Analysis of the Uncertainty in High Level NEON AOP Data Products

Amanda Roberts, Tristan Goulden, Bridget Hass

Introduction and Site Selection

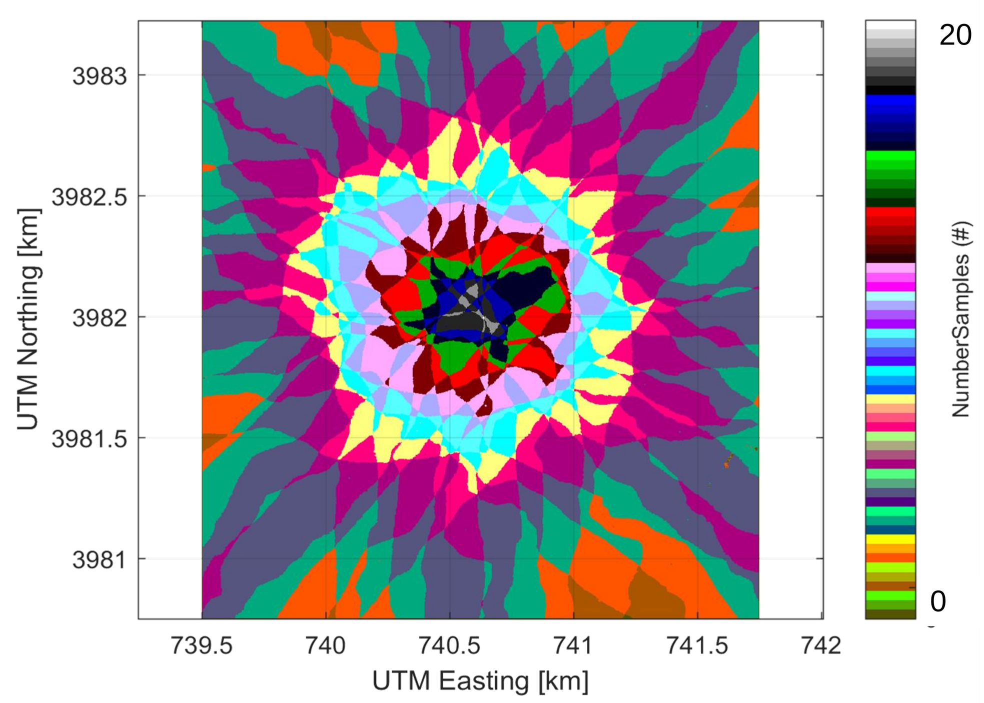

Figure 1. The center of a BRDF flight path at Oak Ridge showing the number of overlapping flight lines

The National Ecological Observatory Network (NEON) uses a LiDAR system along with the NEON Imaging Spectrometer for their remote sensing measurements. These are state of the art instruments and limited analysis has been performed to quantify their levels of uncertainty. To quantify uncertainty, a two part research project was performed; 1) pixel variations that appeared in L2 and L3 products from Bidirectional Reflectance Distribution Function (BRDF) flights, flown in an asterisk shape (Figure 1) were quantified. This allowed for an empirical statistical analysis to estimate uncertainty, 2) a case study focusing on Leaf Area Index (LAI) uncertainty generated a predictive model using Random Forest machine learning approach which related landscape features to LAI uncertainty.

The BRDF flight plan has been acquired at three of the NEON sites. The three sites are 1) Oak Ridge in Tennessee, 2) Soaproot Saddle in California, and 3) San Joaquin Experimental Range also in California. Each site has unique landscape characteristics, allowing for analysis on end member environments.

Uncertainty in Bidirectional Reflectance Distribution Function Flights

The algorithm which empirically quantified uncertainty was applied to twenty data products at each site. The algorithm 1) removed interpolated values along the edges, which would skew the analysis 2) calculated various statistical measures at each pixel and 3) produced maps of mean, standard deviation (uncertainty), relative standard deviation, number of samples, and maximum difference (Figure 2, Figure 3). Additionally, a histogram of the distance between each value at value at a pixel (residuals) and its mean was made.

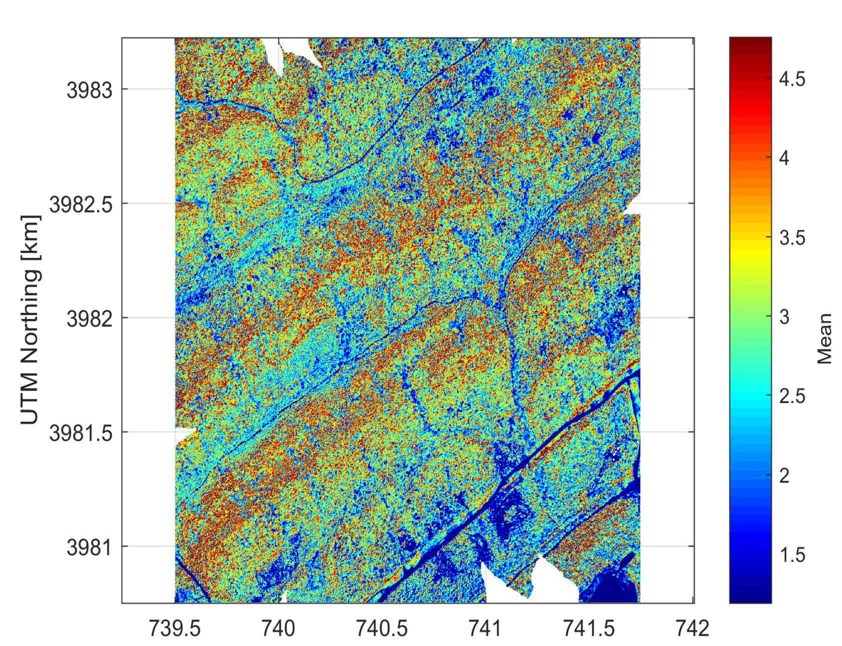

Figure 2. The mean LAI values at Oak Ridge

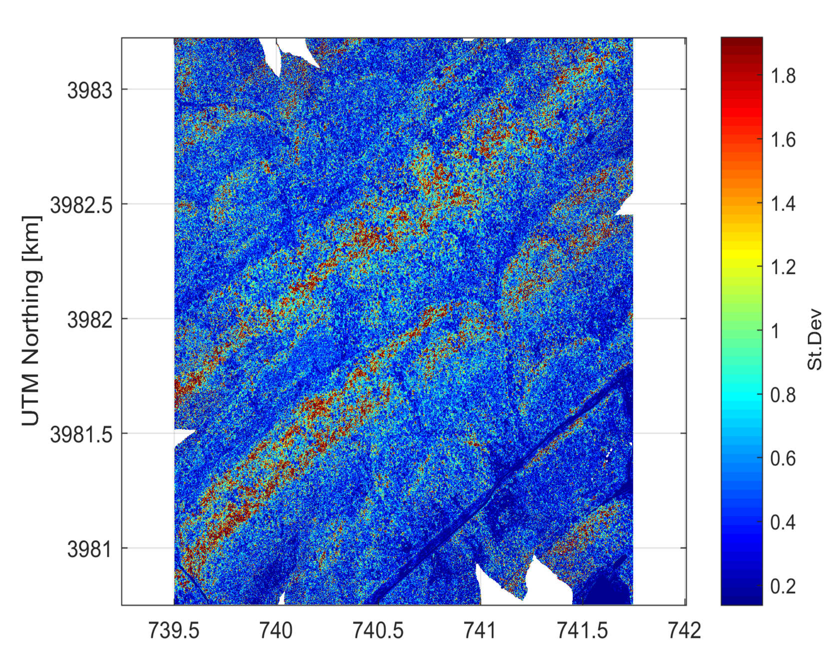

Figure 3. The standard deviations for LAI at Oak Ridge

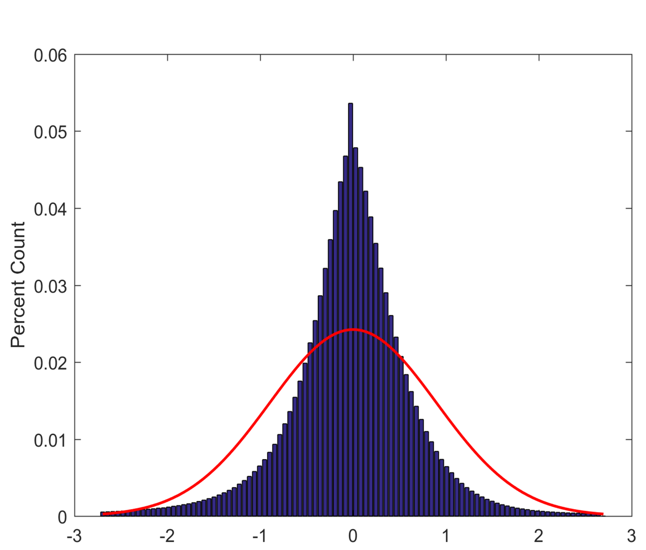

Figure 4. Histogram showing the residuals

Predicting Leaf Area Index Over a Whole Site

Qualitative Methods

LAI at Oak Ridge was selected as a case study because of strong interest in LAI retrievals in a thick, closed canopy. An initial visual analysis which compared the data products to LAI uncertainty was performed to identify which landscape characteristics related to LAI uncertainty. Next, LAI uncertainty was related to landscape characteristics through simple linear regression which provided a slope and r^2 value. The data was then sorted and counted based on how many fell in a particular regression slope range.

Quantitative Methods

The predictive model was determined through a machine learning Random Forest Regressor algorithm written in the Python language. The regressor creates decision trees (Figure 5) and determines the optimal importance of samples for predictions. Training data came from three quarters of the data obtained in the BRDF flight and the model was validated on the final one quarter of data. The model was then applied to the entire Oak Ridge site (Figure 6).

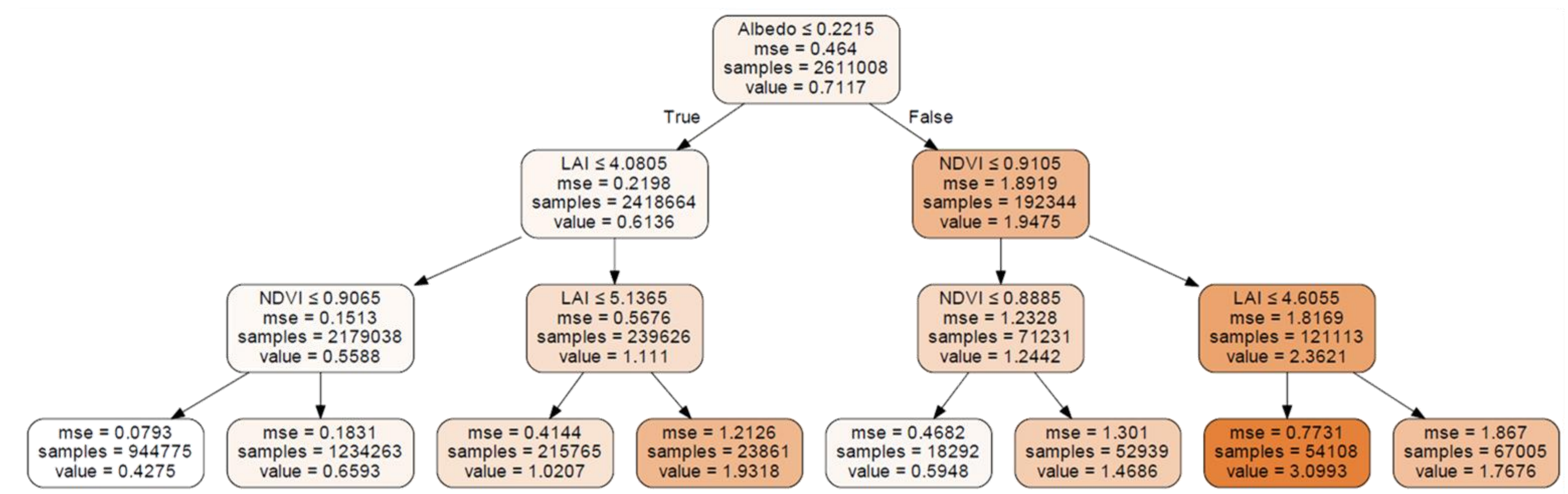

Figure 5. A section of the forest created by the model

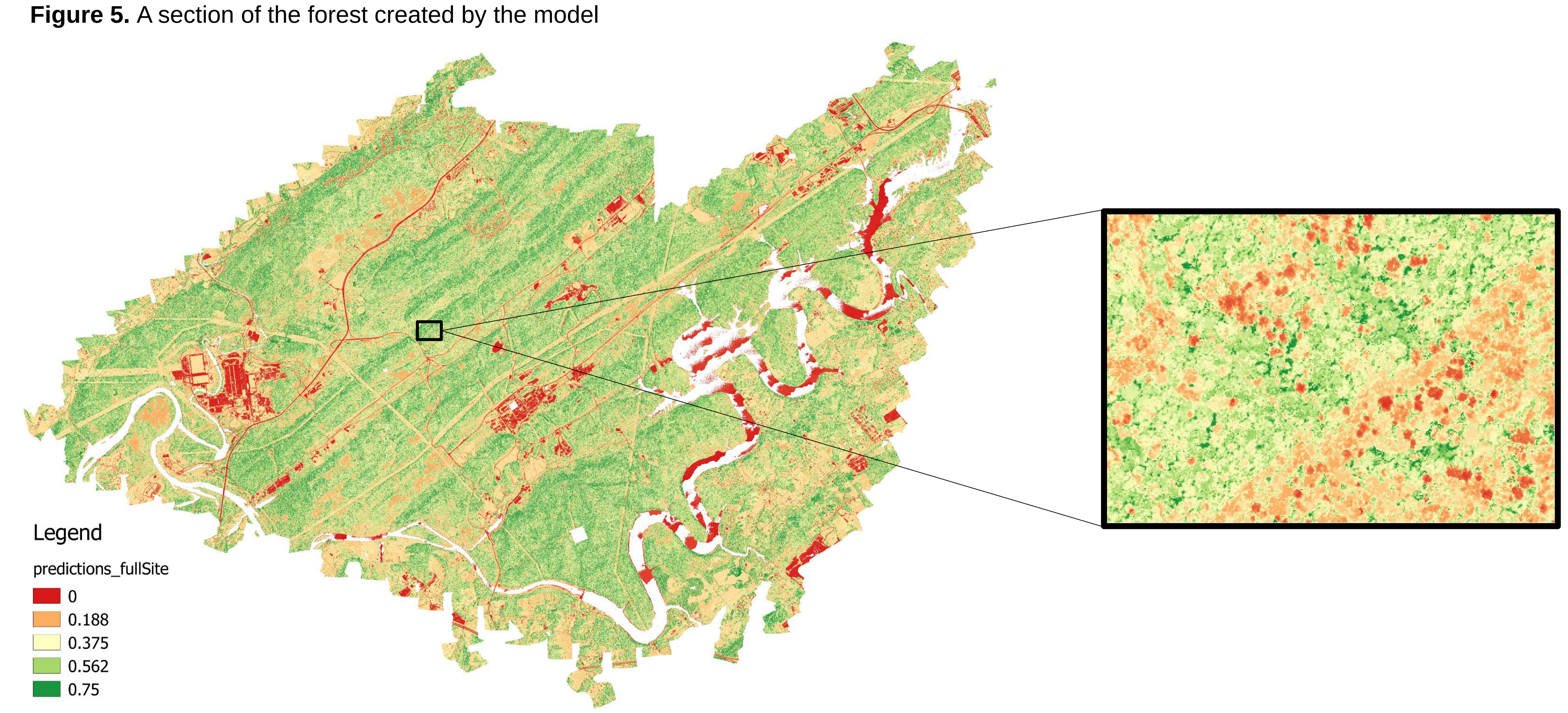

Figure 6. Predicted LAI uncertainty over the Oak Ridge Site

NEON is a project sponsored by the National Science Foundation and proudly operated by Battelle.