Supplement of Earth Syst. Sci. Data, 12, 789–804, 2020
https://doi.org/10.5194/essd-12-789-2020-supplement


*Supplement of*

# Mapping the yields of lignocellulosic bioenergy crops from observations at the global scale

**Wei Li et al.**

*Correspondence to:* Wei Li (wli2019@tsinghua.edu.cn)

**Table R1 Plantation area and maximum and minimum MAI (mean annual increment) of eucalypt, poplar and willow from inventory data compiled in FAO (2006).**

| Species | Area (1000 ha) | MAI min ($m^3$/ha/y) | MAI max ($m^3$/ha/y) | Country |
|---|---|---|---|---|
| *Eucalyptus grandis* | 335 | 21 | 27 | South Africa |
| *Eucalyptus nitens* | 231 | 19 | 26 | South Africa |
| *Eucalyptus spp.* | 473 | 8 | 21 | Sudan |
| *Populus spp.* | 3220 | 9 | 18 | China |
| *Eucalyptus spp.* | 2397 | 8 | 21 | China |
| *Eucalyptus spp.* | 4047 | 8 | 21 | Indonesia, Viet Nam, India |
| *Populus spp.* | 171 | 9 | 18 | India |
| *Populus spp.* | 84 | 9 | 18 | Belgium, Netherlands, Ukraine, Latvia |
| *Populus hybrids* | 83 | 16 | 21 | Italy |
| *Eucalyptus globulus* | 442 | 16 | 25 | Australia |
| *Eucalyptus nitens* | 35 | 19 | 26 | Australia |
| *Eucalyptus dunnii* | 18 | 16 | 18 | Australia |
| *Eucalyptus grandis* | 18 | 21 | 27 | Australia |
| *Eucalyptus pilularis* | 18 | 18 | 18 | Australia |
| *Eucalyptus regnans* | 18 | 18 | 20 | Australia |
| *Eucalyptus spp.* | 3678 | 8 | 21 | Brazil, Chile |
| *Eucalyptus grandis* | 99 | 21 | 27 | Argentina |
| *Populus spp.* | 31 | 9 | 18 | Brazil, Chile |
| *Salix alba* | 23 | 13 | 20 | Argentina |
| *Salix babylonica* | 23 | 20 | 25 | Argentina |
| *Salix babylonica var. sacramenta* | 23 | 20 | 25 | Argentina |
| *Salix hibrids* | 23 | 20 | 25 | Argentina |

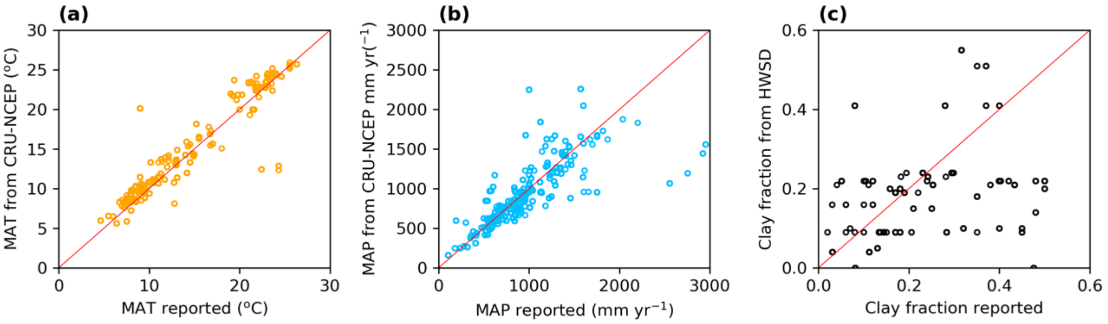



**Figure S1: Comparison of mean annual temperature (MAT), mean annual precipitation (MAP) and Clay fraction (CF) from the global yield dataset with MAT (a) and MAP (b) from CRUNCEP and CF from HWSD (c). The red line is the 1:1 line.**

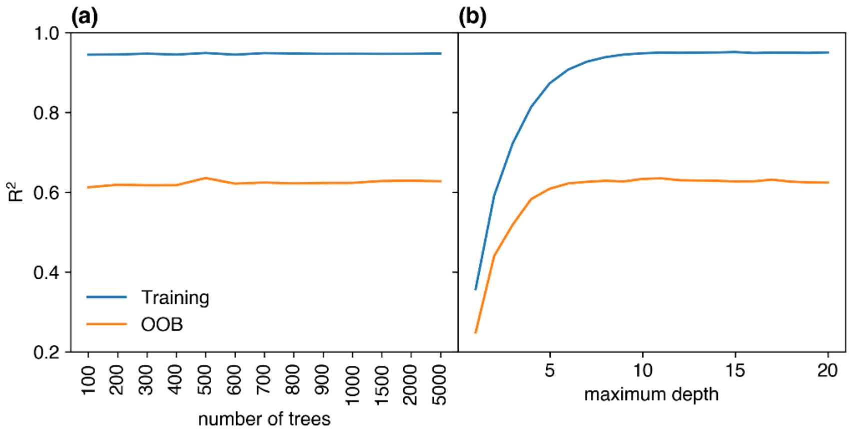



**Figure S2: Coefficient of determination ($R^2$) between predictions and observations in the training data and $R^2$ of OOB validation using different numbers of trees in the forest (a) and different maximum depths of a tree (b). These tests are based on the selected sites after removing the masked sites.**

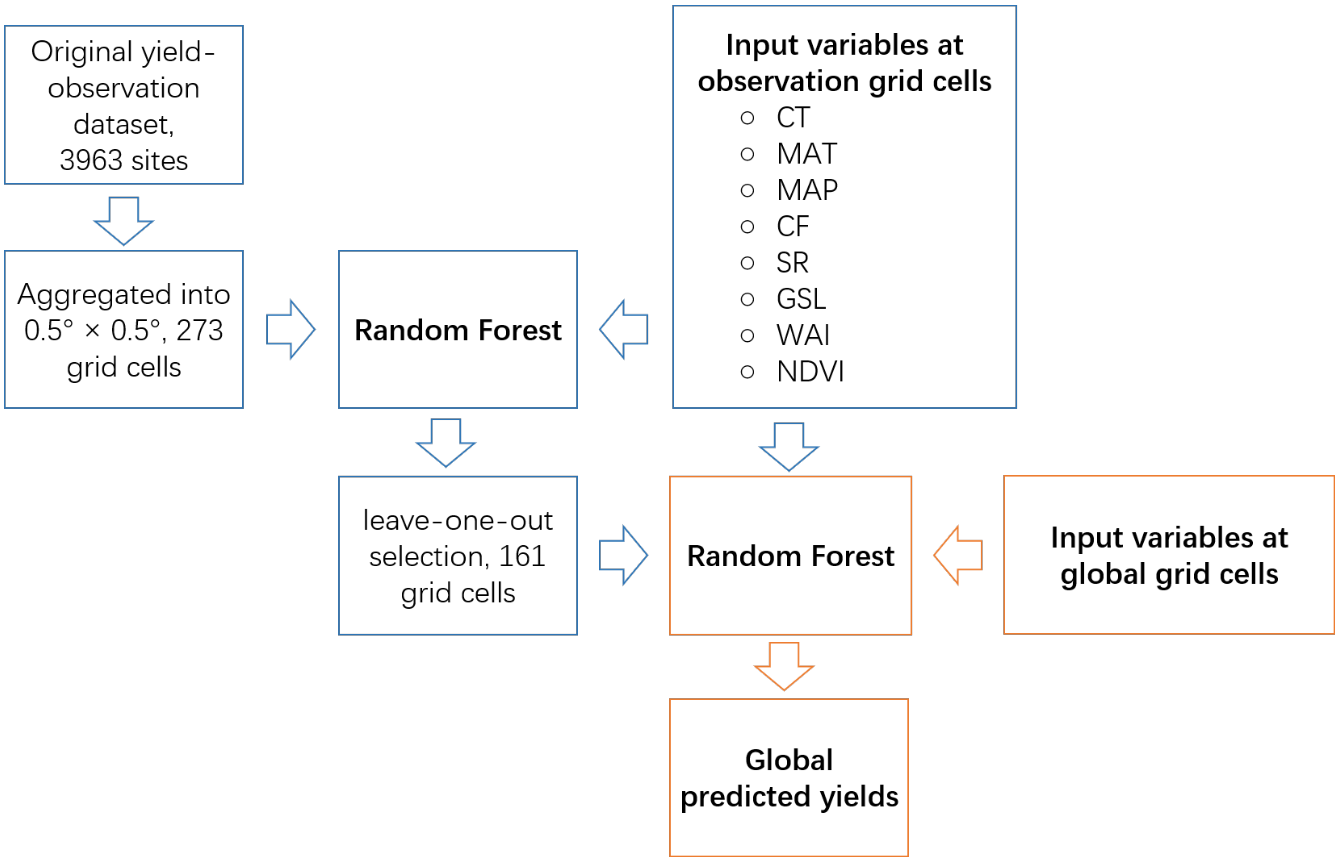



**Figure S3: Workflow of random forest training and predicting in this study. The abbreviations of input variables can be found in Table 1.**

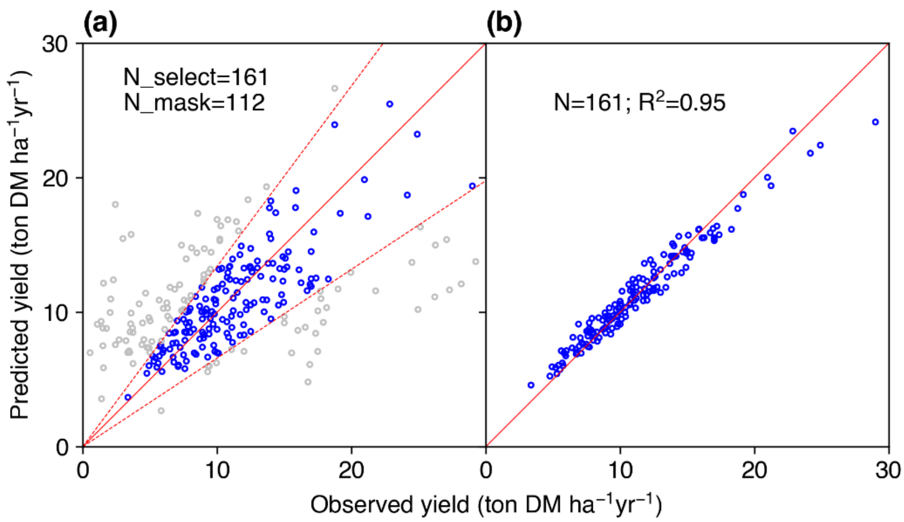



**Figure S4: The comparison between observed and predicted yields. (a) shows results from the leave-one-out method. Grid cells with a relative bias greater than 1-σ were masked (gray dots). (b) shows the yield predictions from the RF model trained using the selected sites. The solid red line is the 1:1 line. The dashed red lines in (a) indicate the 1-σ boundaries.**

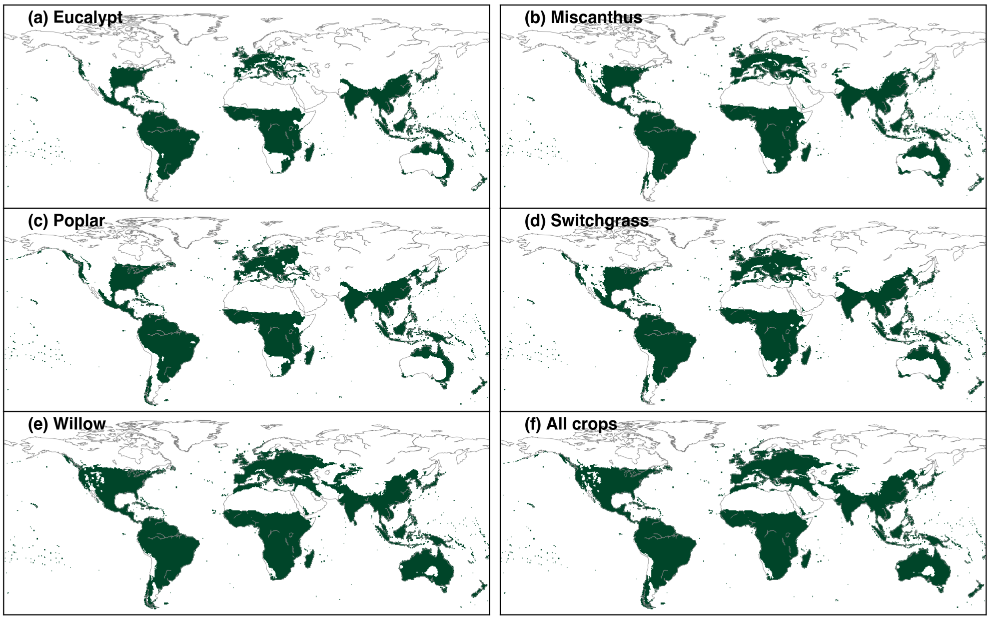



**Figure S5: Grid cells that are adequate for growth defined by minimum MAT and MAP of each bioenergy crop in the training data. "All crops" represents grid cells where at least one bioenergy crop can grow.**

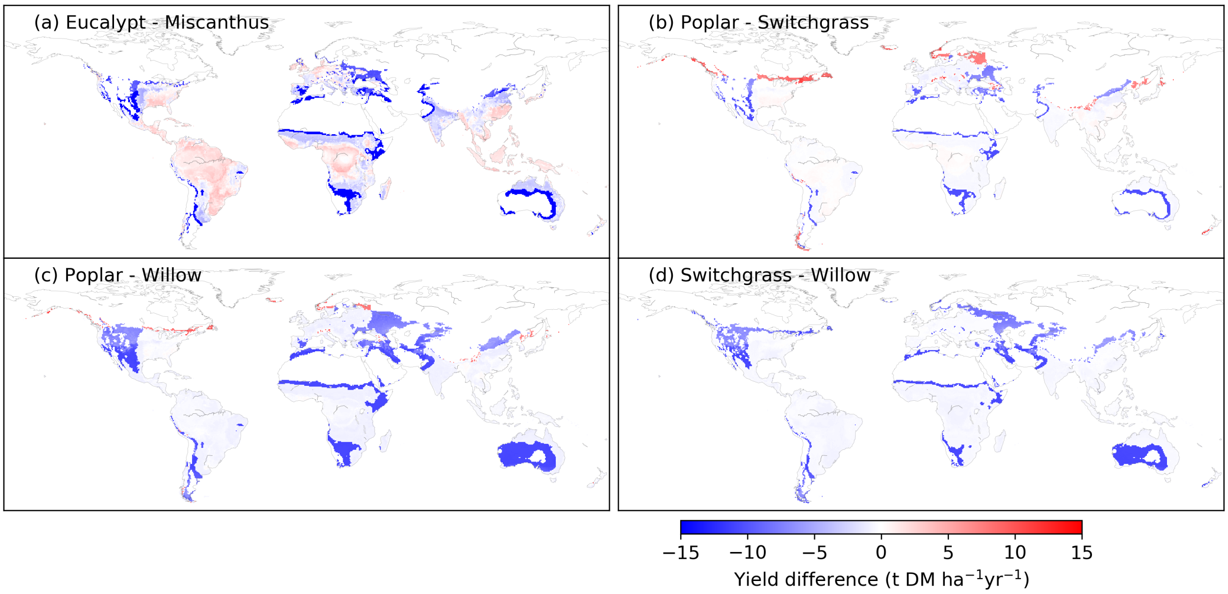



**Figure S6: Difference of predicted yields between various bioenergy crop types.**

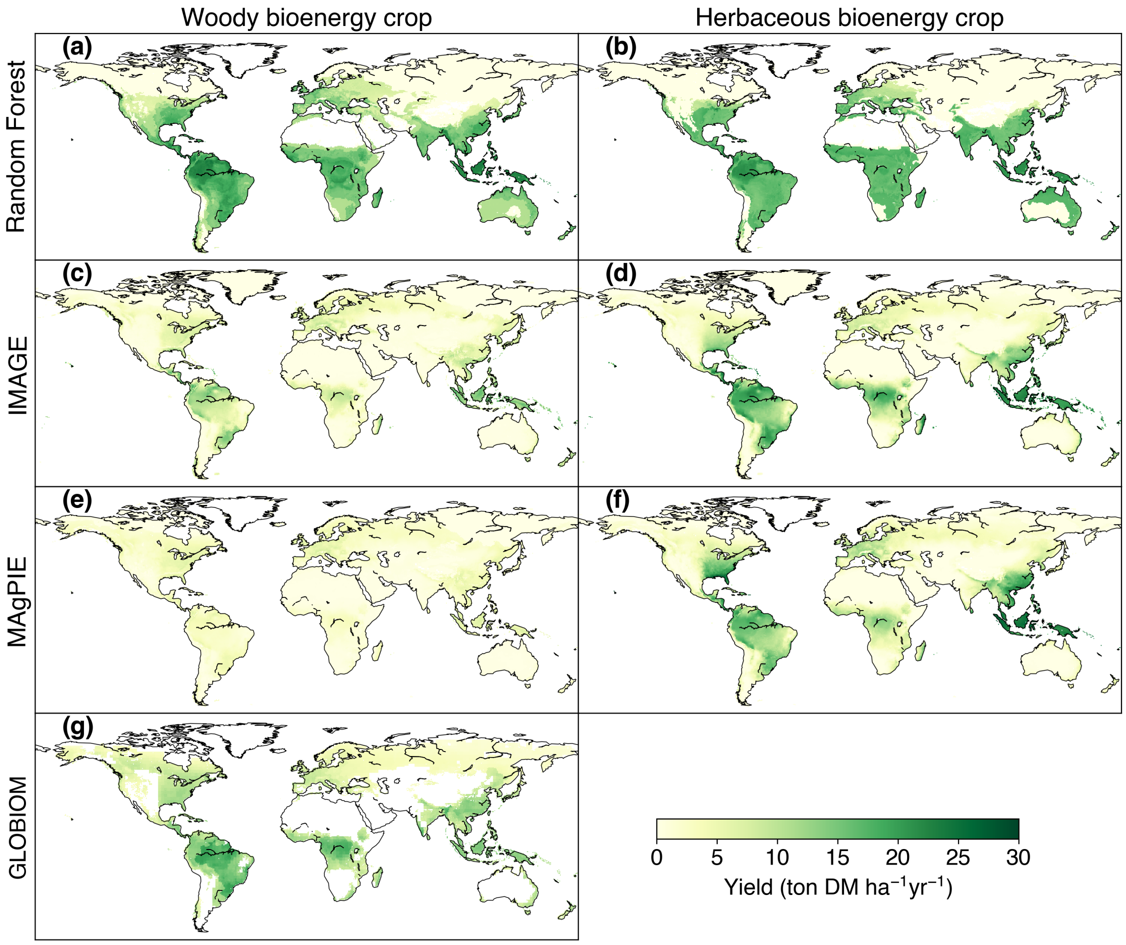



**Figure S7: Woody (left panel) and herbaceous (right panel) bioenergy crop yields from the RF maps and maps used in IMAGE, MAgPIE and GLOBIOM. Because there are three woody crops (eucalypt, poplar and willow) and two herbaceous crops (switchgrass and *Miscanthus*) in the RF maps, only the highest yields in each grid cell are shown for the woody and herbaceous crop categories respectively.**

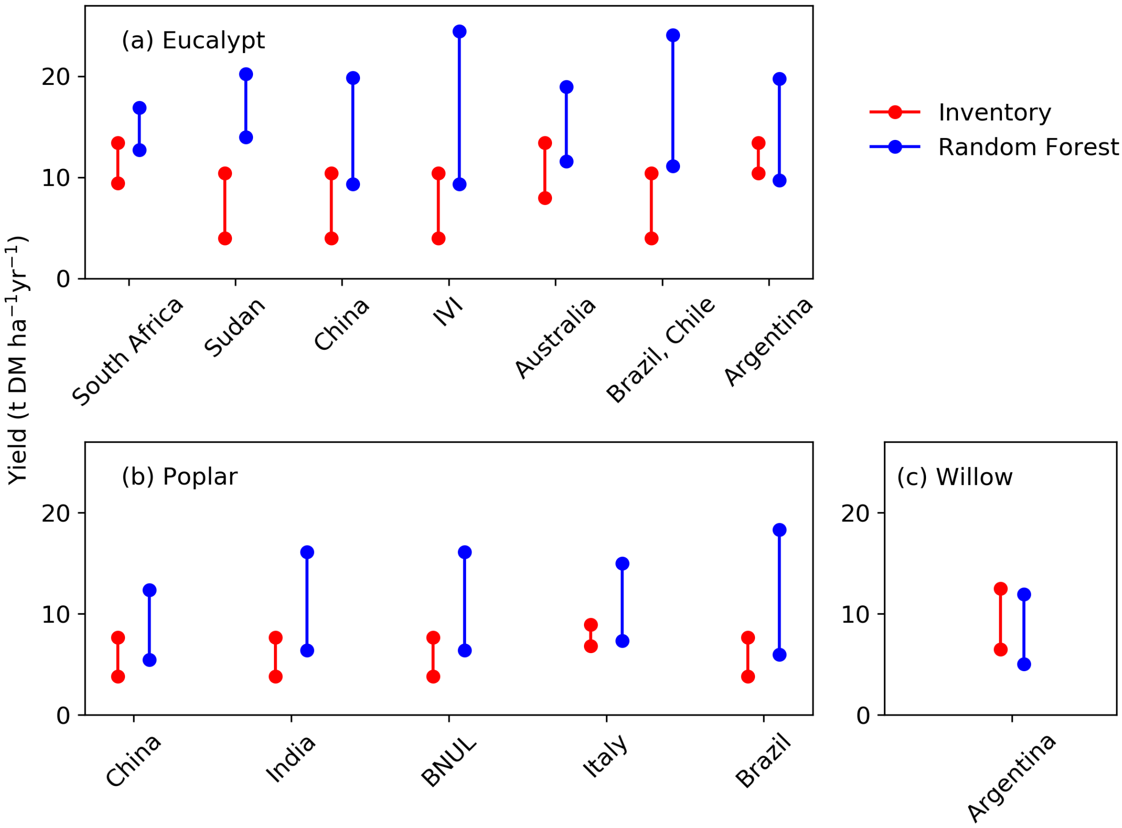



**Figure S8: Yield ranges from (limited) inventory data and our random forest maps at country levels. IVI stands for Indonesia, Viet Nam and India; BNUL stands for Belgium, Netherlands, Ukraine and Latvia.**

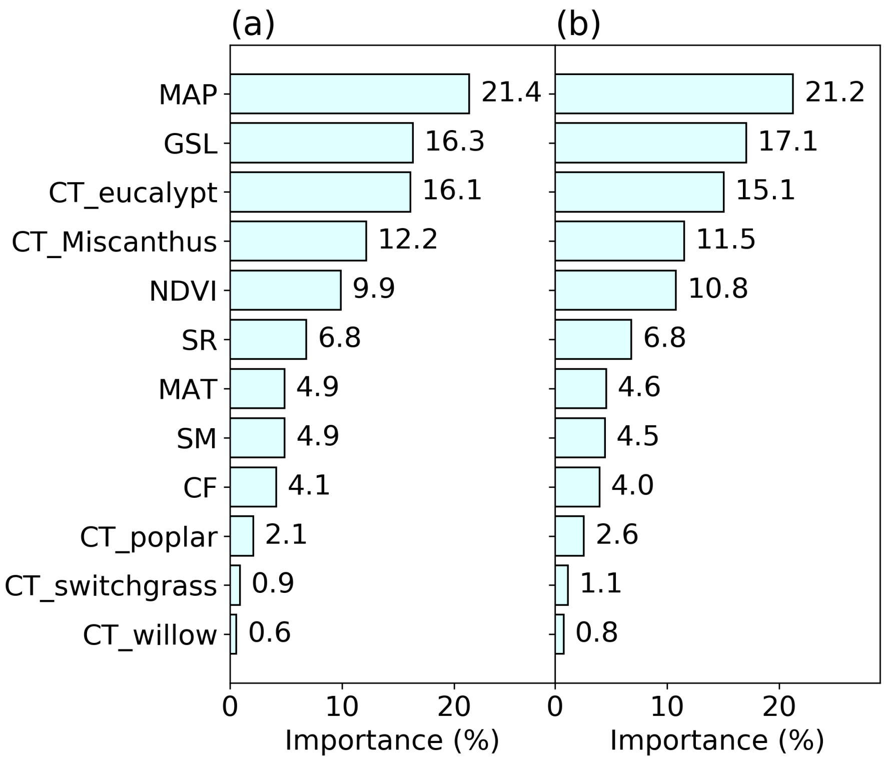
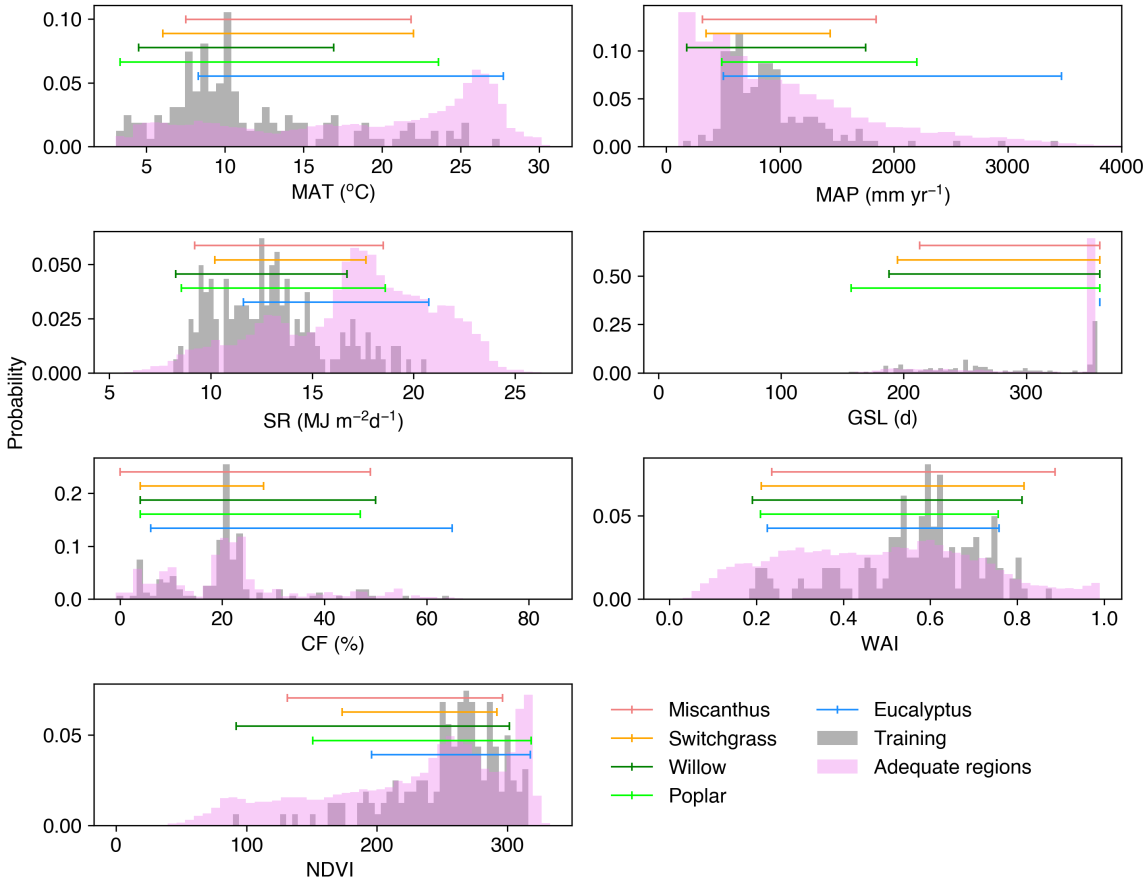



**Figure S9: Variable importance in the trained RF model using soil moisture (SM) data from SMOS (a) and SMAP (b).**




70

**Figure S10: Distributions of explanatory variables in the training data and in the regions that are adequate for bioenergy crop growth. The ranges of variables for each bioenergy crop type in the training data are also shown as lines with different colors.**

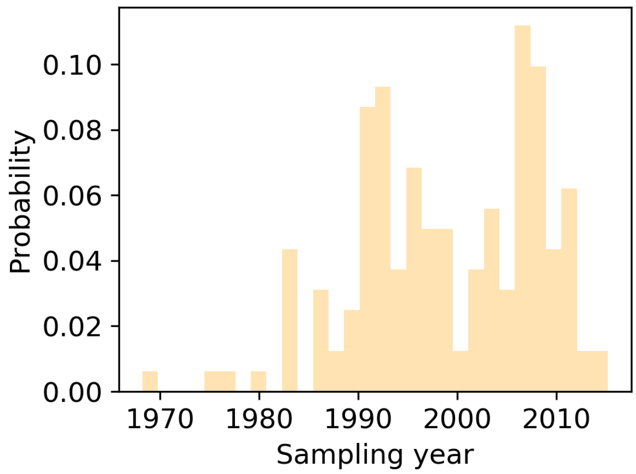



**Figure S11: Histogram of sampling years in the yield observation grid data used for random forest training.**

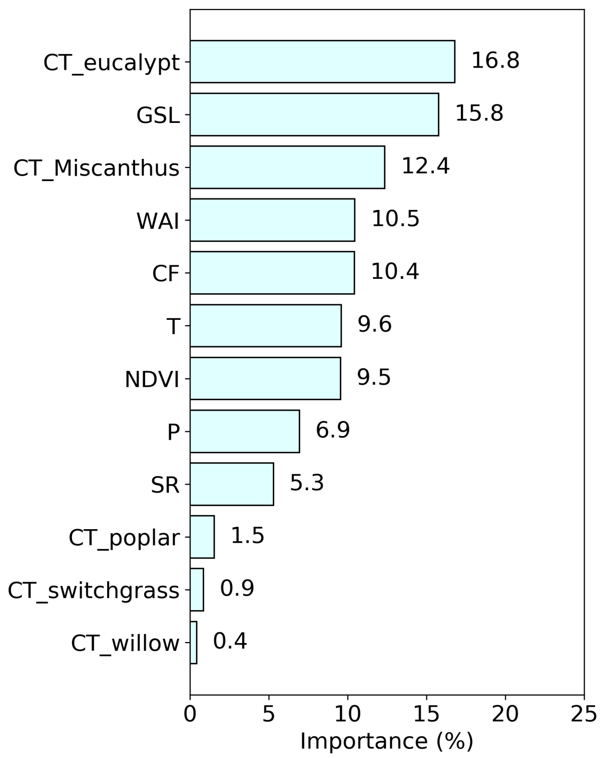



**Figure S12: Variable importance in the trained RF model using climate conditions at the sampling years.**

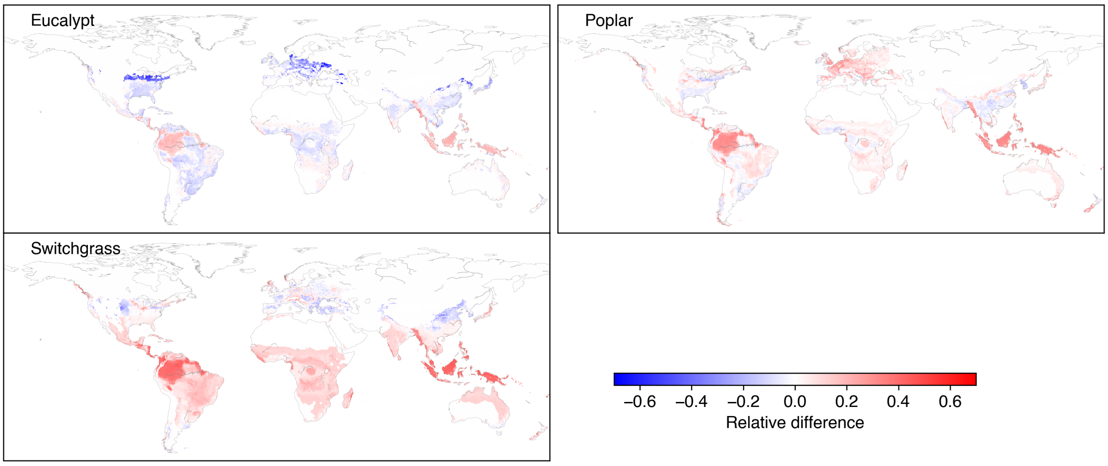
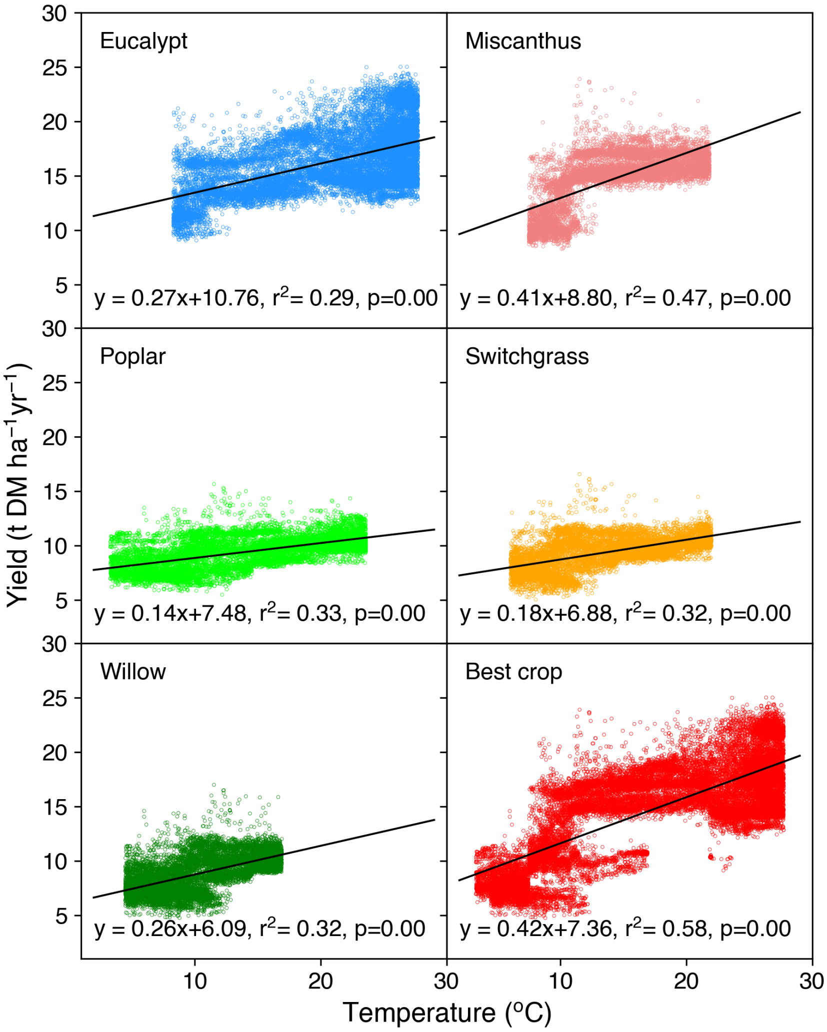
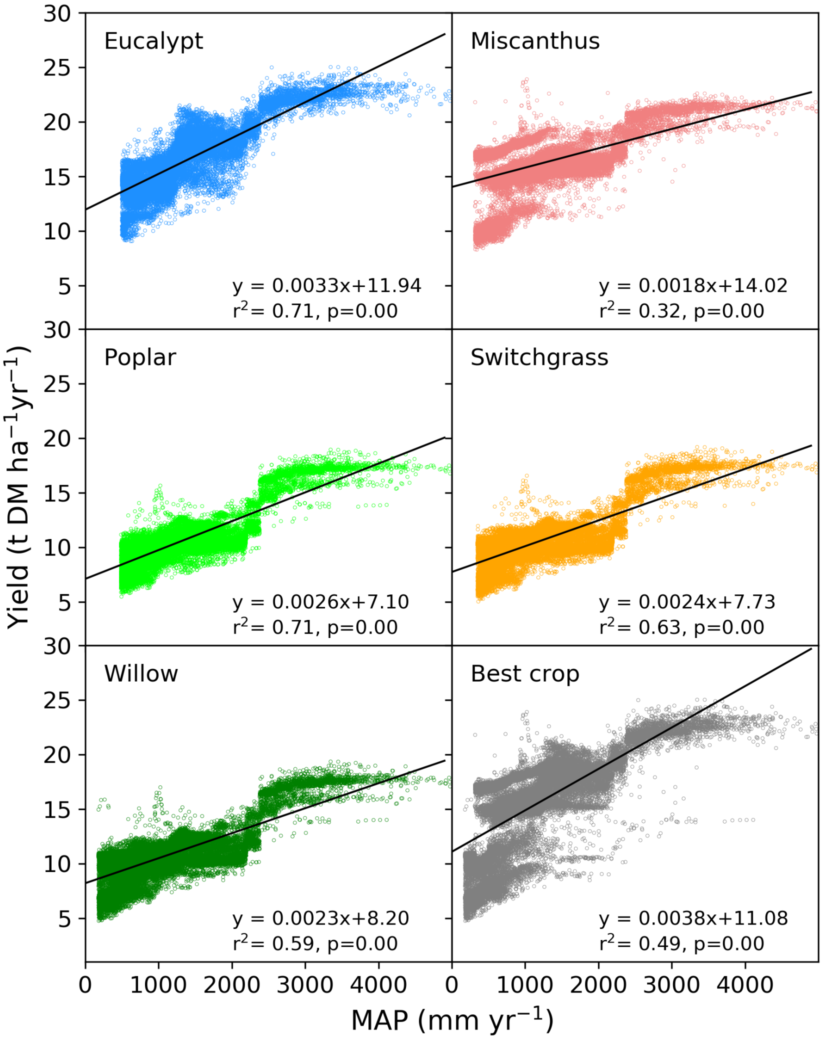



**Figure S13: Relative differences of yields between the original RF predictions using crop type as a categorical variable and the predictions from the RF model trained for each individual bioenergy crop. Red color indicates higher yield predictions from the former method while blue color indicates higher yield from the latter.**




95

**Figure S14: Relationship of bioenergy crop yield with temperature across all grid cells that are adequate for bioenergy crop growth.**




100

**Figure S15: Relationship of bioenergy crop yield with mean annual precipitation (MAP) across all grid cells that are adequate for bioenergy crop growth.**

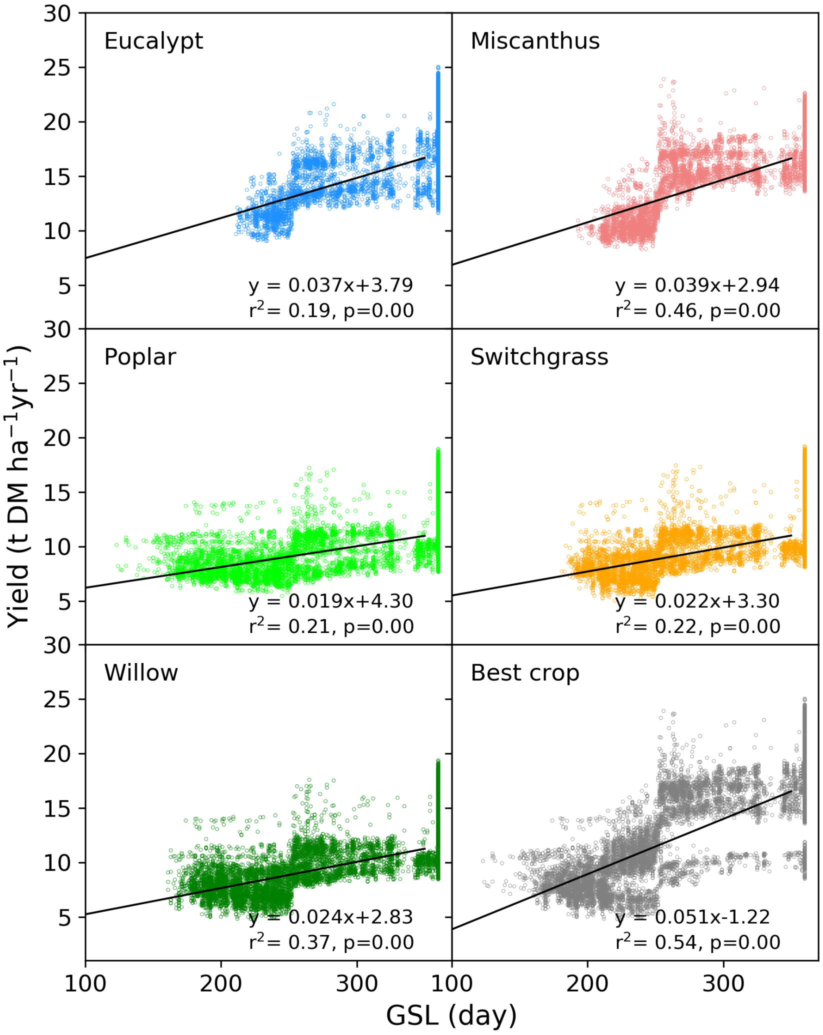
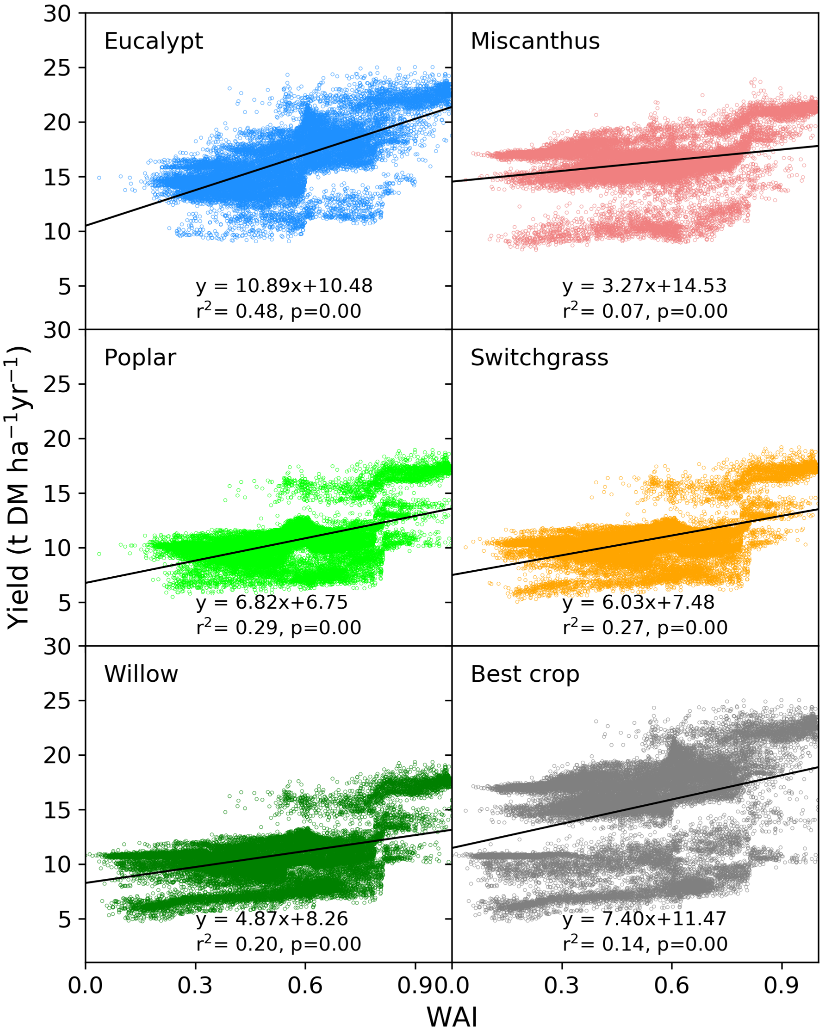



**Figure S16: Relationship of bioenergy crop yield with growing season length (GSL) across all grid cells that are adequate for bioenergy crop growth.**




110 **Figure S17: Relationship of bioenergy crop yield with soil water availability index (WAI) across all grid cells that are adequate for bioenergy crop growth.**

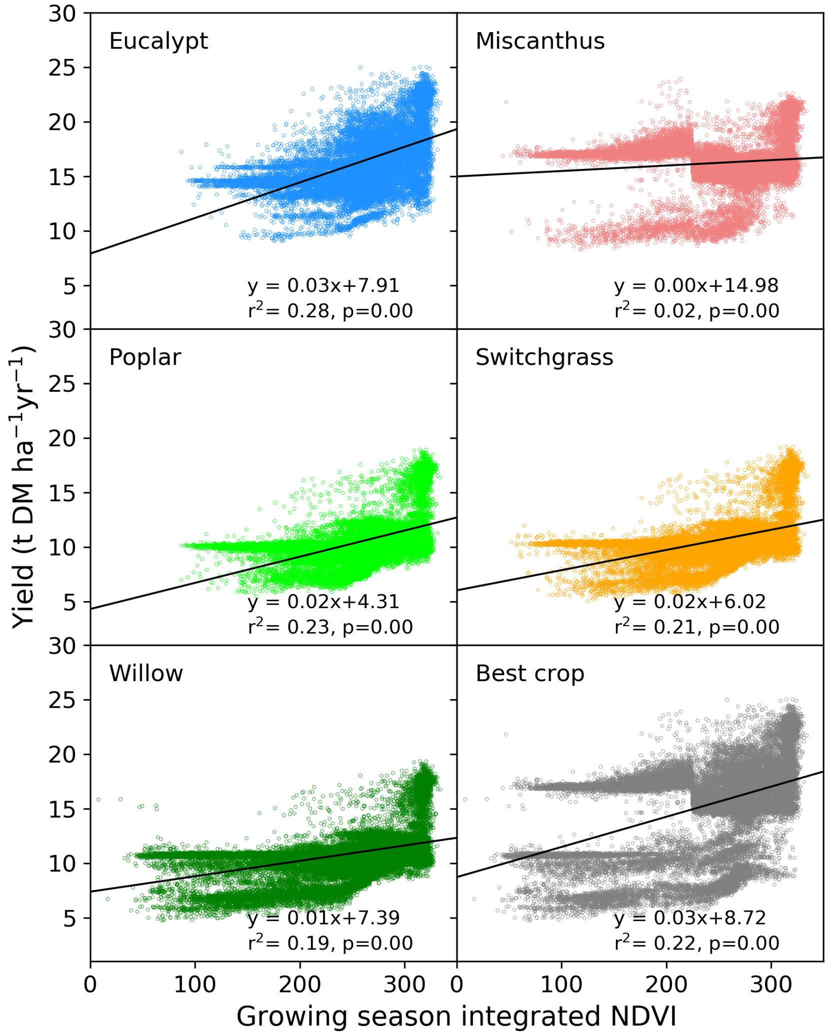



**Figure S18: Relationship of bioenergy crop yield with growing season integrated normalized difference vegetation index (NDVI) across all grid cells that are adequate for bioenergy crop growth.**

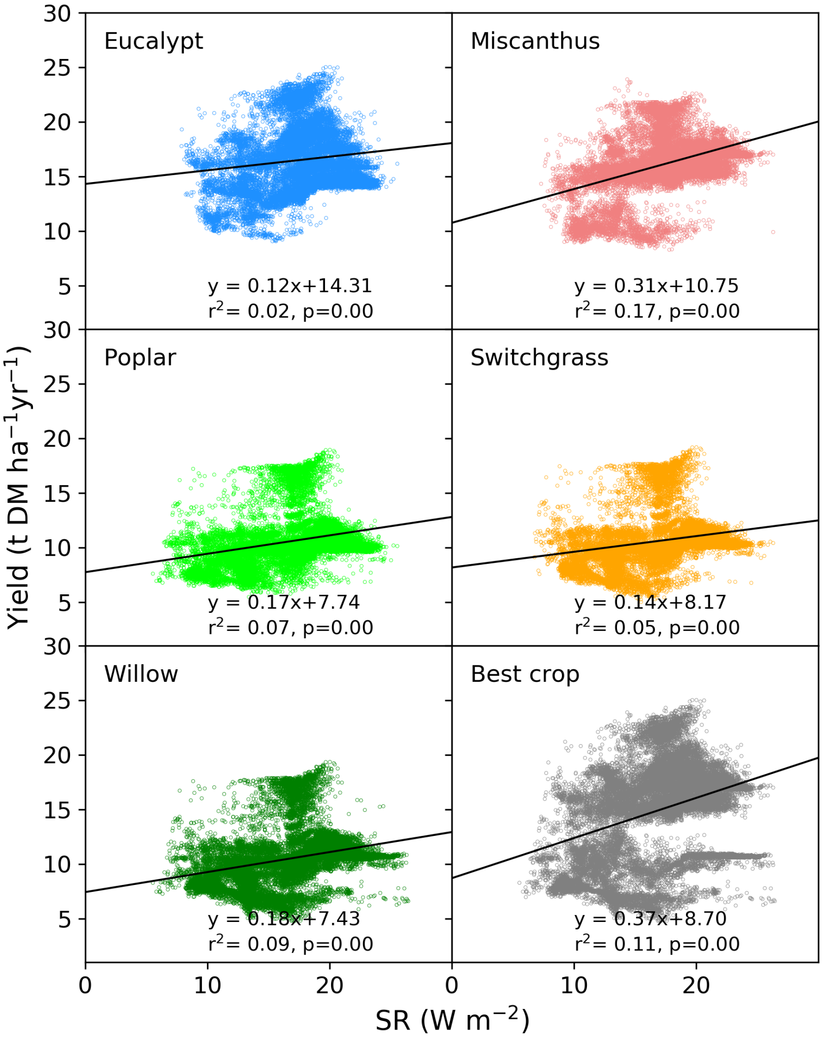



**Figure S19: Relationship of bioenergy crop yield with shortwave radiation (SR) across all grid cells that are adequate for bioenergy crop growth.**

120

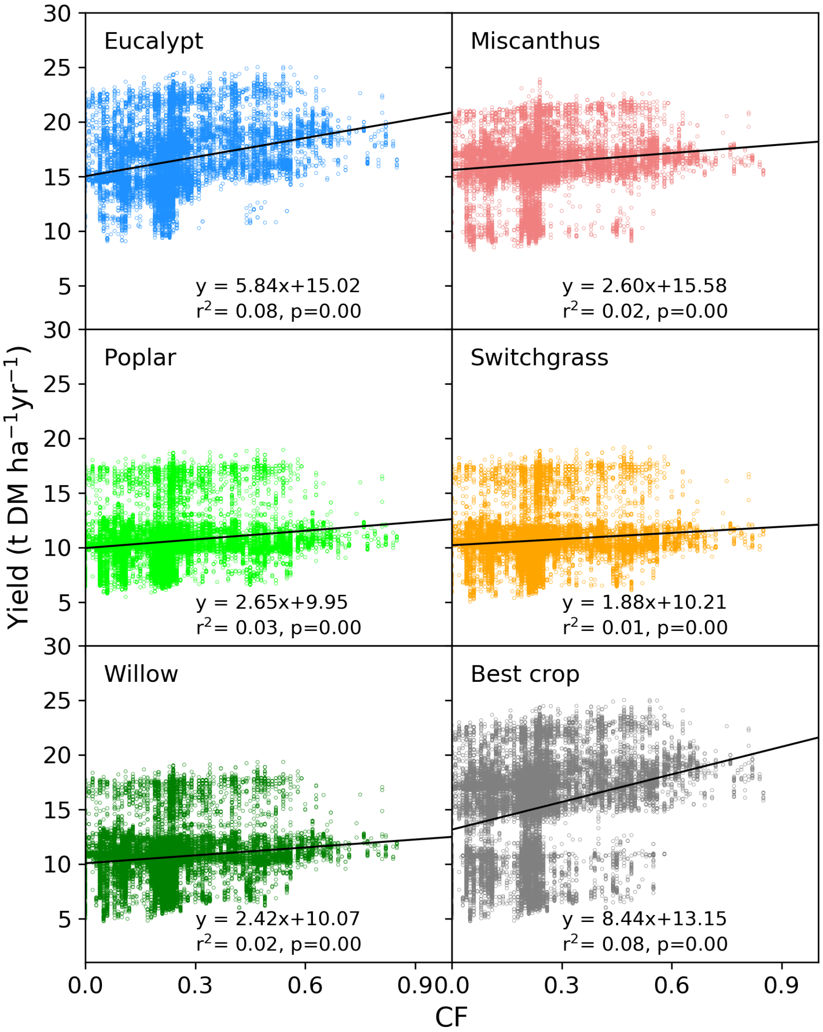



**Figure S20: Relationship of bioenergy crop yield with clay fraction (CF) across all grid cells that are adequate for bioenergy crop growth.**

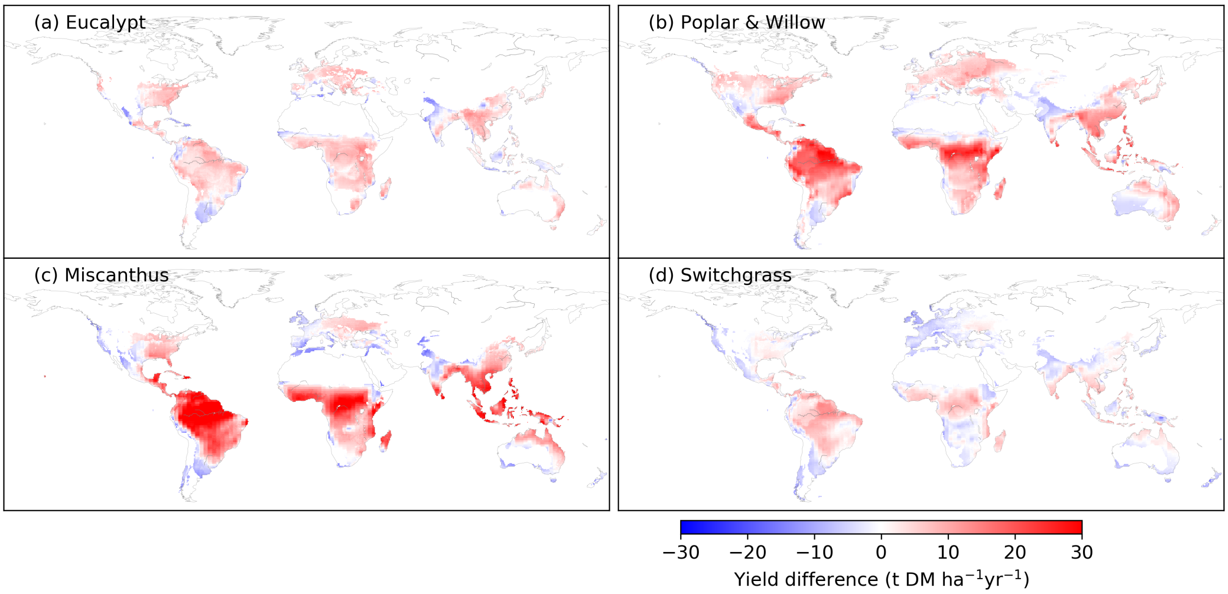



**Figure S21: Comparison of bioenergy crop yields between the RF map and maps simulated by ORCHIDEE (ORCHIDEE yields minus RF yields where yields are available in both paired maps).**