Supplement of

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

Wei Li[1,2], Philippe Ciais[2], Elke Stehfest3, Detlef van Vuuren3, Alexander Popp4, Almut Arneth5, Fulvio Di Fulvio6, Jonathan Doelman3, Florian Humpenöder4, Anna Harper7, Taejin Park8, David Makowski9,10, Petr Havlik6, Michael Obersteiner6, Jingmeng Wang[1], Andreas Krause[5,11], Wenfeng Liu[2]

[1]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, 100084, China
[2]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
[3]Department of Climate, Air and Energy, Netherlands Environmental Assessment Agency (PBL), The Hague, The Netherlands
[4]Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany
[5]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, Germany
[6]International Institute for Applied Systems Analysis, Ecosystem Services and Management Program, Schlossplatz 1, A-2361, Laxenburg, Austria
[7]College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter EX4 4QF, UK. 2 College of Life and Environmental Sciences, University of Exeter, Exeter EX4 4QF, UK
[8]Department of Earth and Environment, Boston University, Boston, MA 02215, USA
[9]CIRED, CIRAD, 45 bis Avenue de la Belle Gabrielle, 94130 Nogent-sur-Marne, France
[10]UMR Agronomie, INRA, AgroParisTech, Université Paris-Saclay, ThivervalGrignon 78850, France
[11]TUM School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany
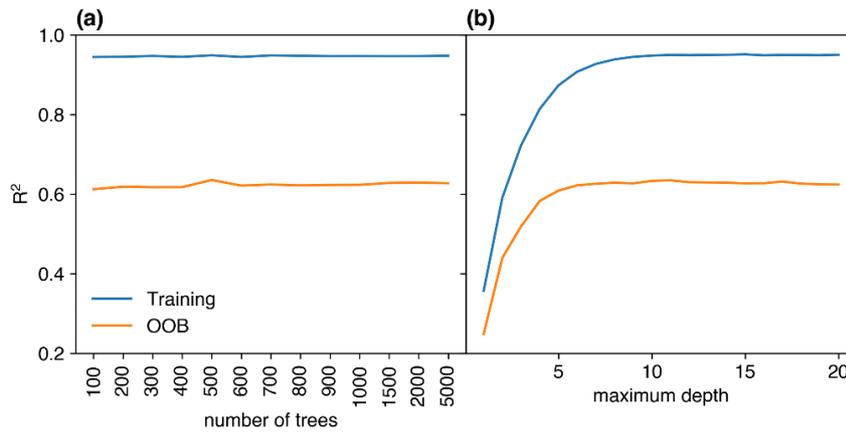
Correspondence to: Wei Li (wli2019@tsinghua.edu.cn)
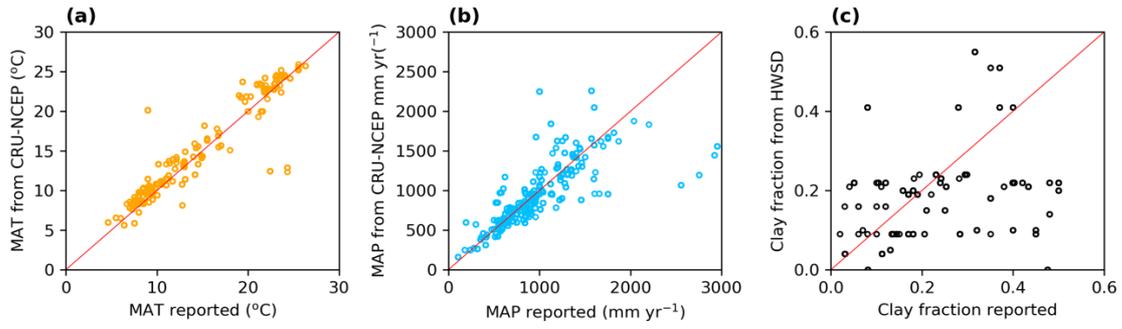
Figure S1: Comparison of mean annual temperature (MAT), mean annual precipitation (MAP) and Clay fraction (CF) from the global yield dataset with MAT (a) and MAP (b) from CRUNCEP and CF from HWSD (c). The red line is the 1:1 line.
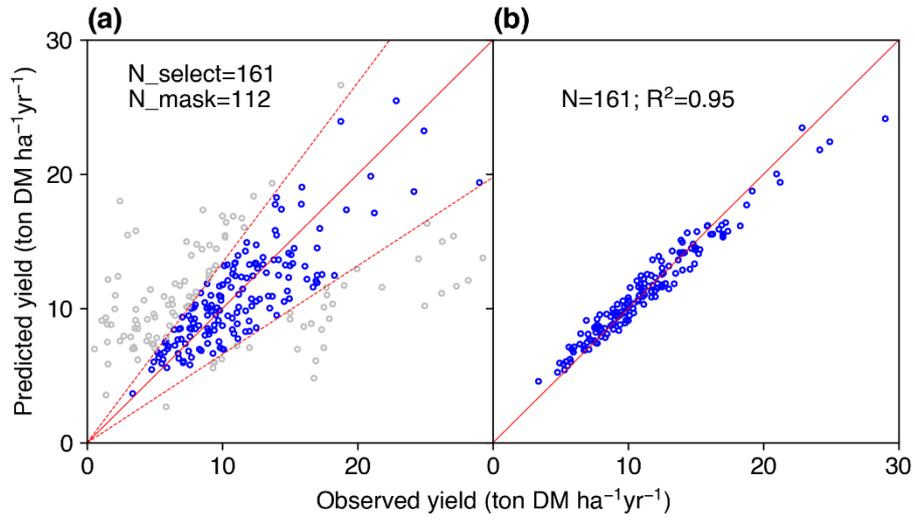
Figure S2: Coefficient of determination (R^2) between predictions and observations in the training data and R^2 of OOB validation using different numbers of trees in the forest (a) and different maximum depths of a tree (b). These tests are based on the selected sites after removing the masked sites.

Figure S3: 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.
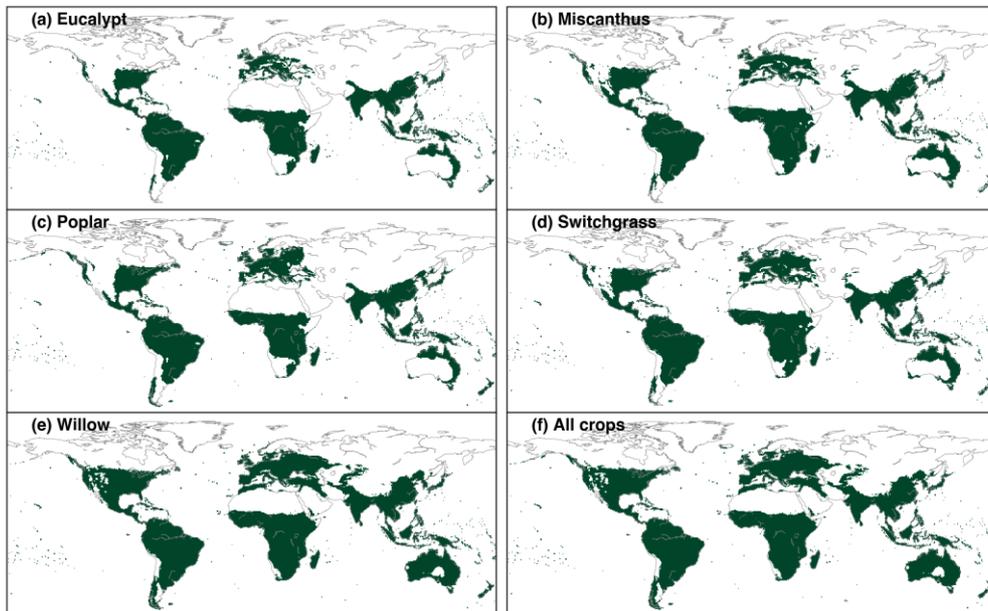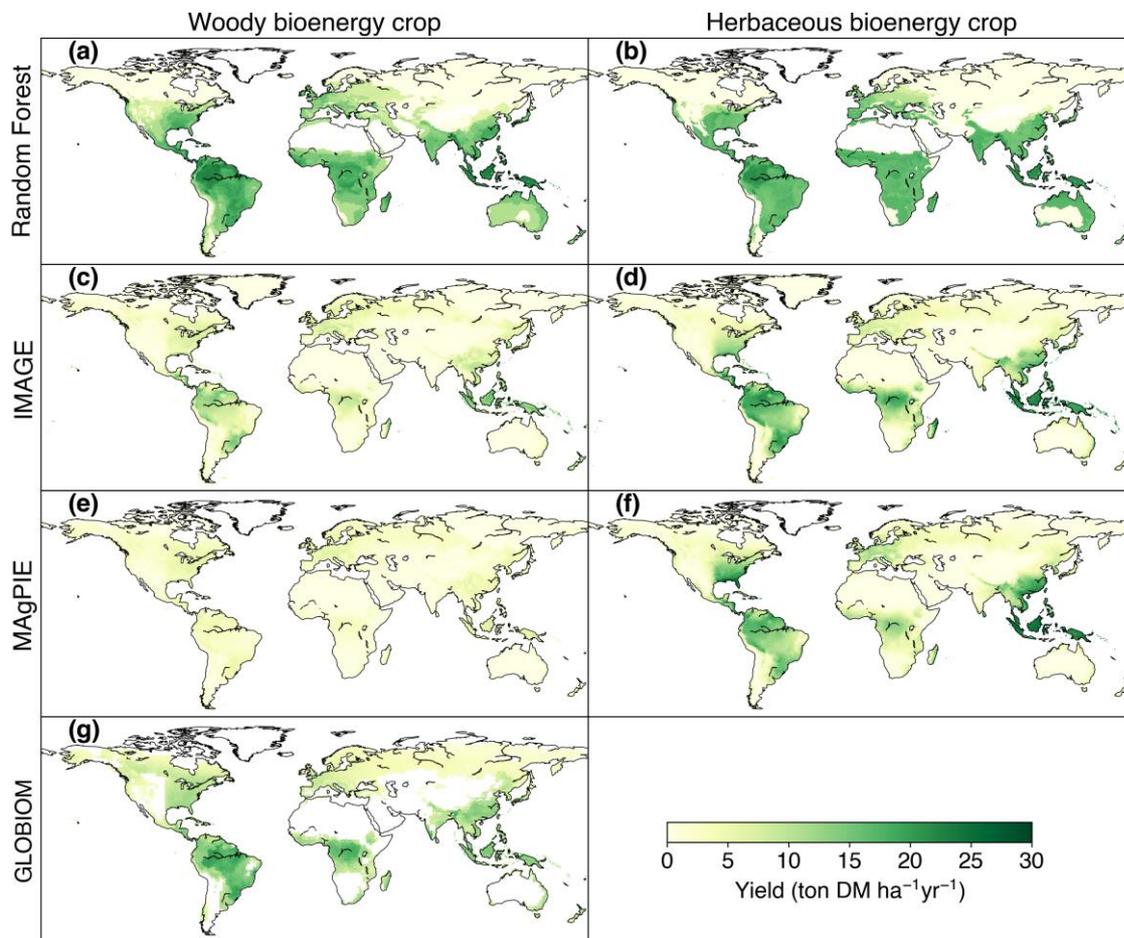

Figure S4: Grid cells that are adequate for growth defined by minimum MAT and MAP of each bioenergy crop in the training data. "All crops" represents grid cells where at least one bioenergy crop can grow.

Figure S5: 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.

65

4
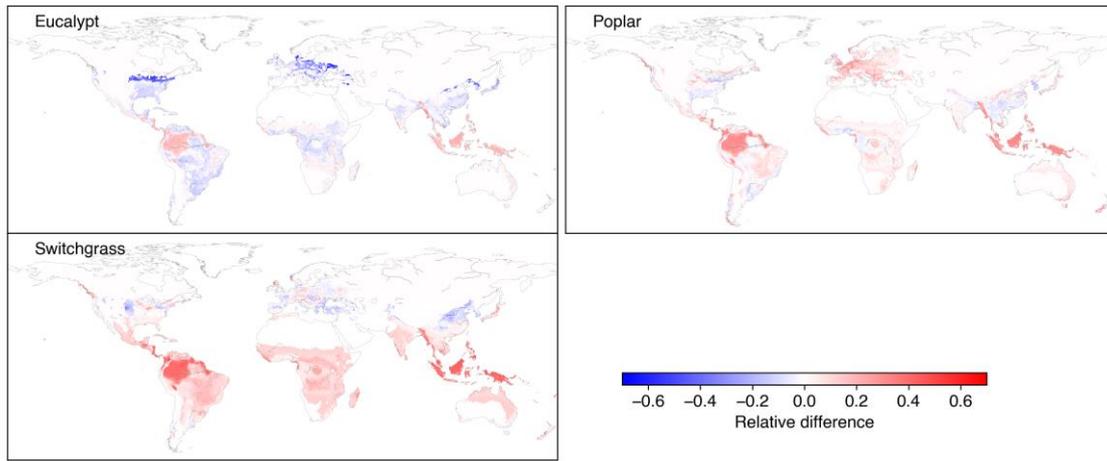
Figure S6: Relative differences of yields between the original RF predictions using crop type as a categorical variable and the predictions from the RF model trained for each individual bioenergy crop. Red color indicates higher yield predictions from the former method while blue color indicates higher yield from the latter.
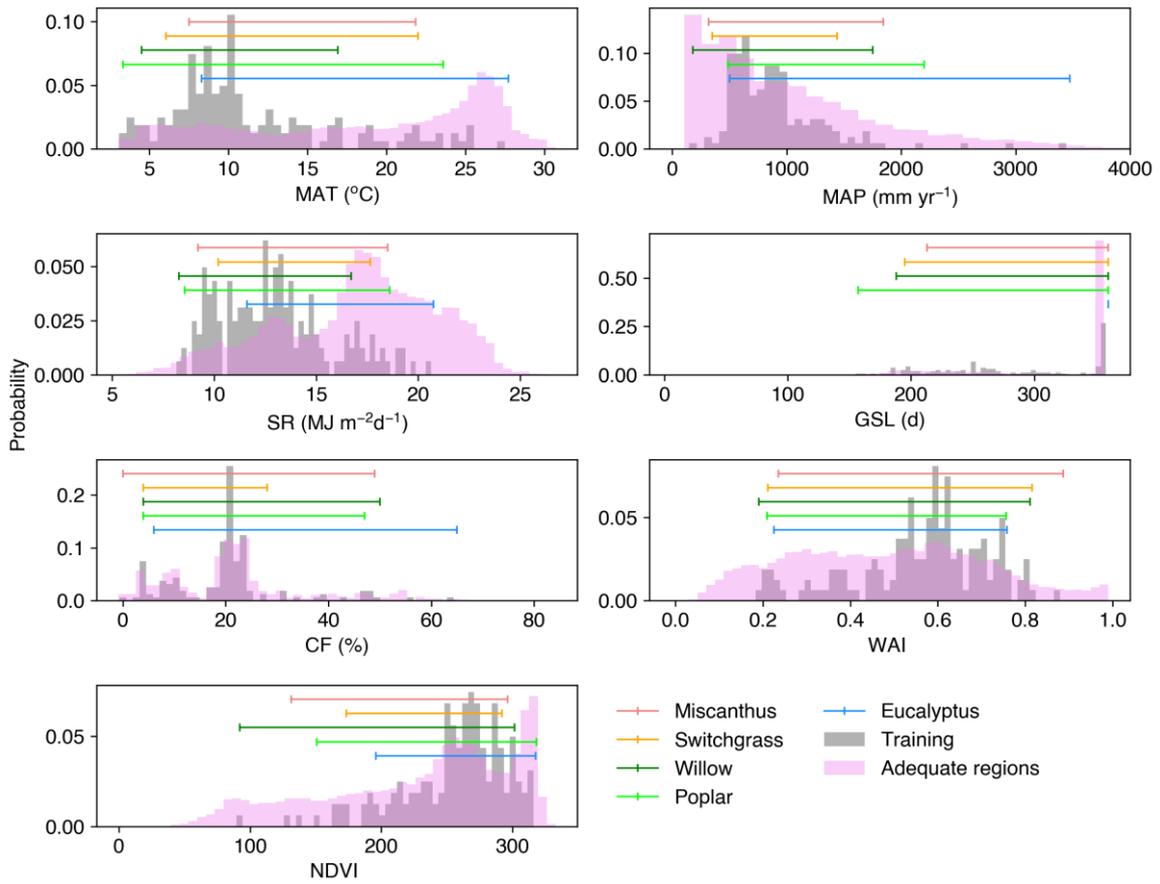
Figure S7: Distributions of explanatory variables in the training data and in the regions that are adequate for bioenergy crop growth. The ranges of variables for each bioenergy crop type in the training data are also shown as lines with different colors.
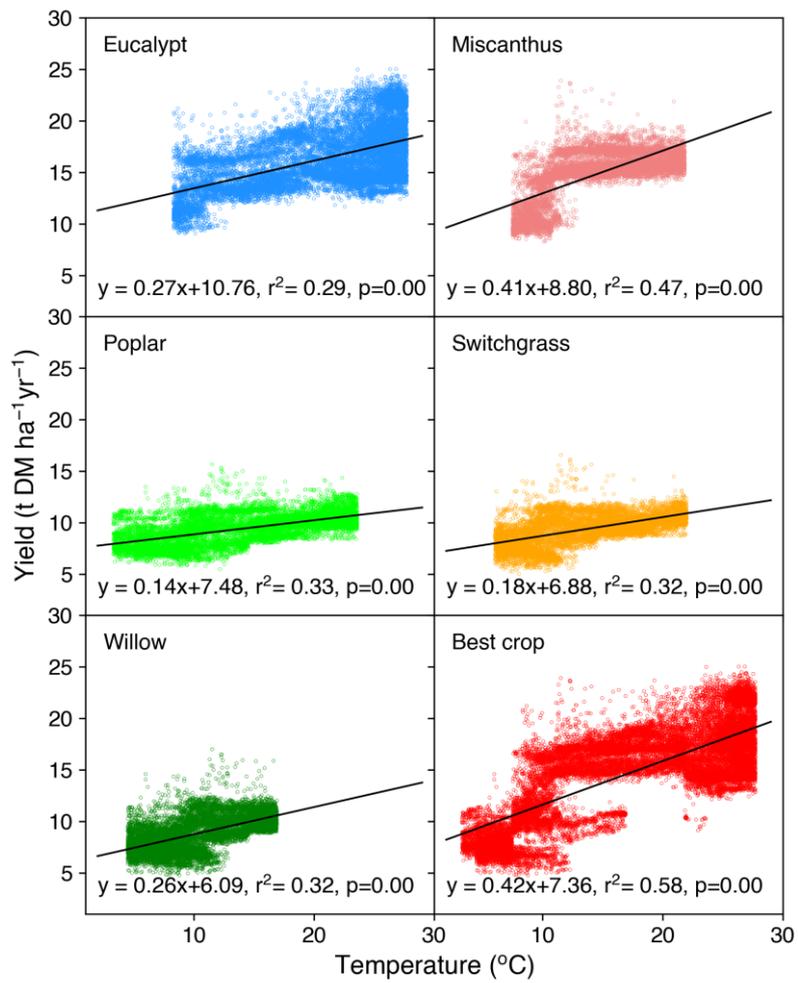
80

Figure S8: Relationship of bioenergy crop yield with temperature across all grid cells that are adequate for bioenergy crop growth.

85