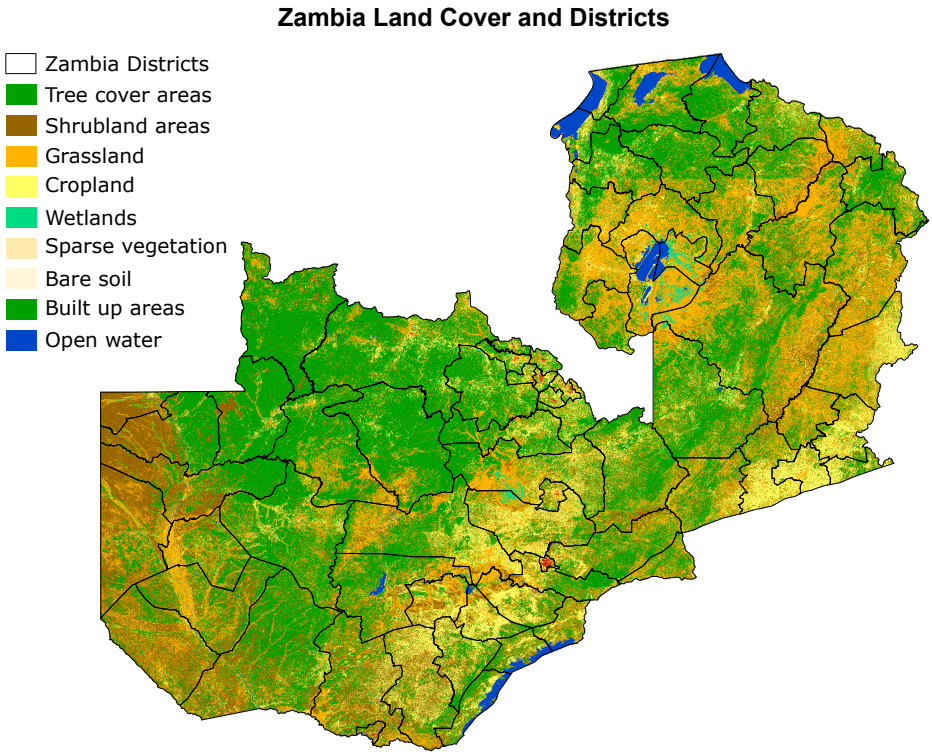

Figure S1. Zambia districts and the land cover distribution at 20-m resolution from ESA-CCI (2016).

Table S1. Implementation of the Recursive Feature Elimination approach.

Recursive Feature Elimination
1. Train the RF model using the entire training predictor set;
2. Calculate R^2 performance on the testing sample and set it as the baseline R^2;
3. Calculate the importance of each predictor by:
– Re-training the RF model without the selected predictor;
– Calculate the new R^2 performance;
– Calculated the change in performance as delta R^2 = baseline R^2 - new R^2. The higher the positive change in delta R^2 the more important is the predictor.
4. Rank the predictors based on their importance (higher delta R^2 to lowest delta R^2);
5. Remove the least important predictor (lowest delta R^2), and update the predictor set;
6. Repeat step 1–5 until stop when the lowest delta R^2 is < 0.001;
7. Compute the importance rank (step 3–4) for the final predictor set.

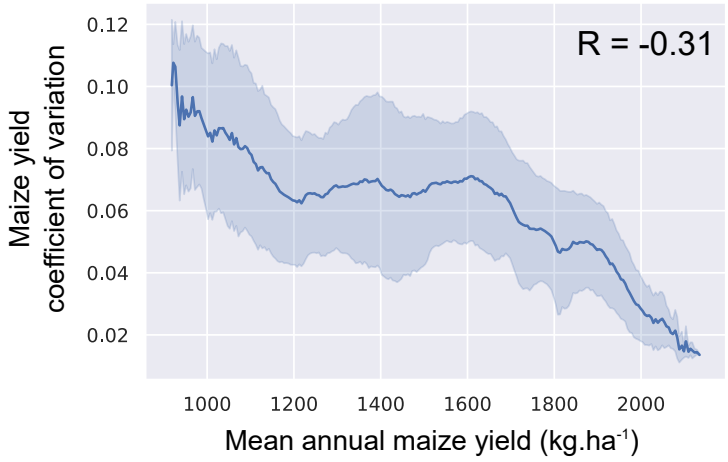

Figure S2. Relationship between field-scale mean annual maize yields and coefficient of variation, as estimated by the random forest model. We observe an inverse relationship (Pearson correlation of -0.31) that shows that locations of lower mean annual yields tend to have higher variation on their inter-annual yields, while locations of consistent high yield productivity tend to show minimal variation.

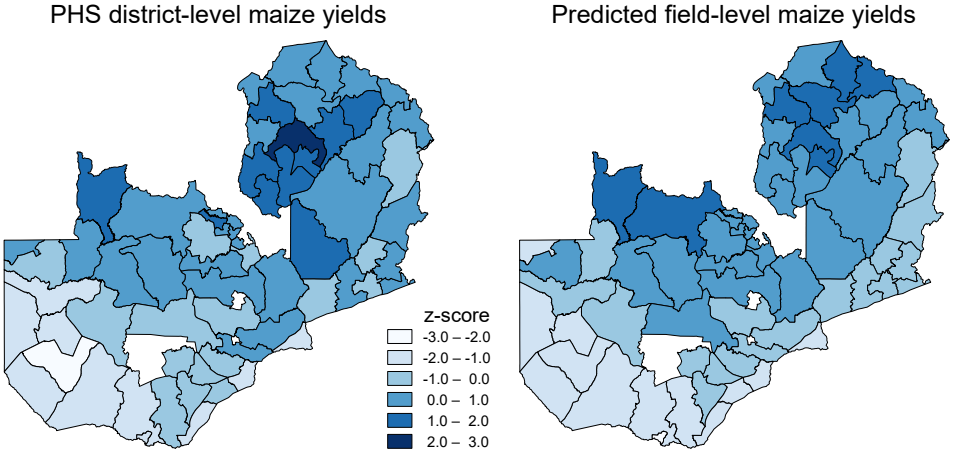

Figure S3. The z-score of the field-scale maize yields aggregate to the district-level in comparison with the z-score from the PHS district-level data.

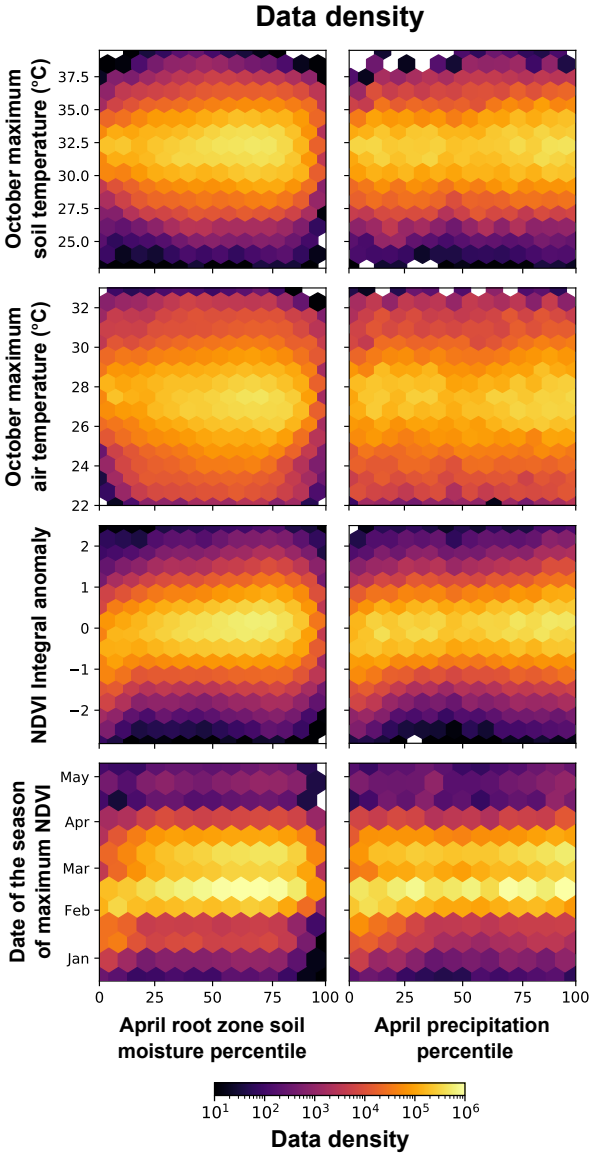

Figure S4. The number of grid cells used to calculate the mean yield anomaly values of each hexbin in Figure 8

.

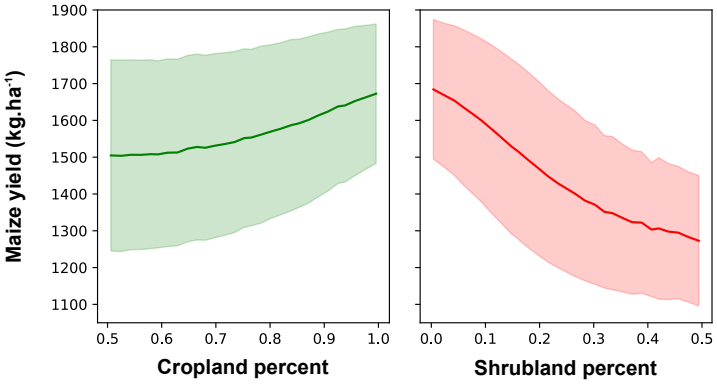

Figure S5. Mean field-scale maize yield for different cropland and shrubland percent. Shrubland percent characterized how much the agricultural area is fragmented, and it showed to be a strong predictor (inversely correlated) to maize yield. The shade shows the standard deviation.