World Bank Big Data Innovation Challenge
Supplementary Material

Jiaxuan You, Xiaocheng Li, Stefano Ermon

November 12, 2016

Contents

1 Introduction

In this supplementary material, we provide quantitative and qualitative performance metrics for our system, looking at historical soybean and corn yields. We select 11 states in the U.S. that account for over 75% of the national soybean production and use data from 2003 to 2015. We compare our results with other existing approaches from the literature as well as the USDA predictions.

2 Visualization of the prediction results

2.1 Prediction before harvest

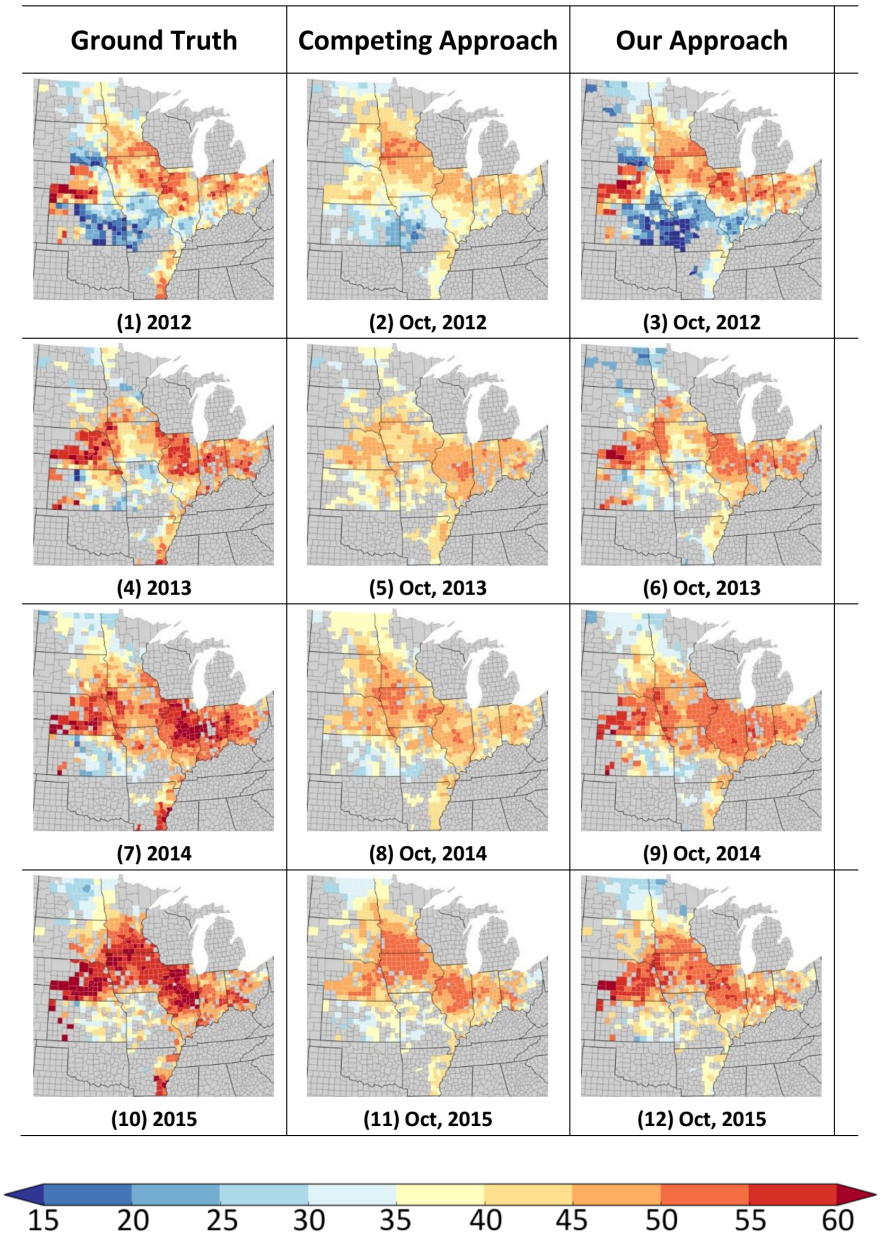


Figure 1: County-level soybean yield maps from 2012 to 2015, measured in bushel per acre. Predictions are made in October. Our predicted yields match the ground-truth data fairly well, and largely outperform competing approaches.

2.2 Real-time prediction throughout the year

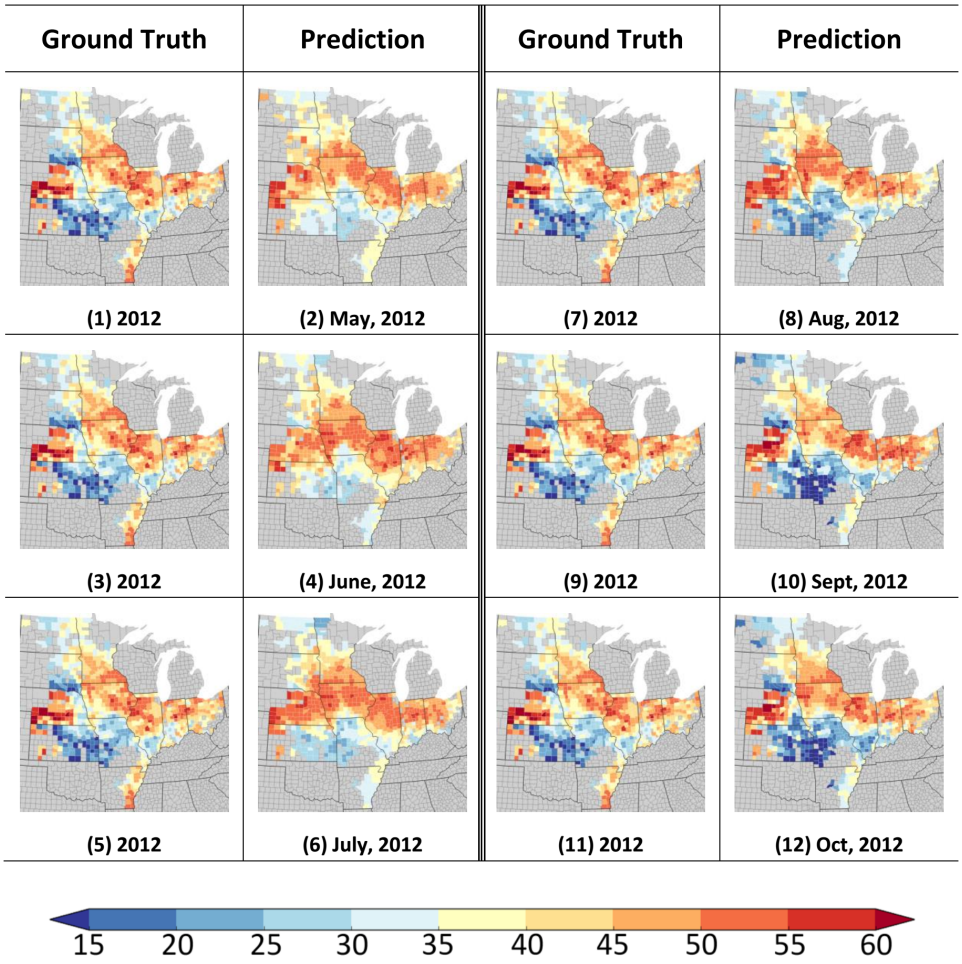


Figure 2: County-level soybean yield maps in 2012, with monthly prediction from May to October. A severe drought occurred in July and August 2012. The maps show how our model gradually identifies the potential yield decrease (as early as in July), and finally makes an accurate prediction.

2.3 Out-of-region predictions

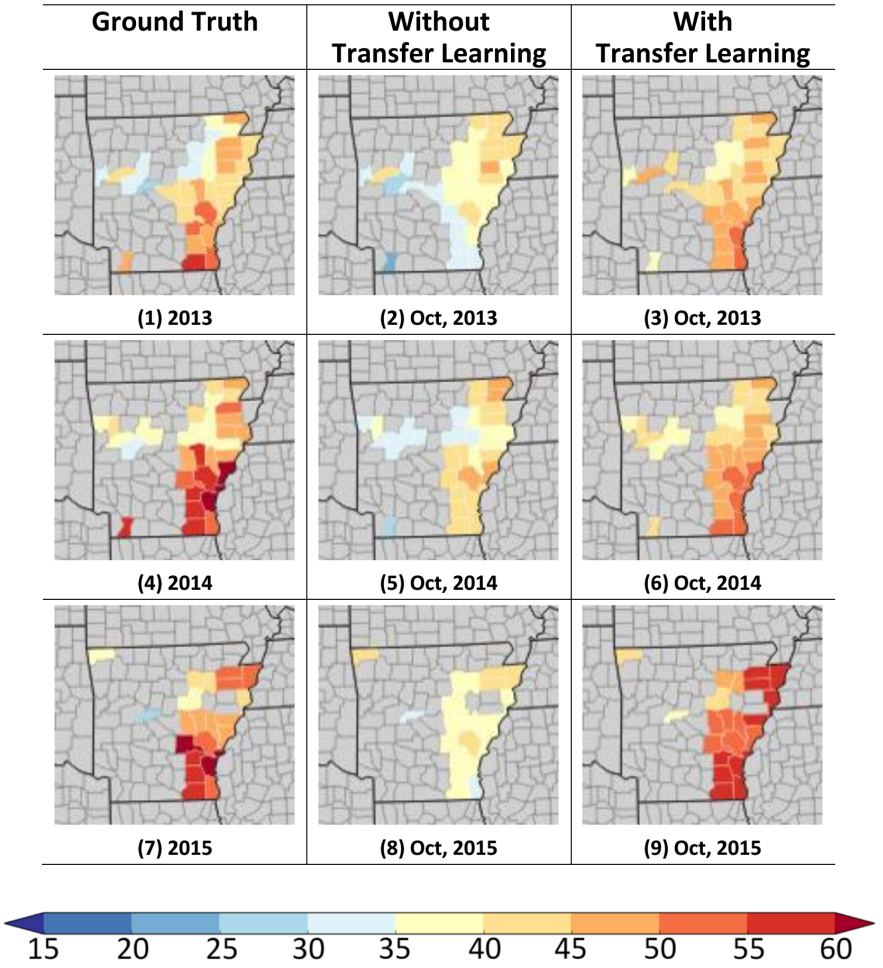


Figure 3: Soybean yield maps from 2013 to 2015 in Arkansas. The figures show how the idea of transfer learning helps our model scale to other regions. Arkansas exhibits a rather unique crop yield pattern, as is shown in the bottom of Fig. 1. After adopting transfer learning ideas, the prediction bias is largely corrected. This indicates that a model trained in one region can be adapted to work in other regions with minimal additional training data.

2.4 Application to Corn Yield Predictions

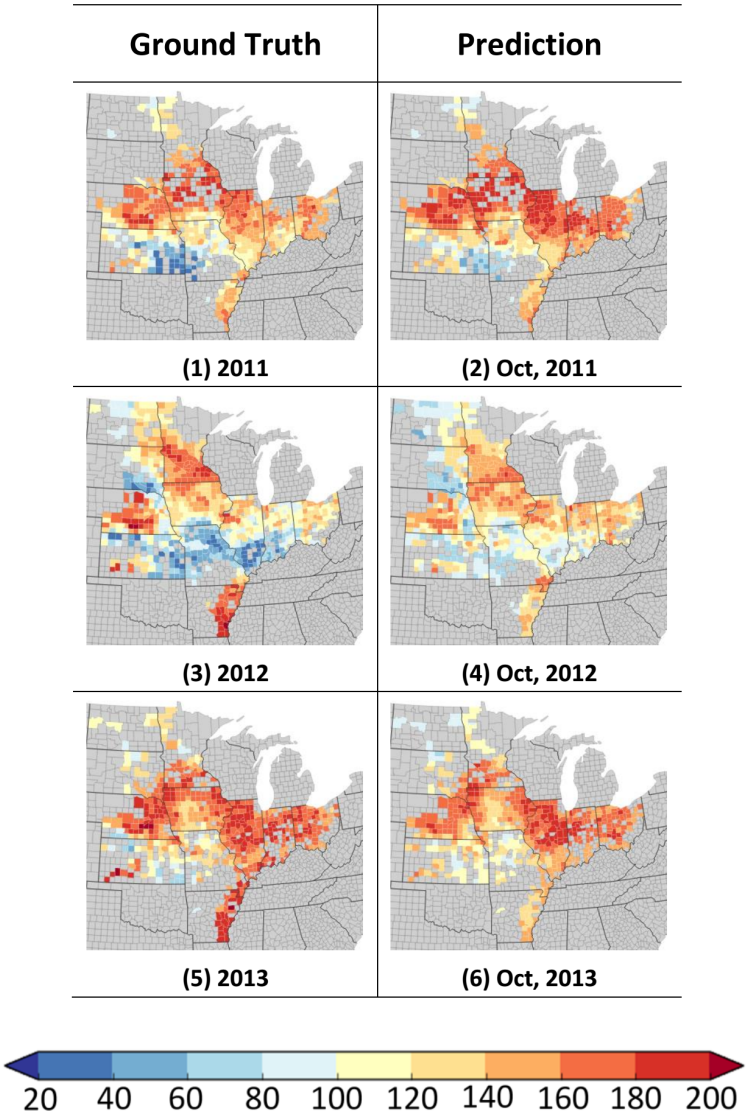


Figure 4: Corn crop yield maps from 2011 to 2013, measured in bushel per acre. The experiment is done without any modification on the model structure, indicating that our model can generalize to different crops.

3 Comparing with traditional approaches and USDA predictions

3.1 RMSE compared with traditional approaches

Year	Ridge	Tree	DNN	**Ours**
2011	9.00	7.98	9.97	**5.76**
2012	6.95	7.4	7.58	**5.91**
2013	7.31	8.13	9.2	**5.5**
2014	8.46	7.5	7.66	**5.27**
2015	8.10	7.64	7.19	**6.4**
Average	7.96	7.73	8.32	**5.77**

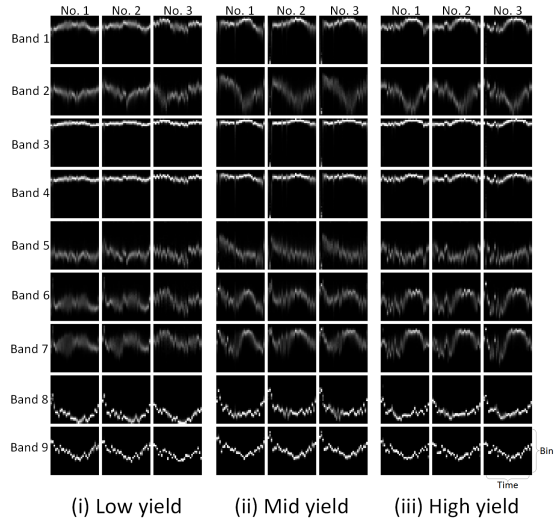
Table 1: Root of Mean Square Errors (RMSE) among all the counties: Our model significantly outperforms traditional approaches by at least 30%. Traditional approaches include ridge regression (Ridge), decision tree (Tree) and Deep Neural Network (DNN) on normalized Difference Vegetation Index (NDVI) features.

3.2 MAPE compared with USDA

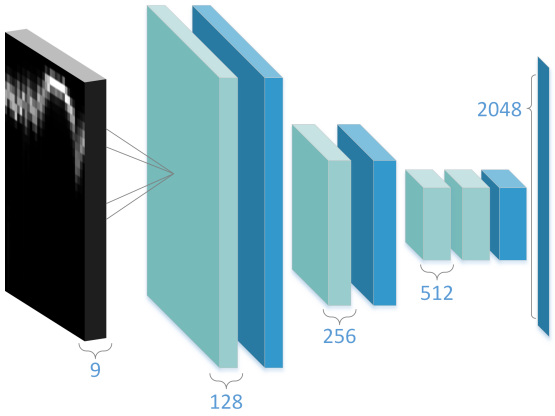
Year	Ours (Jul.)	USDA (Aug.)	Ours (Aug.)	USDA (Sept.)	Ours (Sept.)	USDA (Oct.)	Ours (Oct.)
2009	-4.26	-5.23	**-3.84**	-3.86	**1.12**	-3.64	**1.22**
2010	-7.02	**1.15**	-2.64	**2.76**	-5.93	**2.07**	-3.76
2011	7.22	**-1.43**	6.62	**-0.48**	7.14	**-1.19**	6.90
2012	11.3	-9.75	**1.04**	-11.75	**-1.63**	-5.50	**3.33**
2013	-1.47	-3.18	**3.17**	-6.36	**-2.36**	N/A	**-2.15**
2014	3.53	-4.42	**1.67**	**-0.54**	-4.42	-0.84	**-0.71**
2015	-4.77	**-2.29**	-4.62	**-5.15**	-8.38	**-1.67**	-4.26
Absolute Mean	5.65	3.92	**3.37**	4.14	**3.41**	**2.48**	3.19

Table 2: Mean absolute percentage error (MAPE): our model outperforms USDA predictions by 15% on average in August and September. Note that USDA predictions are survey-based, which can be costly and cannot be easily scaled to other regions, whereas our model only uses publicly available, passively collected data.

4 Model Structure



(a) 3-D histogram visualization



(b) The CNN structure

Figure 5: Visualization of the input data and used convolutional neural network (CNN) architecture. **(a)** Figures of typical 3-D histograms $\boldsymbol{T} \in \mathbb{R}^{b \times m \times d}$ flattened in the band dimension d under (i) low crop yield, (ii) mid crop yield and (iii) high crop yield conditions are shown in the left panel. Each row of squares represents a different spectral band, while each column represents an individual data point. Each square is a slice of $\boldsymbol{T}$, where the x-axis corresponds to the "time" dimension m, and the y-axis to the "bin" dimension b. Brighter pixels indicate higher pixel counts in that bin. There exists distinctive visual differences between high yield and low yield conditions (for example in the second and the seventh bands). **(b)** The adopted CNN structure, where stride 1 convolution layers are in light blue, stride 2 convolution layers are in dark blue and a fully connected layer is attached at the end.