URL

Supporting Information for "Coupling water and carbon fluxes to constrain estimates of transpiration: the TEA algorithm"

Jacob A. Nelson[1], Nuno Carvalhais[1,2], Matthias Cuntz[3], Nicolas Delpierre[4], Jrgen Knauer[1], Jrme Oge[5], Mirco Migliavacca[1], Markus Reichstein[1,6], Martin Jung[1]

Jacob A. Nelsonjnelson@bgc-jena.mpg.de

[1]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, Jena, Germany

[2]Faculdade de Ciências e Tecnologia, FCT, Universidade Nova de Lisboa

[3]INRA, Université de Lorraine, UMR1137 Ecology et Ecophysiologie Forestières, Route d'Amance, 54280 Champenoux, France

[4]Ecologie Systématique Evolution, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91400 Orsay, France

[5]INRA, UMR 1391 ISPA, 33140 Villenave d'Ornon, France

[6]Michael-Stifel-Center Jena for Data-Driven and Simulation Science, Jena, Germany

Contents of this file

Additional Supporting Information (Files uploaded separately)

1. Table of sites used: FileS5.pdf

Introduction This document gives additional figures related to the sensitivities of the TEA algorithm to filtering and Random Forest hyper-parameter, as well as sensitivity of bias in transpiration estimates to plant and environmental factors.

0148-0227/18/$9.00

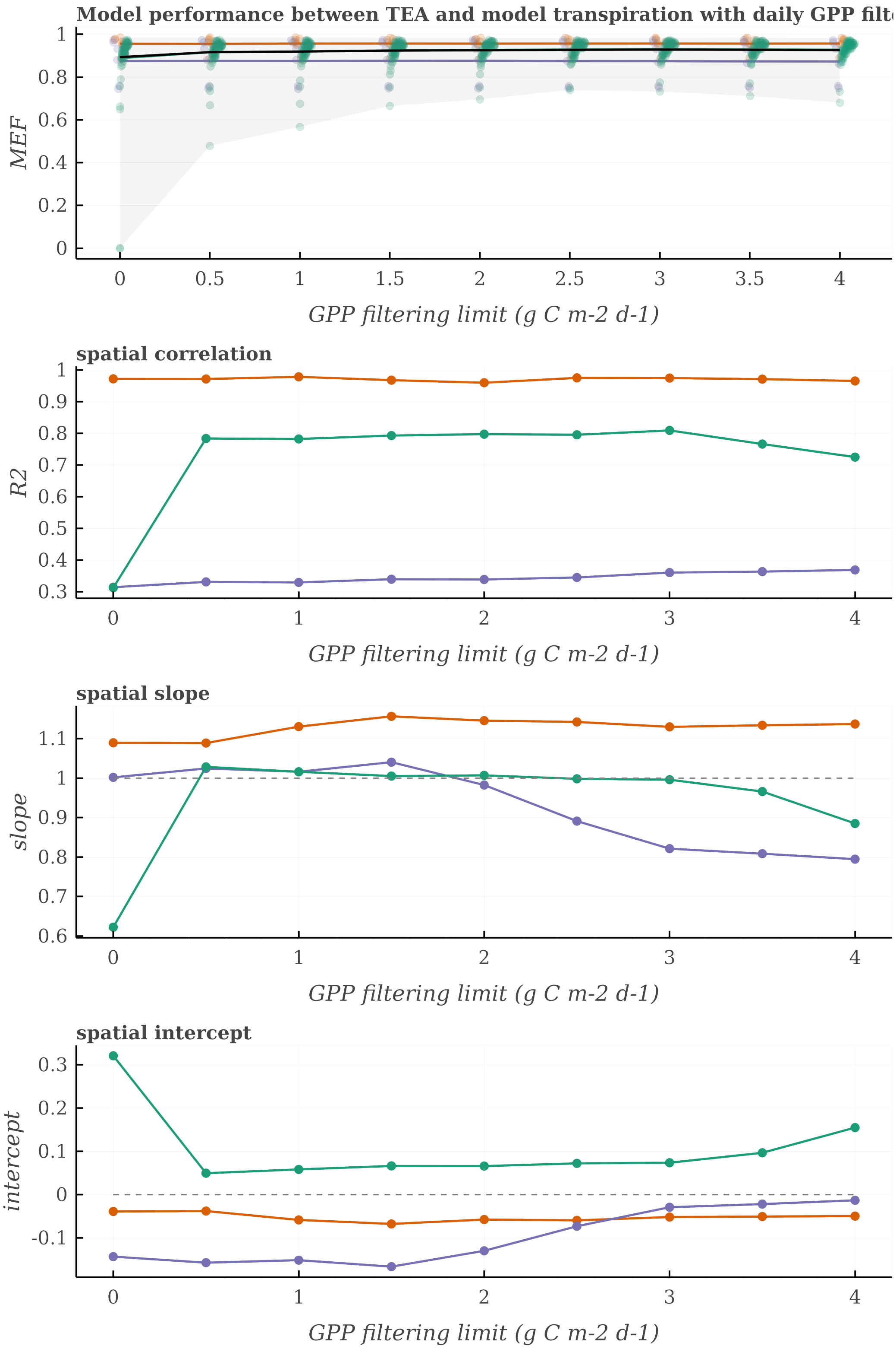

Figure S1. Sensitivity of daily GPP threshold filter to modeling efficiency (MEF) and spatial correlation, slope, and intercept. Filter limits above 0 show significant improvements, particularly with JSBACH spatial performance. Higher limits show degraded performance with MuSICA spatial performance, likely due to decreased training set data in dry sites under high GPP limits.

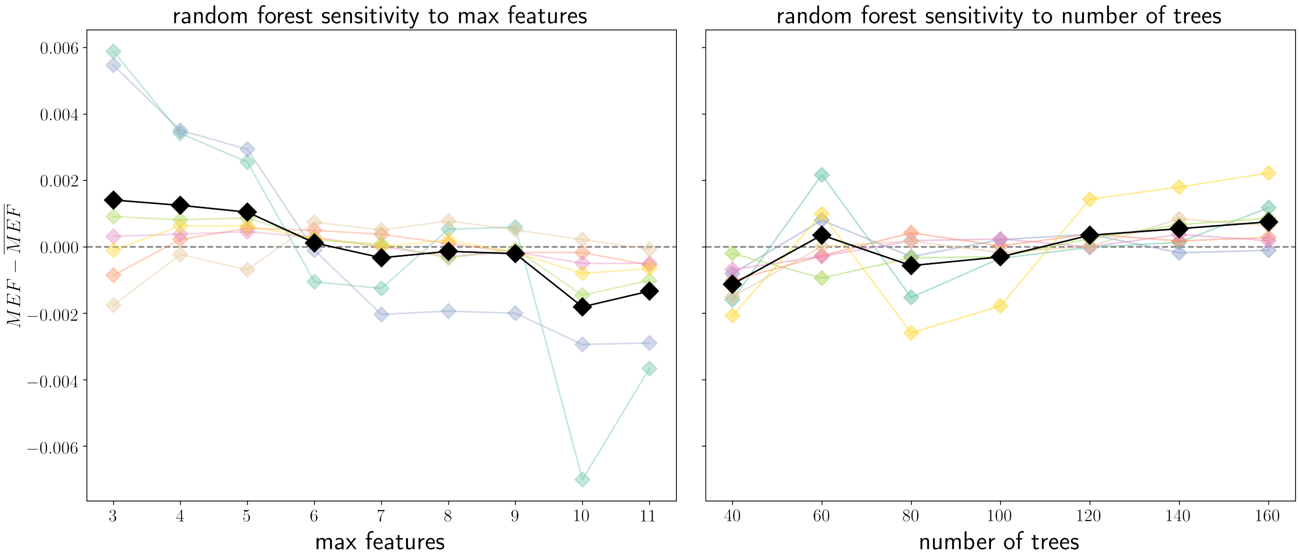

Figure S2. A sensitivity analysis of the number of trees and max number of feature parameters on modeling efficiency (MEF) for seven site from MuSICA model output. The mean MEF across all max feature experiments for each site was subtracted to make sites comparable. Colored lines represent different sites, while the black line represents the mean from the seven sites. Though the number of max features used had little effect, a value of four showed the lowest variability and corresponds to the standard practice of using one third of the number of features. MEF decreased when the number of trees went below 100.

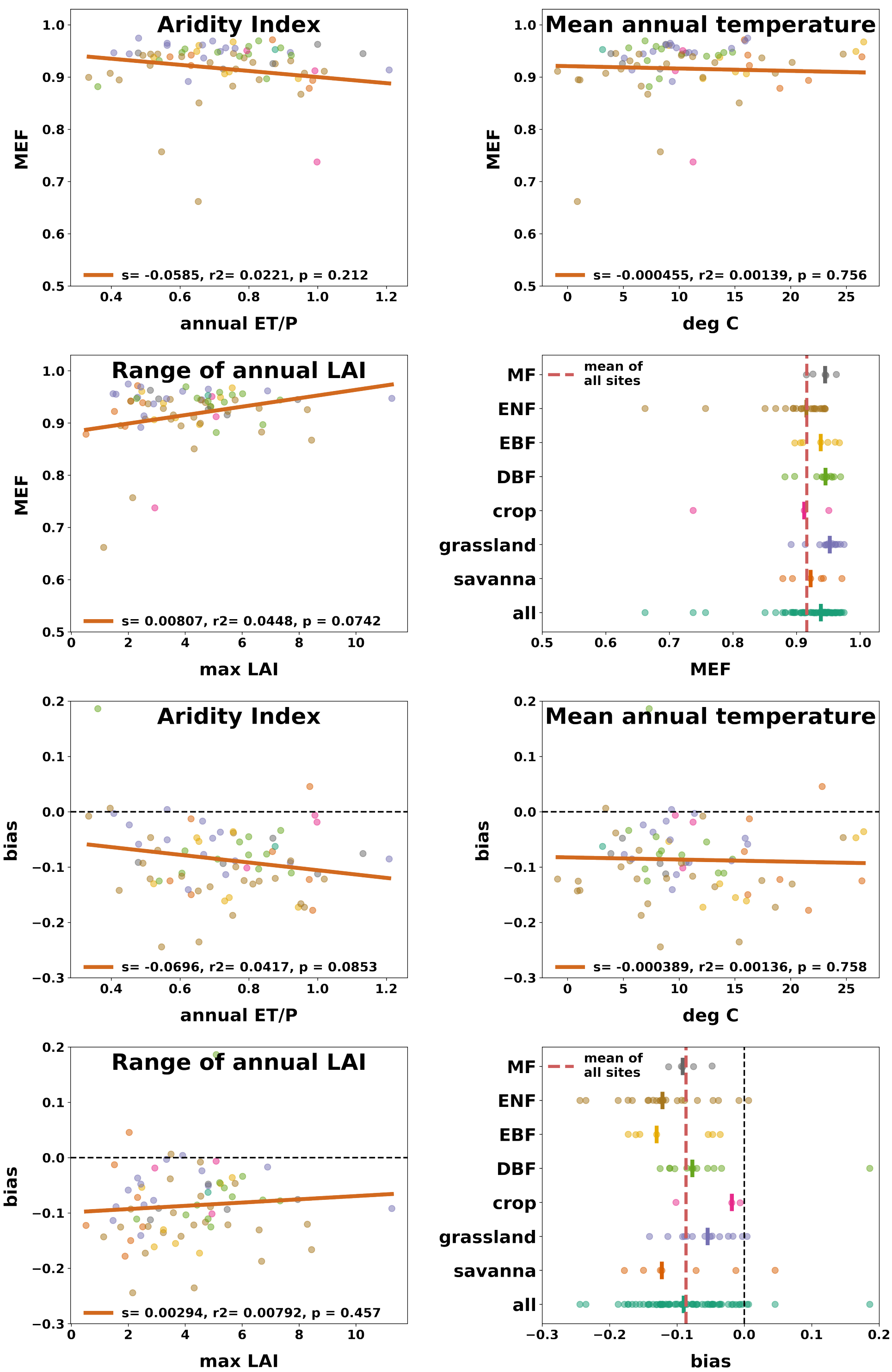

Figure S3. Modeling efficiency (MEF) and relative bias response to plant function and climate across 72 site-runs from JSBACH. Climate variables (aridity index and mean annual temperature) showed no significant effect on MEF. Vegetation parameters show a slight effect, but is primarily driven by three sites.

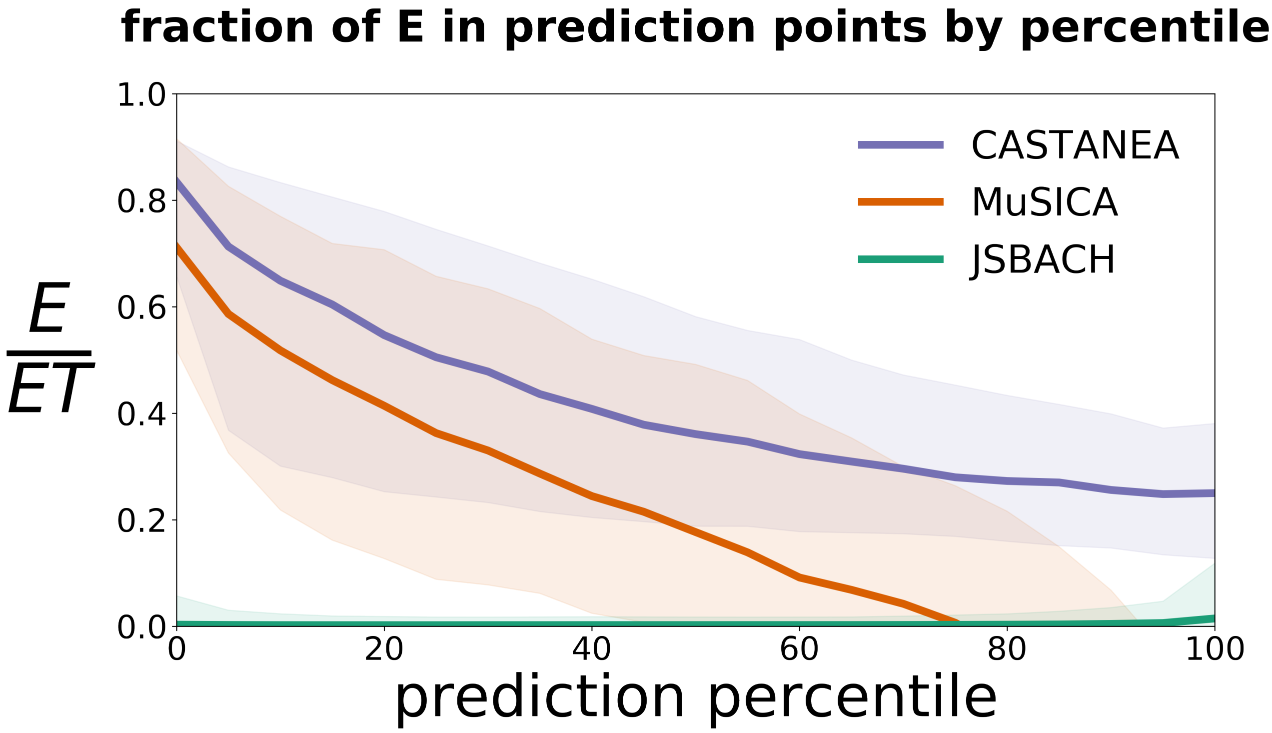

Figure S4. Relationship between the fraction of evaporation (directly from the models) the prediction percentiles used. E/ET is calculated by retrieving the specific half hours corresponding to the predicted WUE as output from the Random Forest from the training dataset. The E/ET is then calculated in these corresponding half hours, giving the actual contamination from the predictive points. Note that this process is done for each half hour of the dataset, giving a distribution of E/ET values which is dependent on the percentile used in prediction (each percentile gives a different corresponding half hour from the training dataset). Lines represent the median across all model runs, with shading representing the interquartile range.

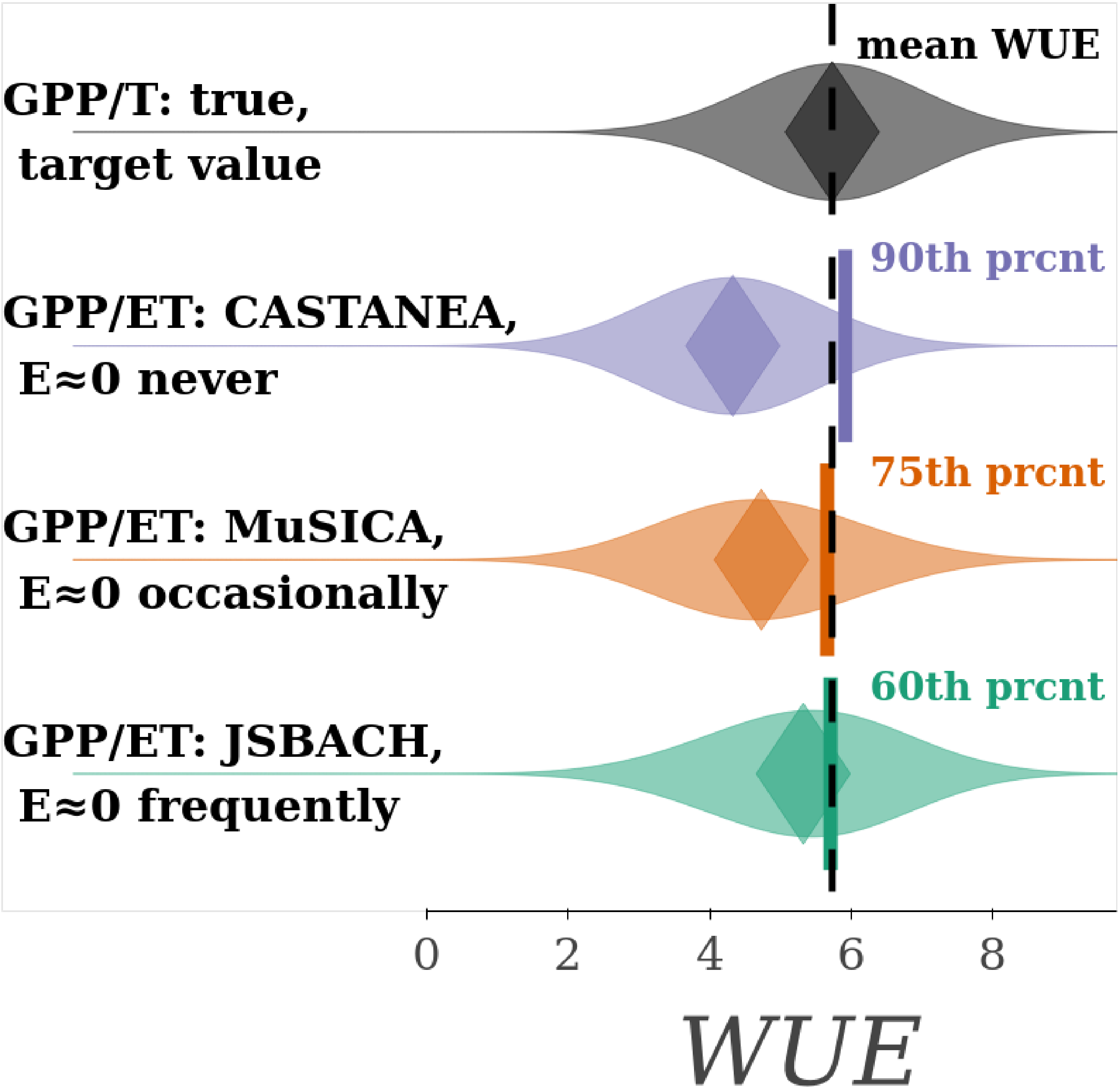

Figure S5. Theoretical diagram showing how the average true WUE (GPP/T, gray lines) can be estimated using the prediction percentile of eWUE (GPP/ET) under three scenarios: "CASTANEA" where evaporation is always 30% of ET, "MuSICA" where evaporation is high but does reach 0 at some points, and "JSBACH" where evaporation is usually 0. In all three scenarios, including when evaporation is never 0, the mean WUE can be approximated using different percentiles, with the caveat that highest percentiles of WUE cannot be estimated in the "CASTANEA" scenario and would then be truncated.

Table S6. Table of sites used: FileS6.pdf