

Lab 4 - Cloud Data

Stat 215A, Fall 2014

Tim Meyers, Jason Poulos, and Zhao Ruan

<https://github.com/jvpoulos/stat215a-lab4>

October 27, 2014

1 Exploratory Data Analysis

Figure 1 spatially plots the expert labels for the presence or absence of clouds, for each of the three MISR images. In these plots, darker colors represent cloud labels, lighter colors represent no label, and white space represents non-cloud labels.

Figure 3 includes spatial plots of the radiances of each angle and Figure 2 is a heat map showing the correlations between radiance angles. The spatial plots of radiance angles are almost indistinguishable. Each radiance angle pair are highly correlated (over 0.5 correlation value), although some angles are more correlated than others (e.g., an is much more correlated with af than it is with df .)

How well do the radiance angles and features (i.e., $CORR$, $NSAI$, and SD) predict the presence or absence of clouds? Figure 4 plots the probabilities from two logistic regression models predicting an indicator variable for the presence of clouds, with radiance angles (left plot) and features (right plot) as regressors. For each regression model, we estimate the predicted probabilities for each pixel, and arrange the probabilities in ascending order. We then plot the probabilities for each pixel (black line) and show which pixels are labeled cloud (red lines). The plots demonstrate that both models fit the data well, given that most of the cloud labels are clustered at the right-hand side of the graphs. However, the features model provides a much better overall fit of the data since all the cloud labels are clustered at the right-hand side.

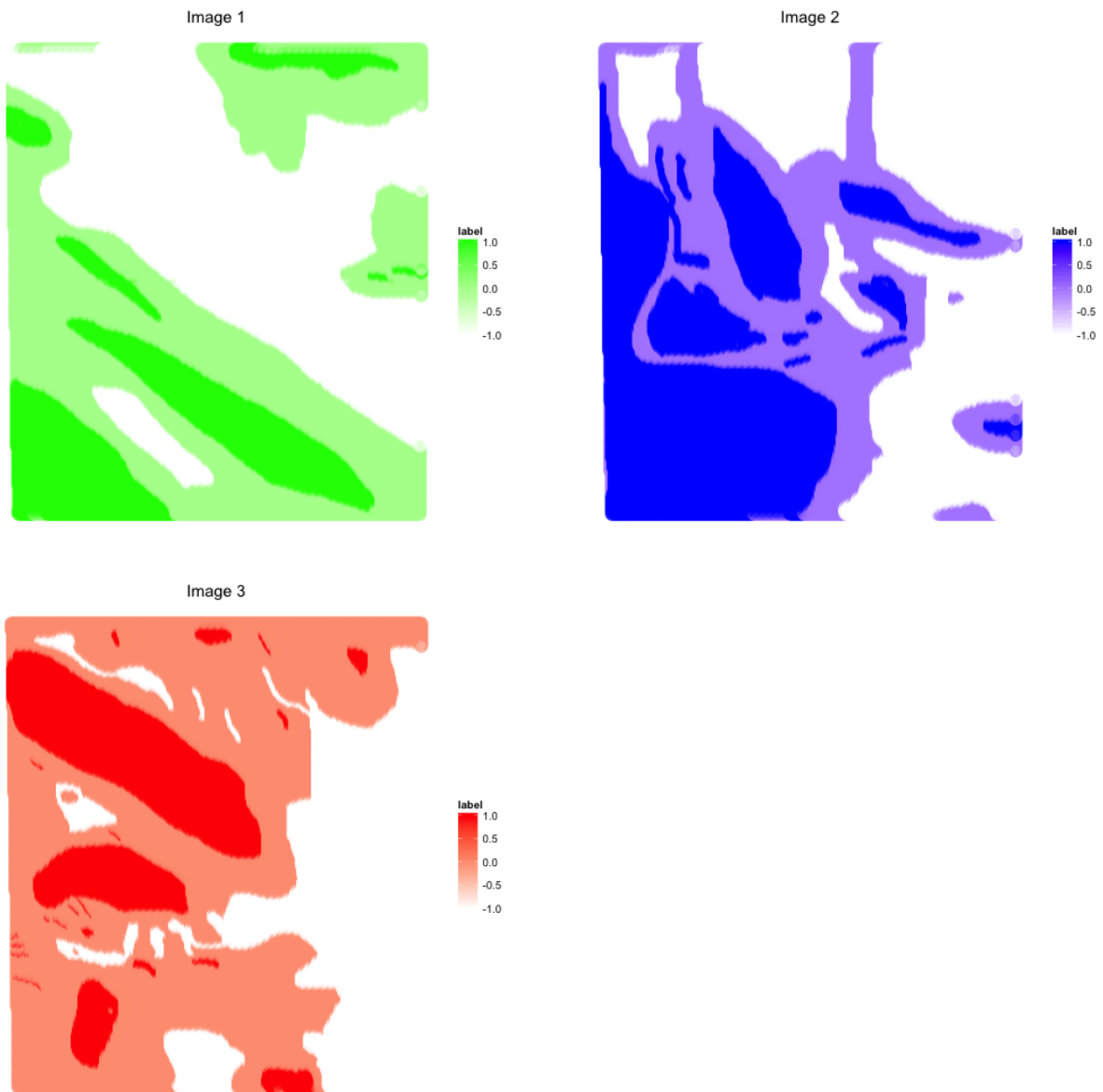


Figure 1: Spatial plot of the expert labels for the presence or absence of clouds, for each MISR image.

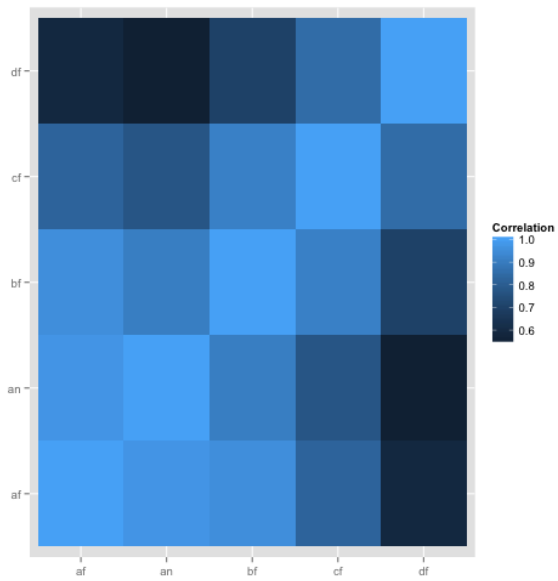


Figure 2: Radiance angle correlations.

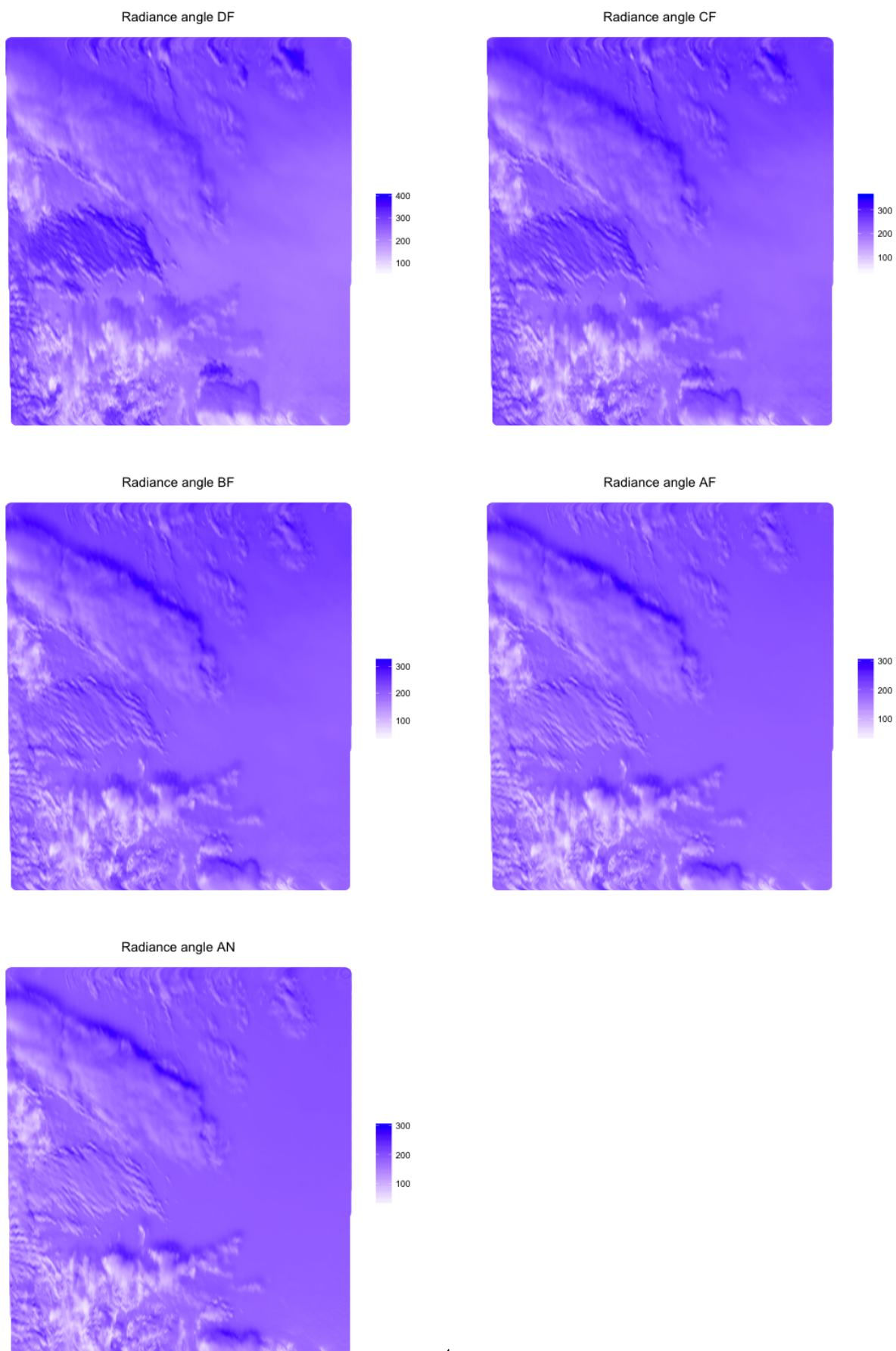


Figure 3: Spatial plot of each radiance angle.

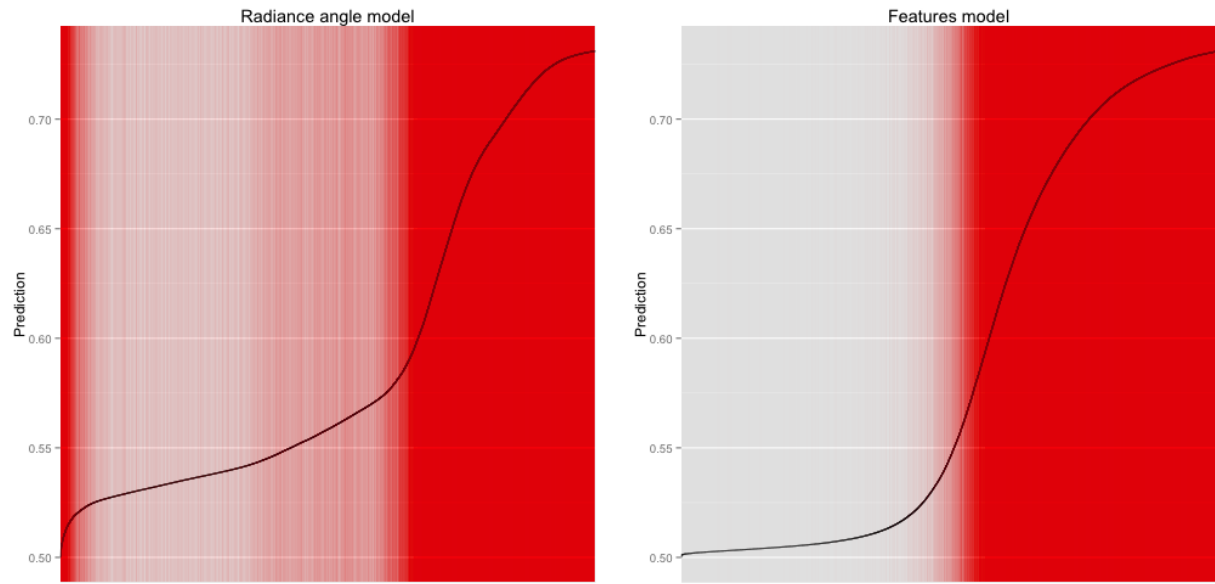


Figure 4: Plot of probabilities from logistic regression model predicting expert label for the presence of clouds, with radiance angles (L) and features (R) as regressors. Predicted probabilities are arranged in ascending order. Red lines correspond to actual cloud labels and the black line represents predicted probabilities for each pixel. The plots are designed to visualize the overall fit of the corresponding model.