year’s crop yields. The experimental results for 2022 crop yield predictions using 2021 growing season data show that all models maintain decent prediction performance. For example, ConvLSTM, CNN-RNN, GNN-RNN, and MMST-ViT achieved average RMSE values of 6.2, 5.4, 5.3, and 4.7,

Table 2: Overall performance for 2022 crop yield predictions, where the yield of cotton is measured in pounds per acre (LB/AC) and those of the rest are measured in bushels per acre (BU/AC).

Method	Corn			Cotton			Soybeans			Winter Wheat		
	RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )	RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )	RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )	RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )
ConvLSTM	19.2	0.795	0.892	56.7	0.834	0.913	5.3	0.801	0.895	6.0	0.798	0.893
CNN-RNN	14.3	0.867	0.923	54.5	0.826	0.899	4.1	0.853	0.915	5.6	0.823	0.906
GNN-RNN	14.1	0.871	0.917	55.1	0.813	0.881	4.1	0.868	0.929	5.3	0.845	0.912
MMST-ViT	13.2	0.890	0.943	50.9	0.848	0.921	3.9	0.879	0.937	4.8	0.864	0.929

respectively, for soybean predictions. MMST-ViT consistently achieved excellent Corr values, averaging 0.922 for corn, 0.890 for cotton, 0.926 for soybeans, and 0.904 for winter wheat predictions. These results are only slightly inferior to those for regular 2022 crop yield predictions, which can be attributed to MMST-ViT’s ability to capture the indirect influence of 2021’s weather conditions on the subsequent year’s crop growth through the use of long-term weather parameters. This further underscores how our CropNet dataset enhances climate change-aware crop yield predictions.

#### 4.4 Improving the Generalization Capabilities of DNNs

Self-supervised learning (SSL) techniques have significantly advanced the generalization capabilities of deep neural networks (DNNs), especially in vision transformers (ViTs). Our CropNet dataset, with over 2 TB of data, benefits both deep learning and agricultural communities by providing large-scale visual satellite imagery and numerical meteorological data for pre-training DNNs. To demonstrate the applications of our CropNet dataset to self-supervised pre-training, we used MMST-ViT for crop yield predictions under three scenarios: MMST-ViT without SSL (w/o SSL), MMST-ViT with SSL in MAE (MAE), and MMST-ViT with the multi-modal SSL technique (MM-SSL). The performance results for four crop types under three metrics (RMSE, R2, and Corr) show that without SSL, MMST-ViT exhibits limitations in generalization capabilities, resulting in suboptimal crop yield prediction performance. Pre-training MMST-ViT with MAE’s SSL technique improves performance compared to the w/o SSL scenario, with decreased RMSE values for corn, cotton, soybeans, and winter wheat predictions. This confirms that our CropNet dataset can improve the generalization capabilities of vision models. Furthermore, MMST-ViT with the multi-modal SSL technique achieved the best performance results under all scenarios, significantly decreasing RMSE values for predicting corn, cotton, soybeans, and winter wheat. The effectiveness of the multi-modal SSL technique may stem from its ability to integrate visual satellite imagery with numerical meteorological data in the CropNet dataset, enhancing the generalization capabilities of the MMST-ViT model by improving its ability to discern the influence of weather conditions on crop growth patterns during pre-training.

#### 4.5 Significance of Each Modality of Our CropNet Dataset

To demonstrate the necessity and significance of each modality in our CropNet dataset, we examined five scenarios. First, we dropped the temporal satellite images (w/o temporal images) by randomly selecting only one day’s imagery data. Second, we discarded the high-resolution satellite images (w/o high-resolution images) by using only one satellite image to capture the whole county’s agricultural information. Third, we ignored the effects of weather variations on crop yields by dropping all meteorological data (w/o WRF-HRRR data). Similarly, w/o short-term data and w/o long-term data represent masking out the daily and monthly meteorological parameters, respectively. We also included prediction results using all modalities of the CropNet dataset (All) for performance comparison. Note that the USDA Crop Dataset provides the label for crop yield predictions, so no ablation study is required for this modality.

Table 3 presents the experimental results under the MMST-ViT model. Discarding the temporal satellite images (w/o temporal images) significantly degrades performance, increasing RMSE values and lowering Corr values for corn and soybean yield predictions. This is because a sequence of satellite images spanning the whole growing season is essential for tracking crop growth. The w/o high-resolution images scenario achieved the worst prediction performance, with the highest RMSE values and lowest Corr values for corn and soybean yield predictions. This is because high-resolution satellite images are critical for precise agricultural tracking. Dropping meteorological parameters (w/o WRF-HRRR data) prevents MMST-ViT from capturing meteorological effects on crop yields, leading to increased RMSE values and decreased Corr values for corn and soybean yield predictions. Discarding either daily weather parameters (w/o short-term data) or monthly meteorological parameters (w/o long-term data) also lowers crop yield prediction performance, as the former is necessary for capturing growing season weather variations, while the latter is essential for monitoring long-term climate change effects. Therefore, each modality in our CropNet dataset is important and necessary for accurate crop yield predictions, especially for crops sensitive to growing season weather variations and climate change.

Table 3: Ablation studies for different modalities of the CropNet dataset, with five scenarios considered and the last row presenting the results by using all modalities

Modality	Scenario	Corn			Soybeans		
		RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )	RMSE ( $\downarrow$ )	R2 ( $\uparrow$ )	Corr ( $\uparrow$ )
Sentinel-2 Imagery	w/o temporal images	22.1	0.758	0.870	5.72	0.773	0.879
	w/o high-resolution images	27.9	0.656	0.810	7.80	0.631	0.794
WRF-HRRR Computed Dataset	w/o WRF-HRRR data	20.6	0.758	0.871	5.78	0.764	0.874
	w/o short-term data	18.6	0.796	0.892	5.04	0.816	0.903
	w/o long-term data	15.3	0.854	0.924	4.72	0.825	0.908
All	—	13.2	0.890	0.943	3.91	0.879	0.937

## 5 The CropNet Package

In addition to the CropNet dataset, we release the CropNet package, which includes three types of APIs available on the Python Package Index (PyPI). These APIs are designed to help researchers develop DNNs for multi-modal climate change-aware crop yield predictions.

**DataDownloader:** This API enables researchers to download CropNet data for specific times and regions of interest on the fly. For instance, given the time and region (e.g., the FIPS code for a U.S. county), the DataDownloader API can be used to download the up-to-date CropNet data.

**DataRetriever:** This API allows researchers to conveniently obtain CropNet data stored locally (e.g., after downloading the curated dataset) for specific times and regions of interest. The requested data is presented in a user-friendly format.

**DataLoader:** This API assists researchers in developing DNNs for crop yield predictions. It allows for the flexible merging of multiple modalities of CropNet data and exposes them through a DataLoader object after performing necessary data preprocessing.

## 6 Conclusion

This work introduces the CropNet dataset, an open, large-scale, and multi-modal dataset specifically designed for county-level crop yield predictions across the contiguous United States. The CropNet dataset comprises three modalities of data: Sentinel-2 Imagery, WRF-HRRR Computed Dataset, and USDA Crop Dataset, containing high-resolution satellite images, daily and monthly meteorological conditions, and crop yield information, aligned both spatially and temporally. This dataset is ready for use in deep learning, agriculture, and meteorology, facilitating the development of new solutions and models for crop yield predictions, considering both growing season weather variations and climate change impacts on crop growth. Extensive experimental results confirm the general applicability of our CropNet dataset to various deep learning models for both timely and one-year ahead crop yield predictions. Additionally, the application of our dataset to self-supervised pre-training scenarios demonstrates its utility in improving the generalization capabilities of DNNs. Alongside the dataset, we have developed the CropNet package, which enables researchers to construct CropNet data on the fly for specific times and regions of interest and to flexibly build deep learning models for climate change-aware crop yield predictions. While the initial goal of creating the CropNet dataset and package was to enhance crop yield prediction accuracy, we believe its future applicability is broad and warrants further exploration, benefiting the deep learning, agriculture, and meteorology communities in pursuing more interesting, critical, and pertinent applications.

## Acknowledgments

The views and opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies.