# Using the Concept of Ecoregions for Large Area Crop Mapping

Venkata Shashank Konduri[1,2], Jitendra Kumar[2], William W. Hargrove[3]
Forrest M. Hoffman[2] and Auroop Ganguly[1],

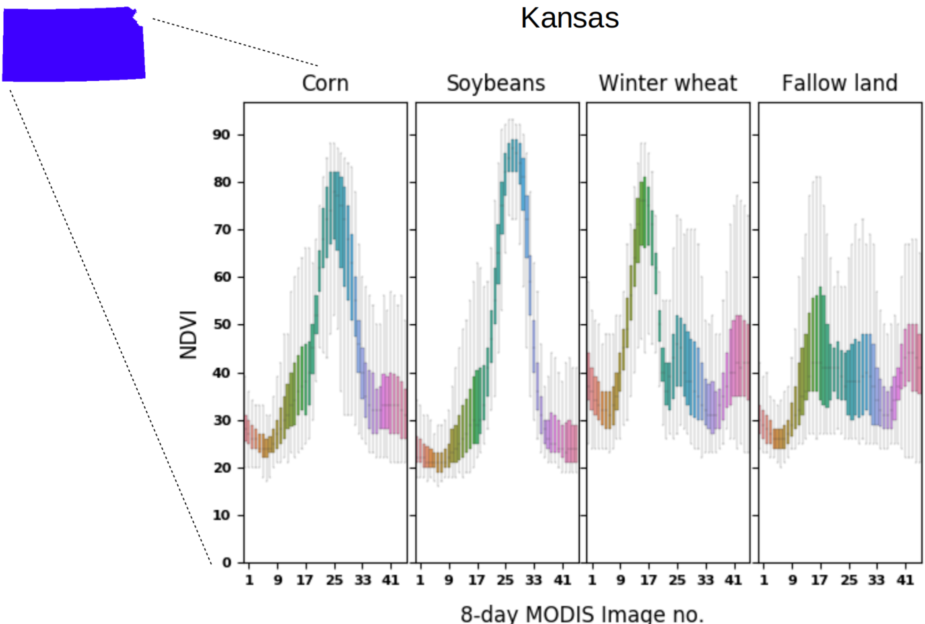
[1]Northeastern University, [2]Oak Ridge National Laboratory; and [3]USDA Forest Service

April 8, 2019

# Introduction: Crop mapping for US

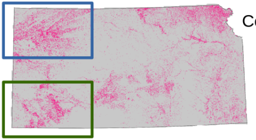
- The United States is a major food producer in the world accounting for about 30% of the world grain exports and crop cultivation accounts for nearly 80% of all water use.
- The Cropland Data Layer (CDL) provided by USDA for CONUS based on extensive ground reference data collected during the mapping year.
- Field data collection for CDL extremely time consuming,expensive and labor-intensive. It is not available for access to the general public.
- The georeferenced raster map is not released until the beginning of the subsequent calendar year for market sensitivity reasons.

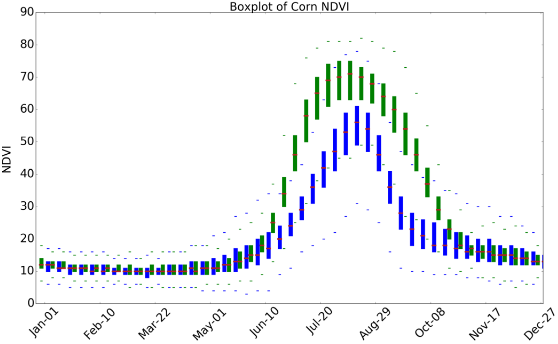
# Crops exhibit wide range of variability in phenology



# Spatial Variability in Phenology

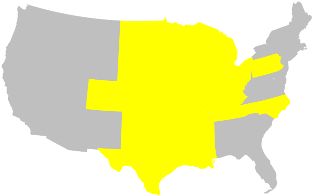
One of the challenges in remote sensing-based large area crop mapping is the **variability across ecological zones, which can result in different timing of crop phenological development**.



# Objectives

1. *Develop a generalized phenology based classification approach to map major crops across the "Extended Corn Belt" region*.



2. *Perform the classification at the scale of ecoregions*.
3. Evaluate model performance using error metrics like Producer Accuracy, User Accuracy, Error Matrix and Percent Deviation.

# Data

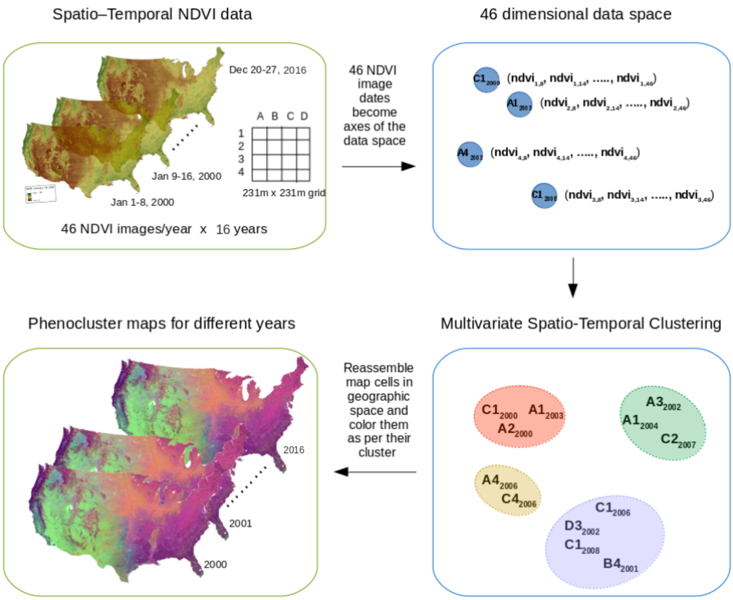
## Remotely Sensed Data

Smoothed and gap-filled **MODIS NDVI** data for the entire CONUS for the period 2000-2016.
(J. Spruce, G. Gasser, and W. Hargrove. MODIS NDVI data, smoothed and gap-filled, for the Conterminous US: 2000-2015)

## Reference Data

The study was performed for the cropland extent for the CONUS defined by the **USDA Cropland Data Layer (CDL)** for the years 2008-2015 at 30m resolution.

# Classification (Step 1) - Creation of Phenoclusters



# Classification (Step 2) - Assigning Crop Labels to Phenoclusters

Crop Data Layer

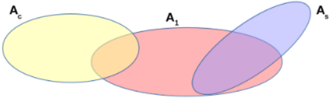
| | | | |
|---|---|---|---|
| Corn | Corn | Soy | Soy |
| Corn | Soy | Soy | Soy |
| Soy | Soy | Soy | Soy |
| Soy | Soy | Soy | Corn |

Assign crop labels to clusters →

Phenocluster map

| | | | |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 1 |
| 1 | 1 | 1 | 2 |
| 1 | 1 | 2 | 2 |

Find Goodness-of-fit (GOF) for every cluster with each crop




GOF for Cluster 1 with Corn = $\frac{A_c \cap A_1}{A_c} * \frac{A_c \cap A_1}{A_1}$

GOF for Cluster 1 with Soy = $\frac{A_s \cap A_1}{A_s} * \frac{A_s \cap A_1}{A_1}$

⋮

For all crops

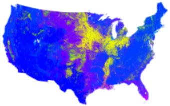
Assign the cluster to the crop with the best fit →

Translate Table

| Cluster | Crop |
|---|---|
| 1 | Soy |
| 2 | Corn |
| 3 | Rice |
| 4 | Sorghum |
| ⋮ | |
| N | Corn |

Reassign the labels geographically →

Phenocluster map reclassed to crop layer categories



# Training and Testing the model

- The model has been tested each year from 2008-2015.
- For each individual state, selected crop progress stages from USDA weekly reports were processed to measure interannual similarity (Zhong et al., 2016).

Table: Selecting the training year for every state based on phenological similarity

| State | Mapping Year | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
| Iowa | 2013 | 2014 | 2011 | 2010 | 2010 | 2008 | 2011 | 2011 |
| Indiana | 2013 | 2011 | 2012 | 2008 | 2010 | 2008 | 2013 | 2013 |

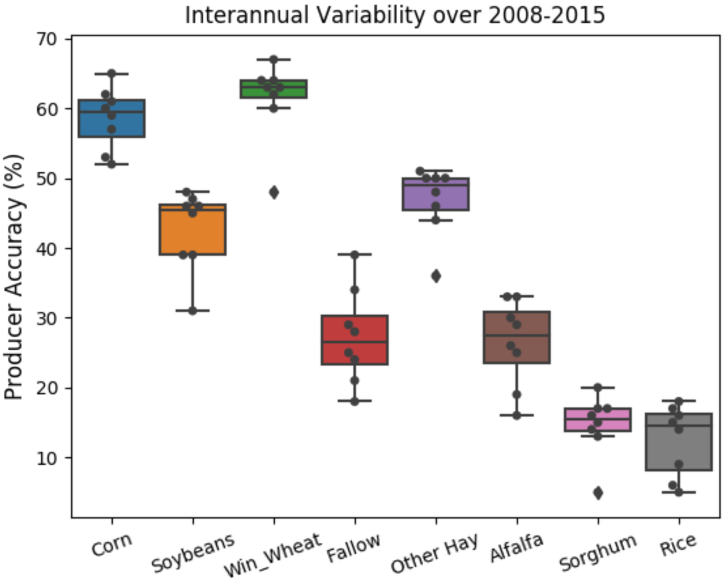
# Data used for ecoregion creation

Table: Environmental variables used for ecoregion delineation. These data are in the form of ~1 km raster grids.

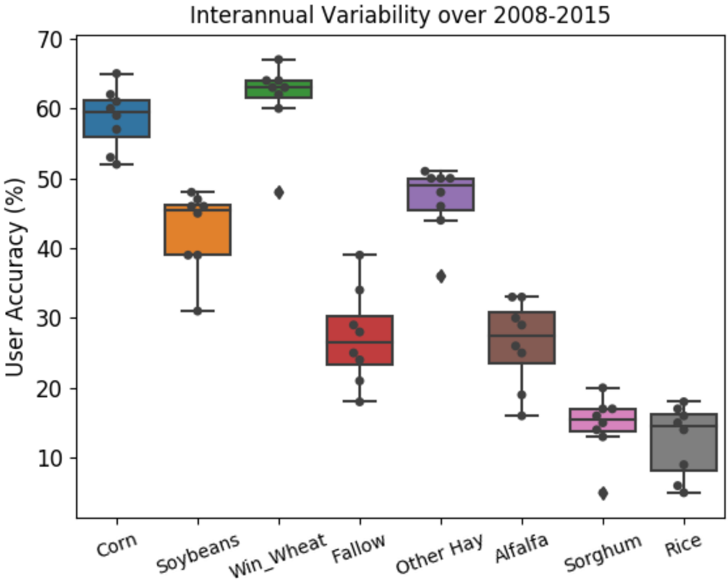
| Variable Description | Units | Source |
|---|---|---|
| **Bioclimatic Variables** | | |
| Annual mean temperature | °C | Fick and Hijmans (2017) |
| Mean diurnal range | °C | Fick and Hijmans (2017) |
| Isothermality | — | Fick and Hijmans (2017) |
| Temperature seasonality | °C | Fick and Hijmans (2017) |
| Mean temperature of warmest quarter | °C | Fick and Hijmans (2017) |
| Mean temperature of coldest quarter | °C | Fick and Hijmans (2017) |
| Annual precipitation | mm | Fick and Hijmans (2017) |
| Precipitation seasonality | mm | Fick and Hijmans (2017) |
| Precipitation during the wettest quarter | mm | Fick and Hijmans (2017) |
| Precipitation during the driest quarter | mm | Fick and Hijmans (2017) |
| **Edaphic Variables** | | |
| Available water holding capacity of soil | mm | Global Soil Data Task Group (2000); Saxon et al. (2005) |
| Bulk density of soil | $g/cm^3$ | Global Soil Data Task Group (2000); Saxon et al. (2005) |
| Soil carbon density | $g/m^2$ | Global Soil Data Task Group (2000); Saxon et al. (2005) |
| Total nitrogen density | $g/m^2$ | Global Soil Data Task Group (2000); Saxon et al. (2005) |
| **Topographic Variables** | | |
| Compound topographic index (relative wetness) | – | Saxon et al. (2005) |

# Dividing CONUS into 500 ecoregions

# Producer Accuracy (Probability that a CDL pixel will be correctly mapped)



## User Accuracy (Probability that a reclassed map pixel matches the CDL)



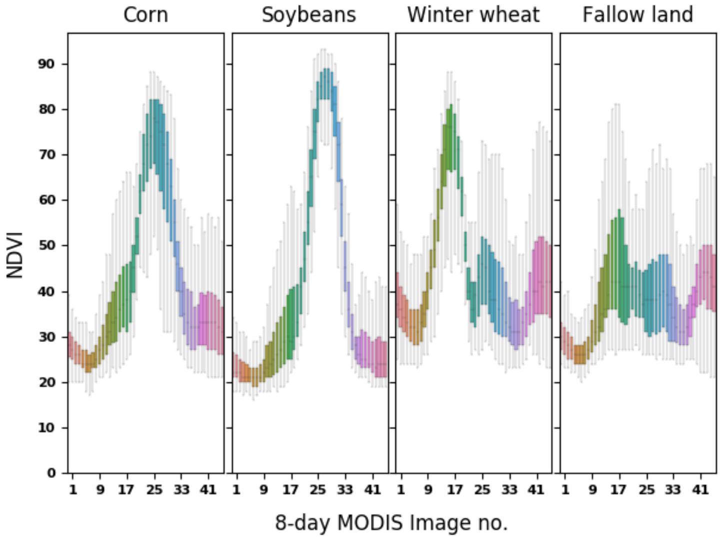
# Crop types getting misclassified

- Corn - Soybeans
- Other Hay with most other crops
- Winter Wheat and Fallow

| | | Crop Data Layer | | | | | | | | User |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Corn | Rice | Sorghum | Soybeans | WinWht | Alfalfa | Other Hay | Fallow | Accuracy (%) |
| Reclassed Map | Corn | 208 | 0 | 5 | 135 | 11 | 15 | 7 | 8 | 51 |
| | Rice | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 20 |
| | Sorghum | 1 | 0 | 5 | 0 | 2 | 0 | 0 | 3 | 35 |
| | Soybeans | 87 | 5 | 3 | 145 | 6 | 5 | 8 | 10 | 47 |
| | WinWht | 12 | 0 | 15 | 4 | 73 | 2 | 3 | 22 | 49 |
| | Alfalfa | 4 | 0 | 0 | 2 | 2 | 10 | 3 | 1 | 37 |
| | Other Hay | 7 | 0 | 0 | 9 | 3 | 3 | 25 | 3 | 42 |
| | Fallow | 2 | 1 | 2 | 3 | 12 | 0 | 2 | 12 | 30 |
| **Producer Accuracy (%)** | | 62 | 5 | 15 | 46 | 63 | 26 | 46 | 18 | |

Table: Error Matrix for CONUS (2015)

# Winter Wheat and Fallow



# Reasons behind misclassification

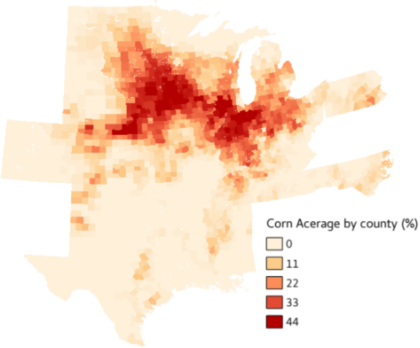
- Similar phenologies for certain crops.
- Lower CDL accuracies for lesser grown crops

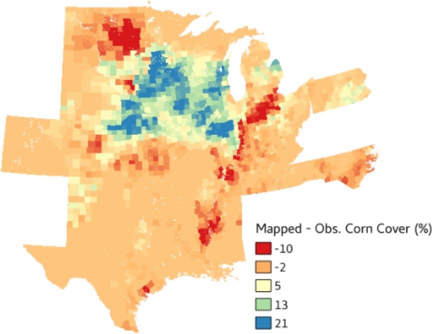
| Crop type | Area (x 1000 ha) | Producer Acc (%) | User Acc (%) |
|---|---|---|---|
| Winter Wheat | 399 | 94.4 | 94.5 |
| Corn | 179 | 93.2 | 93.6 |
| Soybeans | 144 | 92.9 | 92.9 |
| Fallow | 124 | 87.5 | 87.8 |
| Sorghum | 124 | 89.3 | 89.3 |
| **Other Hay/Non Alfalfa** | 50 | **56.1** | 90.4 |
| Double Crop WinWht/Soy | 35 | 85.9 | 85.3 |
| Alfalfa | 22 | 85.9 | 91.2 |
| **Cotton** | 0.82 | **62.7** | **87** |
| **Potatoes** | 0.13 | **75.7** | 95.4 |

Table: Accuracy table for the 2013 CDL for Kansas (Downloaded from the USDA CDL website)

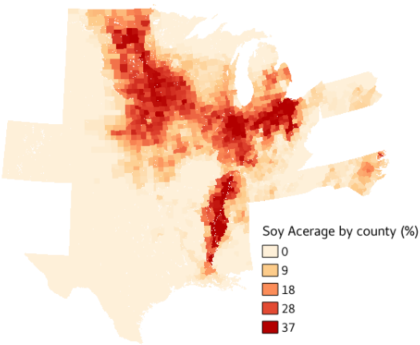
- Mixed pixel effect

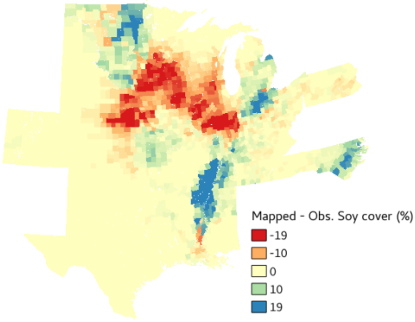
# Corn Acreage aggregated to county



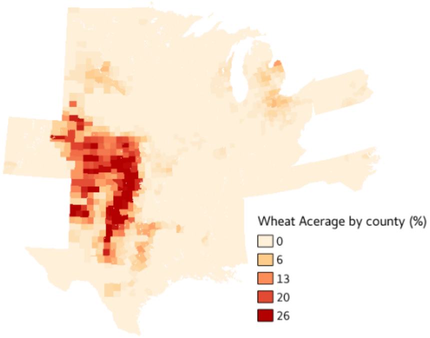



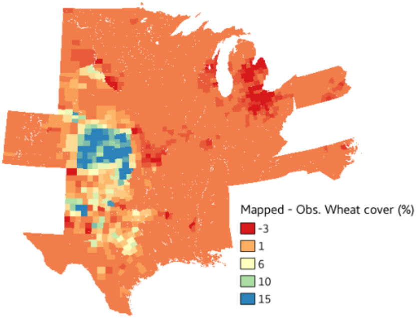
# Soybeans Acreage aggregated to county



# Wheat Acreage aggregated to county



# Conclusions

- Land surface phenology can be used to identify and map crops at continental (to global) scale.
- While accuracy are high for dominant crops, they tend to be be lower for less dominant crops (in part due to mixed pixel effect and limited training data).
- Use of ecoregions helps to reduce crop misclassification by addressing spatial variability and allowing for development of more specific models.
- Interannual variability in phenology can have a significant impact on the accuracies.
- When aggregated to the county scales, there is an over prediction in acreages in the dominant crop growing regions and an under prediction in the less dominant areas to the tune of about 10%.

# Future Work

- Develop a system for continuous tracking and mapping of agricultural ecosystem using near real time remote sensing (similar to USDA Forest Service ForWarn).
- Estimate crop yield based on phenological trajectory/completion metric through the season.

# Acknowledgements

This research was sponsored by the U.S. Department of Agriculture Forest Service, Eastern Forest Environmental Threat Assessment Center (EFETAC) and the U.S. Department of Energy Biological and Environmental Research (BER) program. This research used resources of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

# References

S. E. Fick and R. J. Hijmans. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315, 2017. ISSN 1097-0088. doi: 10.1002/joc.5086. URL http://dx.doi.org/10.1002/joc.5086.

Global Soil Data Task Group. Global gridded surfaces of selected soil characteristics (igbp-dis). [global gridded surfaces of selected soil characteristics (international geosphere-biosphere programme - data and information system)]., 2000. URL http://dx.doi.org/10.3334/ORNLDAAC/569.

E. Saxon, B. Baker, W. Hargrove, F. Hoffman, and C. Zganjar. Mapping environments at risk under different global climate change scenarios. 8(1):53–60, Jan. 2005. doi: 10.1111/j.1461-0248.2004.00694.x.

L. Zhong, L. Yu, X. Li, L. Hu, and P. Gong. Rapid corn and soybean mapping in us corn belt and neighboring areas. *Scientific reports*, 6:36240, 2016.