Estimation of flood-damaged cropland area using a convolutional neural network

Rehenuma Lazin, Xinyi Shen* and Emmanouil Anagnostou
Department of Civil and Environmental Engineering, University of Connecticut Storrs, 159 Discovery Drive (IPB-212), Storrs, CT 06269, United States of America

* Author to whom any correspondence should be addressed.
E-mail: xinyi.shen@uconn.edu
Keywords: flood, remote sensing, crop damage, corn, soybean, CNN, deep learning

Abstract
Flood damage to croplands poses a significant threat to global food security. Effective disaster management to cope with future climate change, especially extreme precipitation, requires a robust framework to estimate such damage. For this study, we develop a model based on a convolutional neural network to estimate the area (in acres) of cropland damaged by flooding at the county level. Then we demonstrate the model’s performance for the period 2008–2019 over corn and soybean fields in the midwestern United States, which suffer frequent damage from recurrent flooding. We fed the network with remote sensing images and weather fields and divide the growing season into two windows, the early season (May–June) and the late season (July–November) for better performance. The results show mean relative error within ±25% and relative root mean square error within 35%–75% in majority of the counties for most years. Finally, we show that the model forced with meteorological variables alone can provide acceptable accuracy, which indicates it can be applied to forecasting crop damage area in the upcoming season or the studying of future climate impact on crop productivity. In principle, the model can also be applied to food security assessment at the global scale using available records.

1. Introduction
Climate change poses significant threats to the global environment (Seddon et al. 2016, Kirchmeier-Young and Zhang 2020). Global warming will tend to increase extreme precipitation and flooding in many parts of the world (Tabari 2020), and the resulting crop damage will pose a significant challenge to global food security (Rosenzweig et al. 2002, Kenyon et al. 2008, Li et al. 2016, Lesk et al. 2016, Yang et al. 2016, Boori et al. 2017, Pacetti et al. 2017). From 2003 to 2013, flooding caused about 86% of the damage to crop production in Asia and 60% in Latin America and Caribbean countries (FAO 2016). Among the different crops, rice production is most encumbered by flooding (Bailey-Serres et al. 2012), since about 90% (Maclean et al. 2003) of the world rice acreage is in flood-prone Asia. Flooding in 2010–2011 also significantly affected wheat production in Canada (a loss of about 21%), Pakistan, and Australia, while in the United States, corn, soybeans, and cotton incurred flood damage along the Mississippi River (World Crop Damage 2011).

Existing approaches to investigating the impact of flooding with regard to crop losses rely on field surveys. These techniques are time and labor consuming and, therefore, not feasible for large-scale application. Globally available remote sensing and meteorological big data images, however, contain abundant information on flood dynamics (Haas et al. 2009, Ogilvie et al. 2015, Mohammadi et al. 2017, Yang et al. 2021). Fine resolution (1 m–1 km) remote sensing imageries related to open water detection can be directly utilized to delineate near-real-time flood extent (Shen et al. 2019a, 2019b), or they can be converted to high-level vegetation indices that represent crop growth. Shrestha et al. (2017) use the weekly Normalized Difference Vegetation Index (NDVI) product from the moderate resolution imaging spectroradiometer (MODIS) to analyze the impact of flooding on corn yield by linear regression. Torbick et al. (2017) analyze the time series
of Sentinel-1 data to monitor inundation during rice growth. Di et al (2017) develop a web-based crop loss assessment system, remote-sensing-based flood crop loss assessment service system (RF-CLASS), driven by the Earth Observing System Data and Information System (EOSDIS) of the National Aeronautics and Space Administration (NASA), to support flood-related crop analysis and insurance decision making. Chen et al (2019) develop crop disturbance indices based on the MODIS enhanced vegetation index (EVI) to monitor flood impact. Synthetic aperture radar (SAR) has also been considered as a potential dataset for detecting flooded vegetation (Tsyganskaya et al 2018).

The main drawback of using remote sensing imageries to capture cropland flooding directly is the relatively large revisiting intervals (around once a week), which can be longer than a flood event. Meteorological data, on the other hand, has coarse spatial but fine temporal resolution. Since flooding is primarily driven by meteorological parameters (Chen et al 2017, Shen and Anagnostou 2017, Lazin et al 2020) including precipitation, temperature, humidity, and evaporation, these data may be informative for assessing the status of flood damage to croplands (Merz et al 2014). Again, multiple studies have investigated how crops respond to flooding of different depths, durations, and flow velocities (Vozinaki et al 2015, Xu et al 2015, Arguello et al 2016, Zhang et al 2016, Wang et al 2017, Chen et al 2019), as well as the recovery mechanisms of submerged crops (Bailey-Serres et al 2012, Yeung et al 2018). Yet, the simulated model findings cannot be applied at field scale in the real world because they lack detailed spatial data indicating actual flood characteristics (Yang et al 2016).

The complex interconnection between crops and the physiology and meteorology of agricultural fields (Siebert et al 2017) makes it harder to estimate losses to flood-damaged crops using physically-based mechanistic approaches than to detect flooded croplands. One difficulty is that different types of crops differ in their resistance to flooding. Reynoso et al (2019) have concluded that among the main crops, only rice is resilient to the waterlogging of roots and submergence of aerial tissue. Soybeans can survive underwater for 48 h quite well, but, beyond this, flooding can reduce the stands and eventually the yield (Scott et al 1989, Tewari and Arora 2016). Waterlogged conditions greatly reduce root growth for cotton, making it more vulnerable to subsequent drought (Hake et al 1992).

Also affecting resistance to flooding is the stage of growth. Crop growth can be divided into two stages, vegetative and reproductive. The vegetative stage starts with the emergence of the plant above the soil surface and ends when the plant stops developing leaf nodes on the stem. The reproductive stage starts with blooming or silking (in the case of corn) and ends with full maturity (Allen et al 2000, Espinoza 2003). When corn reaches the silking stage, shallow flooding will not cause any noticeable damage (Lauer 2008). In the early vegetative stage, however, excess soil moisture retards root development by reducing oxygen concentration and, thus, the roots fail to reach available subsoil water. This makes crops subject to more severe damage during the vegetative stage.

In recent years, the advancement of machine learning techniques has offered a way to model such complex interactions (Palanivel and Surianarayanan 2019, Van Klompenburg et al 2020). Machine learning techniques such as support vector machine (Kuwata and Shibasaki 2015, Kim and Lee 2016) and random forest (RF; Mann et al 2019) have been used to determine crop yields. Most recently, multiple studies have used deep learning algorithms to predict crop yields. You et al (2017) use the 2D convolutional neural network (CNN) to estimate crop yields with remote sensing images and show that it outperforms all statistical models, while Russello (2018) use 3D CNN to improve yield predictions. Wolanin et al (2020) input meteorological and satellite-derived vegetation data to CNN to predict wheat yield, while Newlands et al (2019) apply CNN with climate and weather variables to predict yield and risk.

The successful application of deep learning to account for the complex nonlinear relationships between crop yields and remote sensing and weather-related covariates has motivated us to apply this approach to predicting the area of flood damage to crops, which affects the food supply, market prices, and import and export planning. If accurate, such prediction can minimize the socioeconomical impact of crop loss (Ceglar et al 2018). In this study, we build a network based on CNN and demonstrate its accuracy to predict the county-level area of flood-damaged cropland for corn and soybeans in the midwestern United States. Using remote sensing images and meteorological variables, we consider flood impacts at different crop growth stages. We then test the sensitivity of the temporal resolution of the predictor variables and perform spatiotemporal error analysis of the model performance. The framework we developed in this study can be applied to predicting the loss of crops to flood damage using seasonal weather forecast models and help farmers to mitigate damage. In section 2, we briefly introduce the study area where we applied the designed framework to predict cropland damage. In section 3, we describe the data and the model structure, which is followed by the results analysis of model performance in section 4. In sections 5 and 6, we discuss the model results and limitations of the study and conclude with future work.
2. Study area

In this study, we use data on crop loss due to flood damage over the United States available from the Risk Management Agency at the US Department of Agriculture (USDA). Figure 1 shows the annual crop damage in top corn- and soybean-producing states. The range of damage to croplands for both corn and soybeans in these states is very wide. For some years, it may be close to nil, while in others it can affect as much as 2.5 million acres. The damage also varies significantly in spatial terms for given years. We focus our analysis on the 359 counties of Illinois, Iowa, Minnesota, and Missouri that have the highest corn and soybean production and are frequently affected by flooding (see figure 1). In these counties, corn and soybeans are planted in April through June and harvested in September through November (USDA 2018).

3. Methodology

3.1. Model structure

The CNN (LeCun et al 1998) is a class of deep neural networks used for analyzing 2D or 3D structured data, such as visual images and videos. By convolving across the input through multiple filters, the CNN model can learn important complex features as the activation map. So, the motivations to apply CNN in this study are: (a) CNN optimize the feature selection in the training and process automatically rather than subjectively in a physical process; (b) once a deep learning model is trained in a data rich region, one can use the transferring learning concepts to retrain the model in a data sparse region without changing the low-level feature descriptor.

Prior studies have also successfully applied the CNN model for learning temporal patterns of multispectral remote sensing images and weather variables (You et al 2017, Russello 2018, Wolanin et al 2020). Hence, in this study, we apply CNN over remote sensing images and meteorological weather variables to model the complex interconnection of flood and crop damage. Figure 2 shows the model architecture we use to train our input 3D histogram image (details in section 3.3) to predict the area of flood-damaged cropland.

Inspired by the good performance of the CNN for crop yield prediction (You et al 2017), we also design a CNN model architecture shown in figure 2 for flood damaged cropland area prediction. Following the convolution and activation, we use a stride-2 convolution layer instead of a pooling layer to reduce the size of the intermediate feature maps because the location invariant property of the pooling operator (Lecun et al 2015) does not apply to our input histogram image. Moreover, different locations of the input histogram image indicate different growth status. We select $3 \times 3$ as the kernel size of convolution. Each convolution is followed by batch normalization (Ioffe and Szegedy 2015) to tackle internal covariate shift, with a mini-batch size of 32. We select ReLU as the activation function. Figure 2 shows the stride-1 convolution layers in the darker color and the stride-2 layers in the lighter color, and a fully connected layer is attached at the end. The stride-2 convolution reduces the dimension of the output by half and outputs twice the number of input channels at the end of each convolution (128, 256, etc).

3.2. Predictors and the response variable

The USDA releases county-level monthly data on the number of acres of cropland damaged by natural disasters, including flooding, excess moisture, drought, and heat (USDA RMA 2019). If a cropland area were inundated but not damaged, it is not accounted for. In this study, we include as response variables only the cropland damaged area due to flooding. Due to the wide range of variation (in terms of magnitude) we use the damage data in a logarithm scale. As mentioned above, flooding causes greater crop loss when it occurs early in the growing season (Meyer et al 1987, Kanwar et al 1988, Mukhtar et al 1990, Lizaso and Ritchie 1997). Since corn and soybeans are at the vegetative stage and more vulnerable to excess moisture during May and June, the census data show more flood damage during this period. This consequence might also be contributed by more extreme precipitation occurrence during May and June than from July to November (figures 7(a) and (b)) in the study area. For each crop, therefore, we develop models for two seasons: an early season (May and June) model to predict high crop damage from flooding during the vegetative stage and a model to predict crop loss from late-season flooding (July–November) during the reproductive stage. For early, and late season models, we aggregate the actual monthly crop damage data to the two seasons as response variable.

We have applied the CNN architecture to two model scenarios, using both remote sensing and meteorological predictors (RS + Met) and meteorological predictors only (Met), as configured in table 1. As predictors, we used land surface remote sensing (hereafter referred to as remote sensing) and meteorological data and products, as listed in table 1. We select the predictors that can (a) provide insights on either flood drivers/impacts or related to crop growth or loss, and (b) are dynamically available at national or global scale in gridded format. For instance, precipitation is the main driver of flooding; vegetation indices can provide critical information on crop growth stress; and water fraction can indicate the extent of the flood. In the (RS + Met) we consider the combination of potential RS and meteorological predictors. However, since most selected RS products are available at low revisit frequency (8 d), in the Met scenario we consider the meteorological predictors alone that are related
Figure 1. Annual flood-damaged cropland areas for the major corn and soybean growing states. The green color indicates the states considered in this study.
to the water cycle components (ET, flow) as well as flood indication. The remote sensing predictors contains the level-2 diurnal temperatures (Wan et al. 2015); the EVI (Didan 2015), generated from the Moderate Resolution Imaging Spectroradiometer (MODIS); soil moisture (Kidd 2018), retrieved from microwave remote sensing emissions (Njoku et al. 2003, Shen et al. 2015, Arndt et al. 2020); and water fraction, retrieved from the Advanced Microwave Scanning Radiometer (AMSR; Du et al. 2017). We have extracted the meteorological predictors—precipitation, temperature, humidity, and radiation—from the North American Land Data Assimilation System (NLDAS2; Xia et al. 2012, 2012b) atmospheric reanalysis. In the RS + Met scenario, we have composited all predictors into eight-day intervals (Aggregate the 8 d total hourly precipitation, interpolate the 16 d EVI, and average 8 d daily WF, and soil moisture data) to match the availability of MODIS temperature data. Since Water Fraction data is available until 2018, we train the model for the RS + Met scenario up to this period. In the Met scenario, we composite the variables from hourly \( n \) days interval to reduce the computational cost of training the model, where for early season model, \( n = 1 \) and for late season model, \( n = 3 \). Specifically, we aggregate the \( n \) days total hourly precipitation, and average the \( n \) days hourly data for other Met predictors.

3.3. Input data permutation and model configurations

To consider only the pixels that represent the relevant crop fields in the predictor variables, we first extract pixels for each predictor from planting areas masked by the annual cropland maps available from 2008 to 2019 (Han et al. 2012). Then we compute the histogram of pixels at 32 evenly distributed bins. Figure 3 represents the workflow of generating the 2D histograms generation. At first, we show that for a given timestep, all cropland pixels are extracted from six predictors using the cropland mask. Then for each

![CNN Architecture where stride-1 convolution layers are in darker color and stride-2 convolution layers are in lighter color and a fully connected (FC) layer is attached in the end. For stride-1, output channel dimension remains same and becomes half during stride-2. Output channel number doubles at stride-2 convolutions (128, 256, etc). The output of 1 layer is input to the next layers. (a) Model Framework for histogram input of \( H^{20 \times 32 \times 6}, H^{50 \times 32 \times 6} \) and \( H^{60 \times 32 \times 6} \) and (b) Model Framework for histogram input of \( H^{8 \times 32 \times 6} \).]
Table 1. Predictors in the two model types.

| Model type                        | Predictors                                                                 | Temporal resolution in the model | Temporal extent | Source                                |
|-----------------------------------|-----------------------------------------------------------------------------|----------------------------------|-----------------|---------------------------------------|
| Remote Sensing and Meteorological | 1. Temperature (Day) 2000–2019                                               | 8 d                              | 2000–2019       | MOD11A2 (Wan et al 2015)              |
|                                   | 2. Temperature (Night) 2000–2019                                            |                                  | 2000–2019       | MOD13Q1 (Didan 2015)                  |
|                                   | 3. Enhanced Vegetation Index (EVI) 2000–2019                                |                                  | 1979–2019       | NLDAS2 (Xia et al 2012a, 2012b)      |
|                                   | 4. Precipitation                                                             |                                  | 1979–2019       | (Kidd 2018)                          |
|                                   | 5. Soil Moisture (SM) 1979–2019                                             |                                  | 2003–2018       | Advanced Microwave Scanning Radiometer (AMSR) land surface parameters (Du et al 2017) |
|                                   | 6. Water Fraction (WF; Fractional water over land)                          |                                  |                 |                                       |
| Meteorological                    | 1. Precipitation 1 d (Early season) 1979–2019                               | 1 d (Early season)               | 1979–2019       | NLDAS2                                |
|                                   | 2. Humidity                                                                  |                                  |                 |                                       |
|                                   | 3. Temperature (Min)                                                         |                                  |                 |                                       |
|                                   | 4. Temperature (Max)                                                         |                                  |                 |                                       |
|                                   | 5. Short Wave Radiation                                                      |                                  |                 |                                       |
|                                   | 6. Long Wave Radiation                                                       |                                  |                 |                                       |

Table 2. Histogram dimensions for different time and predictors.

| Time      | Meteorological predictors (Interval) | Remote sensing predictors (Interval) |
|-----------|-------------------------------------|--------------------------------------|
| Early season | \(H^{60 \times 32 \times 6} (1 \text{ d})\) | \(H^{6 \times 32 \times 6} (8 \text{ d})\) |
| Late season | \(H^{50 \times 32 \times 6} (3 \text{ d})\) | \(H^{20 \times 32 \times 6} (8 \text{ d})\) |

predic tor (marked red), we calculate the histogram as pixel counts in 32 intensity bins. Then, we concatenate a predictor’s 1D histograms of all time steps to form one 2D histogram image, \(h^{T \times b}_p\). Finally, stacking the 2D histogram images of all predictors, we get the input 3D histogram images, \(H^{T \times b \times p}\) to our CNN model (figure 2). In using a histogram, we assume the permutation invariance holds—that is, the distribution of intensity of pixels intensity would be more informative than the location of the pixels in estimating the total area of inundated croplands. Thus, for the histogram image of a predictor, \(h^{T \times b}_p\), the superscript \(T\) refers to the time series, the superscript \(b\) refers to the histogram bins, and \(p\) refers to the predictors. By stacking the histograms of all the predictor layers, \(h_p(h_1, \ldots, h_p)\), we obtain the 3D histogram as input to the model, \(H^{T \times b \times p}\). Table 2 shows the dimensions of the 3D histogram in each model scenario.

We have used USDA records from 2008 to 2019, excluding 2012 from training, as it was an extremely dry year in the United States (Anandhi 2016), and we want the model to learn only the features of flood-related crop damage. In the (RS + Met) scenario, since the early-season model has only eight time steps, we reduce the convolutional layers and include an additional dense layer (figure 2(b)).

3.4. Model training and hypergeometric parameter optimization

We use K-fold cross validation to train the model and fine tune the hypergeometric parameters: the dropout rate, the learning rate, training steps, and the number of neurons in the fully connected layer. Since the area of cropland loss varies greatly from one year to the next, we split the datasets based on location (county groups) instead of time (years) to preserve the years with extremely high damage in both the training and validation sets. Specifically, we divided the study area into four folds, as shown in figure 4. We divide the K-fold groups in a way that each group contains 25% randomly selected counties from each state. Such randomness can help prevent the model overfitted to a specific location. To test the model performance for an unseen year, we hold out data for one specific year for all counties as a test set while using the data from all other years for the cross validation (table 3). As shown in figures 7(a) and (b), the damage was comparatively low in 2016 and 2017 for both crops, while high in 2008 and 2014 for corn and 2015 for soybeans. Therefore, we leave each of the 4 years (2008, 2014, 2016, 2017 for corn and 2008, 2015–2017 for soybeans) out to test the model performance over high and low damage. In the cross-validation part, we still use four folds and subsample the training and validation sets based on county groups.

4. Results

In this section, we describe the results of the model cross validation using all the years but leaving one unseen year out.
Figure 3. Input data structure and 2D histogram generation.
Figure 4. Grouped counties for four-fold cross validation. In each fold of training, counties of any three groups are used for training while remaining group counties are used for validation. The test set includes data of an unseen year for all counties while all other years’ data are used during the cross validation.

Figure 5. Model performance evaluation during four-fold cross validation for corn.
4.1. Model performance

4.1.1. Four-fold cross validation by leaving one group of counties out

In the K-fold cross-validation process, each fold is trained for up to around 2500 iterations. The performance, measured by root mean square error (RMSE), is evaluated every 500 iterations. Initially, the RMSE of both training and validation sets is high, eventually decreasing almost monotonically with subsequent training steps to a low converged value (see figures 5 and 6). The iteration stops after around 10,000 iterations, when the RMSE of the validation set starts to increase, to prevent overfitting.

Figures 5 and 6 show the four-fold cross validation for $T_{2017}$. The cross validation of the other trainings ($T_{2008}$, $T_{2014}$, $T_{2015}$, and $T_{2016}$) also provide similar results, as shown in the appendix (figures A.3 and A.4). We have optimized the dropout rate at 0.5 to prevent overfitting after each convolutional layer—that is, it randomly drops 50% of the neurons and set the values to zero to generalize the training.

Note that since the damage is greater from early-season flooding, the RMSE of the early-season prediction is also higher than that of the late season for both crops. For corn (figure 5) in the RS + Met scenario, the RMSE is higher at convergence (around 8000 acres) during the early season than in the Met scenario (6000 acres). Similarly, during the late season, the lowest RMSE at the end of training becomes close to 3000 acres in the RS + Met scenario, while in the Met scenario it is approximately 2500 acres. The better performance of the Met model scenario over the RS + Met scenario in both seasons might be explained by the denser input in the time domain (60 and 50) of the 3D histogram as compared to RS + Met (8 and 20).

For the same reason—lower data frequency—the RMSE of soybeans (figure 6) during the late season is higher at the end of training in the RS + Met scenario (2250 acres) than in the Met scenario.
Figure 7. Evaluation of model performance for an unseen year at all the counties in Iowa, Illinois, Minnesota, and Missouri; actual monthly total damage in each year (a), (b); relative error of the model predictions for the unseen years using RS + Met (c), (d) and using only Met (e), (f), the dotted line (in figures (c)–(f)) represent ±0.25 relative error and the data of each marked year (in figures (a) and (b)) represent the test dataset in each training while data of all other years represent the training set.
(1500 acres), while during the early season, the converged RMSE for both scenarios are close (around 4000 acres).

4.1.2. Predicting an unseen year by leaving one year out

As we subsample the data based on county locations in the cross validation to avoid overfitting, we also evaluate the model performance for an unseen year by holding out data of either a high- or a low-damage year as the test set, as shown in table 3. Figures 7(a) and (b) show the actual observed flood damage for corn and soybeans for the two seasons (early, and late season), while figures 7(c)–(f) show the relative error (calculated using equation (1)) of the predicted damage for the unseen test years. The actual damage figure shows high damage in almost every year during early season for both corn and soybeans. The relative error bars (figures 7(c)–(f)) are higher for extreme-damage years, indicating larger uncertainties in these years. Figures 7(c)–(f) exhibit that, for almost all the test years, the median of relative errors is within ±0.25 (the dotted line). This indicates that for any unseen high- or low-damage year, the model can satisfactorily predict the damage in the study areas for both scenarios.

\[
\text{RE} = \frac{\text{pred} - \text{obs}}{\text{mean(\text{obs})}}. \tag{1}
\]

4.1.3. Predicting all-year mean damage for each county

Figures 8–11 show the spatial performance at the county level in terms of all year mean damage (2008–2019), RMSE, percentage of RMSE (ratio of RMSE to mean damage), and correlation coefficient. Mean damages are usually higher in Iowa, Illinois, lower Minnesota, and upper Missouri (above 5000 acres during early season and above 3000 acres during late season for both crops). In all cases (crops, model scenarios, and seasons), the prediction shows high correlation coefficient (R > 0.8), with especially high values in the counties with high flood damage, indicating the successful prediction of the spatial damage pattern. The models show negative correlation coefficient for the counties where damages are quite low (below 2000 acres during early season and below 1000 during late season). Figure 8 shows the model can predict well in the early season in the counties with areas of flood-damaged corn ranging from 5000 to 12 000 acres, with the RMSE below 8000 acres and the RMSE% below 75%. Figures 9 and 11 show that the late-season models for both crops and scenarios have better predictability (RMSE < 4000 acres).
and RMSE% < 75%) in the counties with mean damage ranging from 2000 to 6000 acres. In figure 10, the early-season soybean models in both scenarios exhibit better performance for the counties of mean damage between 4000 and 10 000 acres.

Overall, the low RMSE% shows that the models perform well for both crops in Iowa and Illinois. The high RMSE% in northern Minnesota and southern Missouri, partially caused by the low mean damage values, indicates poorer performance in these regions.

4.2. Predictor relative importance tests
To quantify the predictor efficiency, we perform a permutation test for RFs following Breiman (2001). Specifically, we consider the effect of randomly permuting the values of a specific predictor while holding the values of other predictors unchanged. Then we compute the RMSE of the model predictions for that shuffled predictor to determine the RMSE change of the predictions compared to the model with actual predictors,

\[ \text{Increase in RMSE} \% = \frac{\text{RMSE}_{\text{shuffled predictor}} - \text{RMSE}_{\text{all actual predictors}}}{\text{RMSE}_{\text{all actual predictors}}} \times 100. \]

The greater the increase in RMSE, the more important the predictor. For each county, we calculate the percentage increase in RMSE using equation (2).

Figure 12 shows the predictor importance in the Met and RS + Met model scenarios. In the Met scenario, based on the extent of the error bar and the median above 0% RMSE increase, precipitation, minimum temperature, and longwave radiation show higher importance for corn during the early-season flooding, while humidity, shortwave radiation, and longwave radiation show higher priority during the late-season flooding. For soybeans, humidity, longwave radiation, maximum temperature, and precipitation are more important in the early season and precipitation and minimum temperature seem less important in the late season. In the RS + Met scenario for corn, daytime temperature, EVI, and precipitation show priority during the early season flooding, while EVI and soil moisture (figure 6) for the late season. For soybeans, both day- and night-time temperature, EVI, and precipitation are important during the early-season flooding, and all predictors except precipitation seem important during the late season (figure 6).

From the importance test, we observe that, not only precipitation but temperature, EVI, and
humidity also play an important role in the severity of flood damage in most cases. Flooding with warmer temperatures can cause plants to deteriorate, as heat causes quick depletion of stored energy (Sullivan et al. 2001). Corn may not survive 24 h if the temperature is greater than 77°F. Soybean survival during a flood is also highly affected by temperature (Wuebker et al. 2001). Remotely sensed EVI indicates the actual greenness state of the crops, and high humidity reduces evapotranspiration and withdrawal of nutrients from the soil, with the result that plants eventually get rotten (Grange and Hand 1987).

5. Discussion

The good performance of the model indicates the potential of extending it to national and global scale, given the availability of historical records on cropland damage. The end-to-end capability of the model avoids the use of detailed inundation and crop flood resistance information in the complex physical process of cropland inundation, which are not usually available.

We find the model to perform satisfactorily in both tested scenarios, Met and RS + Met, although the performance is limited by the low revisiting frequency (8 d) of the remote sensing imagers. Considering its better performance and ability to integrate with climate and seasonal climate datasets, the Met scenario shows broader potential for agricultural risk management and food security applications. The better performance in the Met scenario paves the way to utilize weather variables from seasonal forecast models, such as the Climate Forecast Model (CFS; Saha et al. 2014), to predict risk of cropland flood-damage ahead of the season or to use future climate projections, such as the Coupled Model Intercomparison Project phase 3 (CMIP3; Meehl et al. 2007), to analyze crop losses due to flooding in a future climate. These predictions can timely support agriculture planning for farmers, as well as disaster mitigation and risk estimation by government agencies and the industry (e.g. insurances).

We apply the CNN framework for the histogram images from remote sensing products and meteorological weather variables. Since the damage from flooding is more intense during early season (May and June) than late season (July–November), we break the time window into two seasons for better performance, and we test the model in soybean- and corn-growing counties in the midwestern United States that are rich in census data. The results show that the model can capture the damage trends and spatial patterns satisfactorily and can predict damage areas.

Figure 10. Spatial evaluation of model performance of soybean for the counties in Iowa, Illinois, Minnesota, and Missouri; left column—actual total damage; second column (from left) RMSE; third column (from left) RMSE% which the percentage ratio of the RMSE and mean actual damage; right column correlation coefficient.
Figure 11. Spatial evaluation of model performance of Soybean for the counties in Iowa, Illinois, Minnesota, and Missouri; left column—actual total damage; second column (from left) RMSE; third column (from left) RMSE% which the percentage ratio of the RMSE and mean actual damage; right column correlation coefficient.

Figure 12. Predictor relative importance test for the Met and (RS + Met) model scenario.
for unseen counties and years. The test of predictor importance indicates that precipitation, temperature, EVI, and humidity are important, which is consistent with the mechanistic understanding of crop damage caused by flooding.

The limitations of the model include the high uncertainty for an unseen extreme-damage year and the exclusion of topographic information. Although the models perform well for the counties where damage is in the average range (Iowa and Illinois), it exhibits comparatively biased predictive skills for regions with low damage (Missouri) or extremely high damage for a few years (Minnesota in 2014), which is a common phenomenon in data-driven models. It also indicates that the labeled data used in the training set might not be sufficient, and this uncertainty could be reduced by incorporating more damage records from upcoming years. The other limitation of the model is the utilization of dynamic predictors only. Due to the model structure limitation, we have not included static topographical and geomorphological parameters (Shen et al. 2016, 2017), that might be closely interconnected with flood and crop damage, such as relative elevation (Nobre et al. 2016), slope, proximity to rivers, drainage density, and relief ratio. The incorporation of these static parameters may facilitate learning by the model of geographical variation to capture distinct spatial characteristics of flood damage.

6. Conclusions

In this study, we develop a deep learning model based on CNN architecture to predict the area of damage caused by the flooding of croplands. Methodologically, the advantage of CNN is that it can extract hierarchical features from a 2D image including spectral, geometric, and positional patterns (Rawat and Wang 2017). By inputting these features into a neural network, it is proved very effective in object identification. Since CNN successfully mimics human’s vision-brain logic of identifying targets, applying CNN to traditional quantitative sciences is to utilize human’s vision identification logic in these subjects. In future development, we will (a) modify the model architecture to accommodate more flood-related indicators including topographic and geomorphological variables, (b) utilize the model for seasonal forecast and future project of flood security, and (c) apply the transferring learning concept to retrain the proposed model for data sparse region, which is pretrained in a data rich region (e.g., the US).

Data availability statement

The datasets generated and analyzed that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgement

This work is supported by Eversource Energy Center. The authors would like to thank all the supporting agencies, and the United States Department of Agriculture (USDA) for providing valuable data support during the completion of this research.

Appendix

The CNN models

In a typical neural network (NN), all the input units are connected to all the output units via a fully connected (FC) layer. Each fully connected layer takes a vector as input followed by a nonlinear function usually a rectified linear unit (ReLU) or tanh and finally outputs a vector such that,

\[ c = f(Wx + b) \]

where, \( W \) is the weight matrix, and \( b \) is the bias. However, in a CNN, each output unit is connected to only a subset of the entire input units where the hidden layers convolve with a multiplication or other dot product. The input subsets are known as receptive fields figure A.1 shows the schematic of 2D convolution (\( x \) and \( y \) direction) with padding, that means adding additional layers of zero values. Without padding, information of the middle pixels is captured more during convolution than the corner pixels. So, to prevent losing information of the corner pixels and limit shrinking of the input, padding is done. Convolution can be done with different stride interval of the filters. For stride-1 convolution with padding, output channel dimension is same as input, while for stride-2 it is reduced to half. CNN is usually composed of convolutional layers, nonlinear activation function and pooling layers.

Convolutional layers

A convolutional layer extract different features at different depth in the network by using filters to perform convolution over the input. Each layer learns from the previous layer. The input to a 2D CNN is a dimension of \( c \times w \times h \) where \( c \) is the number of channels (in a RGB image there are three channels), \( w \) is width, and \( h \) is the height. In each convolution there can be \( n \) number of output channels with kernel size of \( k \times k \) to convolve over \( k \times k \) spatial dimension of the receptive fields. Thus, there are \( n \) number of filters of dimension, \( c \times k \times k \) in a convolutional layer. Figure A.2(a) shows the schematic of the workflow of convolutional layers.

Activation function

After each convolution nonlinear activation function is applied to introduce nonlinearity. Most commonly used activation function is ReLU (Nair and Hinton 2010) where negative values are converted to zero. So,
Figure A.1. Schematic of 2D CNN for stride-1 and stride-2 with zero padding (adding additional layers of zero values) to prevent shrinking. For stride-1 output channel dimension is same as input and for stride-2 output channel dimensions reduced to half of the input dimension.

Figure A.2. Schematic of (a) convolutional layer; (b) activation function; (c) max pooling.

ReLU \( (x) = \max(0, x) \). ReLU has overcome the vanishing gradient problem of sigmoid and hyperbolic tangent activation function and allows deep neural models to learn faster and perform better. Again, Leaky ReLU (LReLU; Maas et al 2013) improves gradient flow of ReLU by setting a small positive slope for the negative values. So, \( \text{LReLU} \ (x) = \alpha x(x < 0) \) where \( \alpha \) is a small positive constant. Figure A.2(b) represents the diagram of the activation functions.

Pooling

A pooling layer is applied after the activation function. While performing the convolution with small stride the receptive fields overlap repetitively and the information can be redundant. To filter the most relevant features, usually maximum pooling is used at each convolution layer. Specifically, a small window \( (2 \times 2) \) and a same size of stride is applied along the input and the maximum value in that window is considered in the output. Pooling allows the location invariant property. Figure A.2(c) indicates the max pooling.

Regularization techniques

Due to the complexity of the neural networks, the deep learning model is more prone to overfitting than machine learning models. So, it is necessary to implement regularization to penalize the weight matrices during backpropagation and prevent overfitting. L1 and L2 regularization, dropout, and early stopping are the most common regularization techniques.

In regularization technique, the general cost function is updated by adding another term known as the regularization term such as Cost Function = Loss + Regularization Term. For regression, the loss function can be mean squared error or mean squared logarithm error. Regularization term differs in L1 and L2. In case of L1, Cost Function = Loss + \( \lambda \sum_{i} w^2 \)
and for L2, Cost Function = \text{Loss} + \lambda \sum |w|.

\lambda (\text{lambda}) is the regularization hyperparameter and \text{w} is weights. Whereas dropout method randomly drops neurons by setting values to zero at certain rate to generalize the model. Again, in early stopping, training is stopped when validation error starts to increase although the training error drops with training steps. Moreover, data augmentation, that is a small shifting of the image pixels while training is also an effective way to prevent overfitting.

**Cross validation results for different test years**

The cross-validation results for leaving one year out (2008, 2014, and 2016 for corn; 2008, 2015, 2016 for soybean) are shown from figures A.3–A.6. In figures A.3 and A.4 we show the comparison of Met and RS + Met scenarios. Figures A.9 and A.10 show the comparison of the train and validation for different scenarios. In all cases (crop, seasons, and train/validation), overall RMSE for the RS + Met scenario are higher than the Met scenario.
Model Performance for Soybean

![Figure A.4. Model performance evaluation during four-fold cross validation for soybean for the Met and RS + Met scenario. The solid lines represent the train RMSE, the dotted lines represent the validation RMSE while different colors represent the test sets mentioned in table 3.](image)

Model Performance for Corn

![Figure A.5. Comparison of the train and validation results for the Met and RS + Met scenario for corn. The solid lines represent the Met scenario, the dotted lines represent the RS + Met scenario while different colors represent the test sets mentioned in table 3.](image)
Model Performance for Soybean

Figure A.6. Comparison of the train and validation results for the Met and RS + Met scenario for soybean. The solid lines represent the Met scenario, the dotted lines represent the RS + Met scenario while different colors represent the test sets mentioned in table 3.

ORCID iD

Xinyi Shen
https://orcid.org/0000-0002-7411-003X

References

Allen C, Ashlock L, Baldwin F, Beaty D, Cartwright R and Chapman S 2000 (University of Arkansas Division of Agriculture) (www.uaex.edu/publications/pdf/mp197/mp197cover_toc.pdf) (Accessed 11 March 2021)

Anandhi A 2016 How does the drought of 2012 compare to earlier droughts in Kansas, USA? J. Appl. Serv. Climatol. 9 1

Arguello M N, Mason R E, Roberts T L, Subramanian N, Acuña A, Addison C K, Lozada D N, Miller R G and Gbur E 2016 Performance of soft red winter wheat subjected to field soil waterlogging: grain yield and yield components Field Crop Res. 194 57–64

Arndt D S, Blunden J and Dunn R J H 2020 State of the climate in 2019 Bull. Am. Meteorol. Soc. 101 SI–S8

Bailey-Serres J, Lee S C and Brinton E 2012 Waterproofing crops: effective flooding survival strategies Plant Physiol. 160 1698–709

Boori M S, Choudhary K, Evers M and Paringer R 2017 A review of flood risk and flood risk dynamics in central dry zone area of Myanmar Proc. Eng. 201 231–8

Breiman L 2001 Statistical modeling: the two cultures J. Hortic. Sci. 62 199–231

Ceglar A, Toreti A, Prodhomme C, Zampieri M, Turco M and Doblas-Reyes F J 2018 Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast Sci. Rep. 8 1322

Chen H, Liang Q, Liang Z, Liu Y and Xie S 2019 Remote-sensing disturbance detection index to identify spatio-temporal varying flood impact on crop production Agric. For. Meteorol. 269–70 180–91

Chen H, Liang Z, Liu Y, Liang Q and Xie S 2017 Integrated remote sensing imagery and two-dimensional hydraulic modeling approach for impact evaluation of flood on crop yields J. Hydrol. 553 262–75

Di L, Yu E G, Kang L, Shrestha R and Bai Y Q 2017 RF-CLASS: a remote-sensing-based flood crop loss assessment cyber-service system for supporting crop statistics and insurance decision-making J. Integr. Agric. 16 408–23

Didan K 2015 MOD13Q1—MODIS/terra vegetation indices 16-day l3 global 250m SIN grid NASA Lp Daac. (https://doi.org/10.5067/MODIS/MOD13Q1.006)

Du J, Kimball J S, Jones L A, Kim Y, Glassy J and Watts J D 2017 A global satellite environmental data record derived from AMSR-E and AMSR2 microwave Earth observations Earth Syst. Sci. Data 9 791–808

Dubois P C and Engman T 1995 Measuring soil moisture with imaging radars IEEE Trans. Geosci. Remote Sens. 33 915–26

Espinosa L and Ross J 2003 (University of Arkansas Division of Agriculture) (www.uaex.edu/publications/pdf/mp437/mp437.pdf) (Accessed 11 March 2021)

Food and Agriculture Organization of the United Nations (FAO) 2016 Damages and Losses from climate-related disasters in agricultural sectors (www.fao.org/3/i6486e/i6486e.pdf) (Accessed 11 March 2021)

Grange R I and Hand D W 1987 A review of the effects of atmospheric humidity on the growth of horticultural crops J. Hortic. Sci. 62 125–34

Haas E M, Bartholomé E and Combal B 2009 Time series analysis of optical remote sensing data for the mapping of temporary surface water bodies in sub-Saharan western Africa J. Hydrol. 370 52–63
Hake K, Dan Krieg J L and Oosterhuis D 1992 Cotton physiology today Natl. Cott. Coun. 5 7 (www.cotton.org/tech/physiology/cpt/planphysiology/upload/CPT-July92-REPOF.pdf) (Accessed 11 March 2021)

Han W, Yang Z, Di L and Mueller R 2012 CropScape: a web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support Comput. Electron. Agric. 84 111–23

Ioffe S and Szegedy C 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift 32nd Int. Conf. on Machine Learning, ICML 448–56 (http://proceedings.mlr.press/v37/ioffe15.pdf) (Accessed 11 March 2021)

Kanwar R S, Baker J L and Mukhtar S 1988 Excessive soil water effects at various stages of development on the growth and yield of corn Trans. Am. Soc. Agric. Eng. 31 133–41

Kenyon W, Hill G and Shannon P 2008 Scoping the role of agriculture in sustainable flood management Land Use Policy 25 351–60

Kidd R 2018 Algorithm theoretical basis document soil moisture products from active and passive microwave sensors 3–77

Kim N and Lee Y W 2016 Machine learning approaches to corn yield estimation using satellite images and climate data: a case of Iowa state J. Korean Soc. Surv. Geod. Photogramm. Cartography 34 383–90

Kirchmeier-Young M C and Zhang X 2020 Human influence has intensified extreme precipitation in North America Proc. Natl. Acad. Sci. USA 117 13308–13

Kawata K and Shibasaki R, 2015 The University of Tokyo IIS, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153–805, Japan 2015 IEEE Int. Geosci. Remote Sens. Symp. pp 858–61

Lauer J 2008 Flooding impacts on corn growth and yield Field. Crop. 28 49–56 (http://corn.agronomy.wisc.edu/AA/pdfs/A056.pdf) (Accessed 11 March 2021)

Lazin R, Shen X, Kouokoula M and Anagnostou E 2020 Evaluation of the hyper-resolution model-derived water cycle components over the upper Blue Nile Basin J. Hydrool. 590 125251

Lecun Y, Bengio Y and Hinton G 2015 Deep learning Nature 521 436–44

LeCun Y, Bottou L, Bengio Y and Haffner P 1998 Gradient-based learning applied to document recognition Proc. IEEE 86 2278–24

Lesk C, Rowhani P and Ramankutty N 2016 Influence of extreme weather disasters on global crop production Nature 529 84–7

Li S, Tompkins A M, Lin E and Ju H 2016 Simulating the impact of flooding on wheat yield—case study in East China Agric. For. Meteorol. 216 221–31

Lizaso J I and Ritchie J T 1997 Maize shoot and root response to Lesk C, Rowhani P and Ramankutty N 2016 Influence of extreme weather disasters on global crop production Nature 529 84–7

Lizaso J I and Ritchie J T 1997 Maize shoot and root response to Lesk C, Rowhani P and Ramankutty N 2016 Influence of extreme weather disasters on global crop production Nature 529 84–7

Mann M L, Warner J M and Malik A S 2019 Predicting high-magnitude, low-frequency crop losses using machine learning: an application to cereal crops in Ethiopia Clim. Change 154 211–23

Meehl G A, Covey C, Delworth T, Latif M, McAvaney B, Mitchell J F B, Stouffer R J and Taylor K E 2007 The WCRP CMIP3 multimodel dataset: a new era in climatic change research Bull. Am. Meteorol. Soc. 88 1383–94

Merz B et al 2014 Floods and climate: emerging perspectives for flood risk assessment and management Nat. Hazards Earth Syst. Sci. 14 1921–42

Meyer W S, Barrs H D, Mosier A R and Schafer N L 1987 Response of maize to three short-term periods of waterlogging at high and low nitrogen levels on undisturbed and repacked soil Irrig. Sci. 8 257–72

Mohammadi A, Costello J F and Ryu D 2017 Application of time series of remotely sensed normalized difference water, vegetation and moisture indices in characterizing flood dynamics of large-scale arid zone floodplains Remote Sens. Environ. 199 270–82

Mukhtar S, Baker J L and Kanwar R S 1990 Corn growth as affected by excess soil water Trans. Am. Soc. Agric. Eng. 33 437–42

Nair V and Hinton G E 2010 Rectified linear units improve restricted Boltzmann machines ICML 2010—Proc., 27th Int. Conf. on Machine Learning June 2010 807–14 (www.cs.toronto.edu/~hinton/absps/reluCMLE.pdf) (Accessed 11 March 2021)

Newlands N, Ghabari A, Gel Y R, Lyubchivich V and Mahdi T 2019 Deep learning for improved agricultural risk management Proc. 52nd Hawaii Int. Conf. Syst. Sci. vol 6 pp 1033–42

Njoku E G, Jackson T J, Lakshmhi, V, Chan T K and Nghiem S V 2003 Soil moisture retrieval from AMSR-E IEEE Trans. Geosci. Remote Sens. 41 215–29

Nobre A D, Cuartas L A, Momo M R, Severo D L, Pinheiro A and Nobre C A 2016 HANDB contour: a new proxy predictor of inundation extent Hydrol. Process. 30 320–33

Ohlig J, Bledau G, Delenene C, Bailly J S, Bader J C, Oleksiak A, Ferry L and Martin D 2015 Decadal monitoring of the Niger inner delta flood dynamics using MODIS optical data J. Hydrool. 523 368–83

Oh Y, Sarabandeh K and Ulaby F T 2002 Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces IEEE Trans. Geosci. Remote Sens. 40 1348–53

Pacetti T, Caporalì E and Rulli M C 2017 Floods and food security: a method to estimate the effect of inundation on crops availability Adv. Water Resour. 110 494–504

Panivel K and Surianarayanan C 2019 An approach for prediction of crop yield using machine learning and big data techniques Int. J. Comput. Eng. Technol. 10 110–8

Rawat W and Wang Z 2017 Deep convolutional neural networks for image classification: a comprehensive review Neural Comput. 29 2352–449

Reynoso M A, Kajala K, Bajic M, West D A, Woodhouse M, Rodriguez-mediina J and Sinha N 2019 Evolutionary flexibility in flooding response circuitry in angiosperms Science 1295 1291–5

Rosenzweig C, Tubiello F N, Goldberg R, Mills E and Bloomfield J L 2002 Increased crop damage in the US from excess precipitation under climate change Glob. Environ. Change 12 197–202

Russello H 2018 Convolutional neural networks for crop yield prediction using satellite images (https://esc.fnwi.uva.nl/thesis/centraal/files/f1570224447.pdf) (Accessed 11 March 2021)

Saha S et al 2014 The NCEP climate forecast system version 2 J. Clim. 27 2185–208

Scott H D, DeAngelis J, Daniels M B and Wood L S 1989 Flood duration effects on soybean growth and yield Agron. J. 81 631–6

Seddon A W R, Macias-Fauria M, Long P R, Benz D and Willis K J 2016 Sensitivity of global terrestrial ecosystems to climate variability Nature 531 229–32

Shen X and Anagnostou E N 2017 A framework to improve hyper-resolution hydrological simulation in snow-affected regions Comput. Electron. Agric. 138 113–26

Shen X, Anagnostou E N, Allen G H, Robert Brakenridge G and Ketterer A J 2019 Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar Remote Sens. Environ. 221 302–15

Shen X, Anagnostou E N, Mei Y and Hong Y 2017 A global distributed basin morphometric dataset Sci. Data 4 160124
Shen X, Hong Y, Qin Q, Basara J B, Mao K and Wang D 2015 A
semipysical microwave surface emission model for soil
moisture retrieval IEEE Trans. Geosci. Remote Sens. 
53 4079–90
Shen X, Hong Y, Qin Q, Yuan W, Chen S, Zhao S and Grout T 2011
Orientation angle calibration for bare soil moisture
estimation using fully polarimetric SAR data IEEE Trans.
Geosci. Remote Sens. 49 987–96
Shen X, Mao K, Qin Q, Hong Y and Zhang G 2013 Bare surface
soil moisture estimation using double-angle and
dual-polarization L-band radar data IEEE Trans. Geosci.
Remote Sens. 51 3931–42
Shen X, Vergara H J, Nikolopoulos E I, Anagnostou E N, Hong Y,
Hao Z, Zhang K and Mao K 2016 GDBC: a tool for
generating global-scale distributed basin morphometry
Environ. Model. Softw. 83 212–23
Shen X, Wang D, Mao K, Anagnostou E and Hong Y 2019b
Inundation extent mapping by synthetic aperture radar: a
review Remote Sens. 11 879
Shrestha R, Di L, Yu E G, Kang L, Y Zheng S H A O and Bai Y Q
2017 Regression model to estimate flood impact on corn
yield using MODIS NDVI and USDA cropland data layer J.
Integr. Agric. 16 398–407
Siebert S, Webber H and Rezaei E E 2017 Weather impacts on
crop yields—searching for simple answers to a complex
problem Environ. Res. Lett. 12 081001
Sullivan M, VanToai T, Fauney N, Beuerlein J, Parkinson R and
Soboyejo A 2001 Evaluating on-farm flooding impacts on
soybean Crop Sci. 41 93–100
Tabari H 2020 Climate change impact on flood and extreme
precipitation increases with water availability Sci. Rep. 
10 13768
Tewari S and Arora N K 2016 2—Soybean Production under
Flooding Stress and Its Mitigation Using Plant
Growth-Promoting Microbes Environmental Stresses in
Soybean Production 2 23–40
Torbbck N, Chowdhury D, Salas W and Qi J 2017 Monitoring rice
agriculture across Myanmar using time series sentinel-1
assisted by landsat-8 and PALSAR-2 Remote Sens. 9 119
Tsynganskaya V, Martinis S, Marzahn P and Ludwig R 2018
SAR-based detection of flooded vegetation—a review of
characteristics and approaches Int. J. Remote Sens. 
39 2255–93
USDA RMA 2019 Cause of Loss Historical Data Files
(www.rma.usda.gov/SummaryOfBusiness/CauseOfLoss) 
(Accessed 11 March 2021)
USDA 2018 Crop Progress and Conditions (www.nass.usda.gov/
Charts_and_Maps/Crop_Progress_%26_Condition/2018/
index.php) (Accessed 11 March 2021)
Van Komenenburg T, Kasahamu A and Catil C 2020 Crop yield
prediction using machine learning: a systematic literature
review Comput. Electron. Agric. 177 105709
Vozinazi A E K, Karatzas G P, Sibetheros I A and Varouchakis E A
2015 An agricultural flash flood loss estimation
methodology: the case study of the Koliasis basin (Greece),
February 2003 flood Nat. Hazards 79 899–920
Wan Z, Hook S and Hulgy L 2015 MOD11A2 MODIS/terra land
surface temperature/emissivity 8-day L3 global 1km SIN
grid V006 NASA EOSDIS L Process. DAAC. (https://doi.org/
10.5067/MODIS/MYD11A2.006)
Wang X, Deng Z, Zhang W, Meng Z, Chang X and Lv M 2017
Effect of waterlogging duration at different growth stages on
the growth, yield and quality of cotton PLoS One 
12 e0169029
Wolanin A, Matao-Garcia G, Camps-Valls G, Gmez-Chova L,
Meroni M, Duveiller G, Liangzhi Y and Guanter L 2020
Estimating and understanding crop yields with explainable
deep learning in the Indian Wheat Belt Environ. Res. Lett. 
15 024019
World Crop Damage 2011 Global Crop Damage Timeline
(www.ritterfarms.com/cropdamage.html) (Accessed 11
March 2021)
Wuebker E F, Mullen R E and Koehler K 2001 Flooding and
temperature effects on soybean germination Crop Sci. 
41 1857–61
Xia Y et al 2012a Continental-scale water and energy flux analysis
and validation for North American land data assimilation
system project phase 2 (NLDAS-2): 2. Validation of
model-simulated streamflow J. Geophys. Res. Atmos. 
117 03110
Xia Y et al 2012b Continental-scale water and energy flux analysis
and validation for the North American land data
assimilation system project phase 2 (NLDAS-2): 1.
Intercomparison and application of model products J.
Geophys. Res. Atmos. 117 03109
Xu M, Ma H, Zeng L, Cheng Y, Lu G, Xu J, Zhang X and Zou X
2015 The effect of waterlogging on yield and seed quality at
the early flowering stage in Brassica napus L Field Crop Res. 
180 238–45
Yang J, Hao Z, Wu L, Wang T and Zhang G 2016 Indicator-based
evaluation of spatiotemporal characteristics of rice flood in
Southwest China Agric. Ecosyst. Environ. 230 221–30
Yang Q, Shen X, Anagnostou E N, Mo C and Eggleston J R 2021 A
high-resolution flood inundation archive (2016—the
present) from sentinel-1 SAR imagery over CONUS Bull.
Am. Meteorol. Soc. 1–40
Yeung E et al 2018 A stress recovery signaling network for
enhanced flooding tolerance in Arabidopsis thaliana Proc.
Nat. Acad. Sci. USA 115 E6085–94
You J, Li X, Low M, Lobell D and Ermon S 2017 Deep Gaussian
process for crop yield prediction based on remote sensing
data 31st AAAI Conf. Artif. Intell. AAAI 2017 pp 4559–65
(https://aaai.org/ocs/index.php/AAAI/AAAI17/
paper/view/14435/14067) (Accessed 11 March 2021)
Zhang Y, Chen Y, Lu H, Kong X, Dai J, Li Z and Dong H 2016
Growth, lint yield and changes in physiological attributes of
cotton under temporal waterlogging Field Crop Res. 
194 83–93