Using Machine Learning to Develop a Predictive Model for Future Fire Seasons

Andrew T. White[1], Christopher R. Hain[2], Christopher J. Schultz[2], Jonathan L. Case[3], Kristopher D. White[4]

[1]UAH/NASA MSFC SPoRT

[2]NASA MSFC SPoRT

[3]ENSCO, Inc./NASA MSFC SPoRT

[4]NWS/NASA MSFC SPoRT

THE UNIVERSITY OF ALABAMA IN HUNTSVILLE

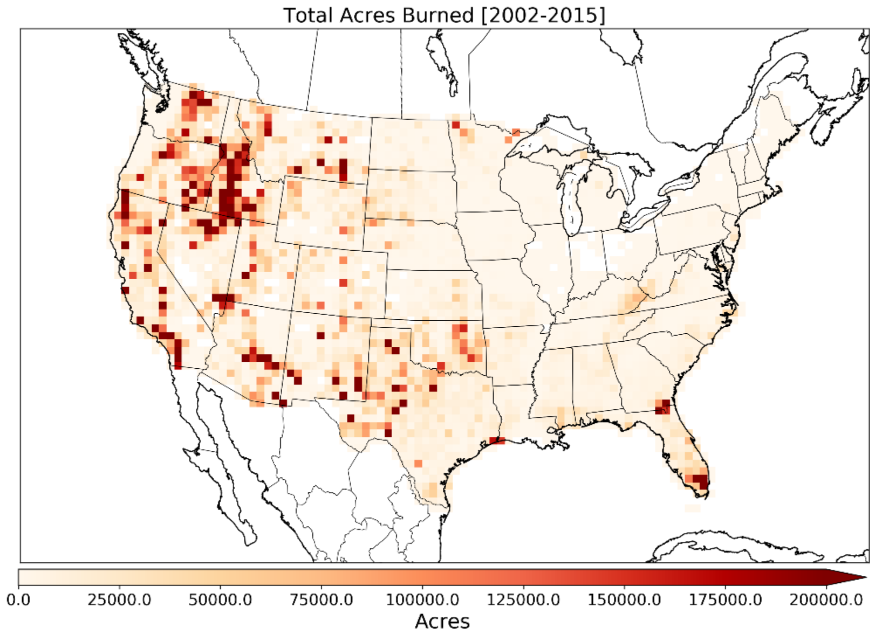
Short-term Prediction Research and Transition Center

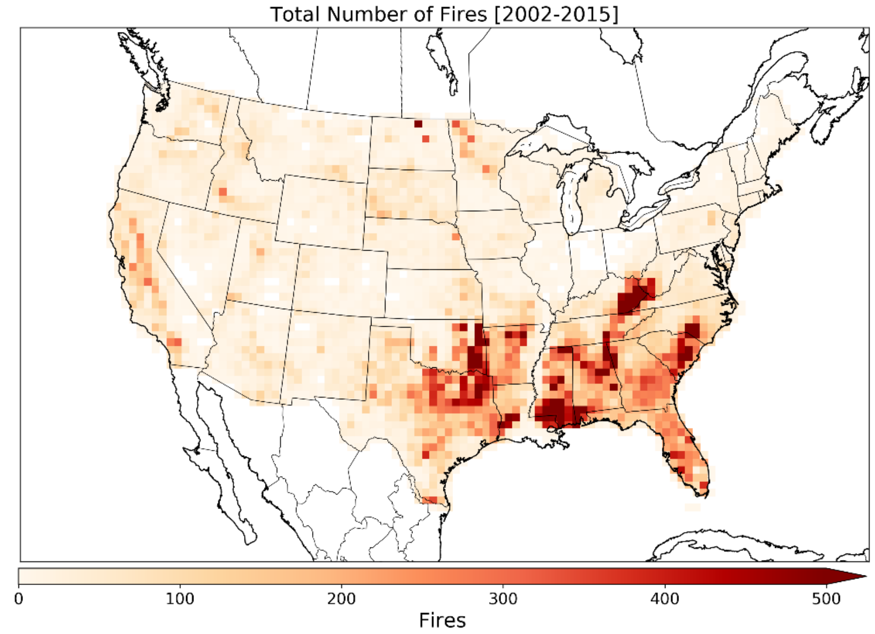
Wildfires Overview

- Wildfires can have devastating social and economic impacts.
  - Loss of life
  - Air quality impacts
  - Yearly costs are rising

- Expansion of the wildland-urban interface (WUI) is leading to increased fire risk.

- While much of the wildfire focus is in the west, a significant number of fires do occur in the eastern U. S.

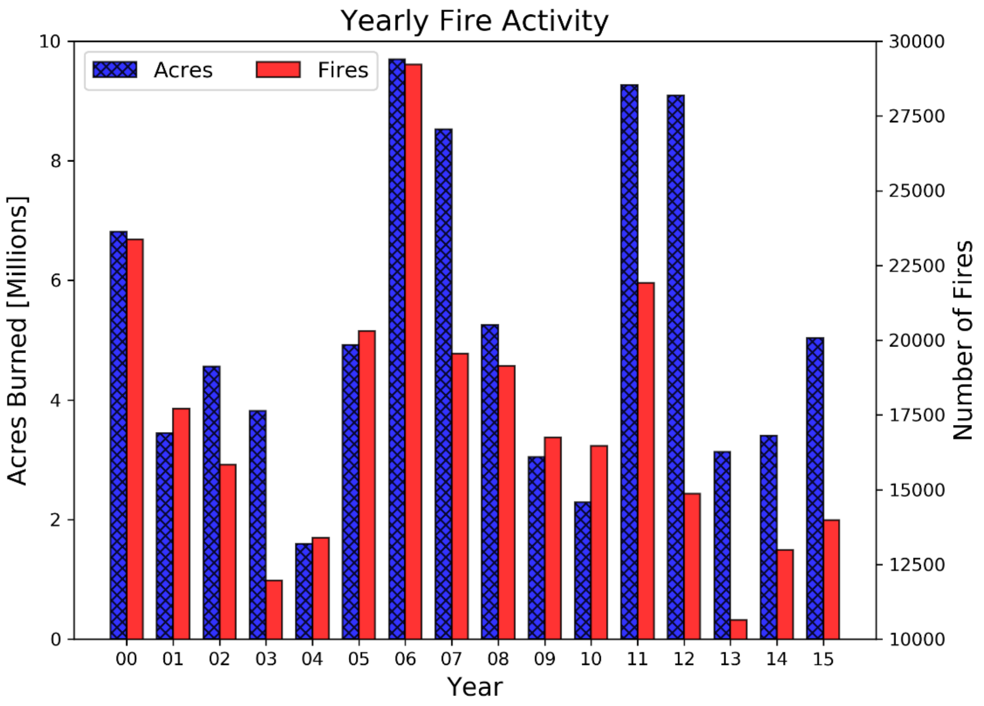




Yearly Variation

- ➢ Fire activity is highly variable from year to year.

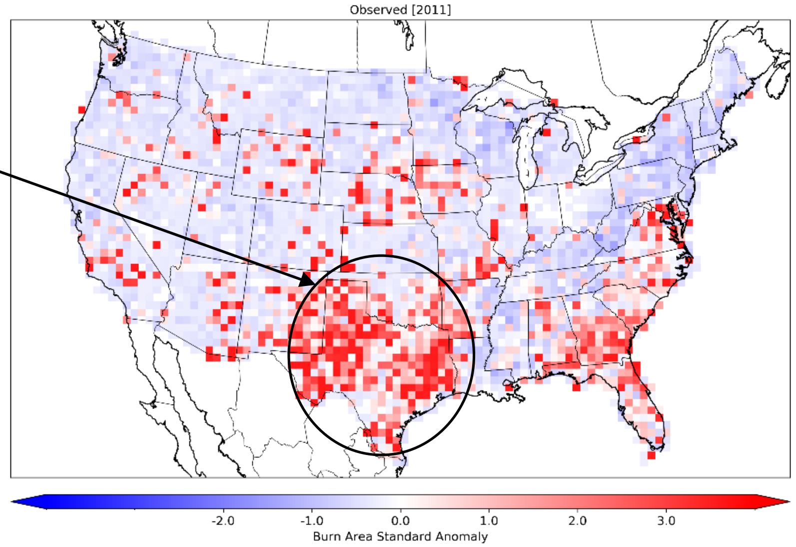
- ➢ Yearly changes in fire activity are related to changes in both atmospheric and land surface conditions.
  - ➢ Numerous amounts of available data related to fire potential (i.e. dead fuel moisture, soil moisture, precipitation, temperature, moisture, etc.)
  - ➢ Antecedent conditions provide an indication about potential fuel availability and dryness.



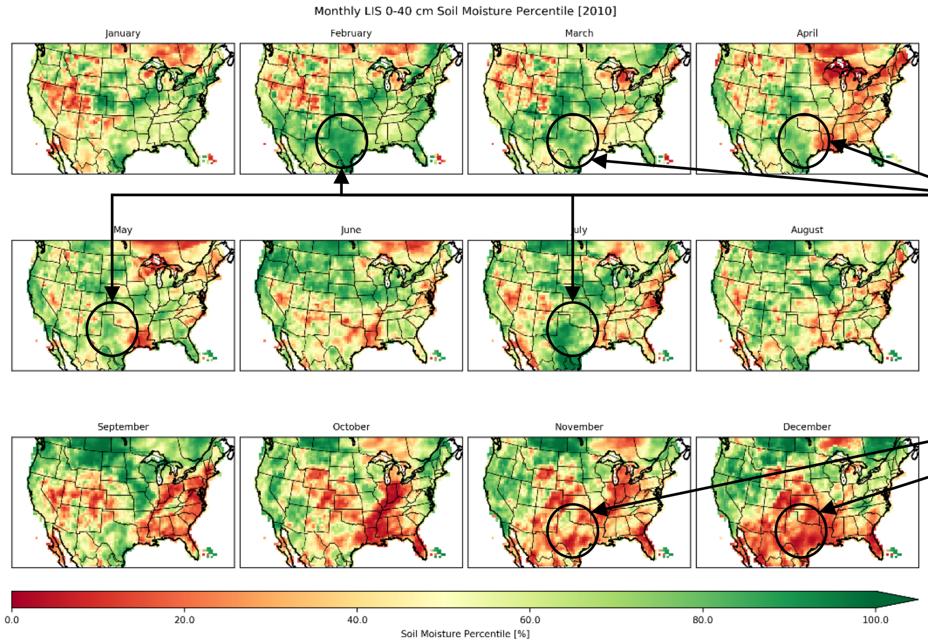
Yearly number of fires and acres burned across the CONUS domain. Indicates high year to year variability.

Antecedent Relationships

- Standardized burn area anomaly for 2011 shows anomalous wildfire activity over much of Texas.
- SPoRT LIS 0 – 40 cm Soil Moisture percentile is high for much of the previous year (2010), especially over the growing season.

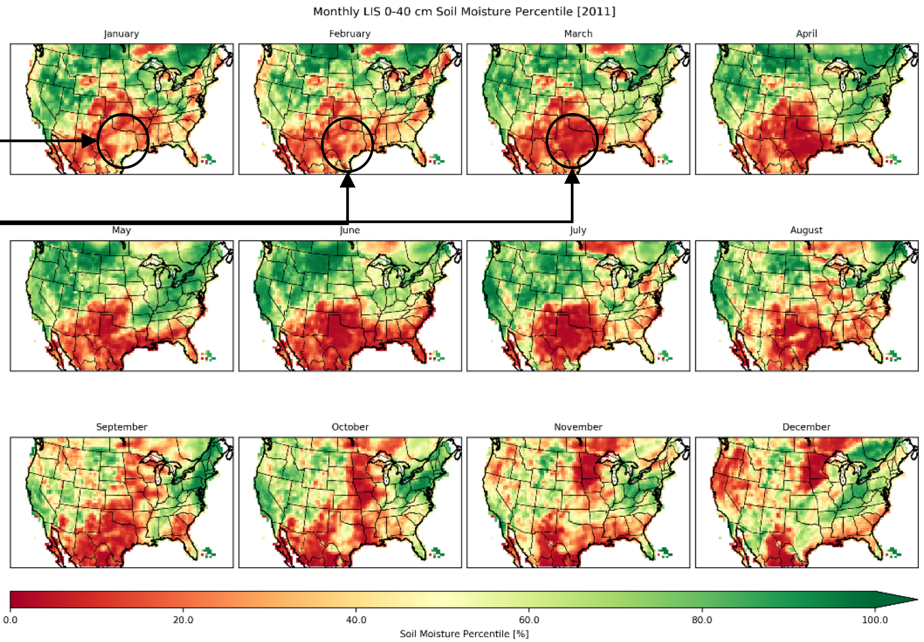


- Drying then occurred from late fall 2010 and continued through 2011.
- High antecedent soil moisture during growing season can lead to a build up of fuel.
- Low soil moisture leading up to fire season continually dries the available fuel.



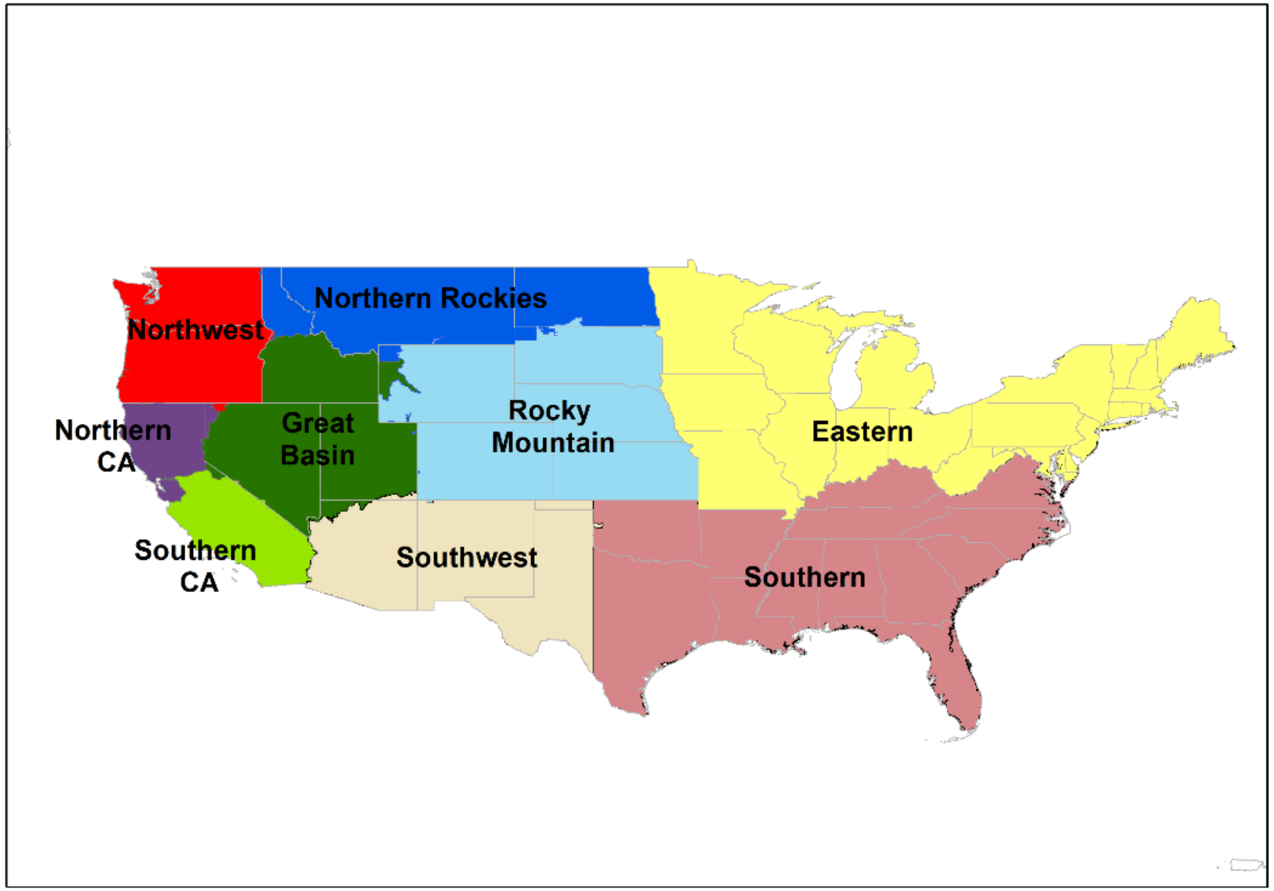
High antecedent (2010) soil moisture during growing season.

Low antecedent soil moisture during late Fall 2010 into early Spring 2011.



Project Overview

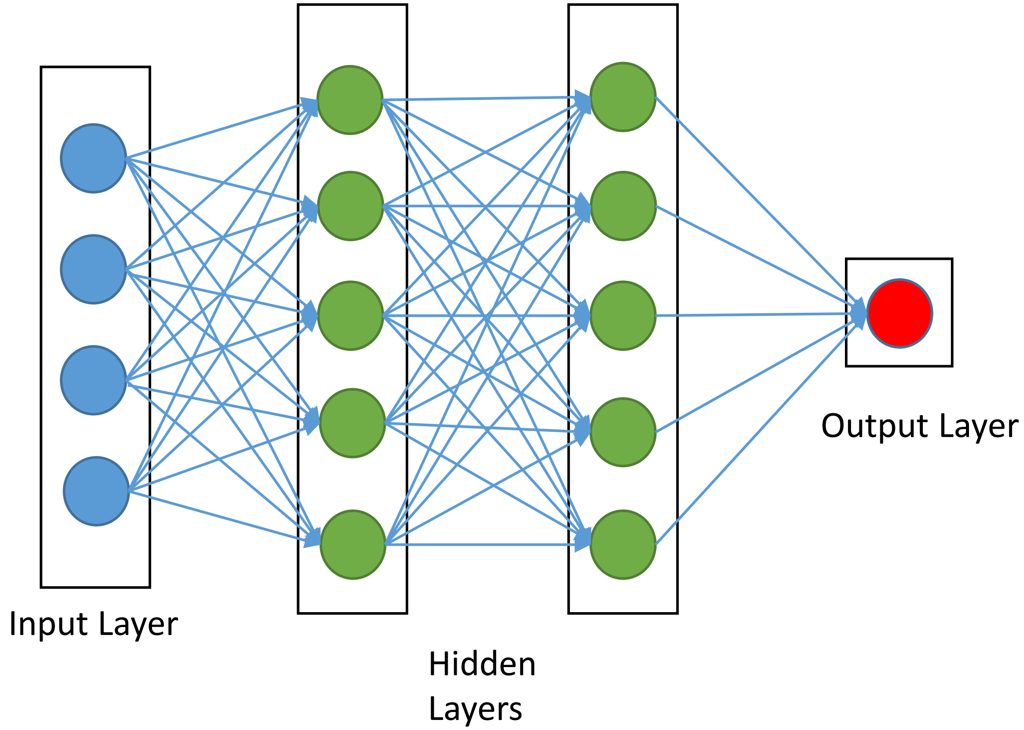
- ***Goal:*** Use deep learning to develop a predictive model for yearly number of fires and acres burned across each Geographic Area Coordination Centers (GACC) region.
  - Study predictability by region.
  - Study the predictor variable importance by region.
  - Study time-lag importance by region.



GACC Regions

Deep Learning

- Deep Neural Network (DNN)
  - Learn representations from the data through hierarchical layers.
  - Works by determining the weights which effectively map the inputs to their targets.
- Each DNN was built using Keras with the tensorflow backend.
  - 5 layers (4 hidden layers and the output)
  - 500 nodes per hidden layer.



Input Data

- A plethora of data corresponding to the antecedent land surface, atmospheric and fuel conditions are used.

- The 4th edition Fire Program Analysis – Fire Occurrence Database (FPA-FOD) is used as the truth dataset (Short 2017).
  - Point data is gridded and smoothed to represent a continuous wildfire truth dataset.

Input Features	
SPoRT LIS Volumetric Soil Moisture (0 – 10 cm, 10 – 40 cm, 40 – 100 cm)	SPoRT LIS Soil Moisture Percentiles (0 – 10 cm, 0 – 40 cm, 0 – 100 cm)
Dead Fuel Moisture (100-hr and 1000-hr)	Precipitation
Daily Minimum and Maximum Temperature	Daily Mean Vapor Pressure Deficit
Daily Minimum and Maximum Relative Humidity	Energy Release Component
Daily Average Downwelling Shortwave Radiation	Wind Speed
Planned Addition MODIS LAI/GVF	Planned Addition Evaporative stress index

Model Training

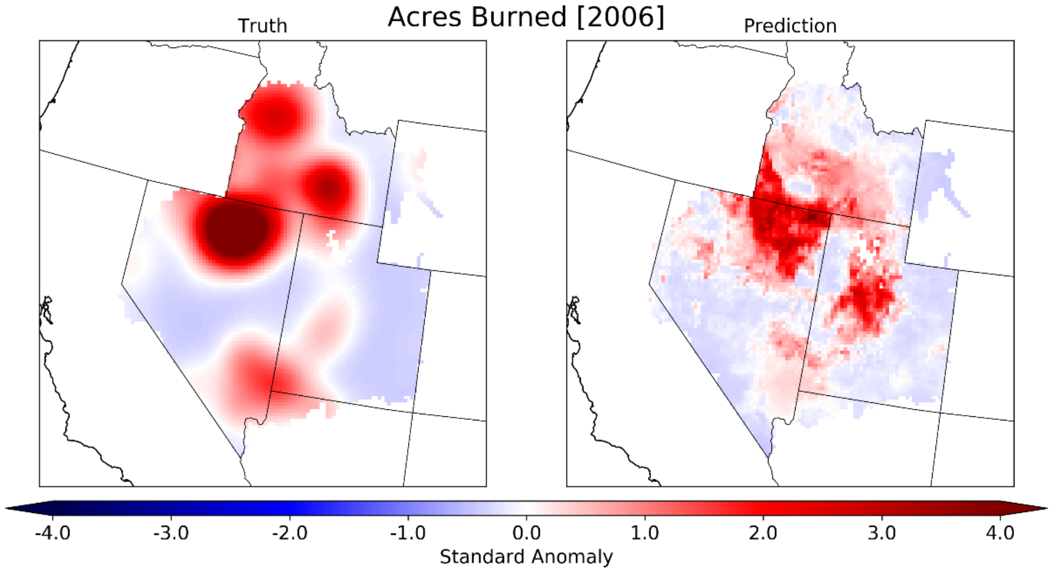
- Monthly averages of each feature from the previous three years through the start of fire season were used.
  - Feature scaling was completed on a GACC region basis.
- DNN models were trained over the 2002 – 2015 time period.
  - Individual years were held out for testing.

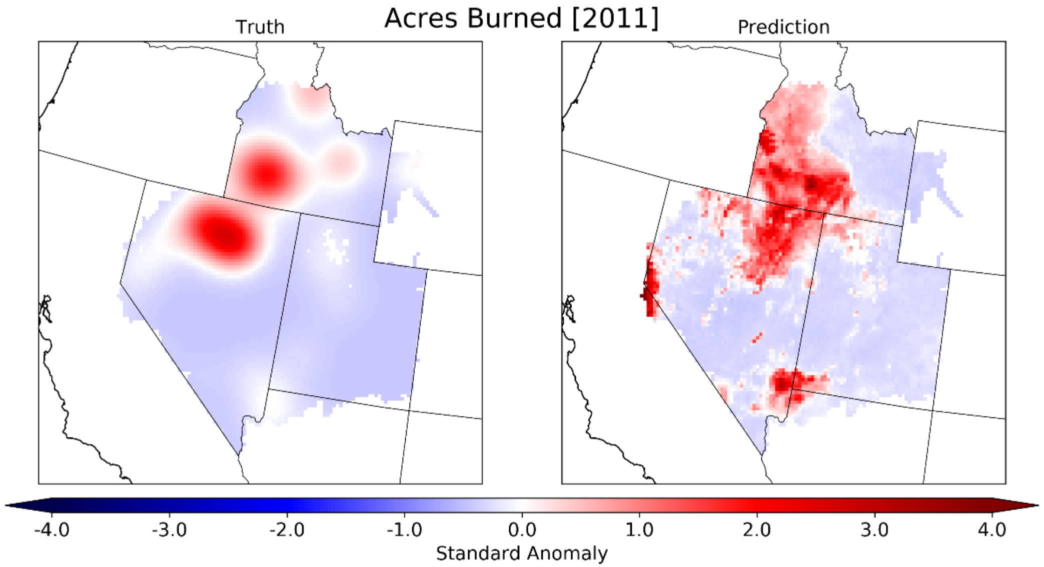
GACC	Start Date (5%)	Median (50 %)	End Date (95%)
Eastern	47 (Feb)	103 (Apr)	317 (Nov)
Northern CA	127 (May)	206 (Jul)	297 (Oct)
Northern Rockies	90 (Mar/Apr)	211 (Jul/Aug)	284 (Oct)
Northwest	102.9 (Apr)	211 (Jul/Aug)	286 (Oct)
Rocky Mountain	43 (Feb)	181 (Jun/Jul)	305 (Nov)
Southern CA	101 (Apr)	196 (Jul)	296.7 (Oct)
Southern	20.5 (Jan)	95 (Apr)	332 (Nov/Dec)
Southwest	28 (Jan/Feb)	158 (Jun)	310 (Nov)
Great Basin	128 (May)	205 (Jul)	270 (Sep)

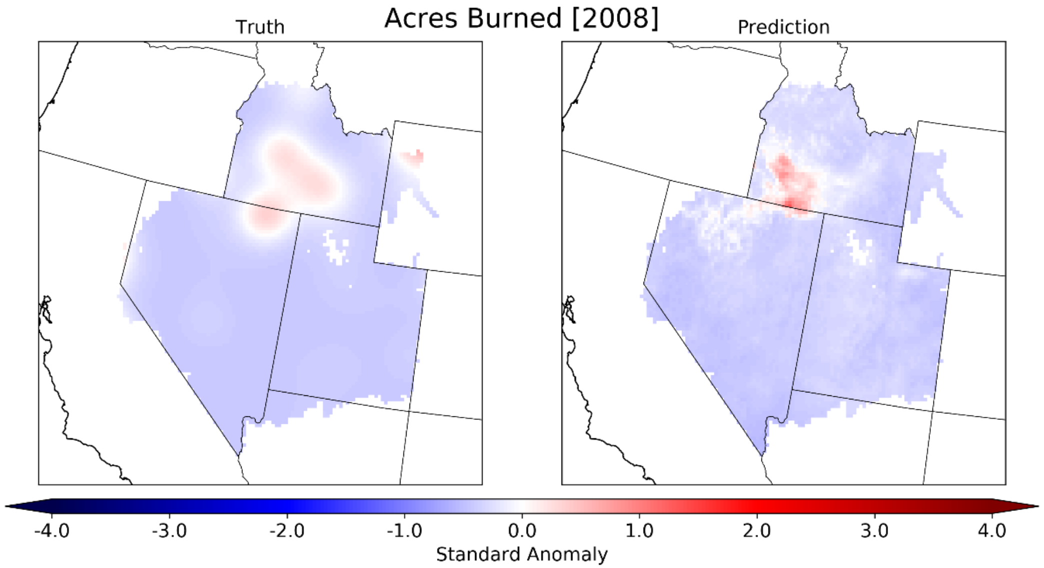
Initial Results for Different GACC Regions

Great Basin

➢ The model shows skill in capturing the yearly variability in acres burned across the region.

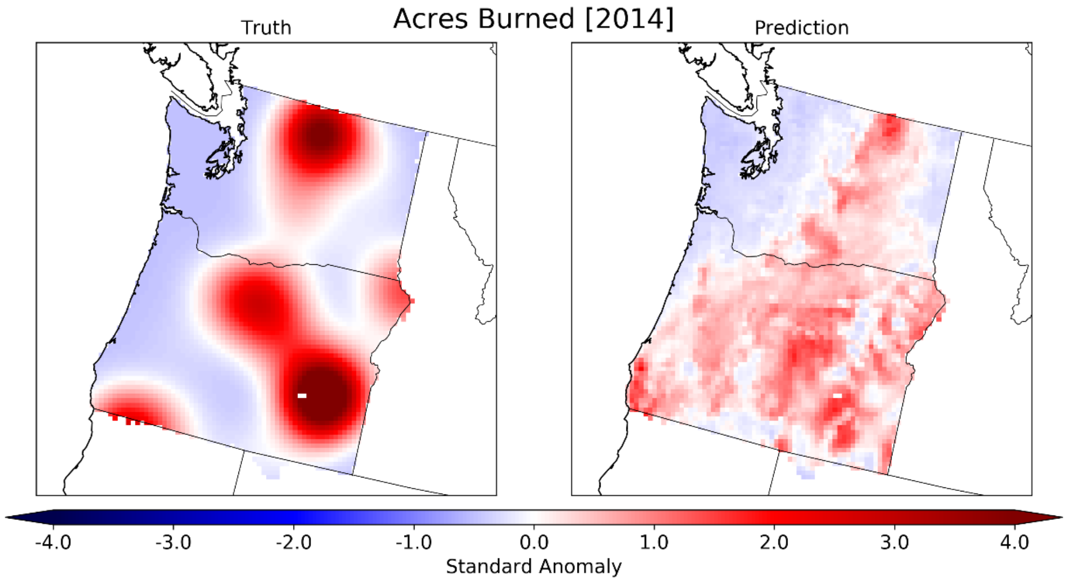


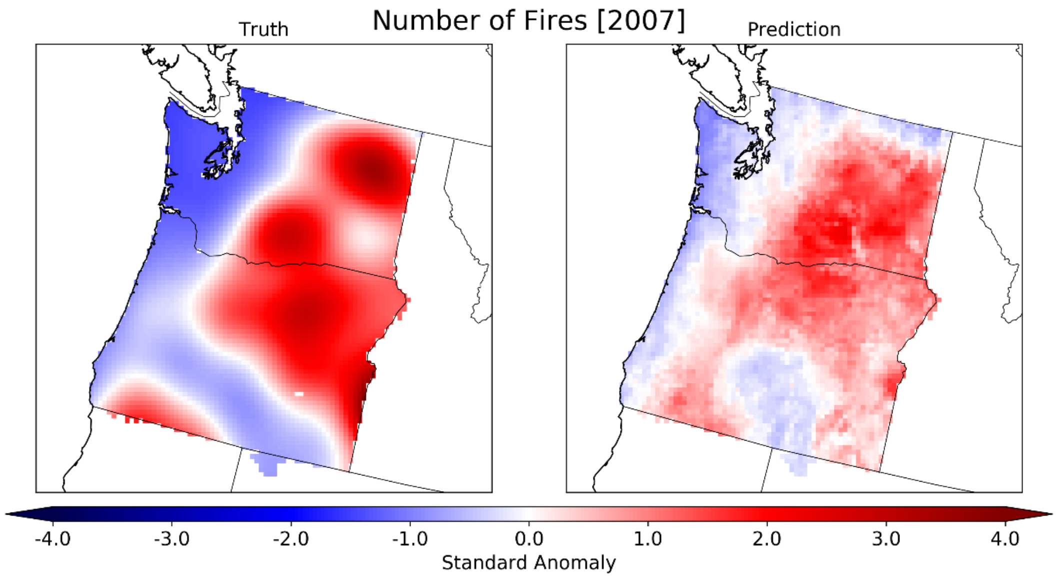


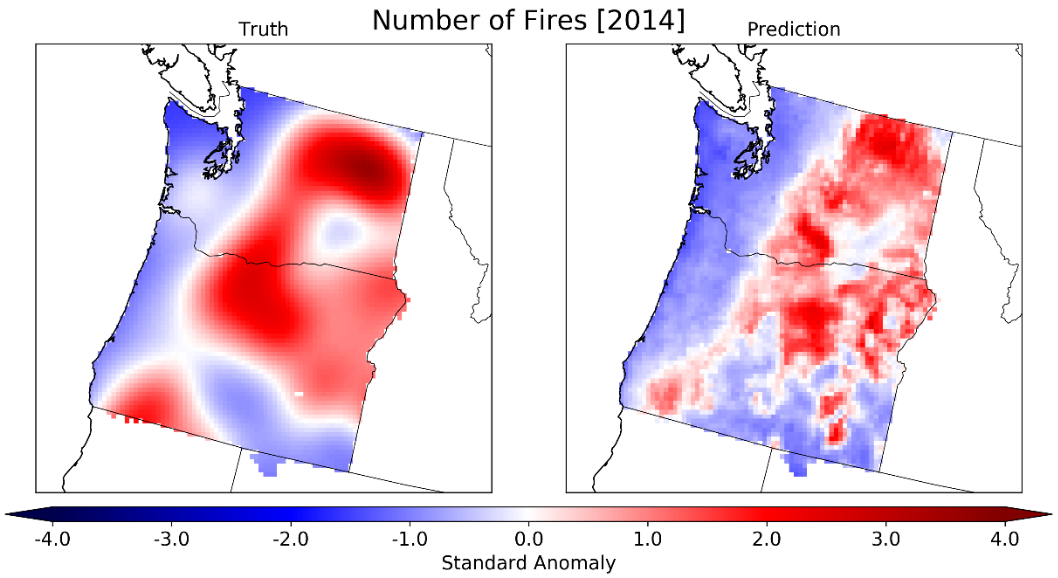


Northwest

- The model currently shows more skill at predicting the number of fires over acres burned.

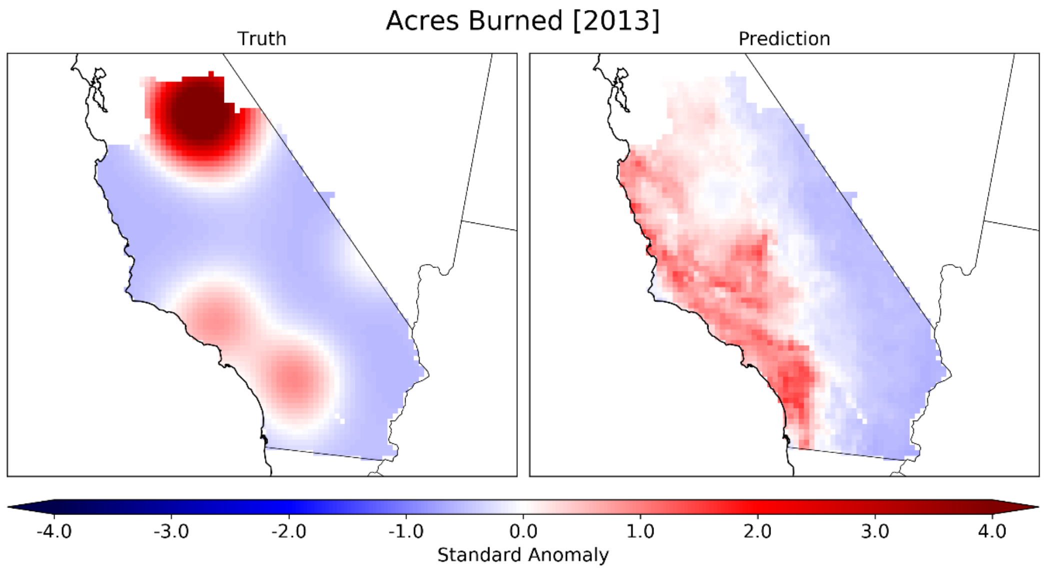


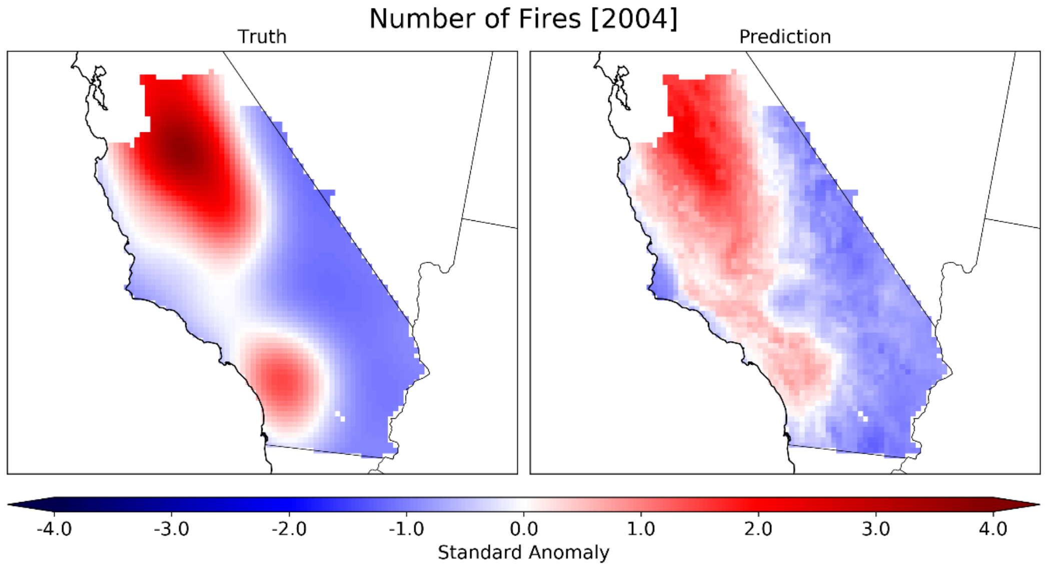


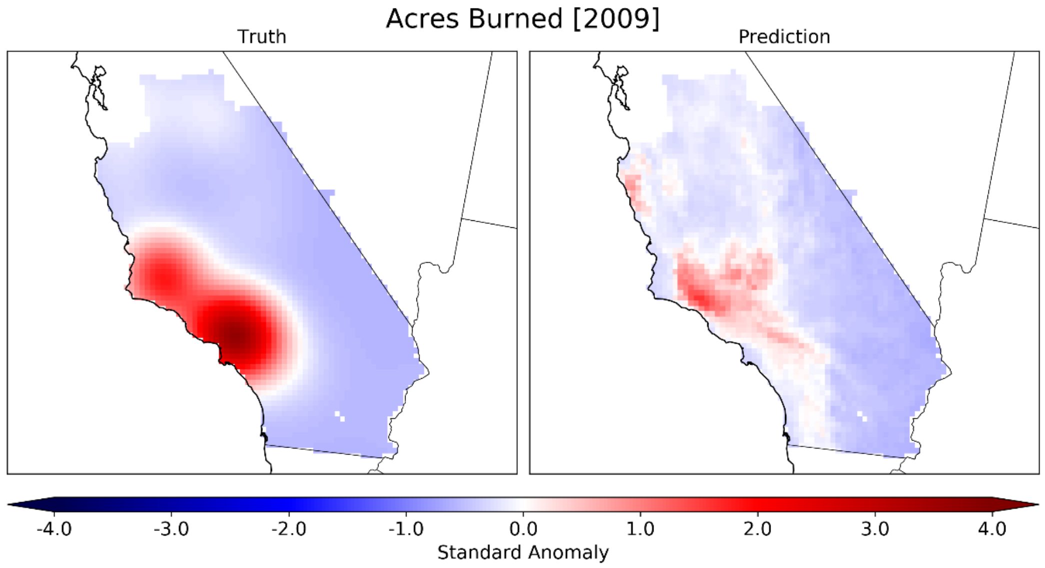


Southern CA

- ➢The model shows better agreement when predicting the number of fires.

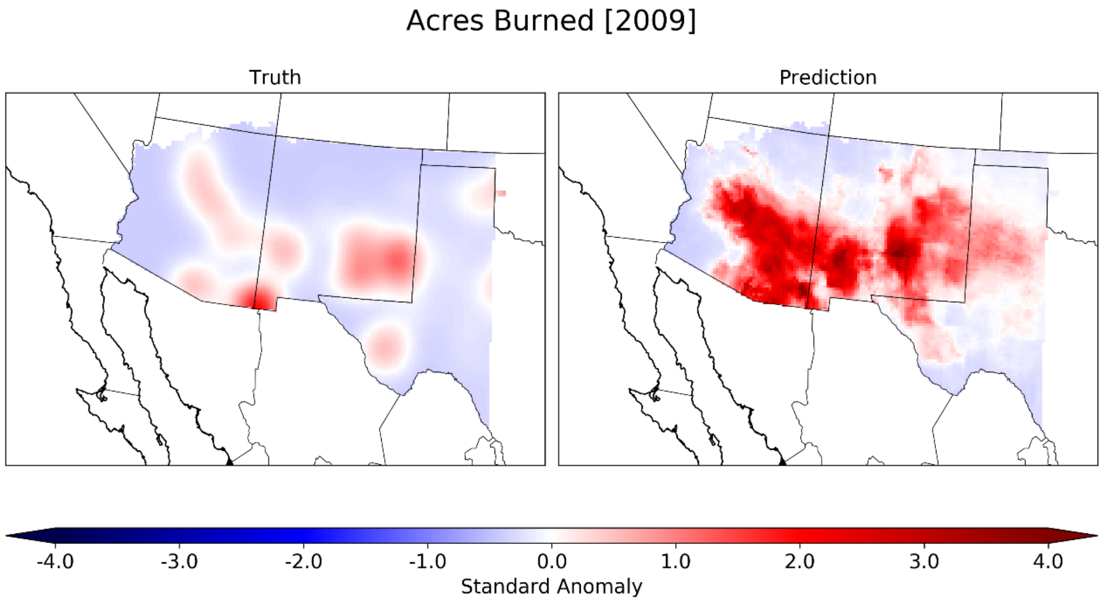


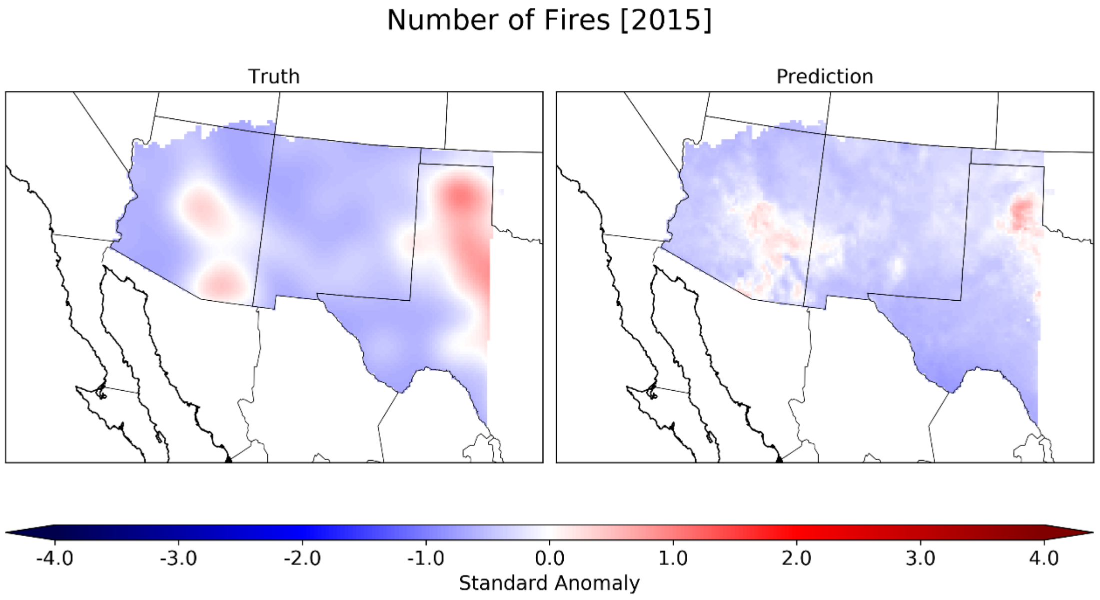


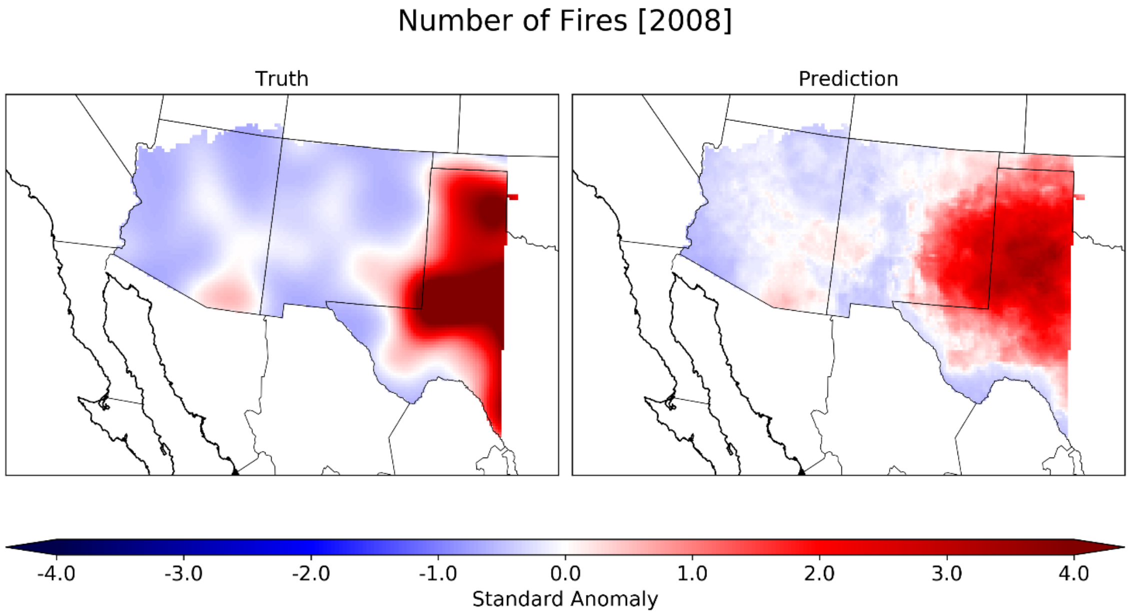


Southwest

- ➢In the Southwest region, the developed model shows more skill in predicting number of fires over the acres burned.

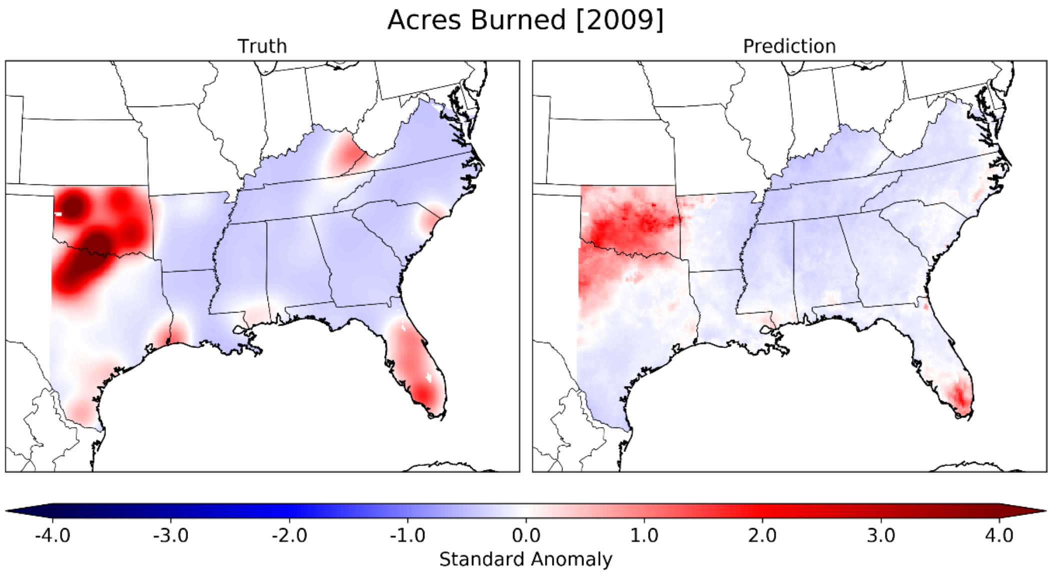


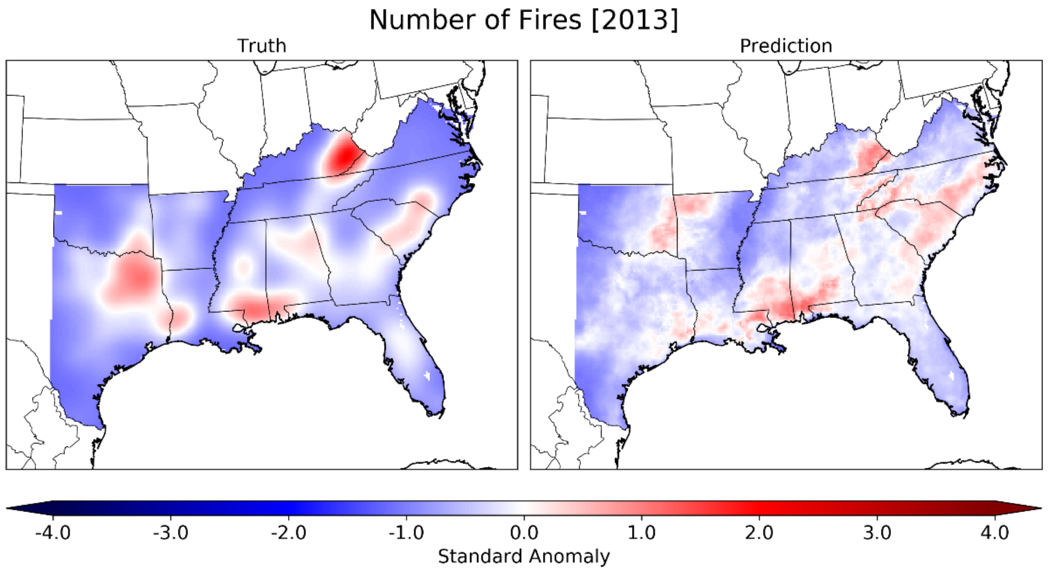


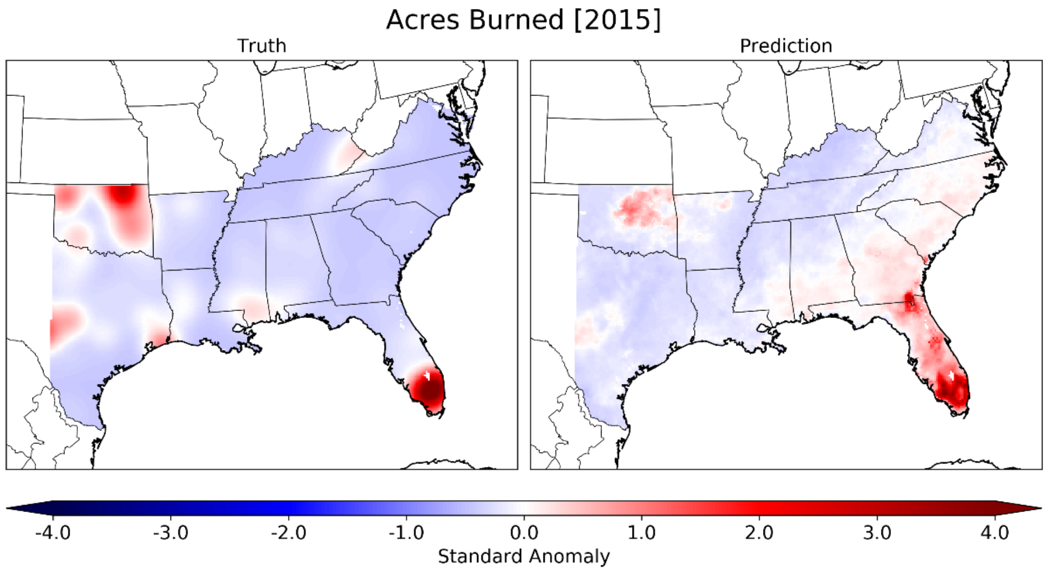


Southern

- For acres burned, the model tends to capture the spatially larger anomalies, while missing the smaller ones.
- For number of fires, the model shows better overall spatial agreement.

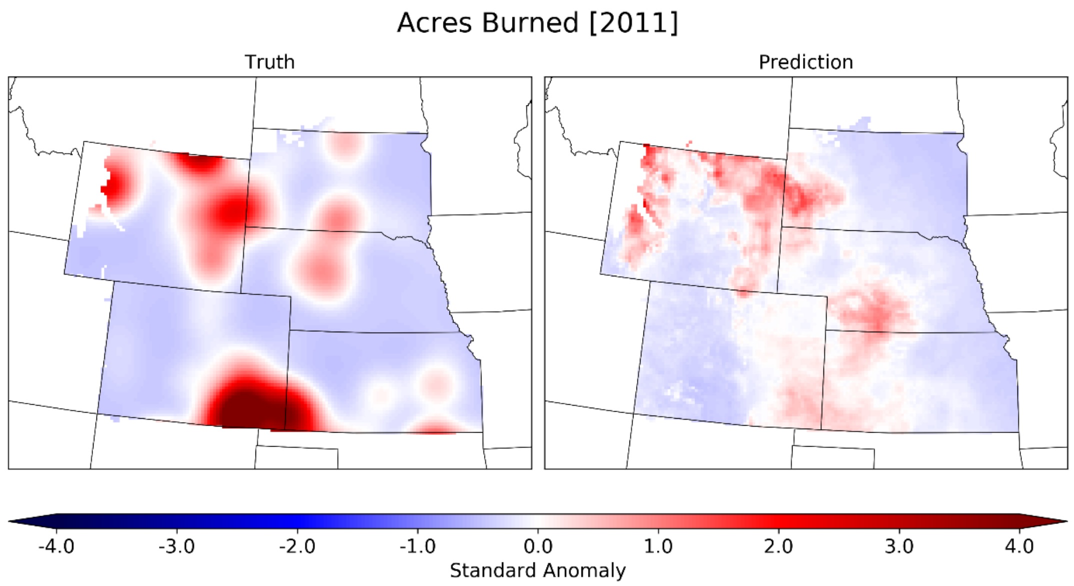


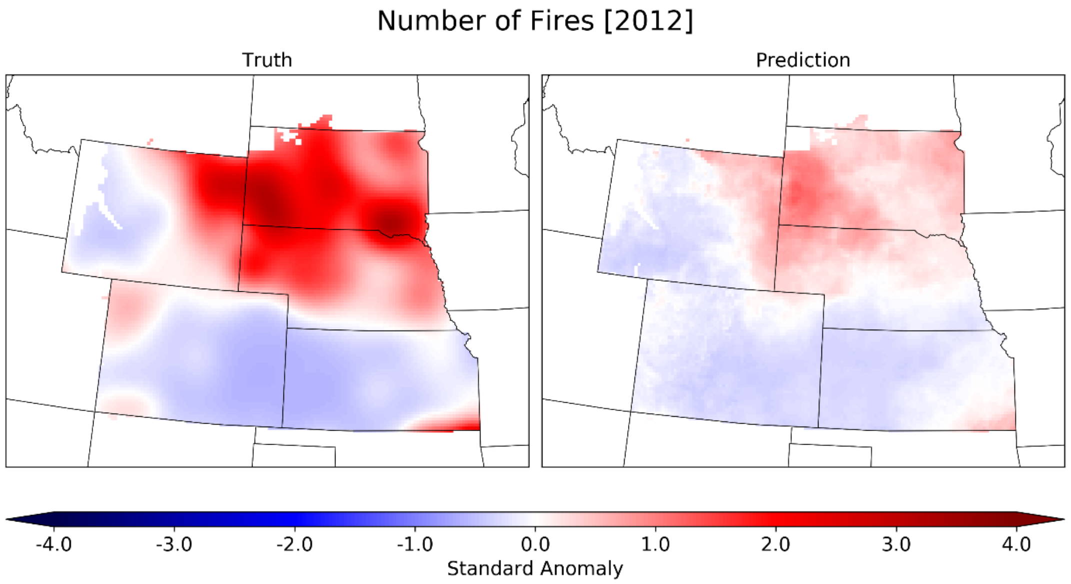


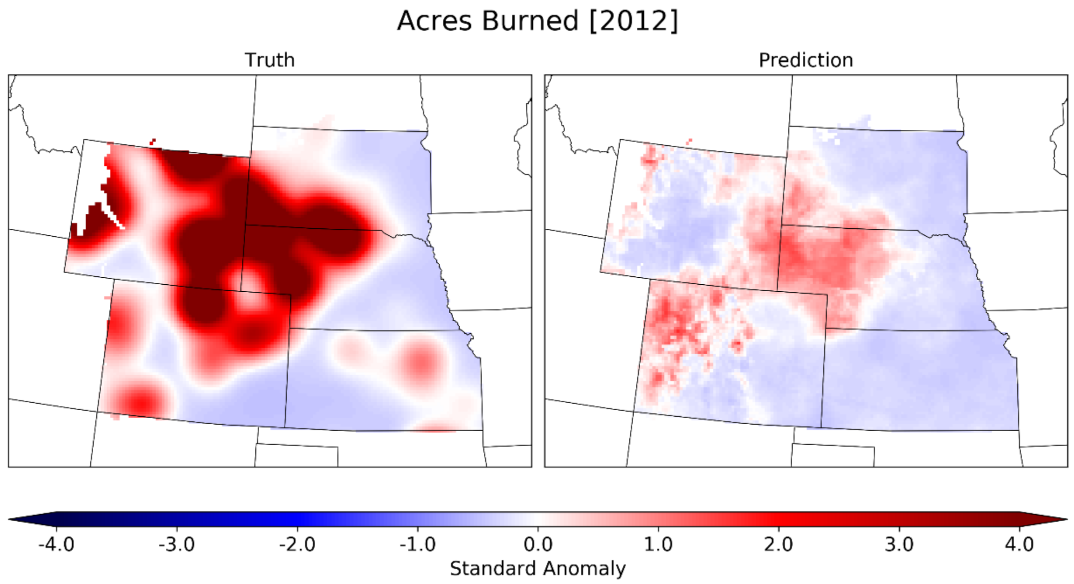


Rocky Mountain

- ➢The model does reasonably well at predicting the locations of the positive anomalies even though the magnitudes are under-predicted.







Summary

- The deep learning model shows promise for predicting areas of high wildfire potential.
  - Full evaluation of the model performance is ongoing.

- Currently, the developed deep learning model is better overall at predicting the number of fires over the acres burned.
  - Acres burned is dependent on location, suppression plan, and current conditions.

- Antecedent conditions are only one piece of the equation.
  - In-season changes are not accounted for.
  - An ignition source is required, which further complicates the model training and prediction.