Exploring the use of 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

Background

Wildfires in the United States can be extremely costly, both economically and through the loss of life. The cost of fire suppression efforts alone surpassed $2B in 2017, which is greater than 50% of the U.S. Forest Service's total budget. With high costs and a disproportional amount going to fire fighting efforts, predictive estimates of fire season severity (i.e. number of fires or acres burned) could be beneficial to management officials. Fire severity is dependent on both the current and antecedent atmospheric and land surface conditions which is highly variable year to year. Partly in response to changing antecedent conditions, fire season variables such as number of fires and acres burned vary year to year as well (Figure 1).

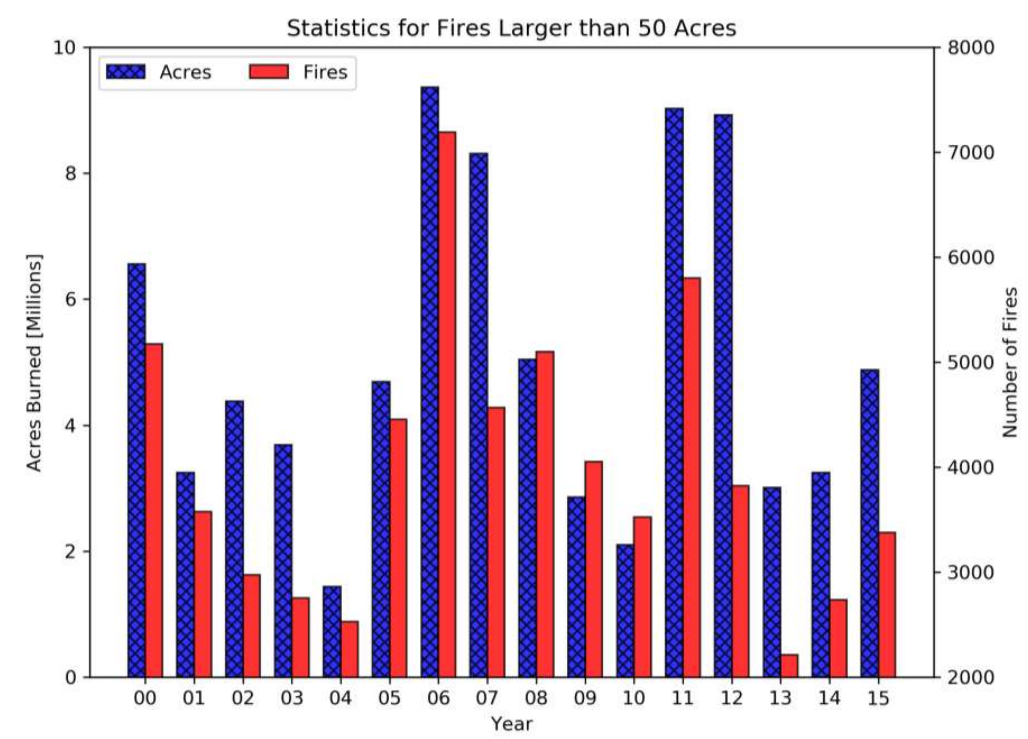


Figure 1: Yearly number of fires and acres burned across the CONUS domain.

Antecedent Relationships

- Previous studies have indicated the importance of antecedent conditions on wildfires (Westerling et al. 2003; Crimmins and Comrie 2004; Morton et al. 2013; Nauslar et al. 2018).
- Much of the focus has been on variables such as precipitation, temperature, and the Palmer Drought Severity Index (PDSI) while very little explicit focus on the actual soil moisture conditions.
- The SPoRT Land Information System (LIS) model data can be used to characterize the instantaneous and climatological soil moisture providing additional information on fuel conditions.

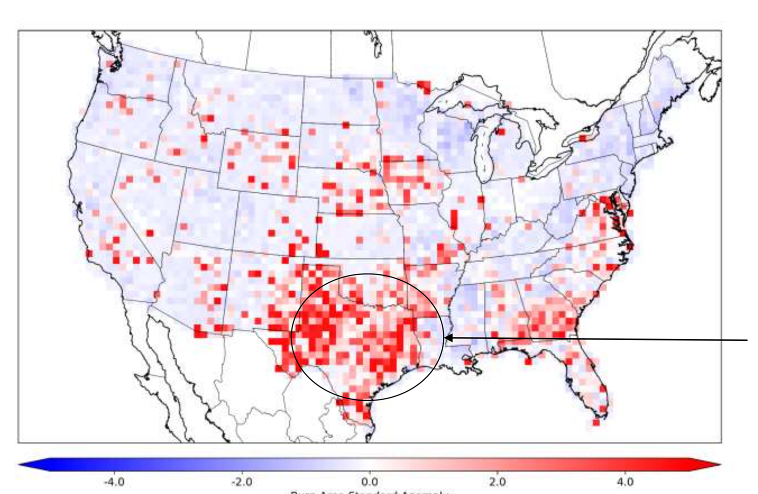


Figure 2: Burn area standardized anomalies for 2011.

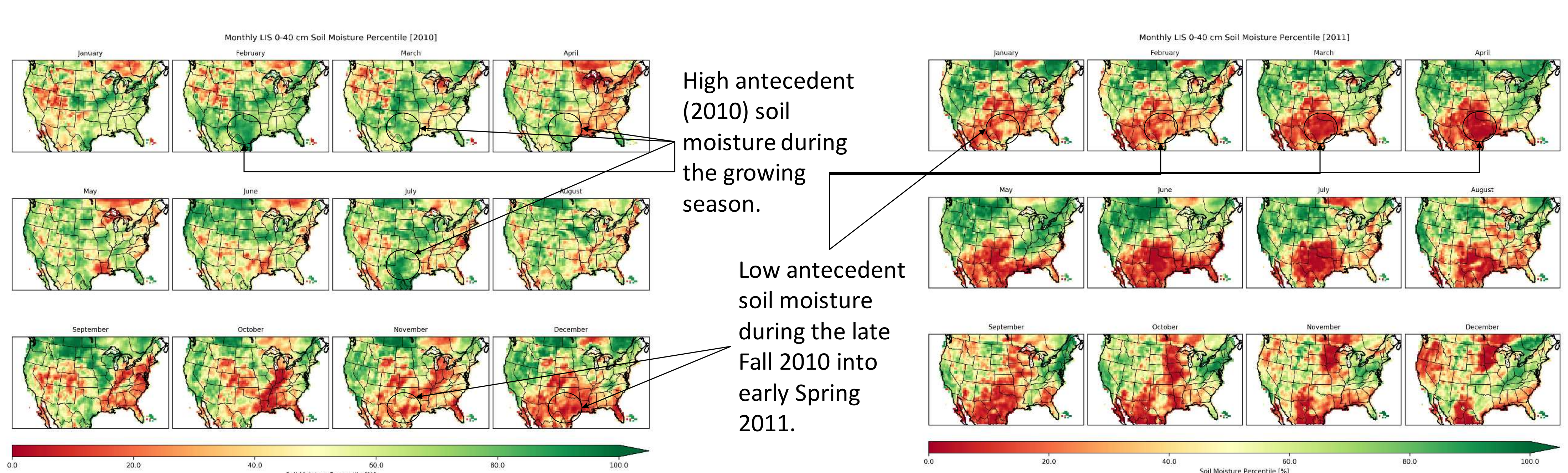


Figure 3: SPoRT-LIS 0 – 40 cm monthly average soil moisture percentile for 2010 (left) and 2011 (right).

Methodology: Random Forest Regression

- The burn area across the U.S. is not uniform and is generally split by east/west climate differences (Figure 4). These spatial differences result in non-uniform relationships between antecedent conditions and fire potential.
  - Therefore, all datasets were transformed to standardized anomalies.
- Due to the numerous amounts of available data related to fire potential, the use of a random forest (RF) machine learning algorithm to develop a burn area predictive model was explored.
  - Monthly averages of each dataset (Table 1) from the previous year through March of the current fire season were used as RF features.
- The RF model was trained using 11 years worth of data (fire years: 2002 – 2012). The model was then tested on 3 different years (fire years: 2013 – 2015). An example RF tree in the model is shown in Figure 5.

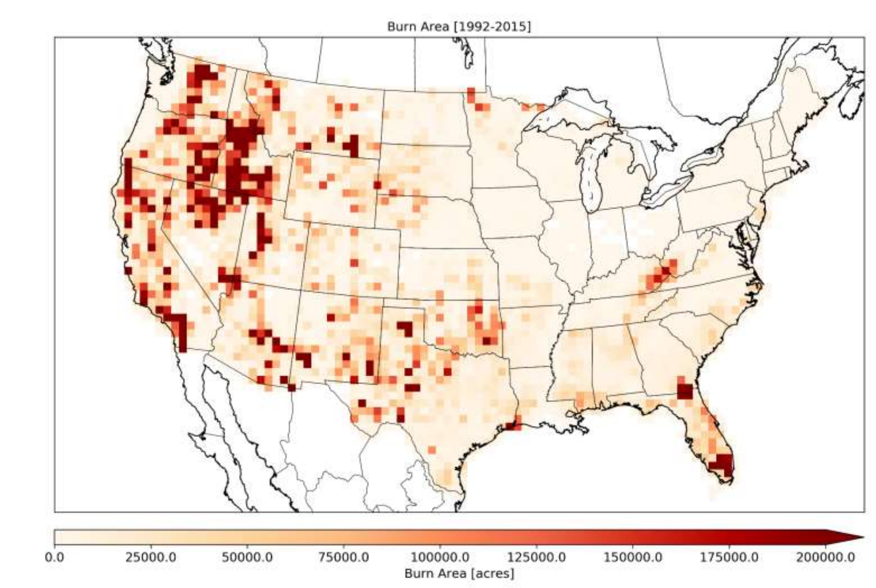


Figure 4: Total acres burned (1992 – 2015)
[Data from the 4th edition US Forest Service database (Short 2017)]

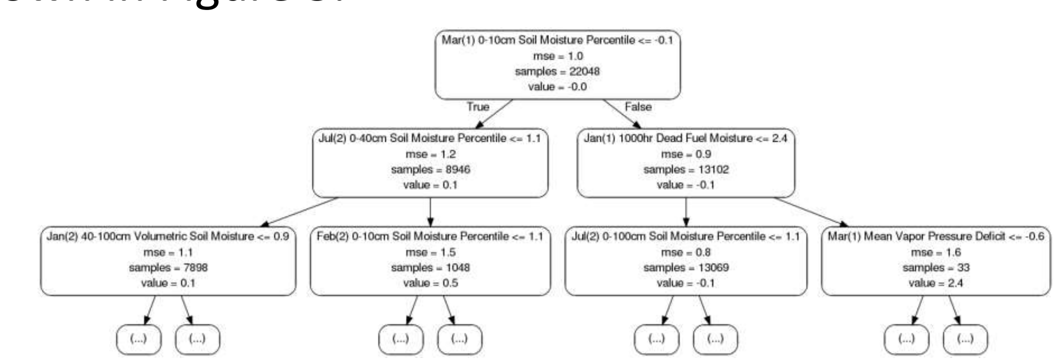


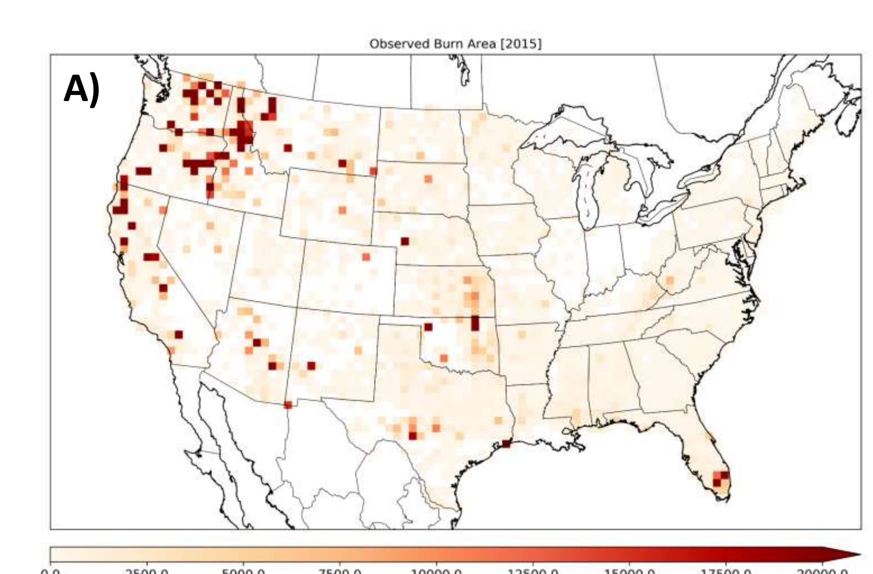
Figure 5: Example RF Tree

Table 1: Random forest features.

RF 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) [Abatzoglou 2013]	Precipitation [Abatzoglou 2013]
Daily Minimum and Maximum Temperature [Abatzoglou 2013]	Daily Mean Vapor Pressure Deficit [Abatzoglou 2013]

Results/Future Work

- At the current state of development, the RF model shows some promising results (Figure 6).
  - While the predicted burn area does not exactly match the observed, the overall patterns are similar.
  - The model is only providing information on the yearly burn area potential based on the antecedent conditions alone.
- The current model does not account for different ignition sources.
  - Human-related ignitions add uncertainty into the model as typical antecedent relationships in human ignition events are different than those in lightning initiated wildfires (Balch et al. 2017; Nagy et al. 2018).
  - Since the number of fires started by humans is much larger than natural ignitions, optimal antecedent conditions alone might not explain the variation in acres burned from season to season.
- Ignition specific exploration with the RF algorithm has been planned. While lightning started events are significantly less frequent, they contribute a large percentage to the area burned at ~56% (Balch et al. 2017).
  - The goal would be to produce probabilistic lightning wildfire potential.



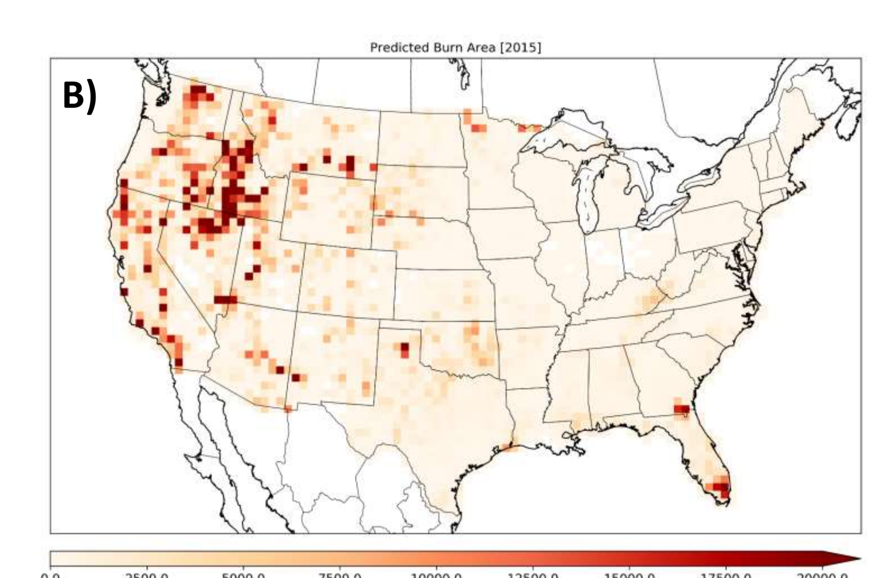


Figure 6: Observed acres burned (A) and predicted (B) for 2015.

99th Annual American Meteorological Society (AMS) Conference on January 6 – 10, 2019, Phoenix, AZ

Contact Information: andrew.t.white@nasa.gov