POTSDAM INSTITUTE FOR
CLIMATE IMPACT RESEARCH

Leibniz
Association

Potsdam Institute for Climate Impact Research

PIK

ISIMIP3 round - global gridded crop model validation

Cornelia Auer, Kobe de Maeyer

22.04.2024 ISIMIP conference

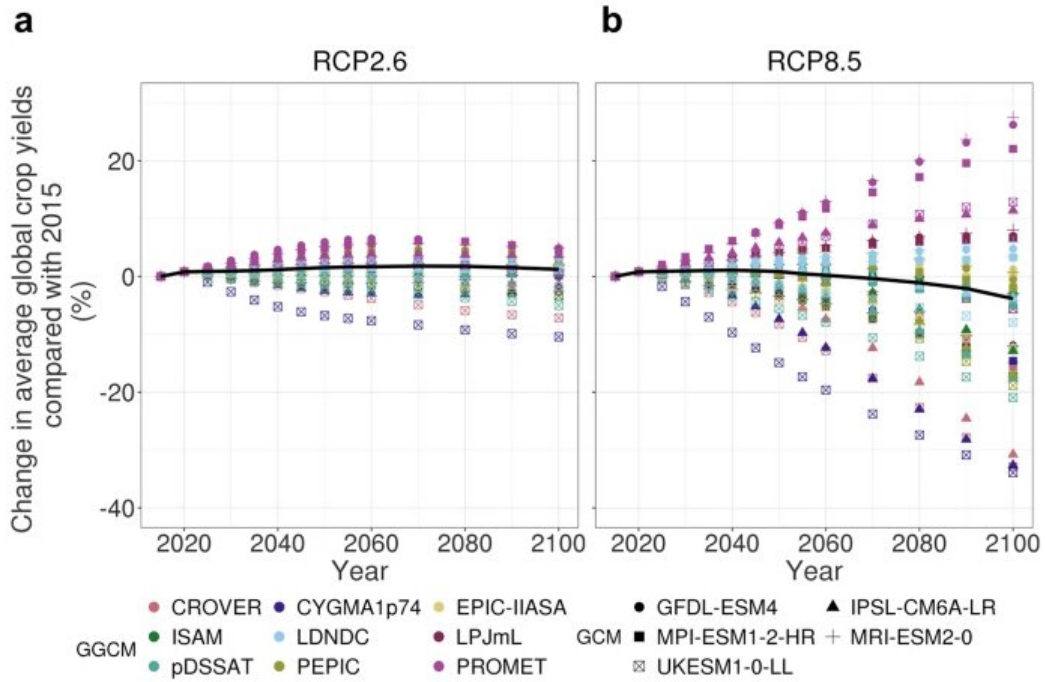
Leibniz Association

Research questions and approach

Why

› Strong divergence
› Hard to communicate to policy and decision makers
› Impression from median
  → no problem, nothing much to do

Fig. 1: Global climate change impacts on crop yields under two different emission scenarios.



Source: Bacca et al. 2023

Research questions

- *1. Which models is the best for which region for which crop?*
- *2. Can we build an Ensemble projector from this analysis?*
  - *weighted model ensemble*

- *3. "What are the regional patterns in model performance and what model characteristics explain these patterns?"*

Approach - Data

› *ISIMIP3a historic runs (11 models)*
  › full irrigation, no irrigation combined according to HYDE 3.2 (Klein et al), harmonized by Hurtt et al.

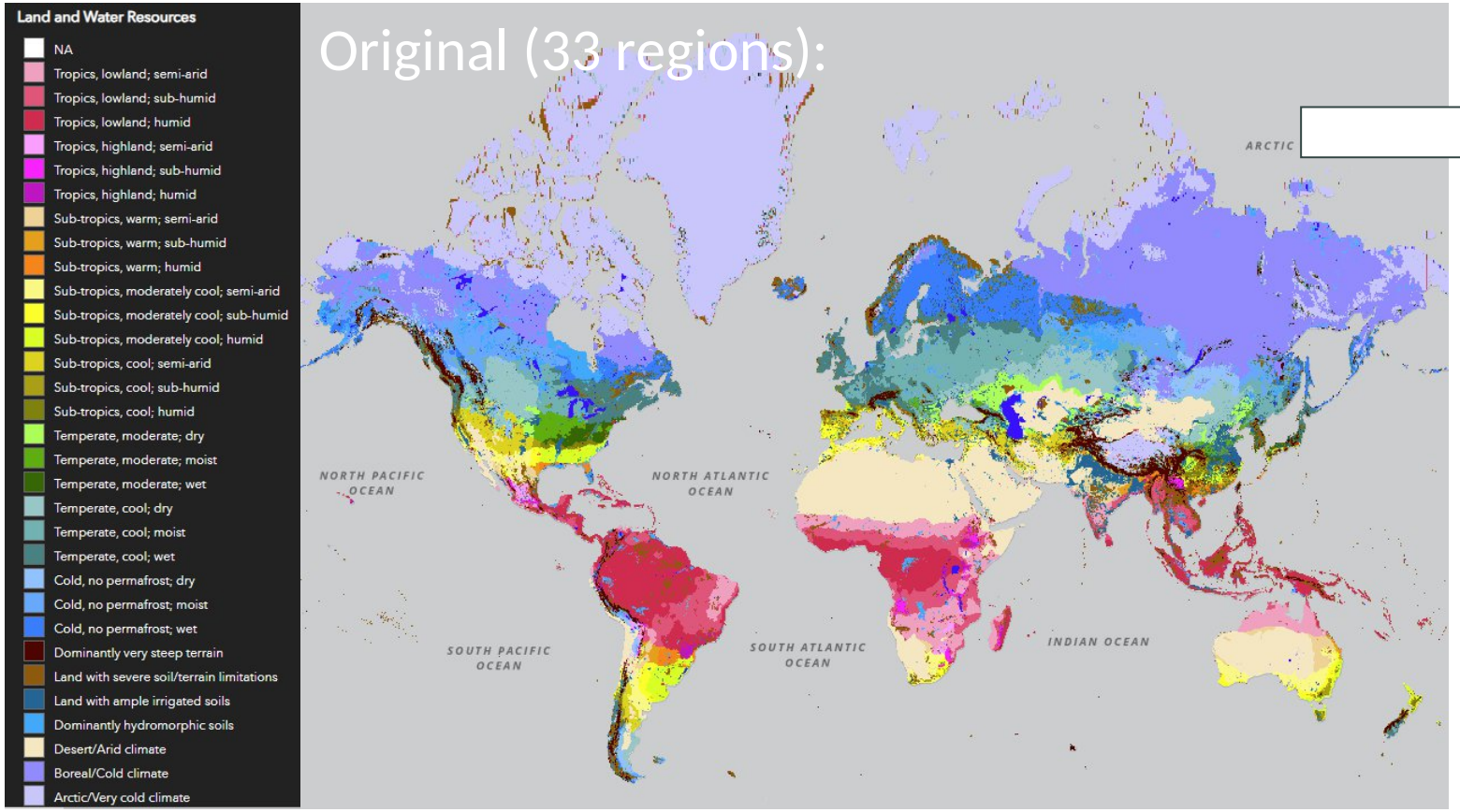
› *Benchmark data:*
Iizumi, Toshichika (2019): Global dataset of historical yields v1.2 and v1.3

0.5-degree grid-cell annual yield estimates
maize, rice, wheat and soybean
period 1981-2016

› *Preprocessing*
Quadratic detrending of yields
No yearly shift yet

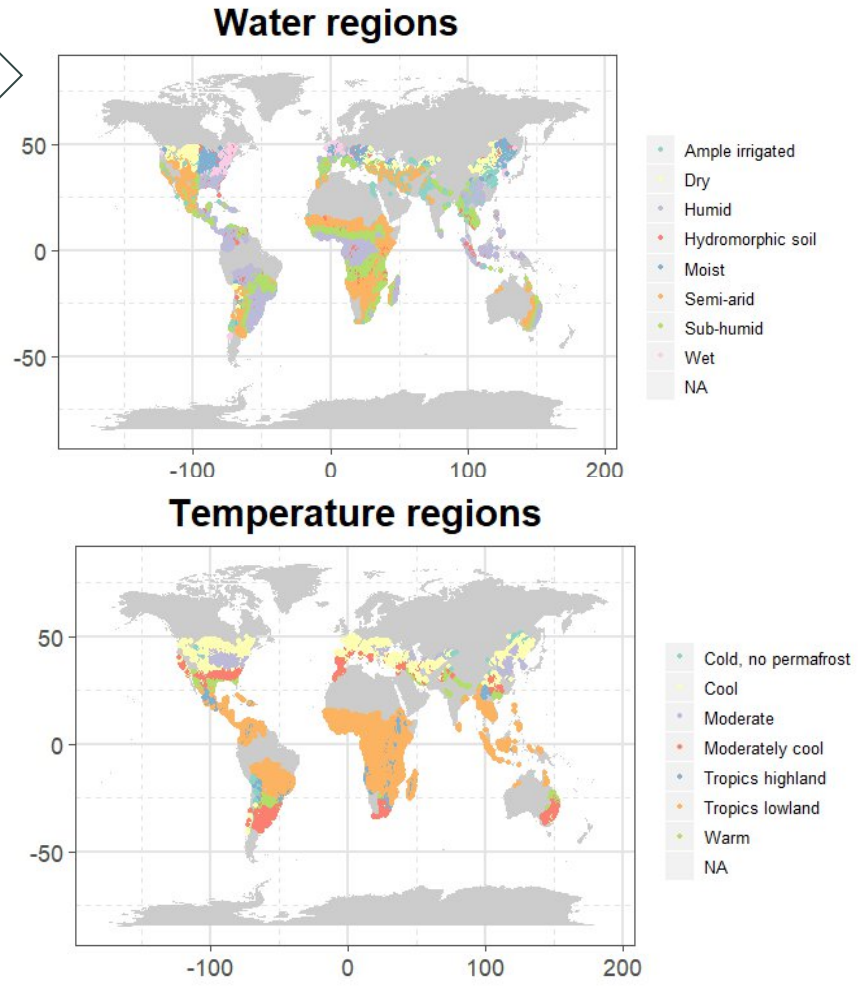
Approach - regions

› Analysis clustered to FAO GAEZ regions, currently to hydrological OR temperature characteristics



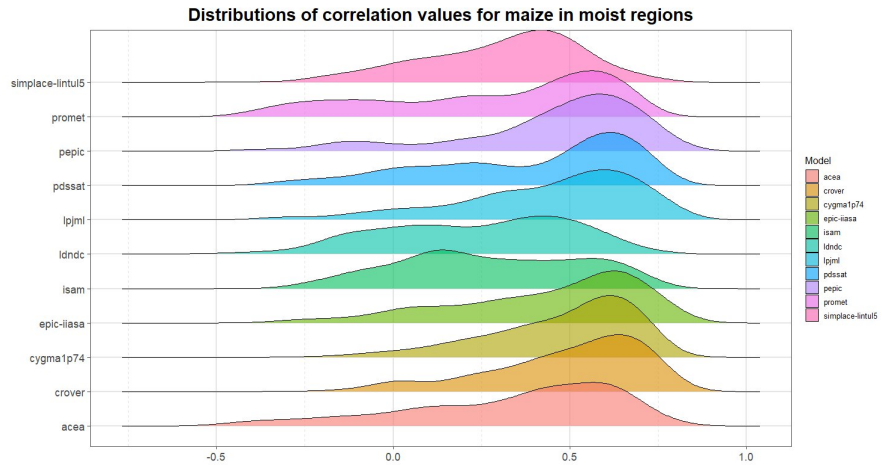
source: GAEZ portal

Our aggregation (7-8 regions):



Approach - Statistical analysis: (Bayesian) regression

- Different performance measures:
  - Pearson correlation
  - Root Mean Square Error
  - Median Error
  - [opt]: Spearman Correlation



- Bayesian Model:
  Response variable: performance measure (currently Pearson Correlation)
  Predictor variables: 88 = 11 * 8 model - region groups
  -> why Bayesian: results with uncertainty, insights in distribution

  Not reflected for now: spatial spillover effects, unequal variances, tails of the distributions

Bayesian model

$$P(\beta|y,X) \propto P(y|X,\beta) \cdot P(\beta)$$

$$P(y|X,\beta) \sim N(X\beta, \sigma^2 I)$$

$$P(\beta) \sim uniform(-1,1)$$

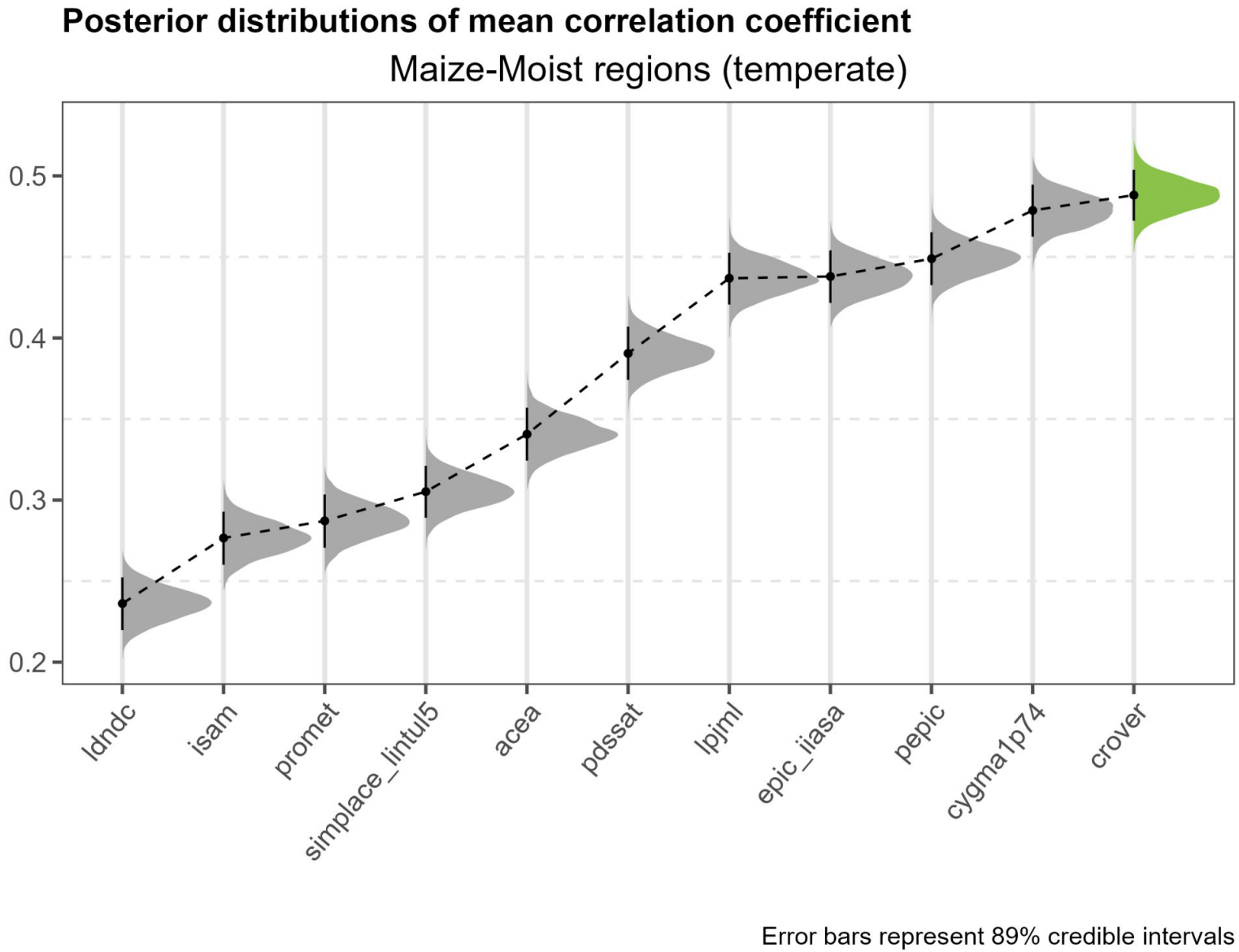
- y is the data of performance measures (e.g. correlation values).
- X is a matrix of categorical predictors with 88 different region-model indicator (0 or 1) variables.
- β is the vector of coefficients associated with each combination of region and model.
- $P(\beta|y,X)$ is the posterior distribution of the coefficients β given the data y and the categorical predictors X.
- $P(y|X,\beta)$ is the likelihood function following a normal distribution with a mean representing the predicted values of y based on the predictors X and the coefficients β and a covariance structure $\sigma^2 I$ representing uncorrelated errors with mean zero and variance σ^2.
- $P(\beta)$ is the flat, uninformative prior distribution for the β coefficients

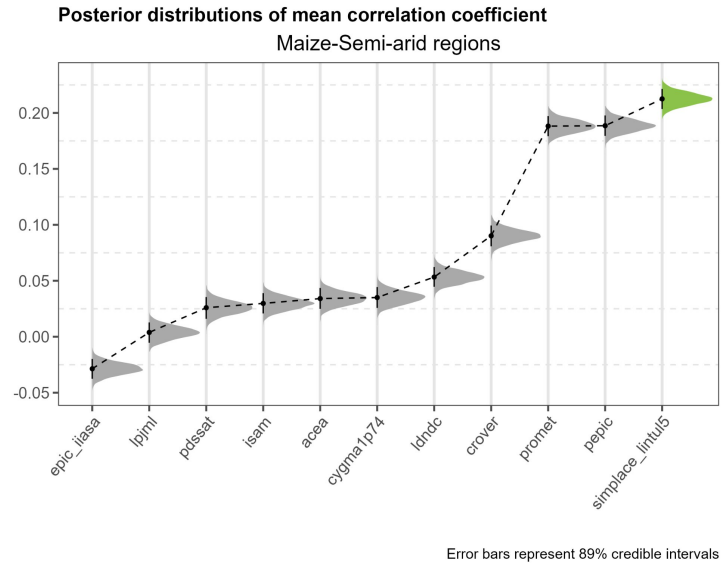
More intuitive representation of model

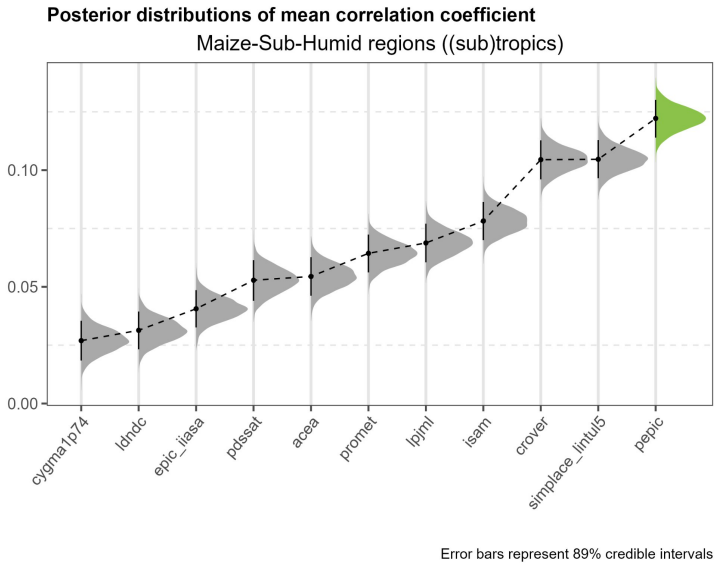
$$y = X\beta + \varepsilon$$

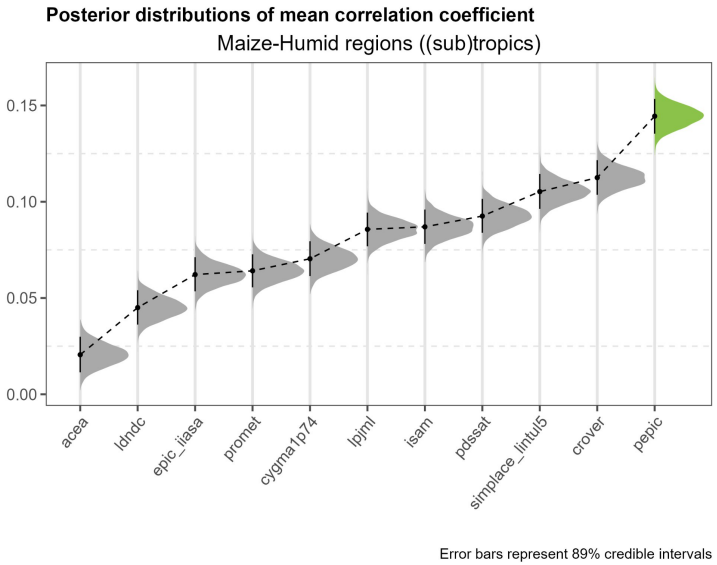
- y is the vector of estimated model performance values (e.g. correlation).
- X is a matrix of categorical predictors with 88 different region-model indicator (0 or 1) variables.
- β is the vector of coefficients associated with each combination of region and model.
- ε is the vector of error terms which are normally distributed with mean zero and variance σ^2.

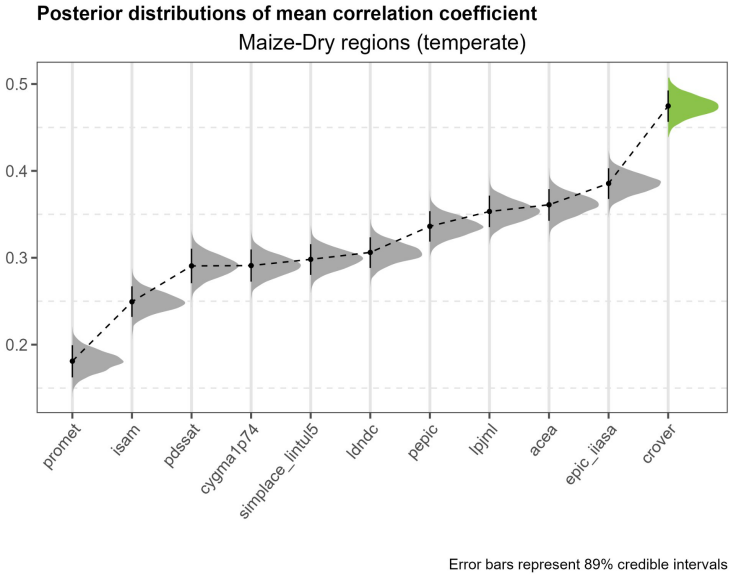
Results

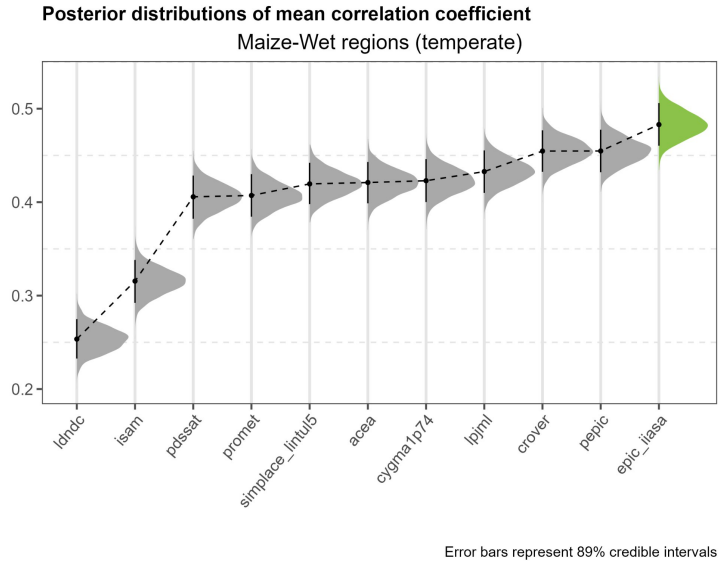


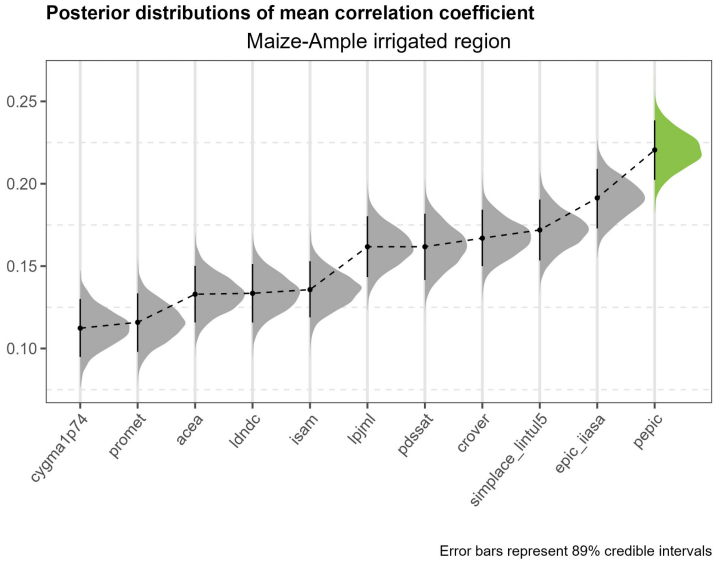


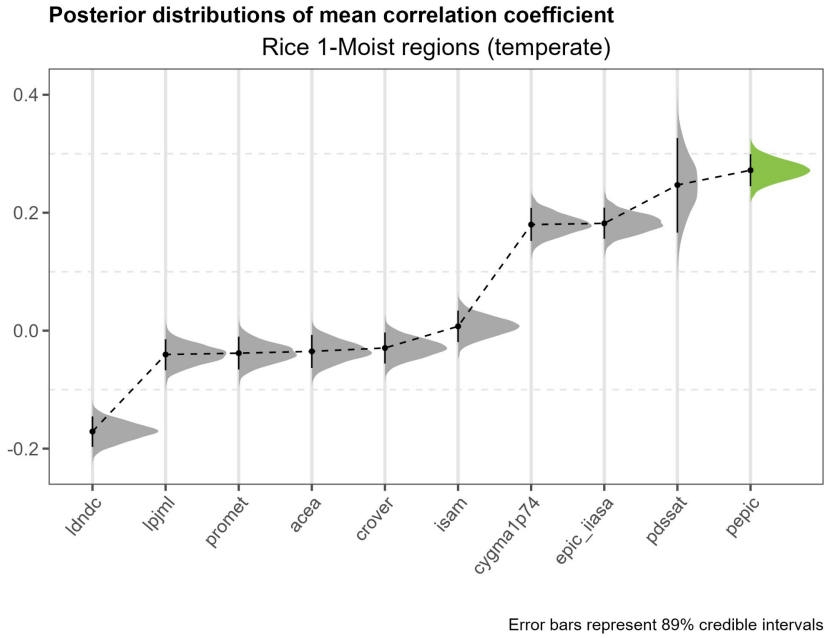


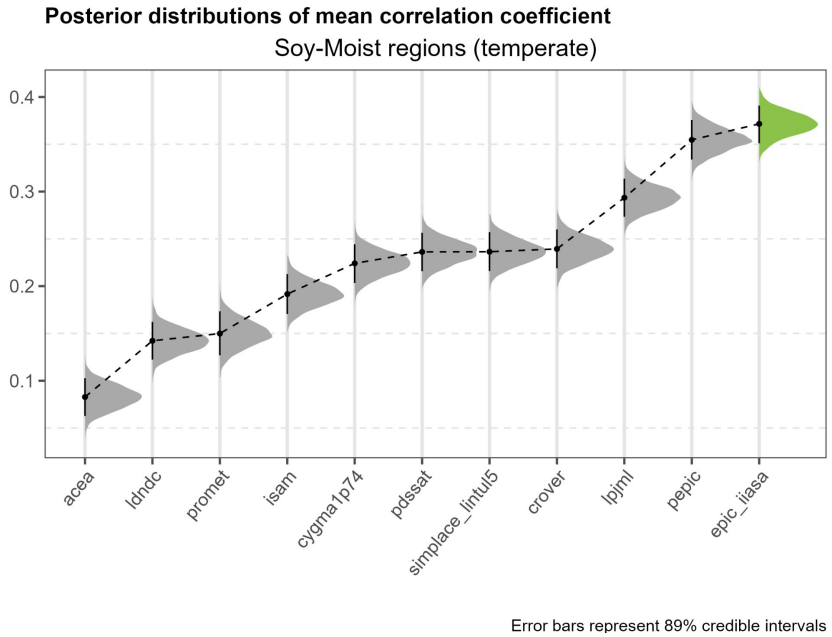


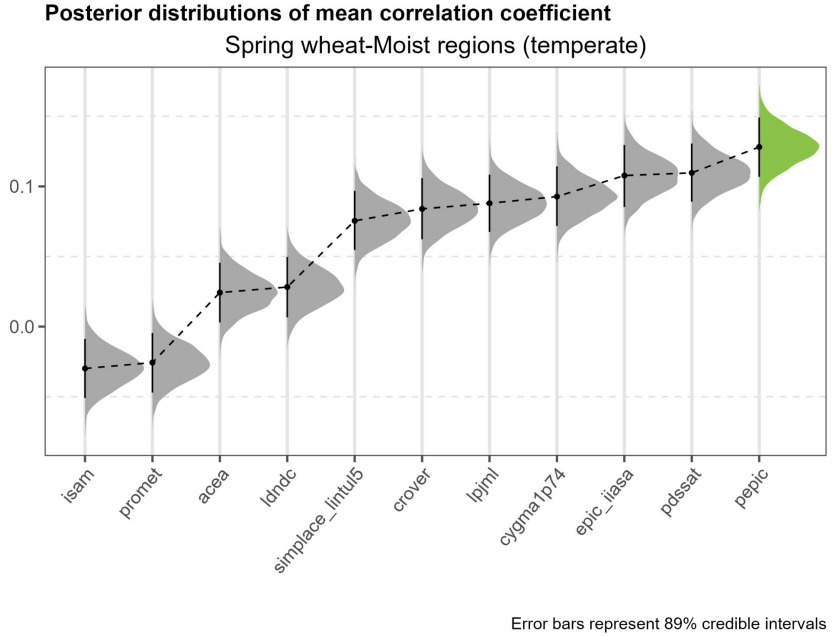


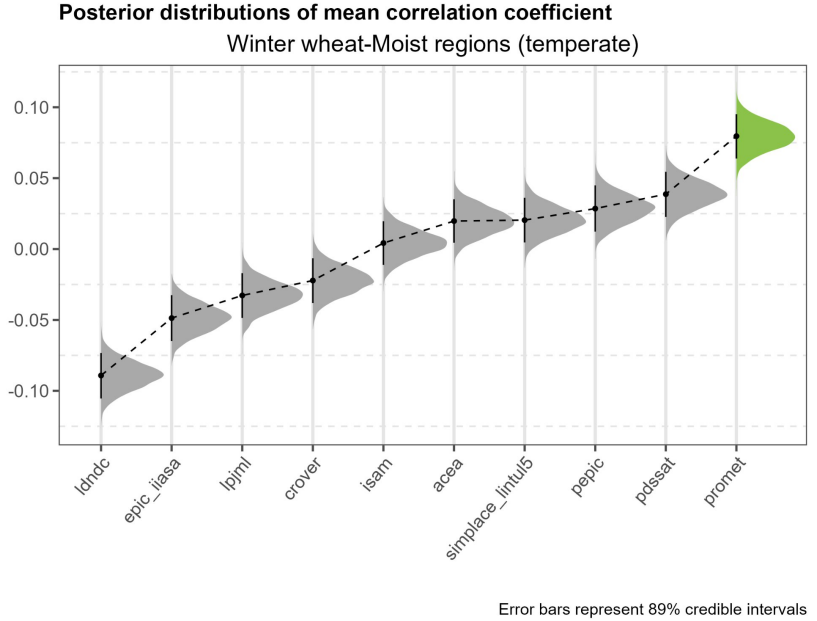


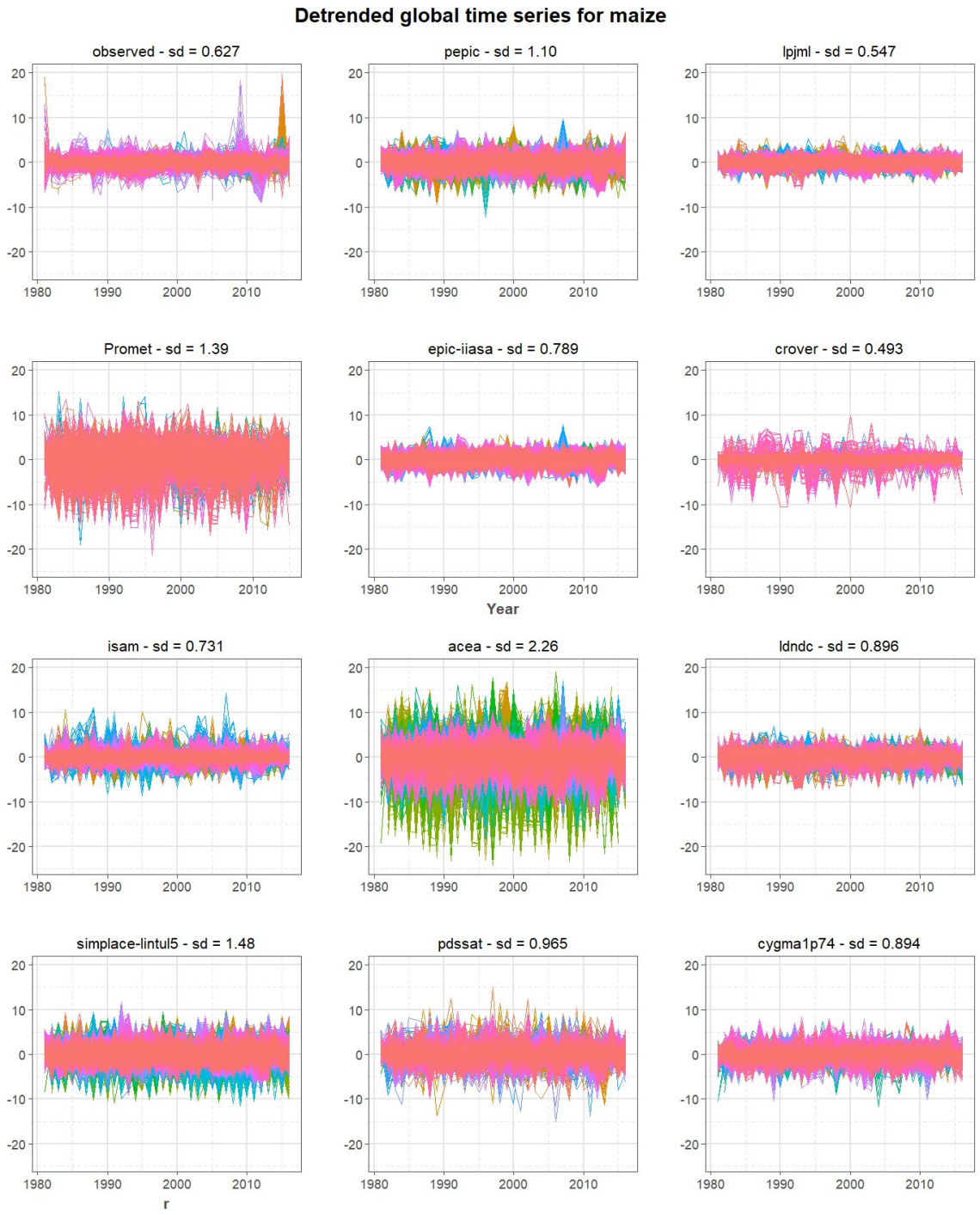










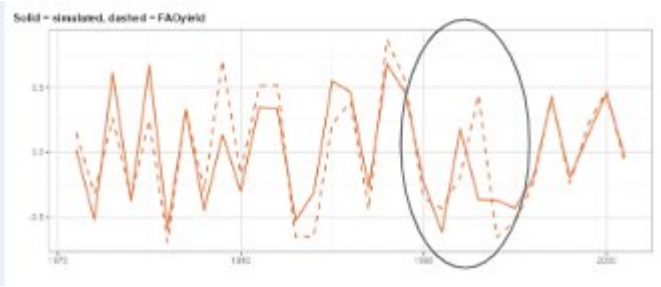


Variance differs strongly
→ filtering for climate extremes

GGCM

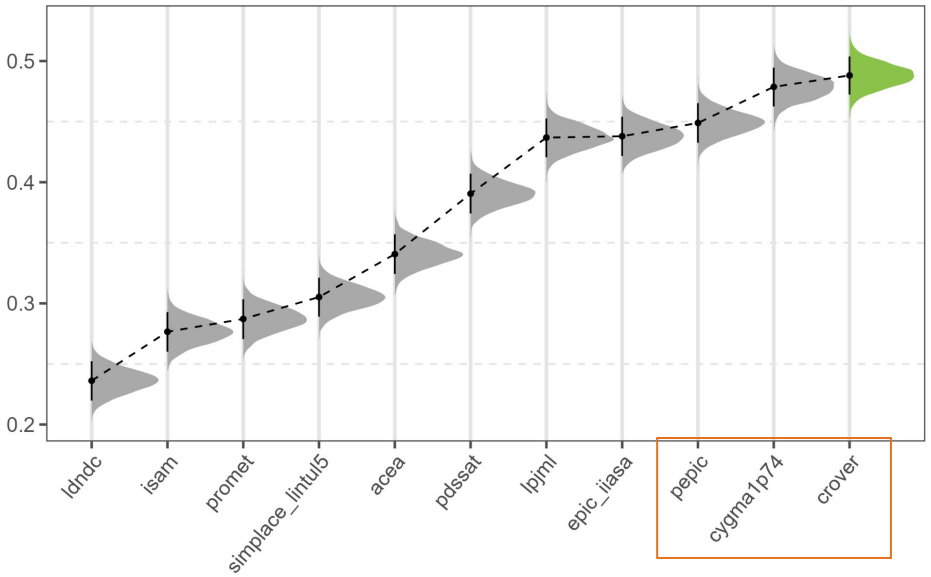
Right performance measure

› RMSE, Median error (over- and underestimation)
RMSE tricky because it rewards models with little reactivity (SD)
› Pearson correlation / Spearman correlation



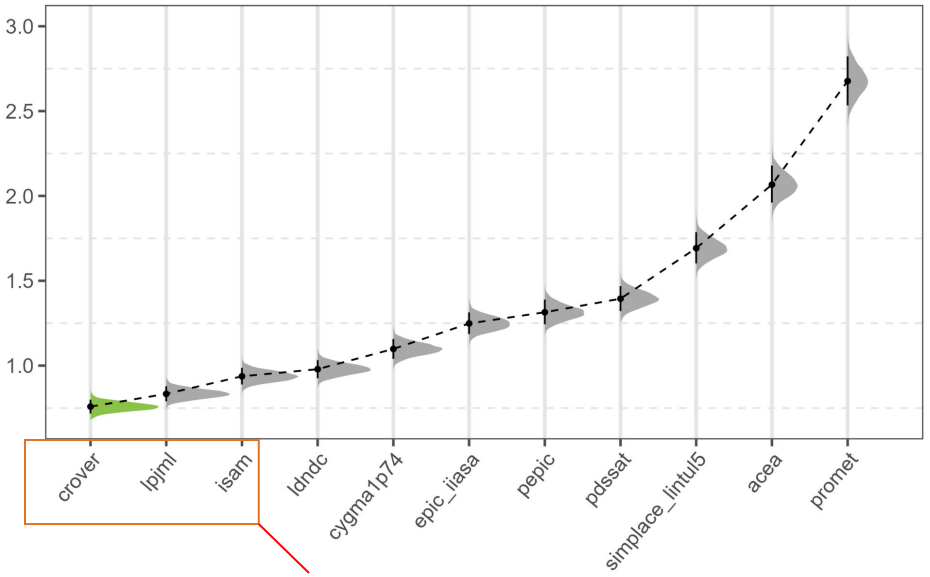
Posterior distributions of mean correlation coefficient

Maize-Moist regions (temperate)



Posterior distribution of mean RMSE

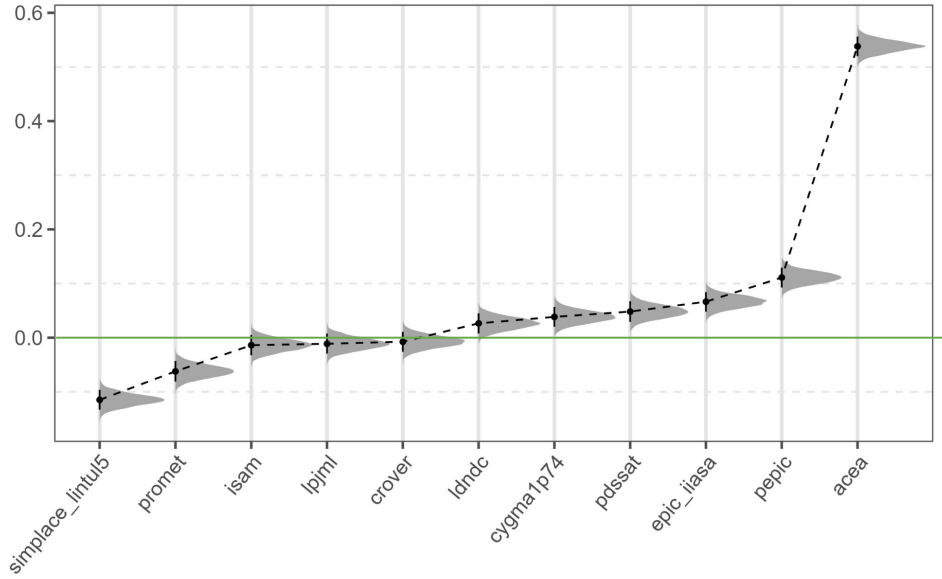
Maize-Moist regions (temperate)



Best are those with the lowest standard deviation

Posterior distribution of mean bias

Maize-Moist regions (temperate)



overestimation

underestimation

Positive = overestimation; negative = underestimation. Error bars represent 89% credible intervals

Key model characteristics

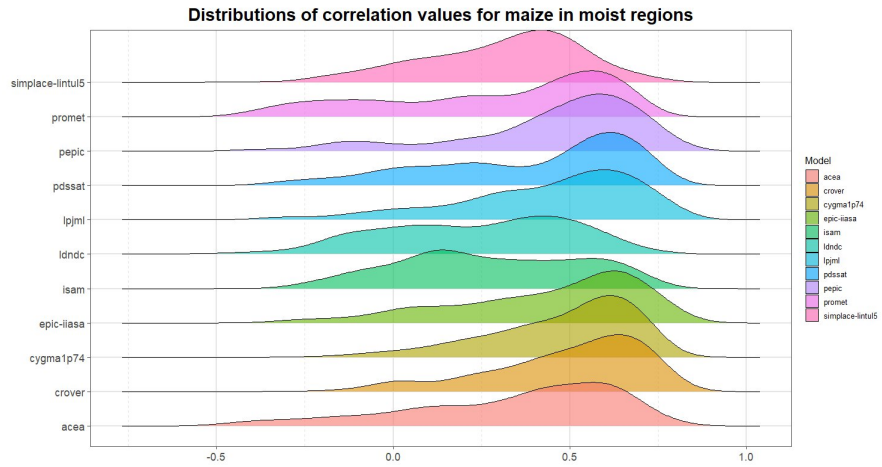
Regression of model characteristics - based on common table, THX!

Impact Model	Leaf area development	Light interception	Light utilization	Yield formation	Crop phenology	Root distribution over depth	Stresses involved
	(P) phenology, (S) stresses, (B) biomass, (D) density	NA, Beer type (Beer according to Monsi and Saeki (1953) or Lamber Beer), Foliage structure, Inclination angle (ellipsoidal inclination angle distribution model by Campbell)	NA, Photosynthesis (P), Radiation use efficiency (RUE)	Harvest index (HI), dynamic harvest index (DHI), dynamic allocation (DA)	Single stage (SST) vs Multiple stage (MST), temperature driven (TD) vs multiple drivers (MD) combined category: e.g. SST x MD	NA, phenology (P), soil depth (sigmoidal, exponential, etc) (SD), linear(L), stresses (S), biomass (B), temperature (T), water use (W), soil strength (SStr)	Heat stress (H), cold stress (C), water deficit stress (WD), aeration (waterlogging) stress (WL), salinity (SA), aluminium toxicity (AT), nitrogen deficit (ND)
Dear all, thanks for contributing -> please	Key model processes	Key model processes	Key model processes	Key model processes	Key model processes	Key model processes	Key model processes
ACEA (Agriculture, ISIMIP3b)	(P), (S), (D)	NA - Not applicable	NA - Not applicable	Dynamic Harvest index (DHI) - Estimated based on accumulated biomass and final harvest index. Biomass accumulation is calculated daily as a function of transpiration and water productivity. Harvest index reflects the harvestable fraction of biomass and is subject to water and heat stresses.	MST x TD Plant development includes several growing stages defined by the growing degree day thresholds.	(L) For annual crops, it increases linearly from the planting date (min depth) until a defined growing degree day threshold (max). The growth can be limited by water stress or water logging.	H, C, WD, WL Heat stress, cold stress, water deficit stress, aeration (waterlogging) stress
ACEA (Agriculture, ISIMIP3a)	(P), (S), (D) Expressed in dynamic green canopy cover ranging from 0 to 100% depending on phenology, heat & water stress, and planting density.	NA - Not applicable	NA - Not applicable	Dynamic Harvest index (DHI), Estimated based on accumulated biomass and final harvest index. Biomass accumulation is calculated daily as a function of transpiration and water productivity. Harvest index reflects the harvestable fraction of biomass and is subject to water and heat stresses.	MST x TD Plant development includes several growing stages defined by the growing degree day thresholds.	(L) For annual crops, it increases linearly from the planting date (min depth) until a defined growing degree day threshold (max). The growth can be limited by water stress or water logging.	H, C, WD, WL Heat stress, cold stress, water deficit stress, aeration (waterlogging) stress
CROVER (Agriculture, ISIMIP3b)	(B) biomass Leaf area index on a given day is calculated based on leaf biomass.	Beer type - Beer according to Monsi and Saeki (1953)	Photosynthesis (P) - Photosynthesis routine is based on the approaches of Farquar.	Harvest index (HI), Estimated based on accumulated biomass and harvest index.	SST x TD Crop development is modelled as a fraction of the accumulated growing degree days relative to the crop thermal requirements.	(L) Dynamic	H, WD Heat stress, cold stress, water deficit stress
CROVER (Agriculture, ISIMIP3a)	(B) biomass Leaf area index on a given day is calculated based on leaf biomass.	Beer type - Beer according to Monsi and Saeki (1953)	Photosynthesis (P) - Photosynthesis routine is based on the approaches of Farquar.	Harvest index (HI), Estimated based on accumulated biomass and harvest index.	SST x TD Crop development is modelled as a fraction of the accumulated growing degree days relative to the crop thermal requirements.	Dynamic	H, C, WD Heat stress, cold stress, water deficit stress

https://docs.google.com/spreadsheets/d/14Z9ObJIPnwWCCiO2G_yaU0zr7HasspsJ/edit#gid=504398012

Approach - Statistical analysis: (Bayesian) multivariate regression

- Different performance measures:
  - Pearson correlation
  - Root Mean Square Error
  - Median Error
  - [opt]: Spearman Correlation

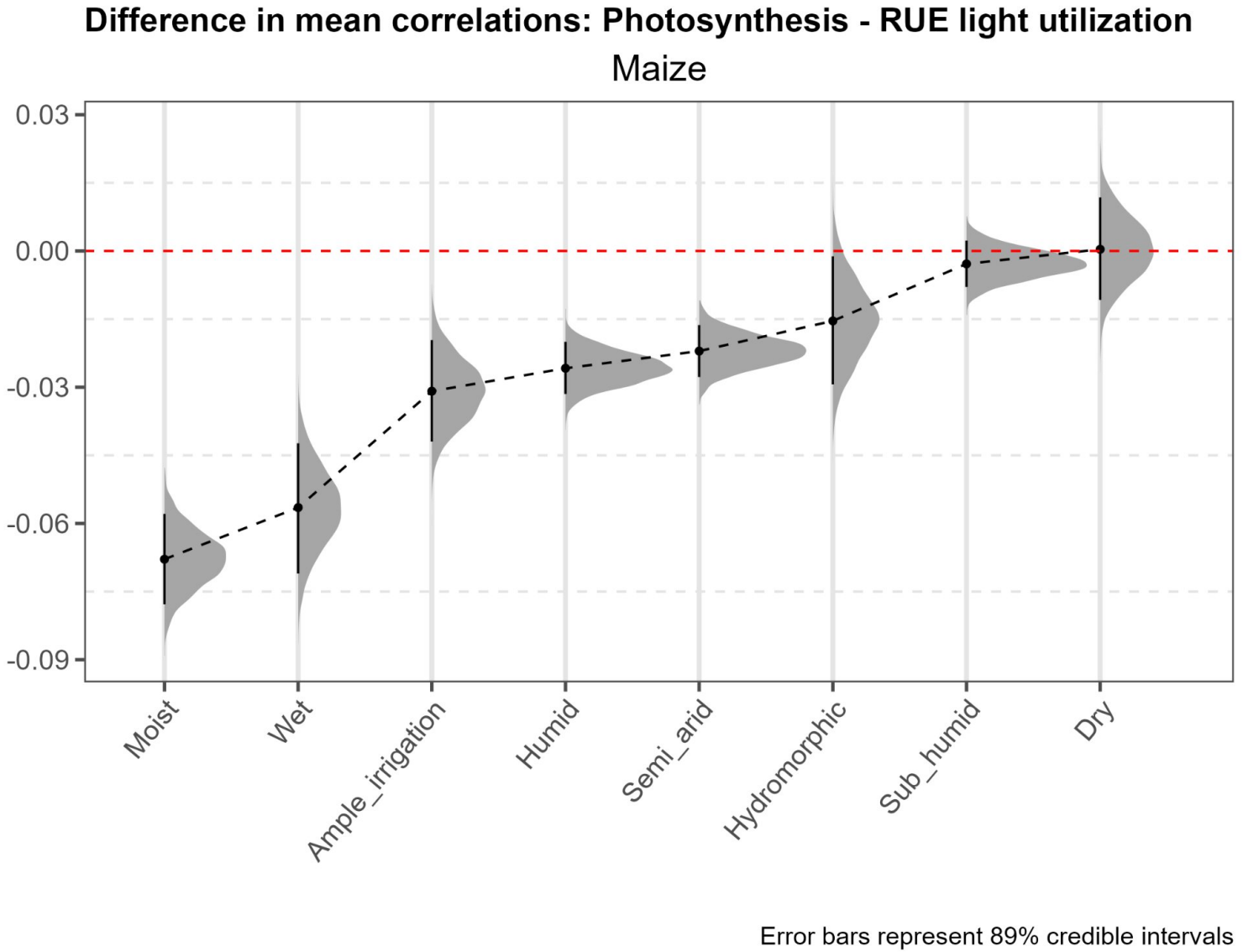


- Bayesian Model:
  Response variable: performance measure (currently Pearson Correlation)
  Predictor variables: 88 = 11 * 8 model -region groups
  16 = 2 * 8 **characteristics** – region groups
  Binary groups of groups – simple vs complex
  e.g. light utilization (RUE vs photosynthesis), water stress (water logging yes / no)

Results – light utilization

- RUE stands for Radiation-use efficiency
- Photosynthesis approach is supposed to be more dynamic
- acea was not considered here (NA in table)

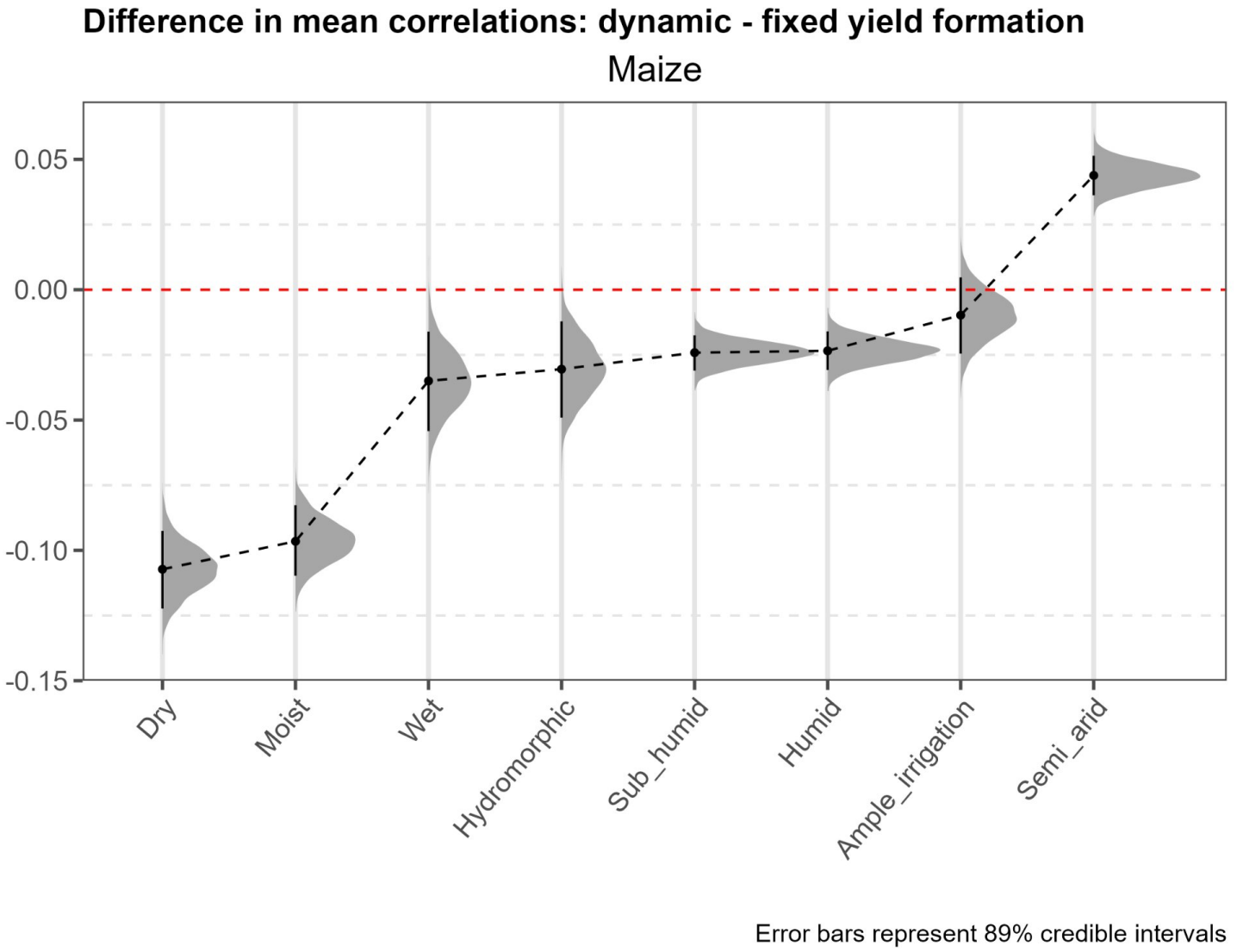
☾ On average the “simple” approaches perform better



Results – yield formation

- Fixed yields: harvest index (HI)
- Dynamic: both dynamic allocation (DA) and dynamic harvest index (DHI)

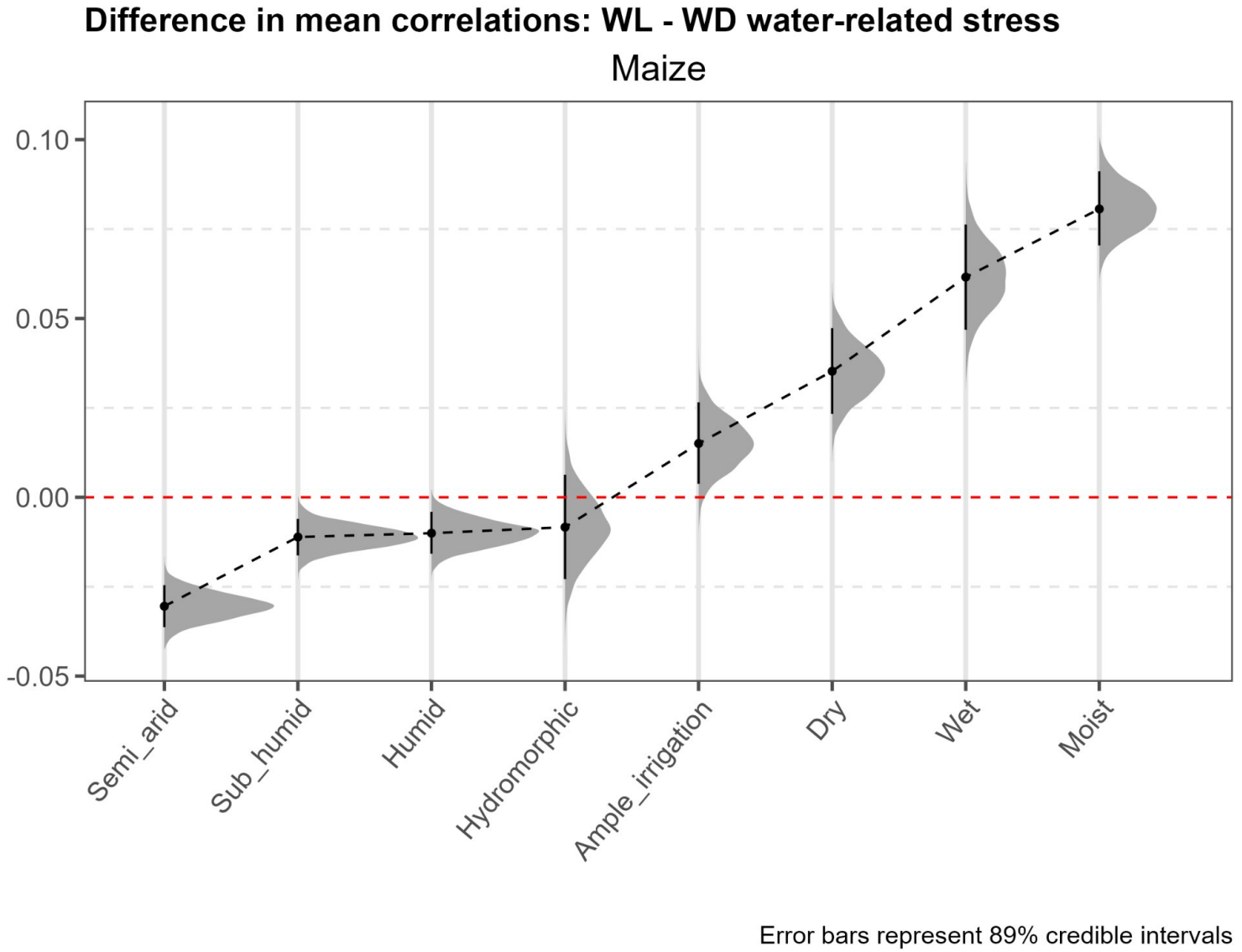
☾ On average the “simple” approaches perform better



Results – water stress

- WD means models that only have a water deficit stress
- WL are the models that also have waterlogging stress

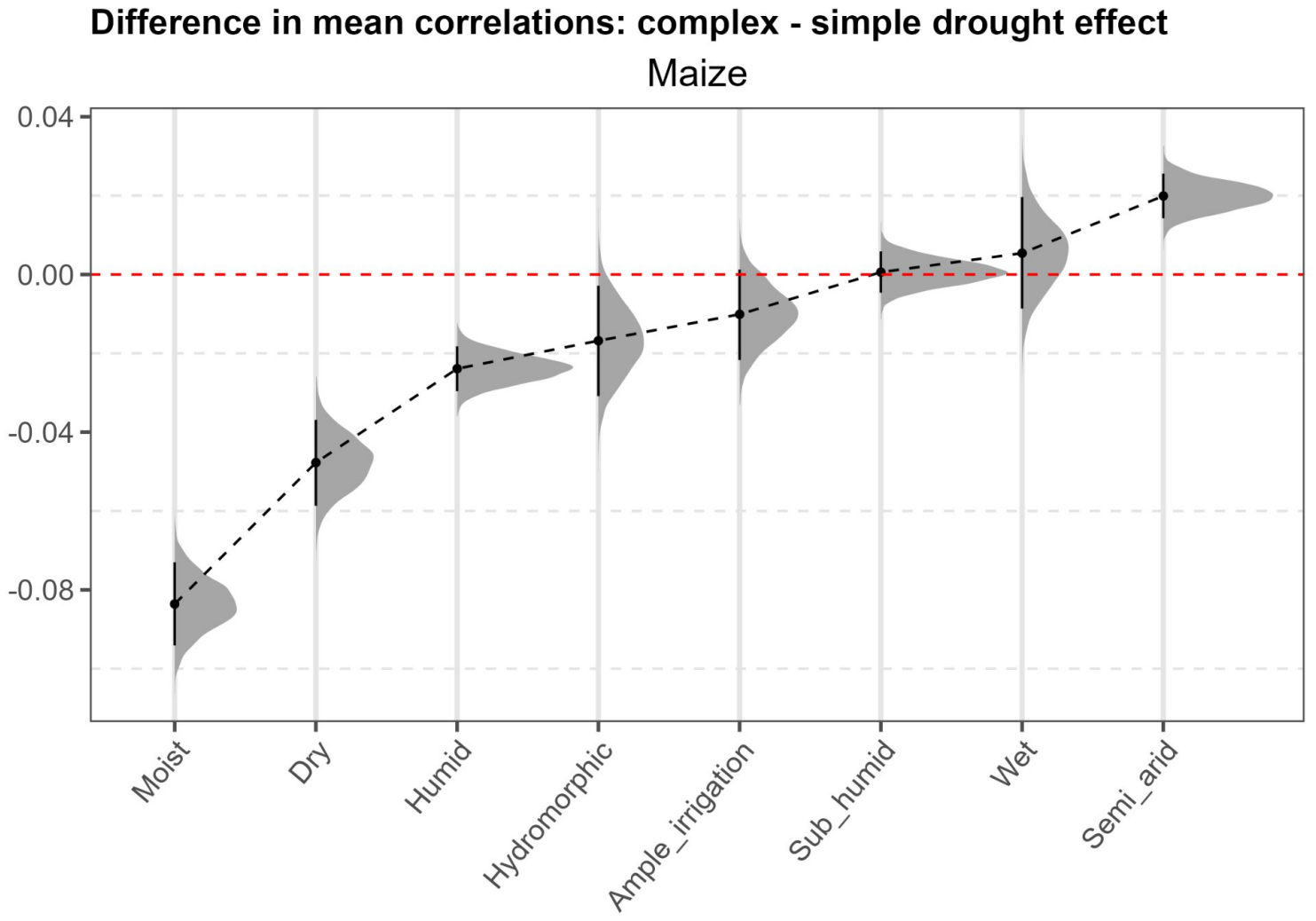
☾ Including waterlogging increases correlation for up to 1/3 (moist regions)



Results – drought effect

- “Simple” models only have or photosynthesis effect, biomass effect or both
- Complex have at least one other effect as well (canopy, leaf temperature,…)

☾ Again simple approach seems to have a slight advantage



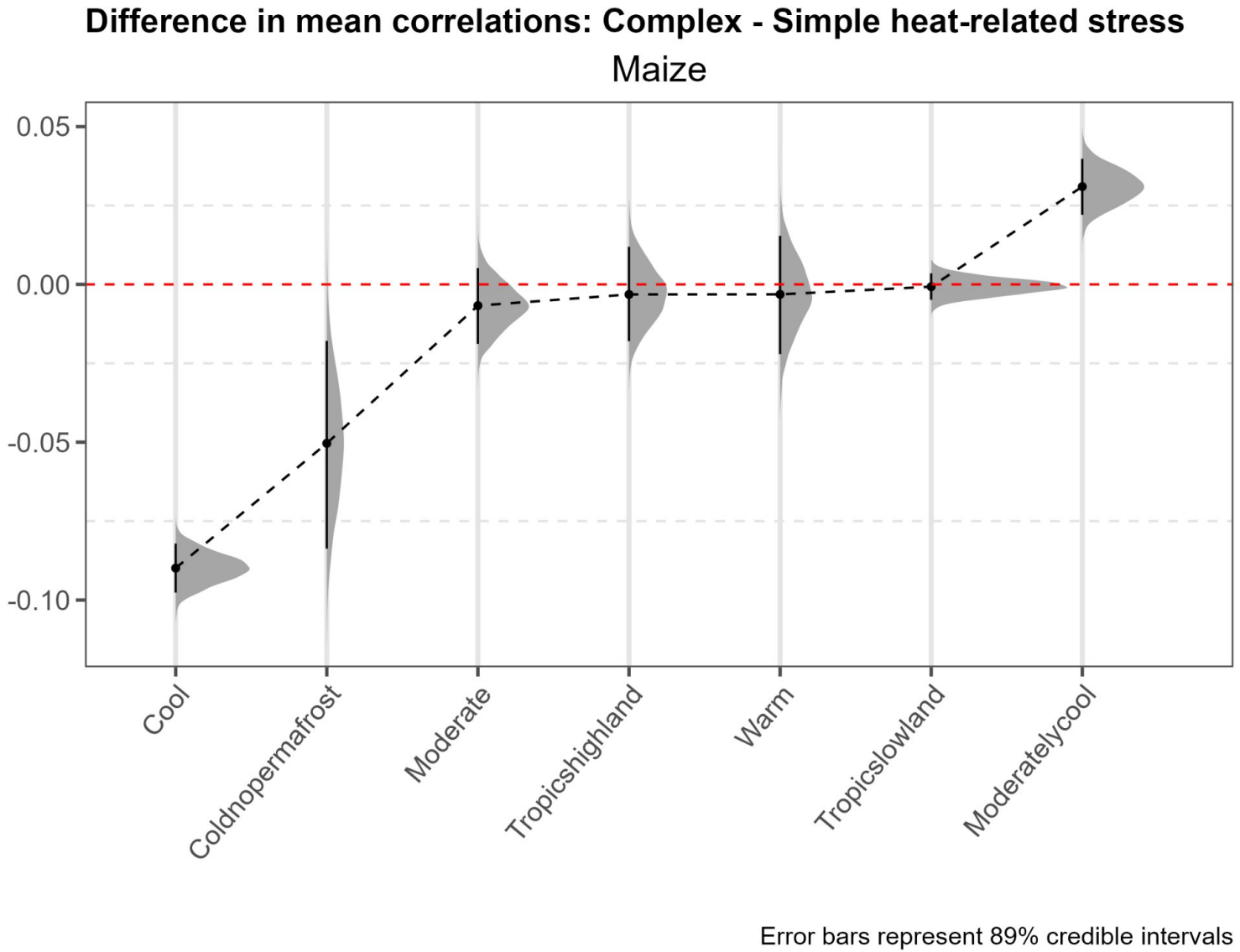
Error bars represent 89% credible intervals

Results – heat related stress

- Simple means that the models only affect how much was grown via photosynthesis and/or biomass
- Complex models have an additional effect on phenology, pollination and yield formation .

à Some results also don't seem to make too much sense?

à Your input on characteristics matters!



Next steps and future work

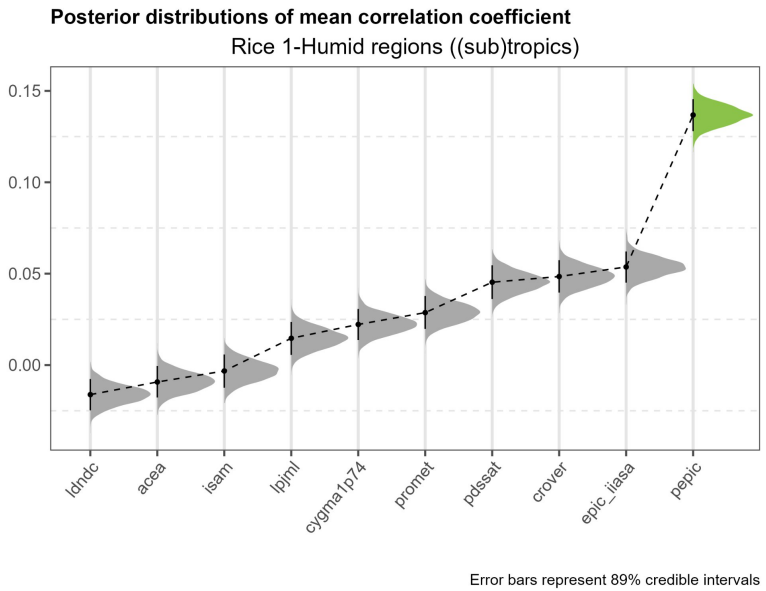
Questions to audience

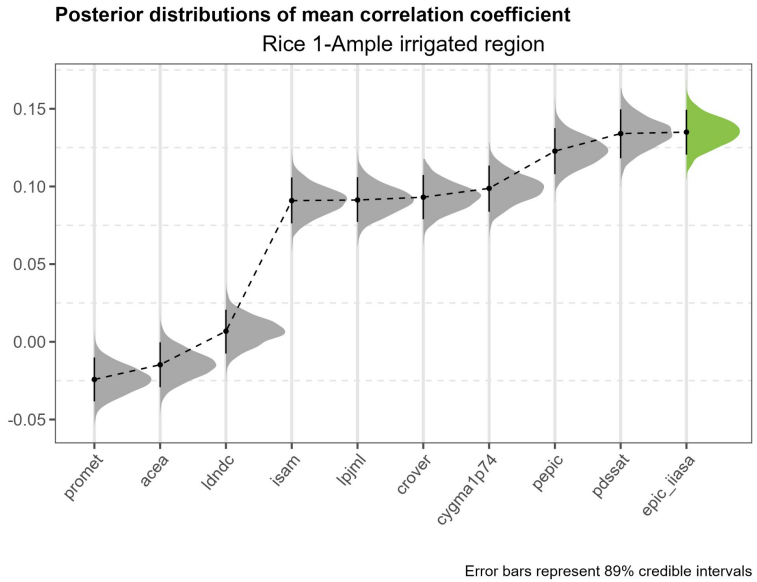
- Which key model characteristics most interesting?

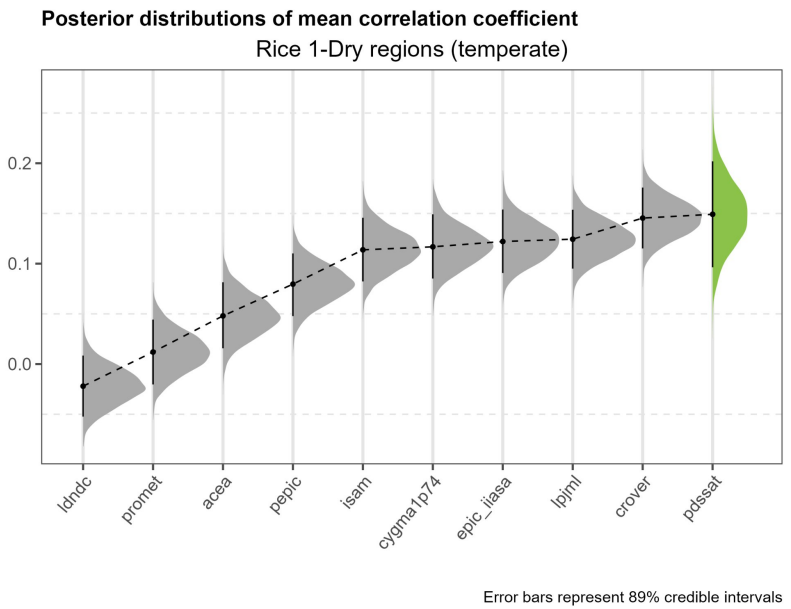
- Filter climate extremes (droughts and heatwaves)?
  - Temperature variability, count of wet days, 99th percentile of wet days,…or CDO indicators
  - How to bring things together again?

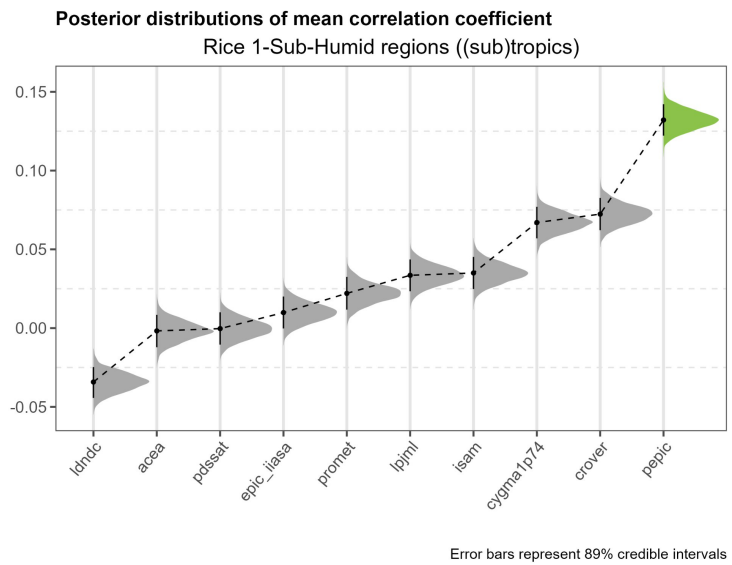
APPENDIX
Additional figures

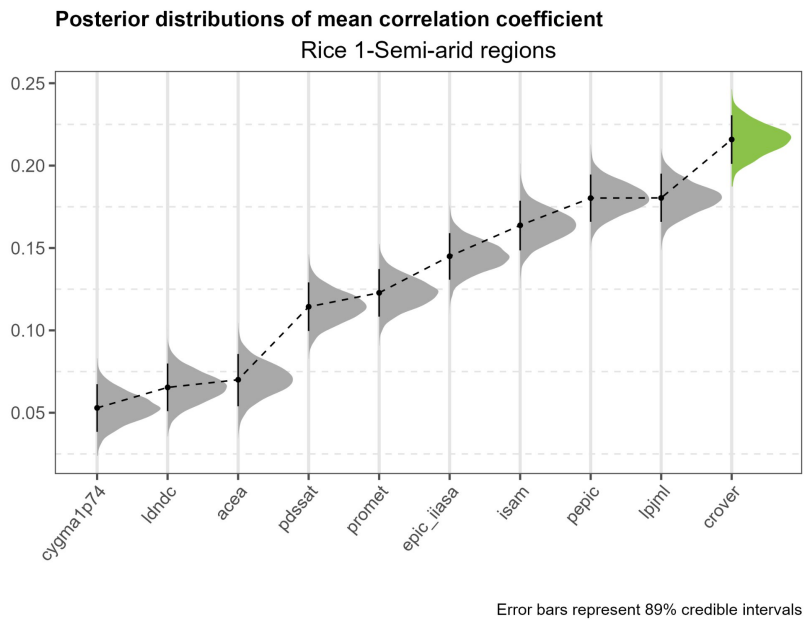
1. Results for other crops

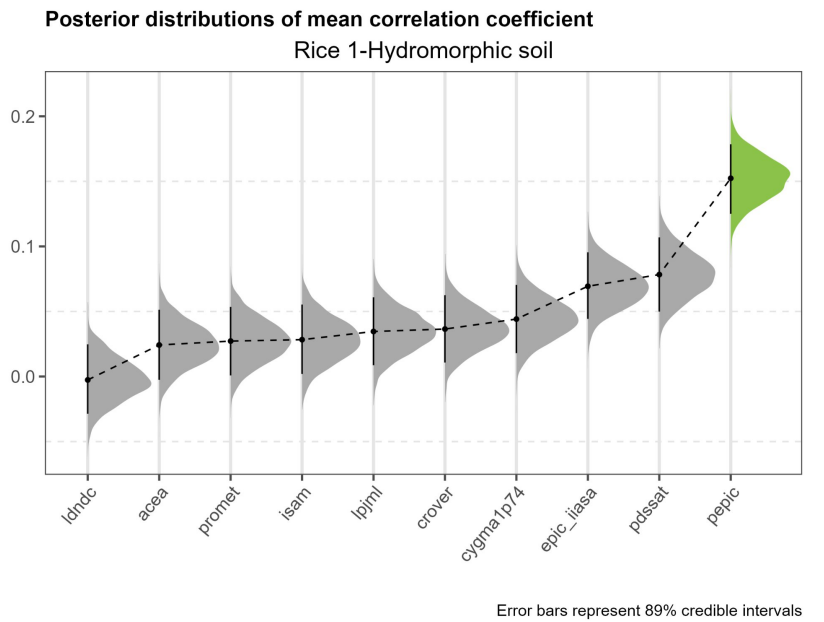

Posterior distributions of mean correlation coefficient
Rice 1-Humid regions ((sub)tropics)
0.15
0.10
0.05
0.00
ldndc
acea
isam
lpjml
cygma1p74
promet
pdssat
crover
epic_iiasa
pepic
Error bars represent 89% credible intervals

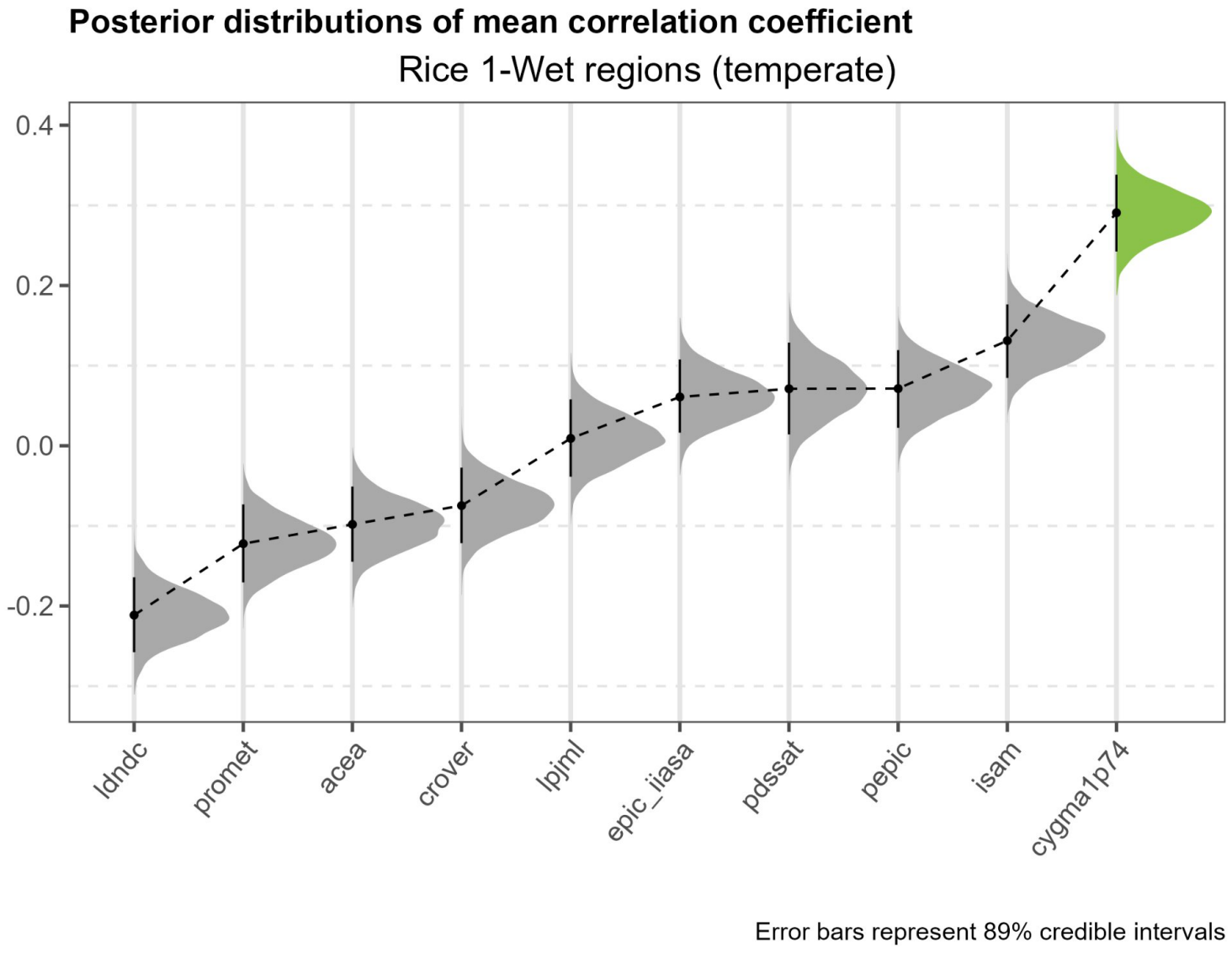

Posterior distributions of mean correlation coefficient
Rice 1-Ample irrigated region
0.15
0.10
0.05
0.00
-0.05
promet
acea
ldndc
isam
lpjml
crover
cygma1p74
pepic
pdssat
epic_iiasa
Error bars represent 89% credible intervals


Posterior distributions of mean correlation coefficient
Rice 1-Dry regions (temperate)
0.2
0.1
0.0
ldndc
promet
acea
pepic
isam
cygma1p74
epic_iiasa
lpjml
crover
pdssat
Error bars represent 89% credible intervals


Posterior distributions of mean correlation coefficient
Rice 1-Sub-Humid regions ((sub)tropics)
0.15
0.10
0.05
0.00
-0.05
ldndc
acea
pdssat
epic_iiasa
promet
lpjml
isam
cygma1p74
crover
pepic
Error bars represent 89% credible intervals

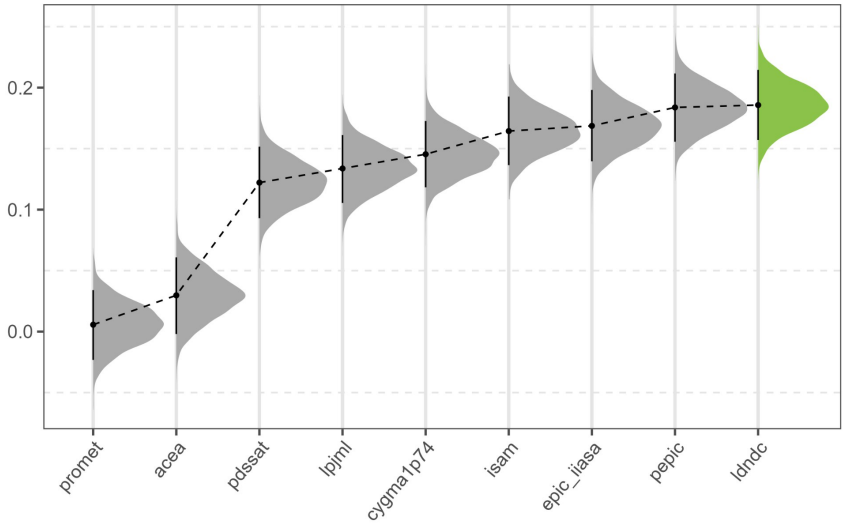

Posterior distributions of mean correlation coefficient
Rice 1-Semi-arid regions
0.25
0.20
0.15
0.10
0.05
cygma1p74
ldndc
acea
pdssat
promet
epic_iiasa
isam
pepic
lpjml
crover
Error bars represent 89% credible intervals


Posterior distributions of mean correlation coefficient
Rice 1-Hydromorphic soil
0.2
0.1
0.0
ldndc
acea
promet
isam
lpjml
crover
cygma1p74
epic_iiasa
pdssat
pepic
Error bars represent 89% credible intervals



Posterior distributions of mean correlation coefficient

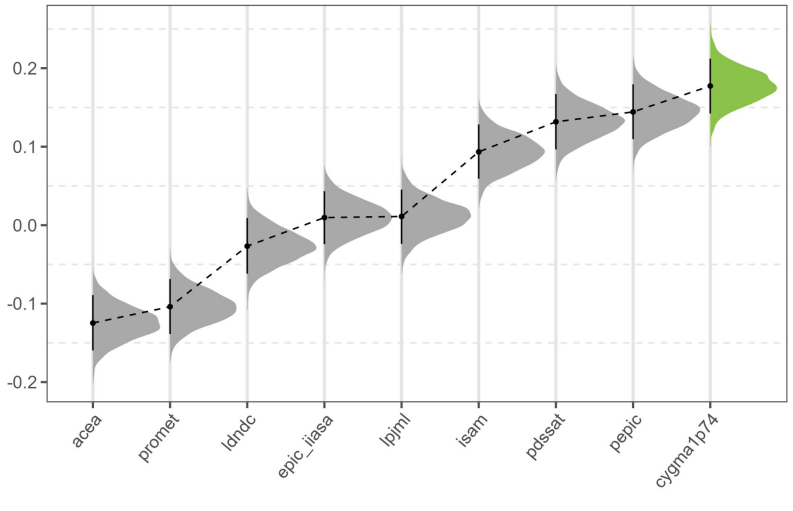
Rice 2-Sub-Humid regions ((sub)tropics)



Error bars represent 89% credible intervals

Posterior distributions of mean correlation coefficient

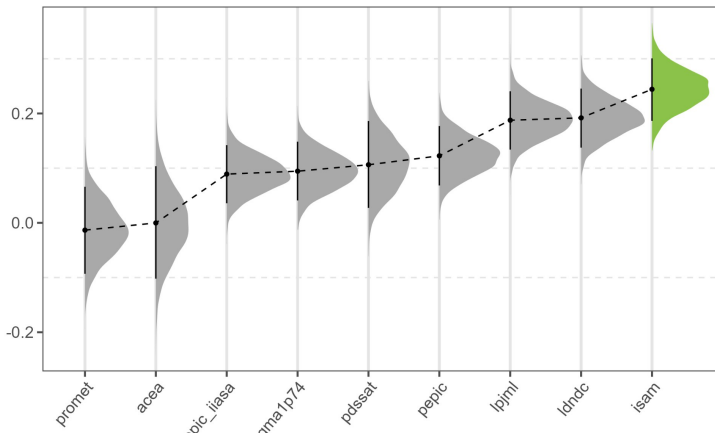
Rice 2-Humid regions ((sub)tropics)



Error bars represent 89% credible intervals

Posterior distributions of mean correlation coefficient

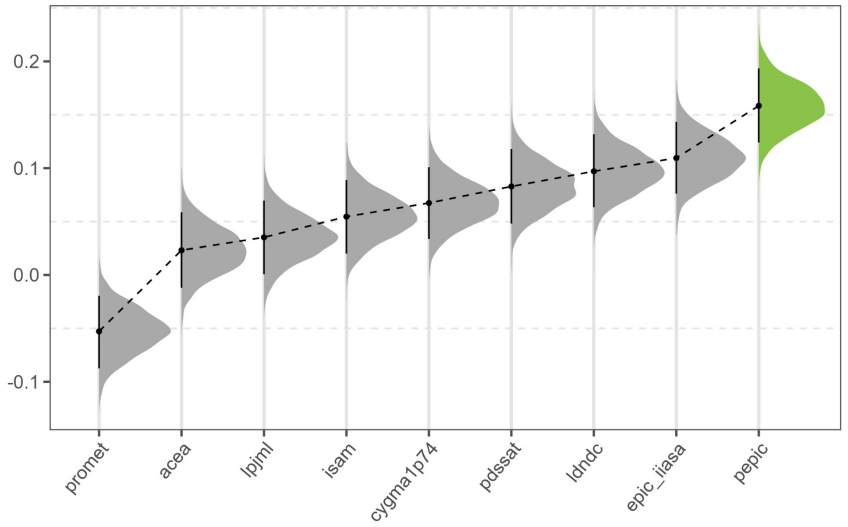
Rice 2-Semi-arid regions



Error bars represent 89% credible interval

Posterior distributions of mean correlation coefficient

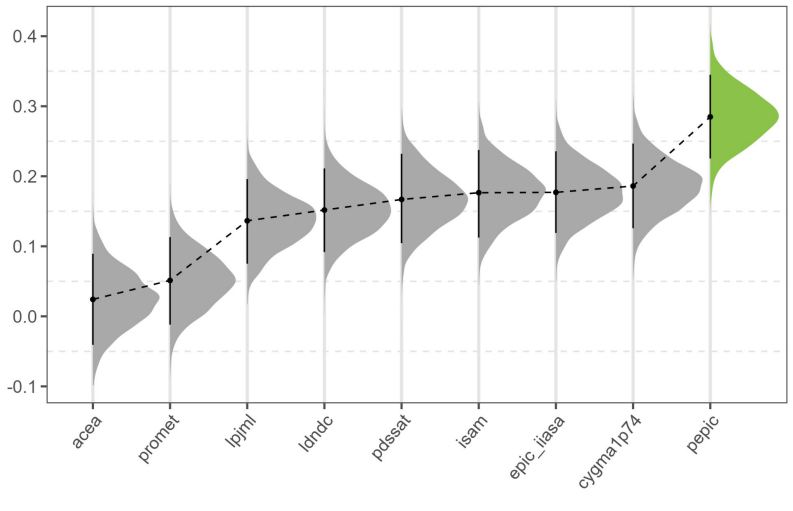
Rice 2-Ample irrigated region



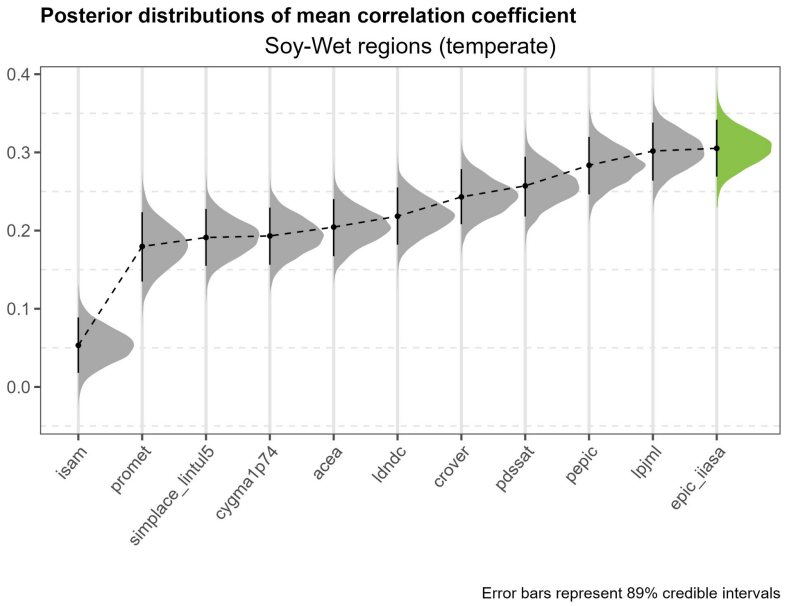
Error bars represent 89% credible intervals

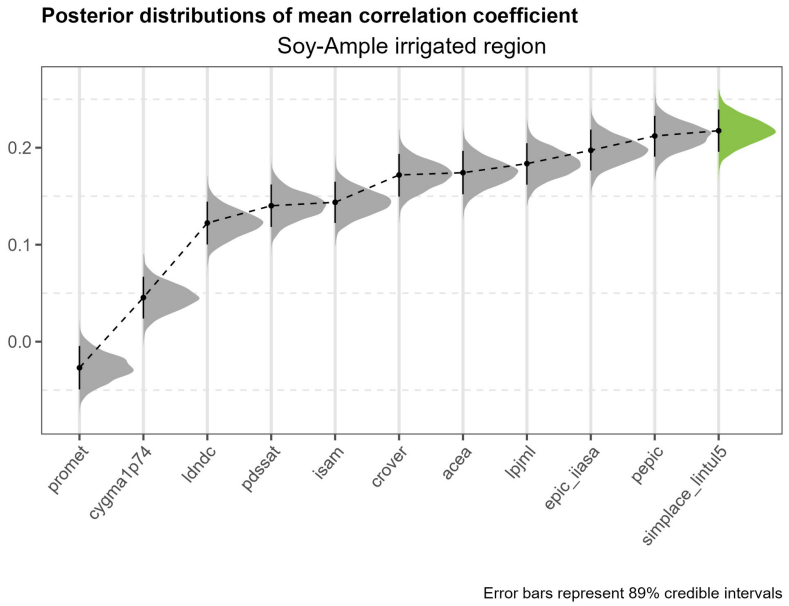
Posterior distributions of mean correlation coefficient

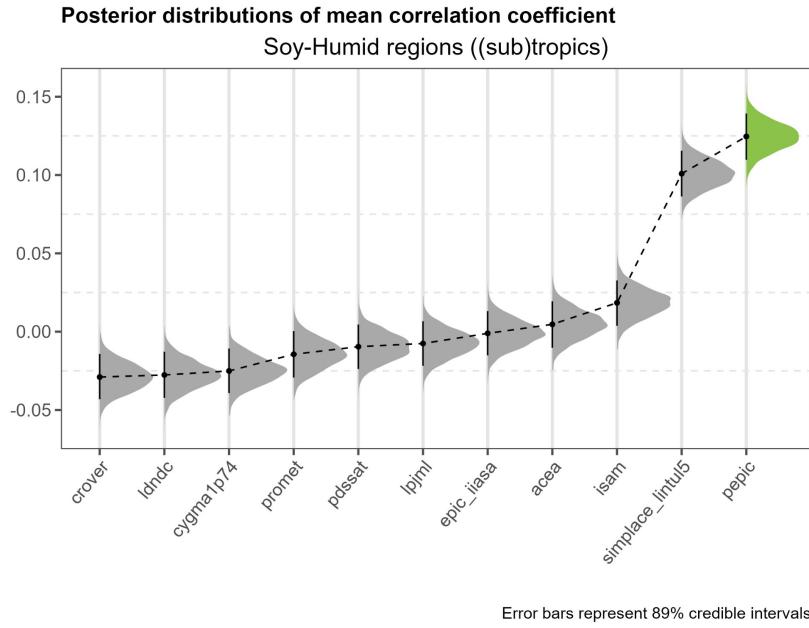
Rice 2-Hydromorphic soil

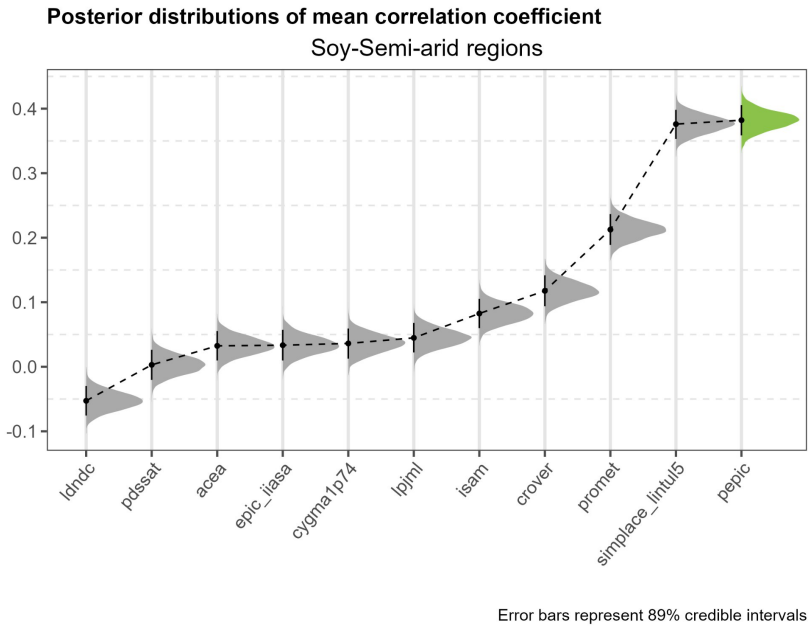


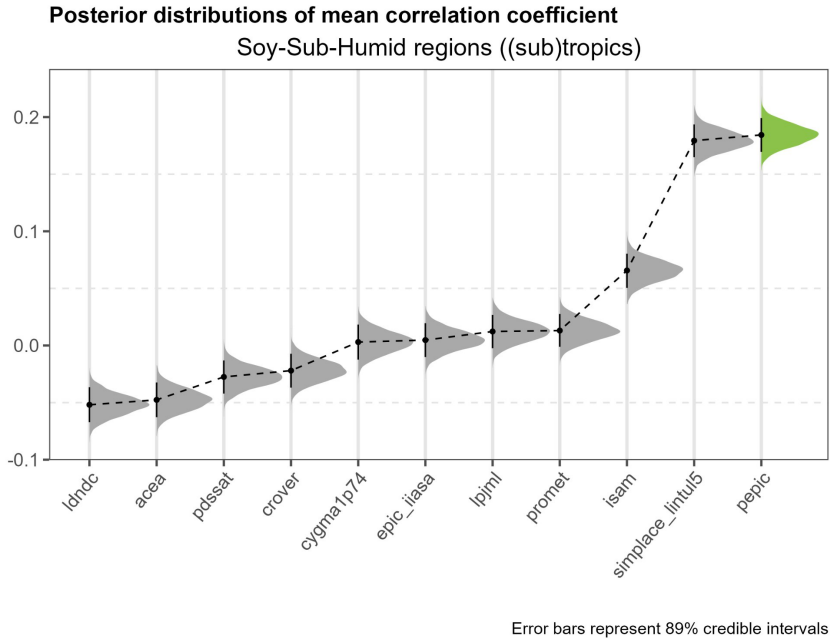
Error bars represent 89% credible intervals

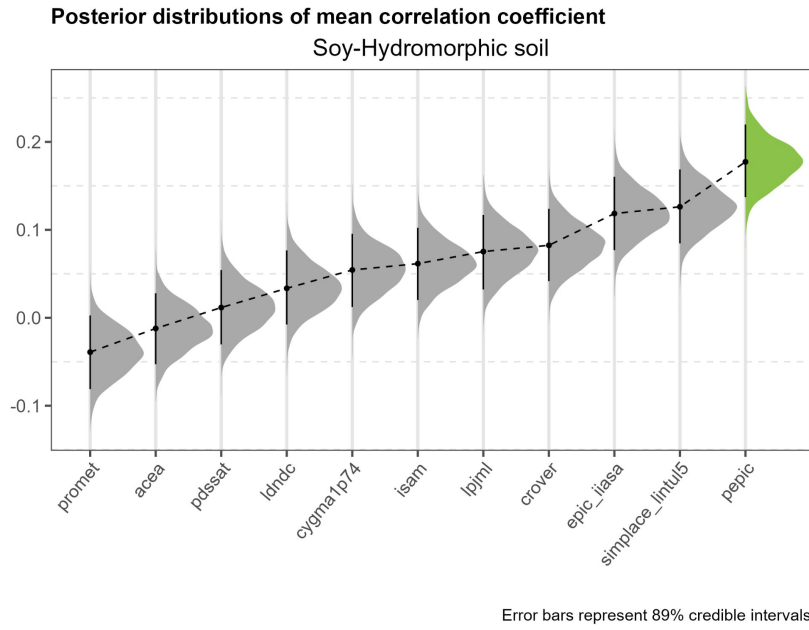


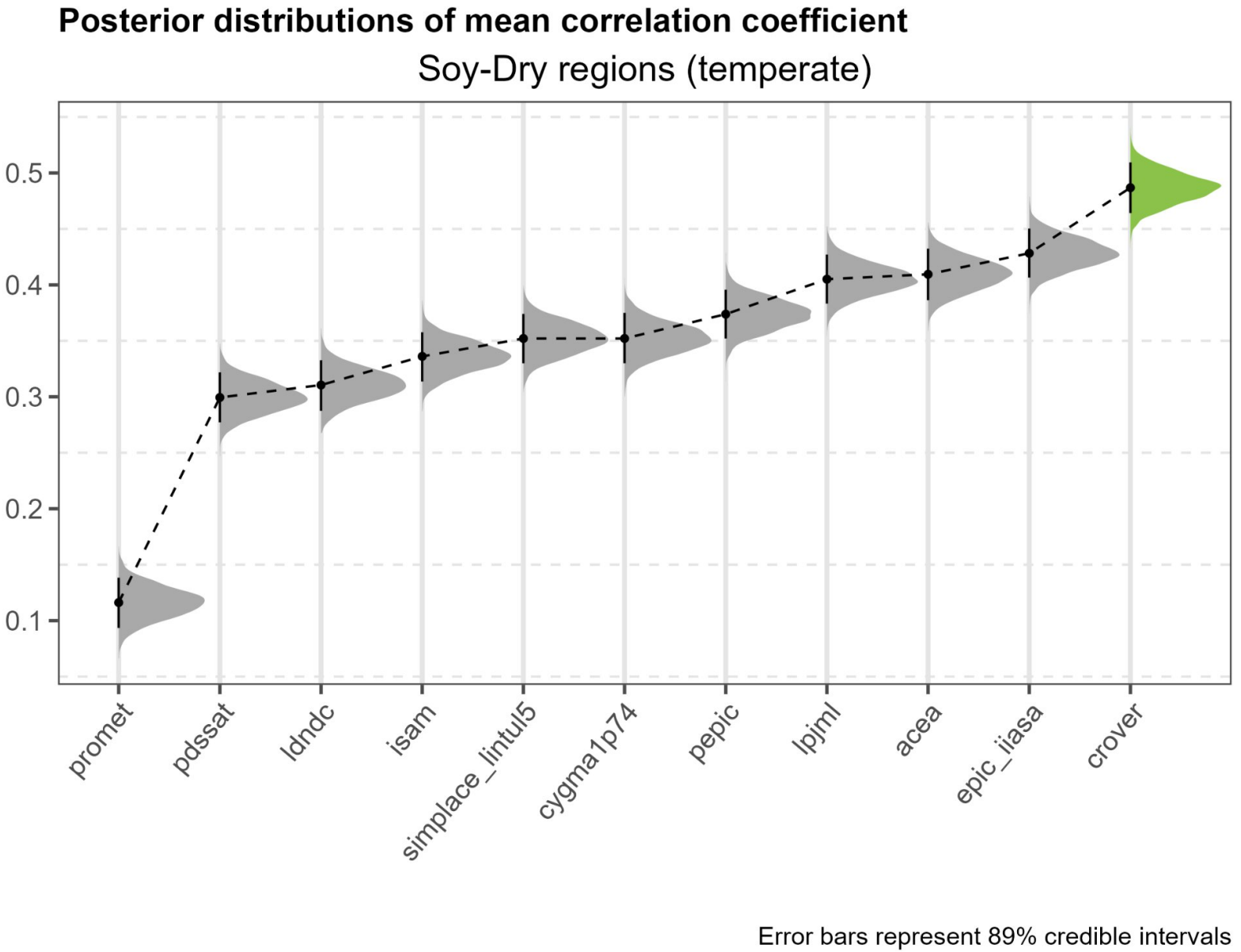


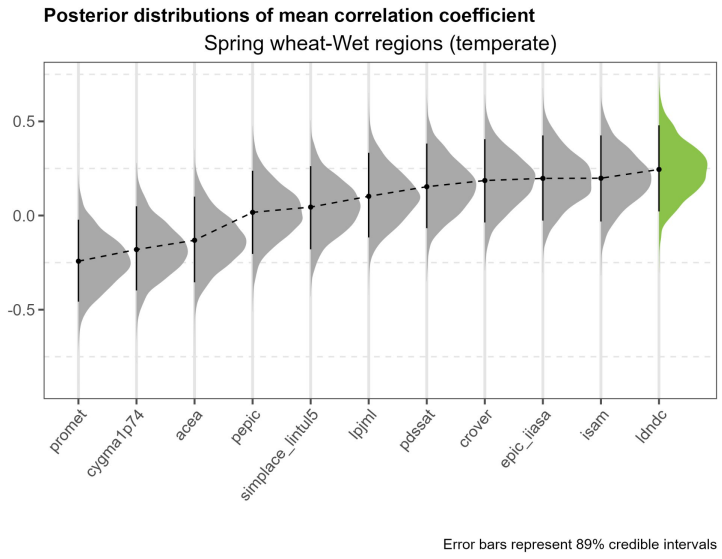


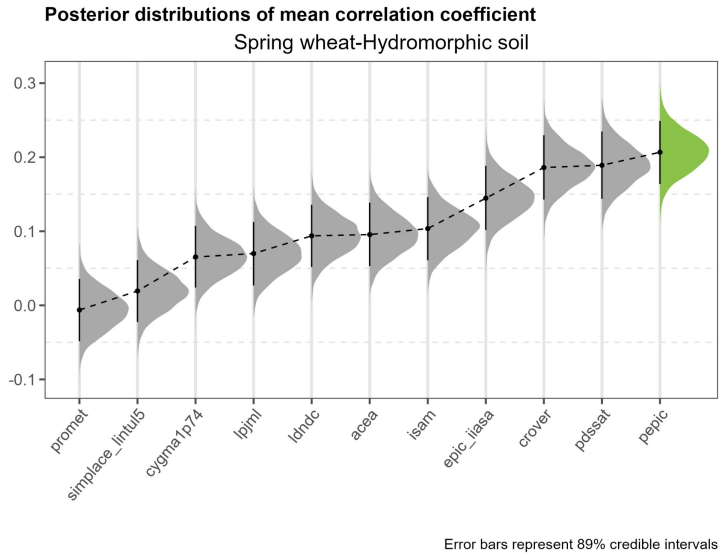


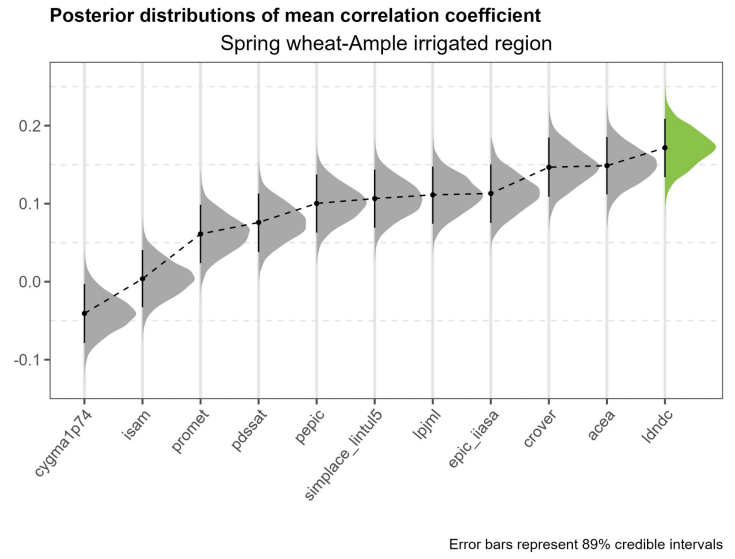


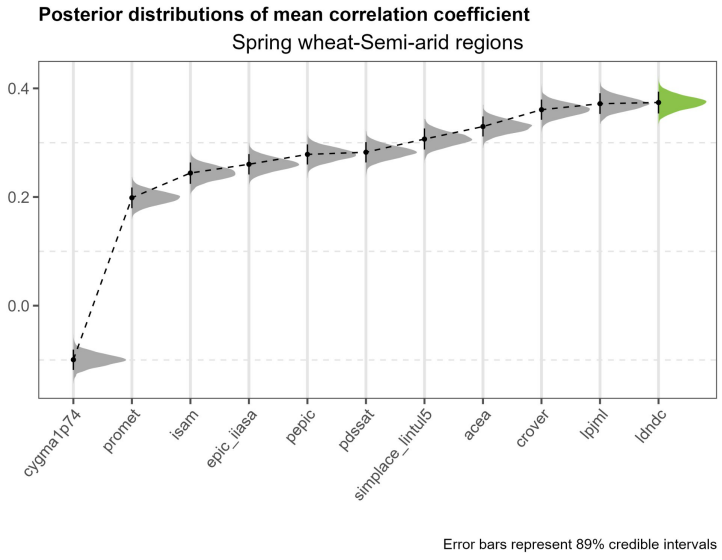


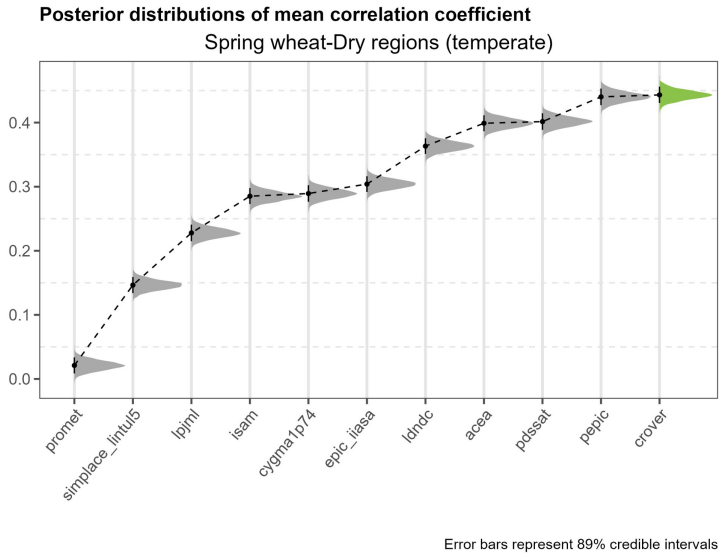


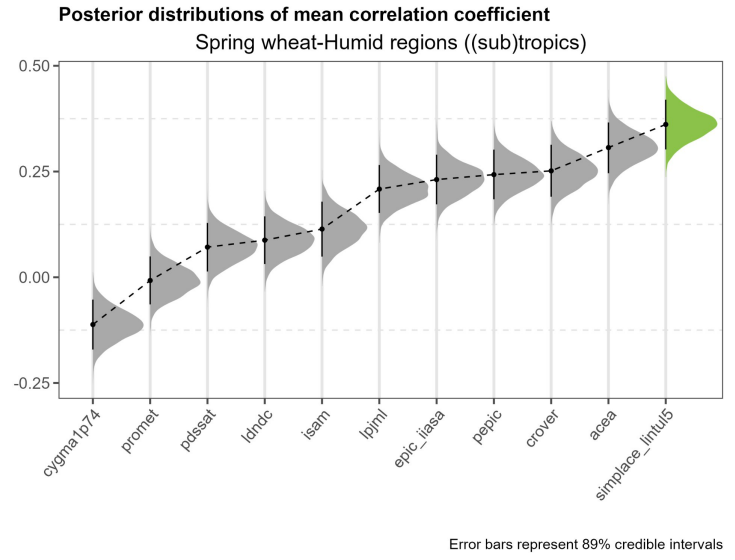


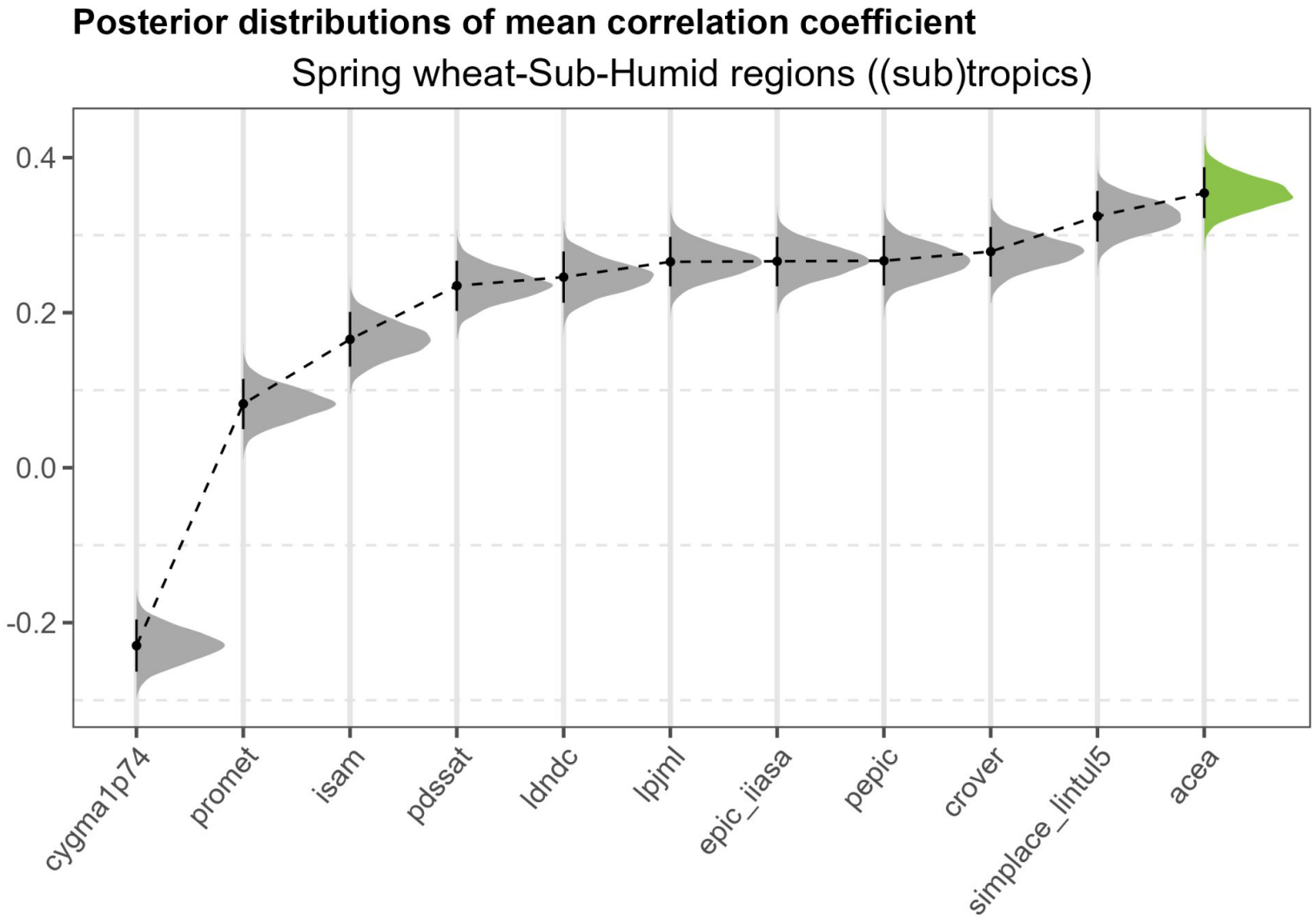


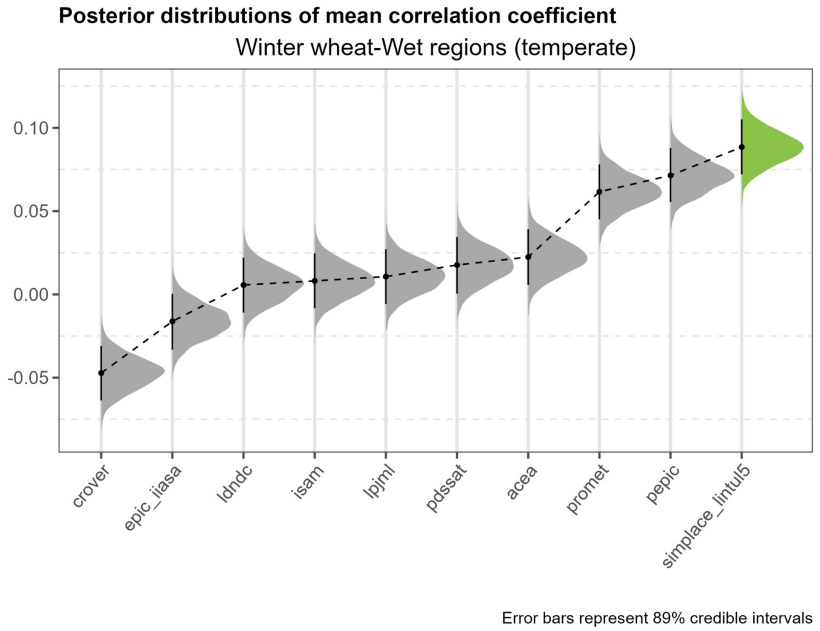


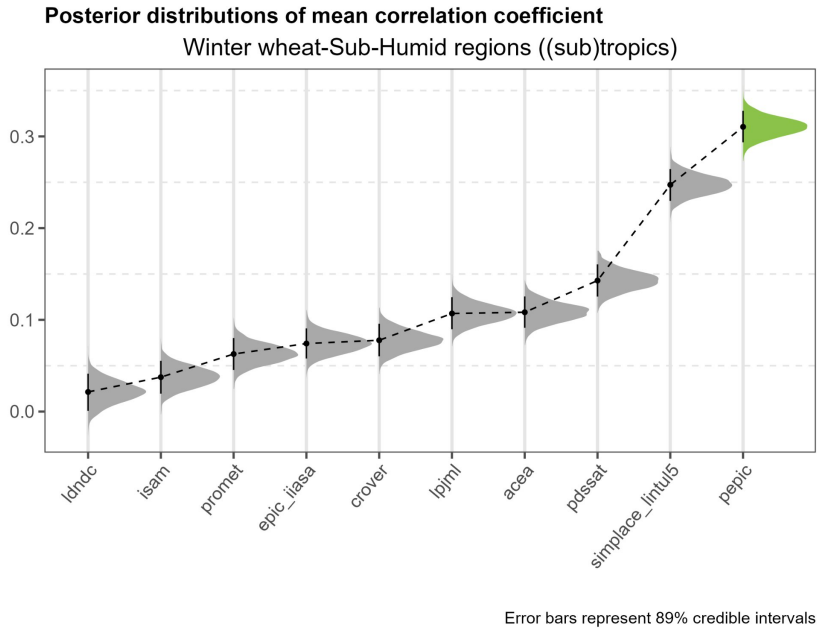


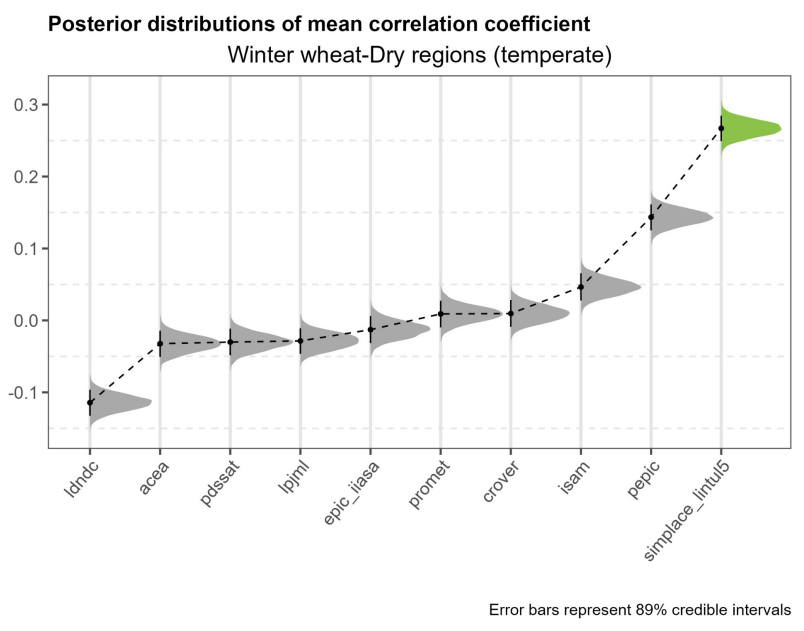


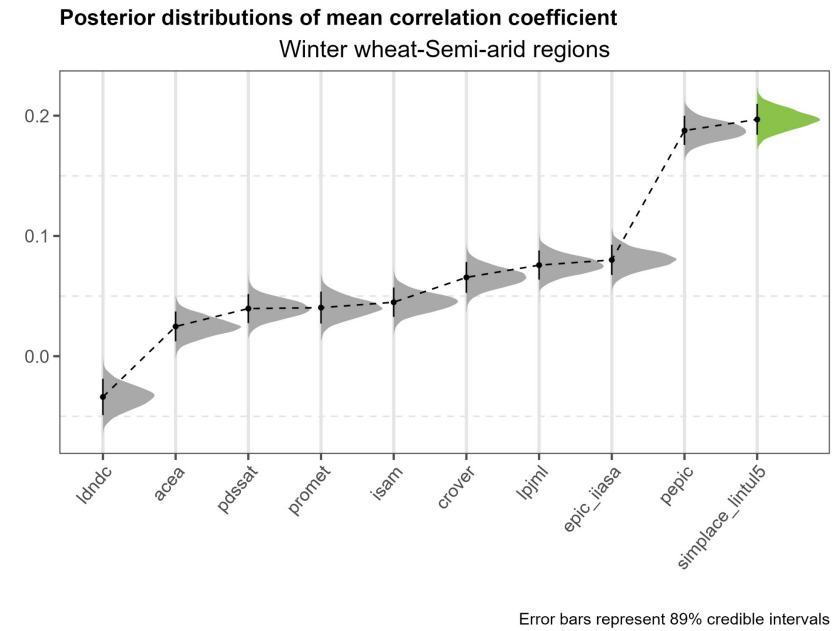


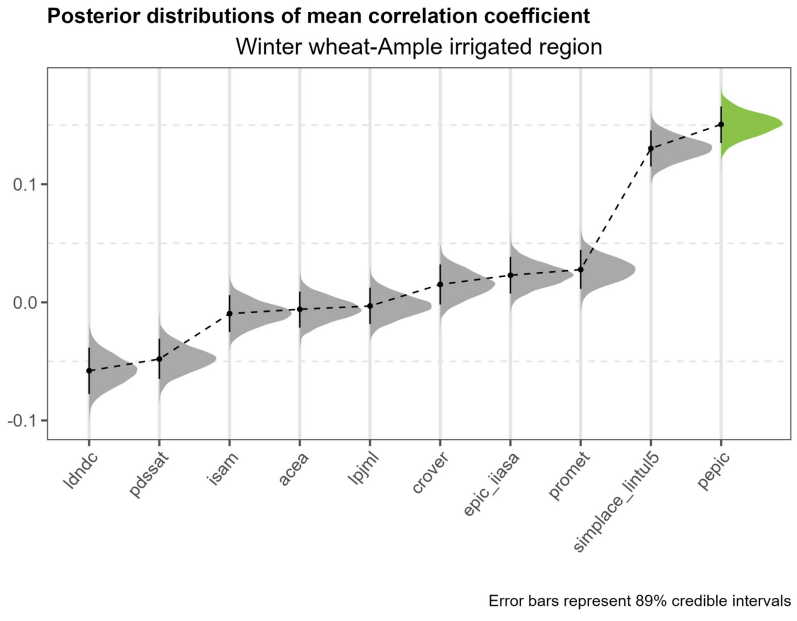


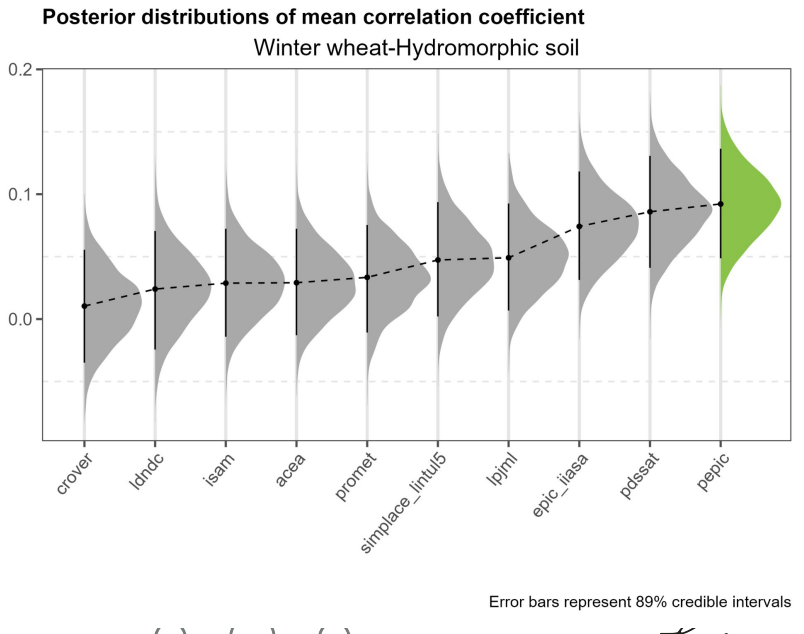


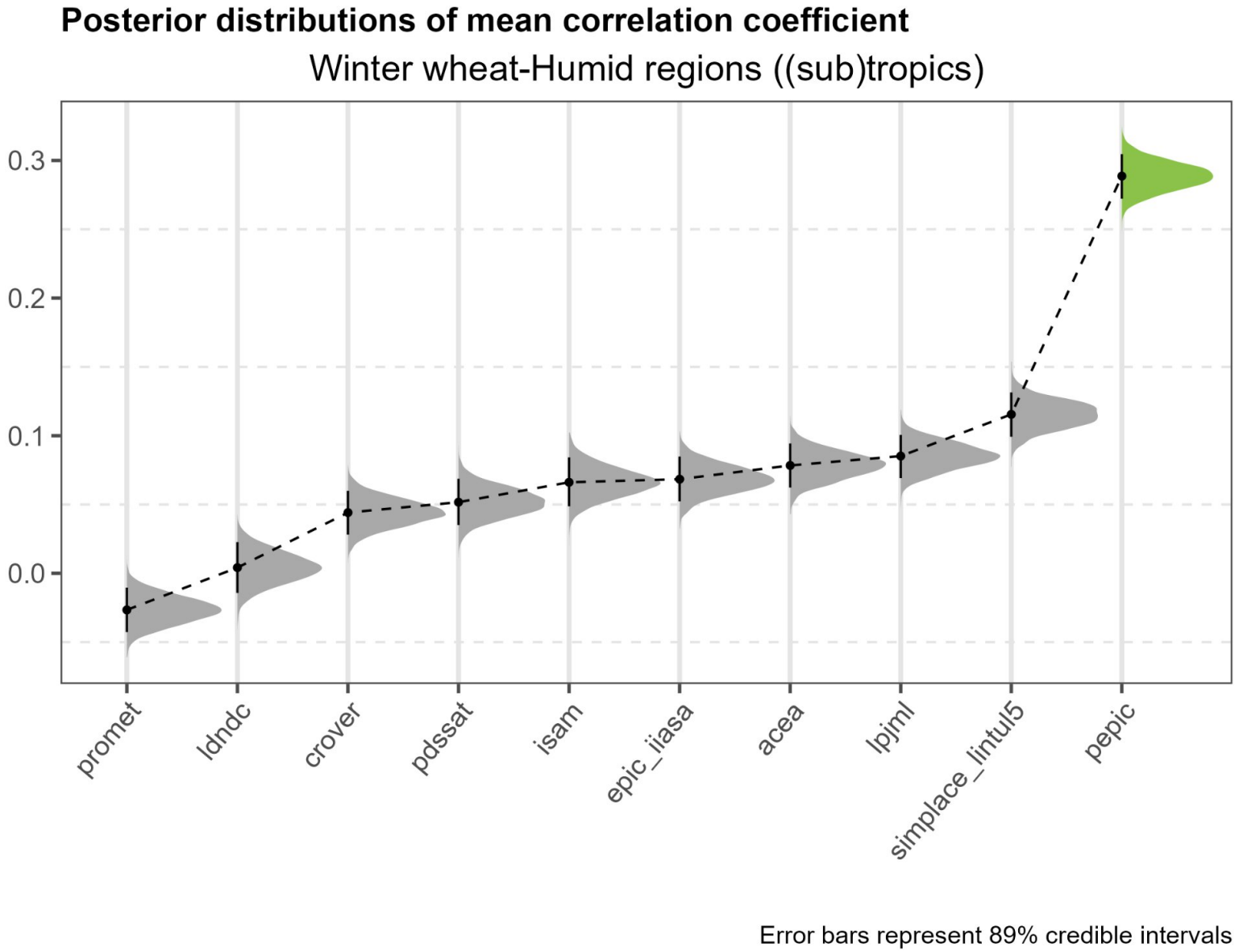




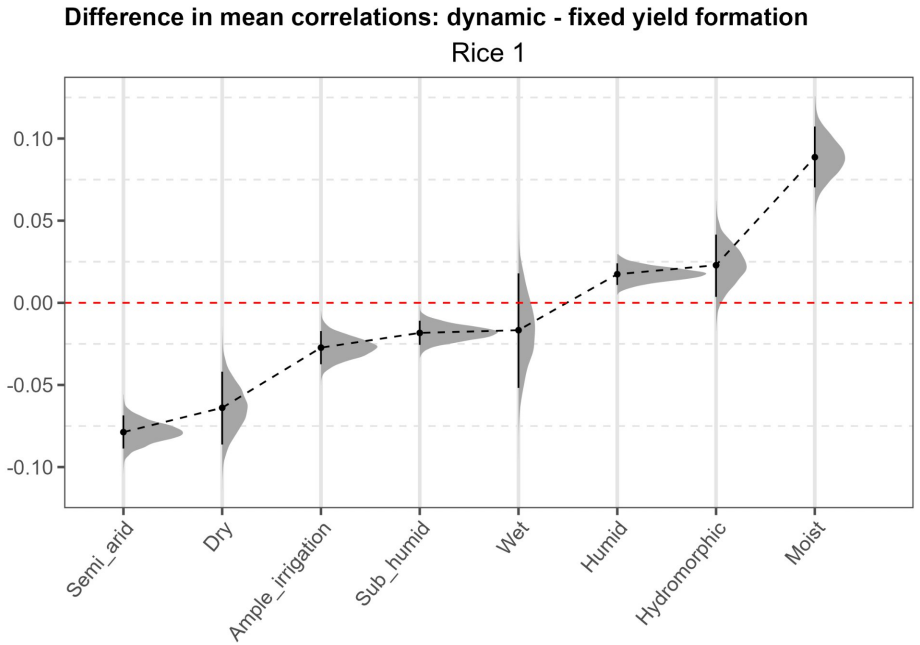


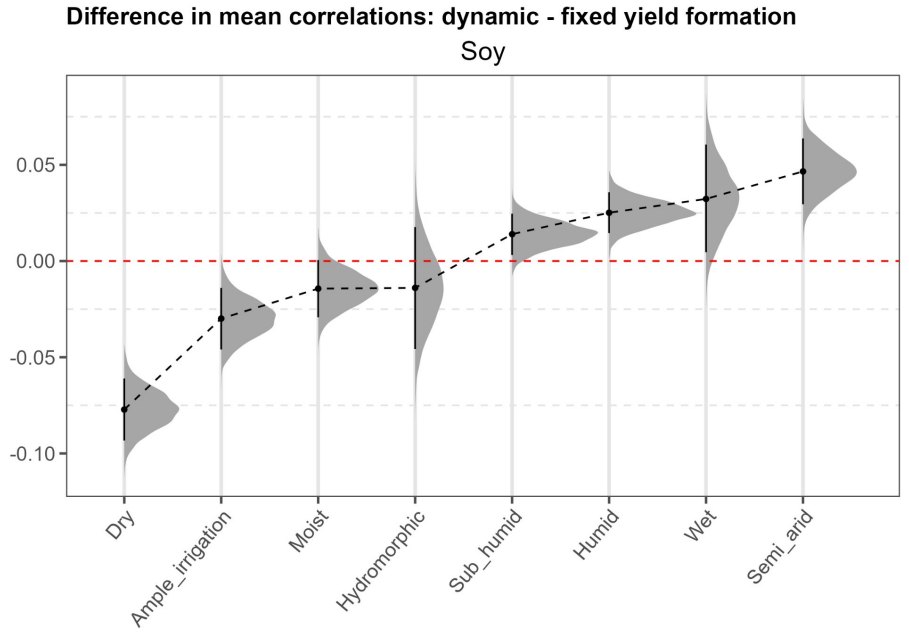


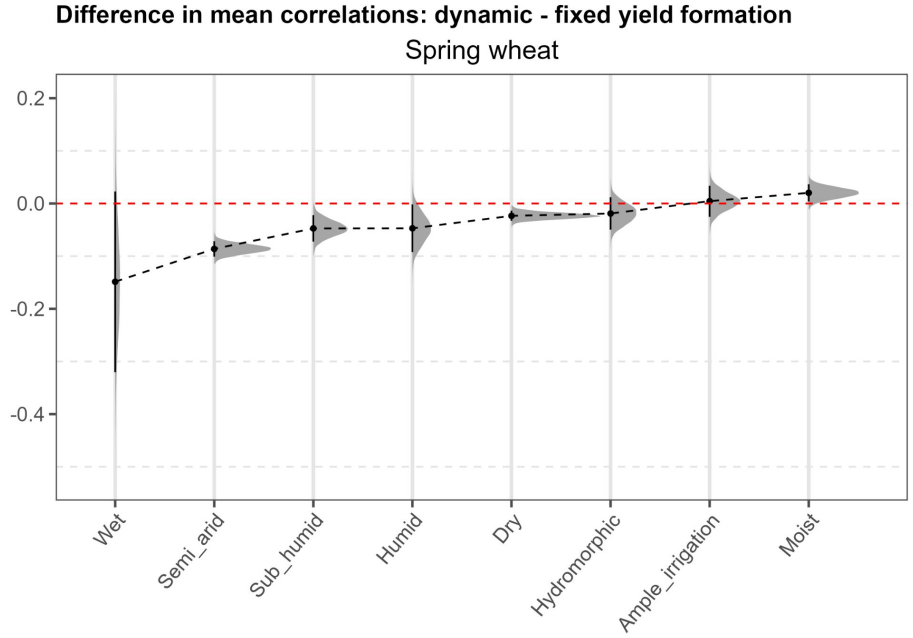


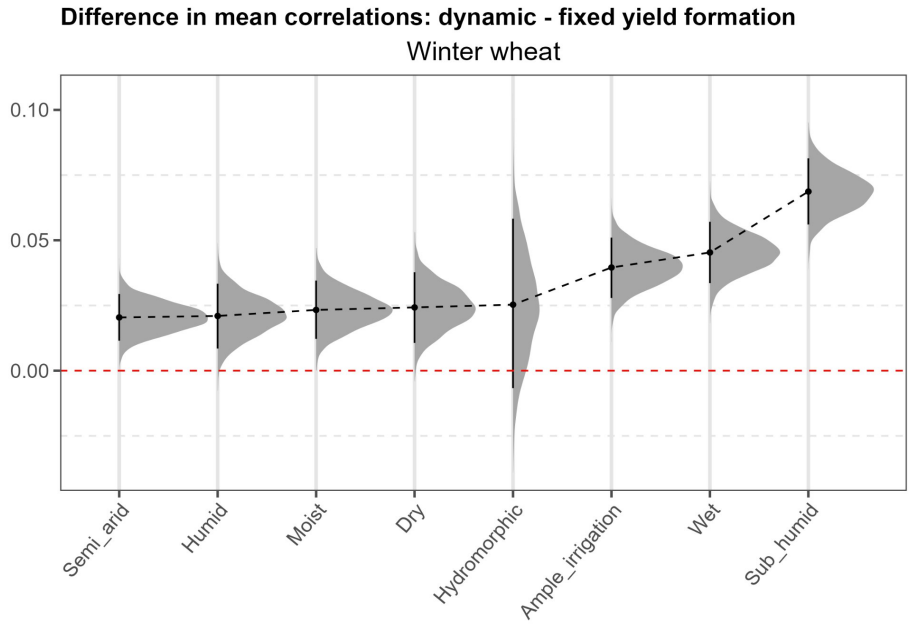


2. Model characteristics: results for other crops









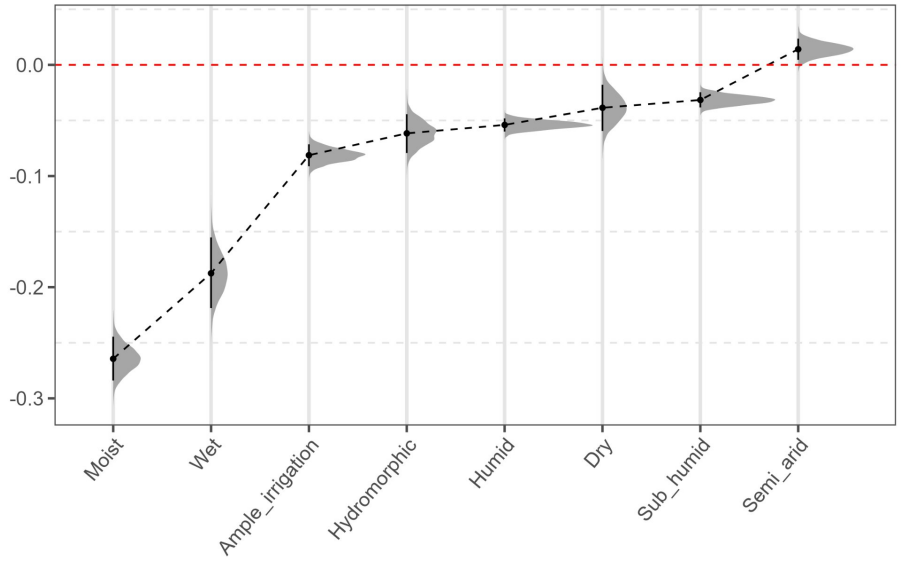
GGCM valida

Error bars represent 89% credible intervals

Error bars represent 89% credible intervals

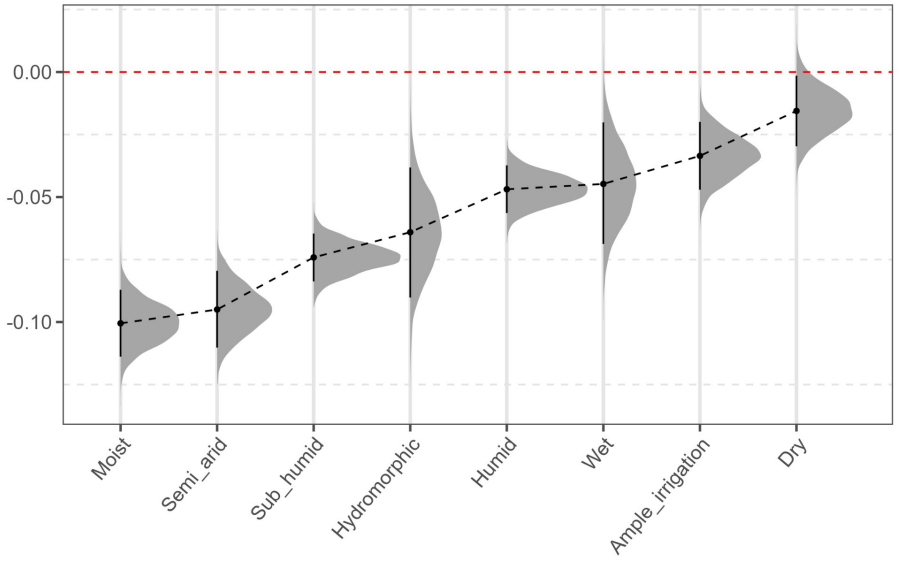
Difference in mean correlations: Photosynthesis - RUE light utilization

Rice 1



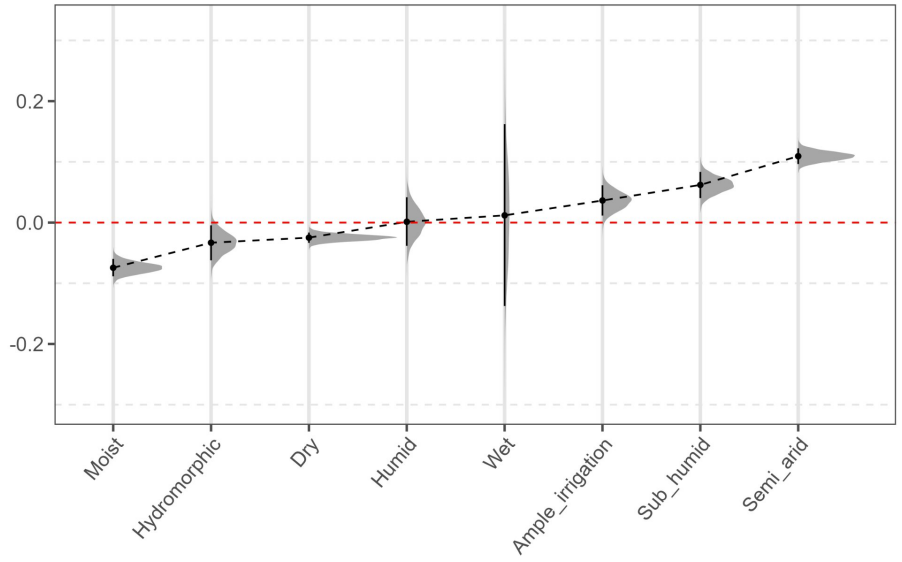
Difference in mean correlations: Photosynthesis - RUE light utilization

Soy



Difference in mean correlations: Photosynthesis - RUE light utilization

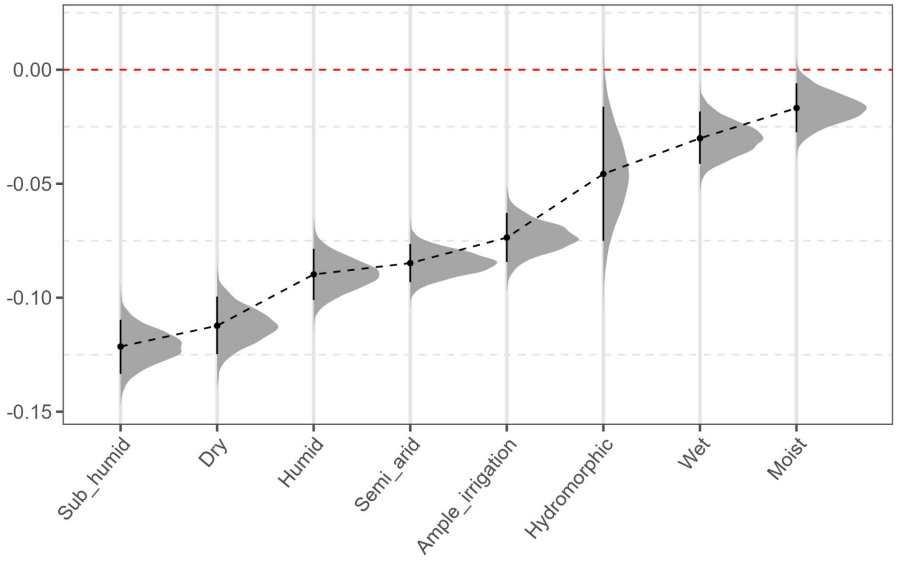
Spring wheat



Error bars represent 89% credible intervals

Difference in mean correlations: Photosynthesis - RUE light utilization

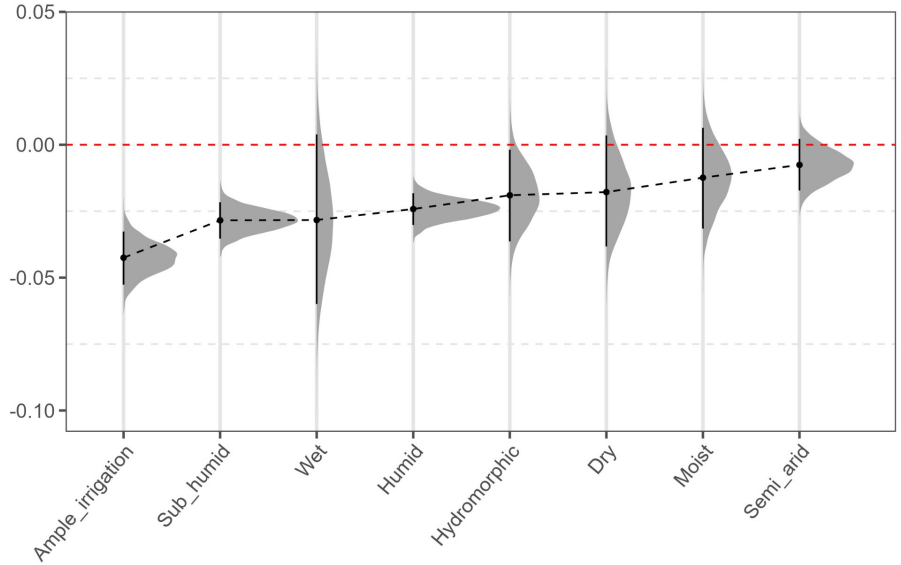
Winter wheat



Error bars represent 89% credible intervals

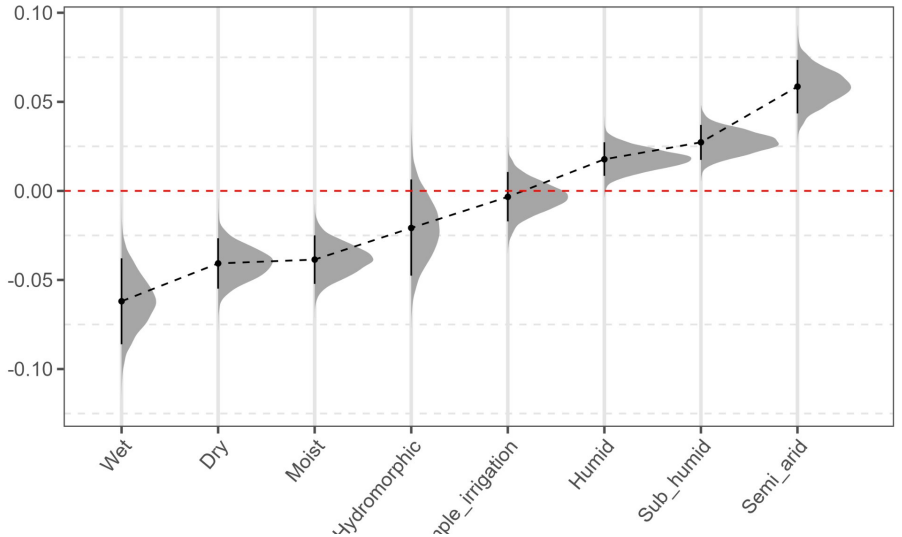
Difference in mean correlations: complex - simple drought effect

Rice 1



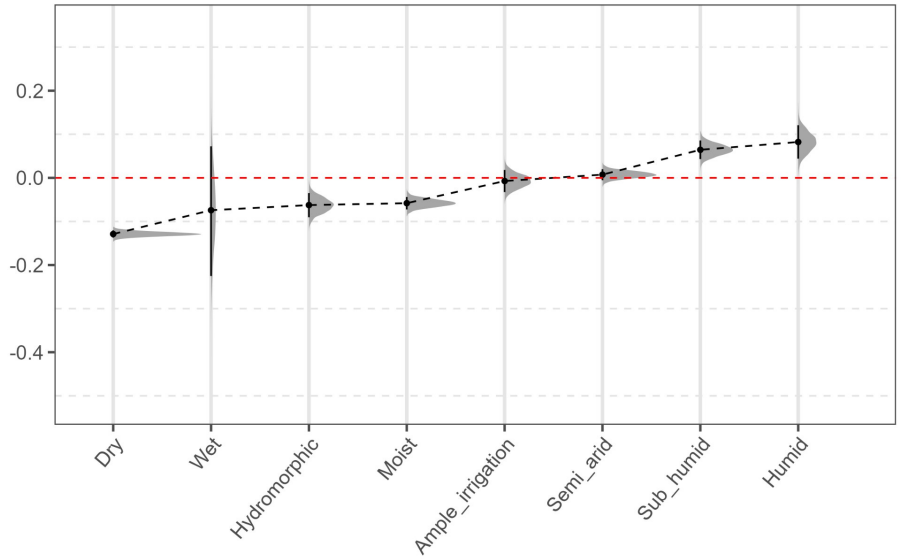
Difference in mean correlations: complex - simple drought effect

Soy



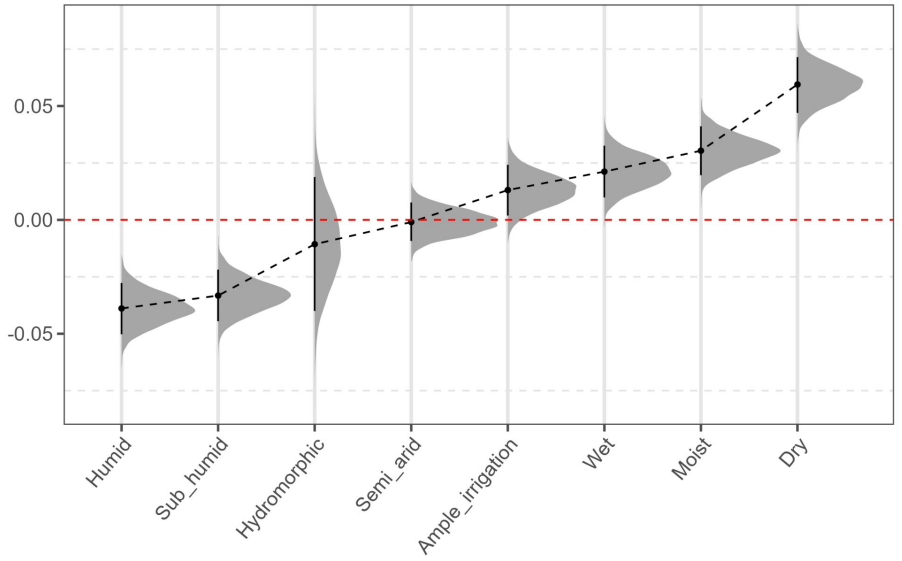
Difference in mean correlations: complex - simple drought effect

Spring wheat



Difference in mean correlations: complex - simple drought effect

Winter wheat

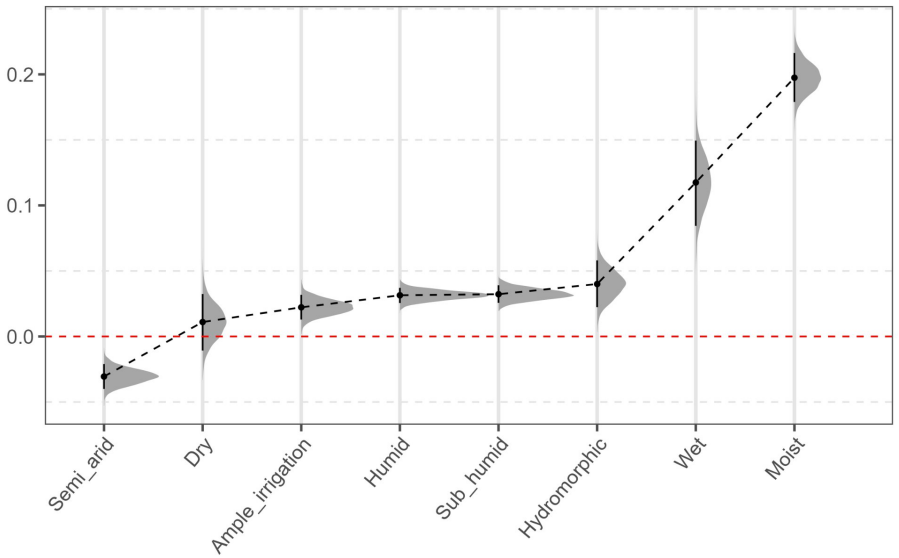


Error bars represent 89% credible intervals

Error bars represent 89% credible intervals

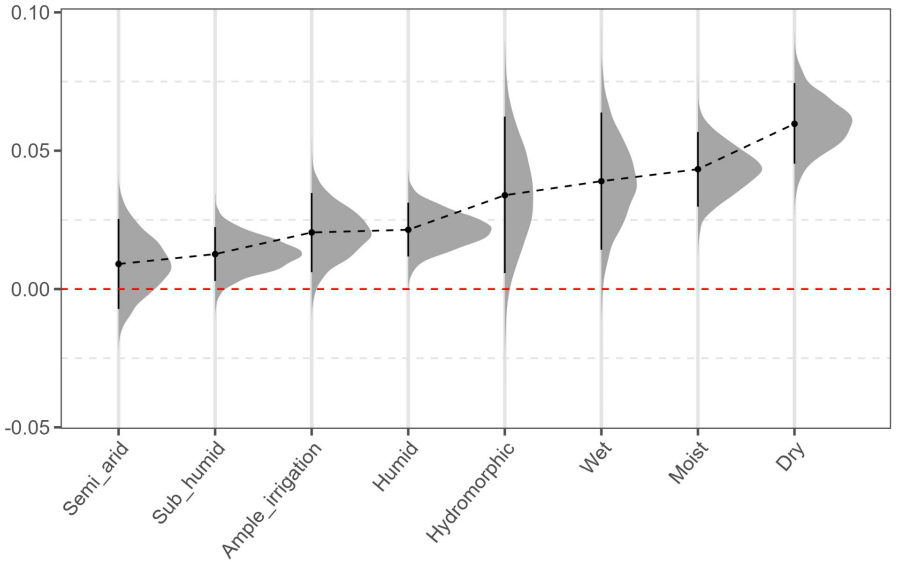
Difference in mean correlations: WL - WD water-related stress

Rice 1



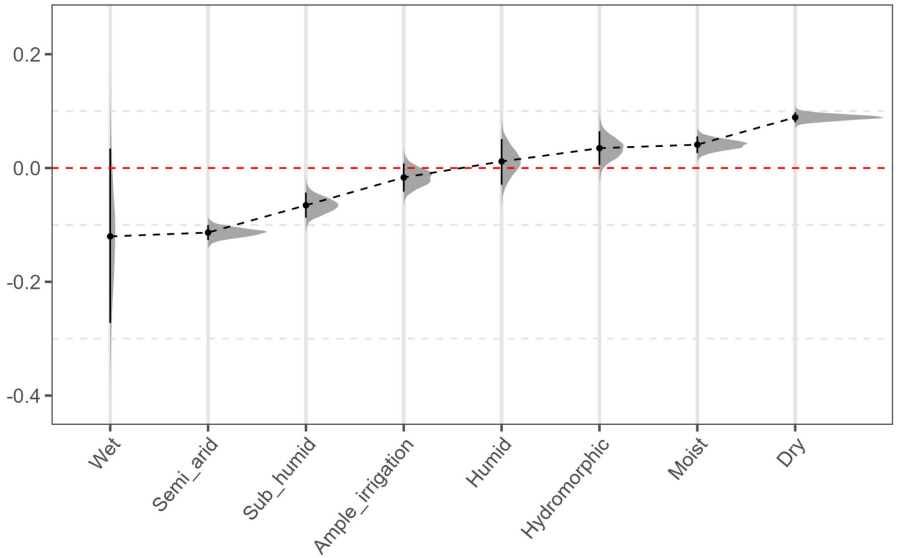
Difference in mean correlations: WL - WD water-related stress

Soy



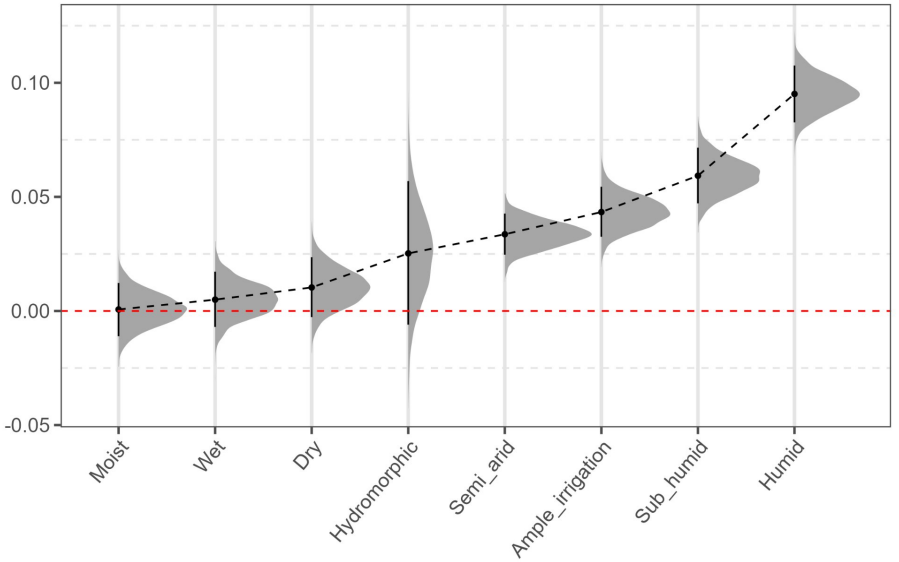
Difference in mean correlations: WL - WD water-related stress

Spring wheat



Difference in mean correlations: WL - WD water-related stress

Winter wheat



GC

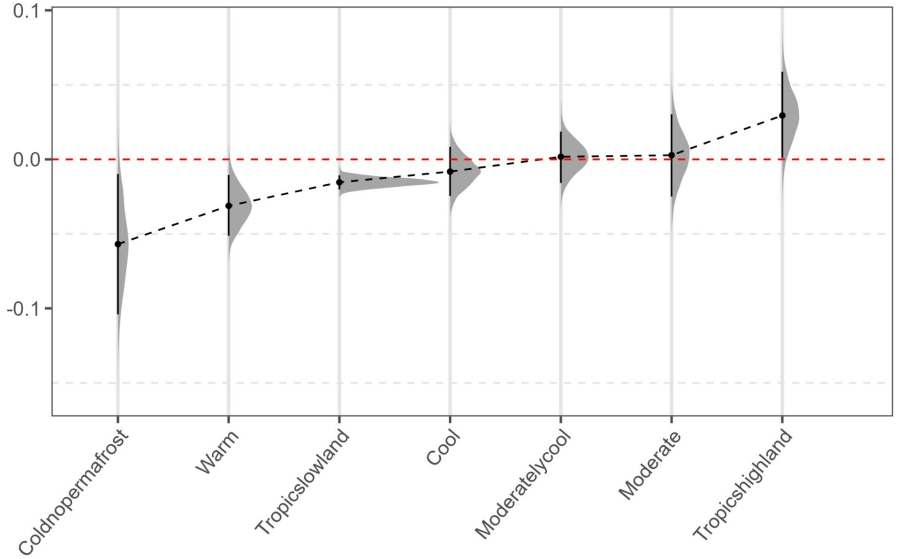
Error bars represent 89% credible intervals

Error bars represent 89% credible intervals

Leibniz Association

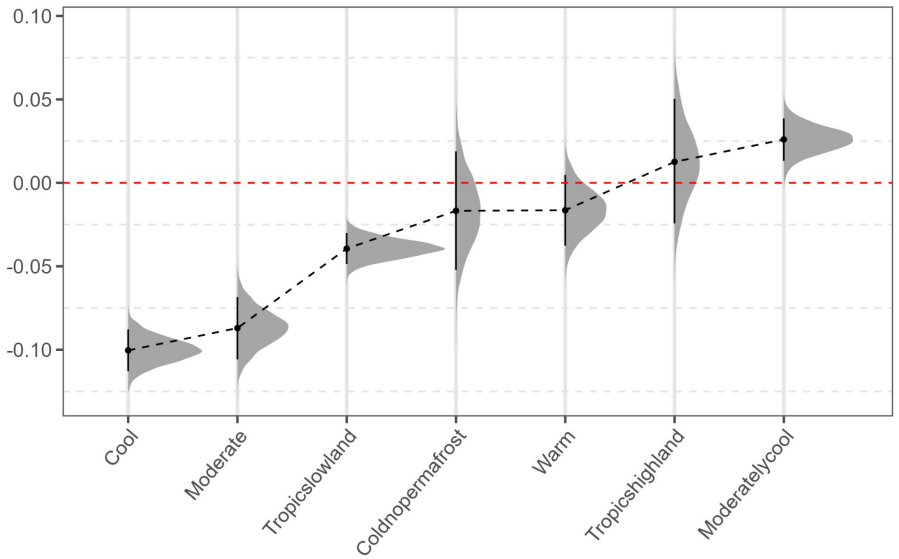
Difference in mean correlations: Complex - Simple heat-related stress

Rice 1



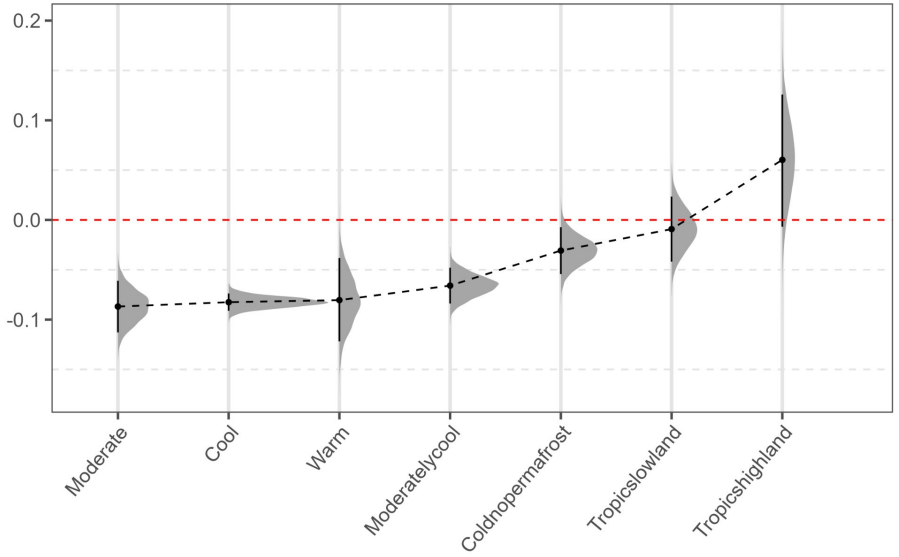
Difference in mean correlations: Complex - Simple heat-related stress

Soy



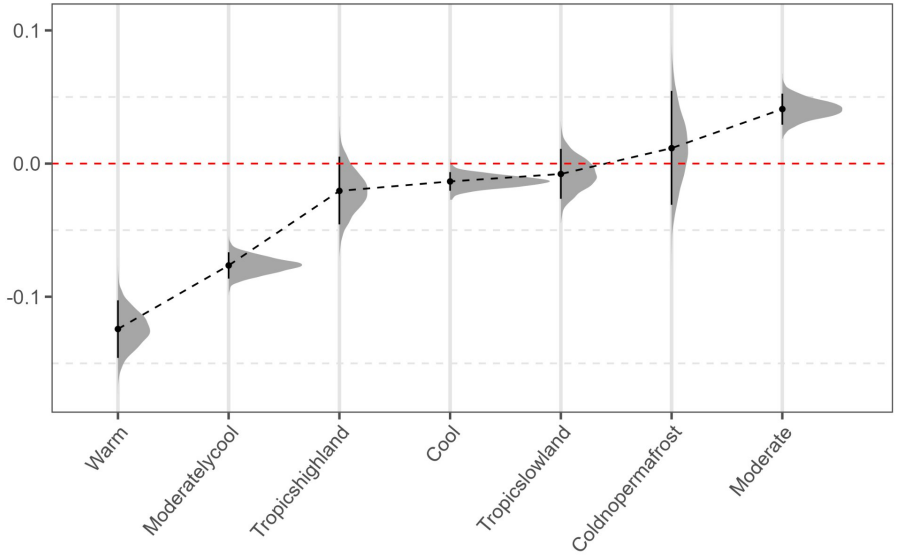
Difference in mean correlations: Complex - Simple heat-related stress

Spring wheat



Difference in mean correlations: Complex - Simple heat-related stress

Winter wheat



GGC

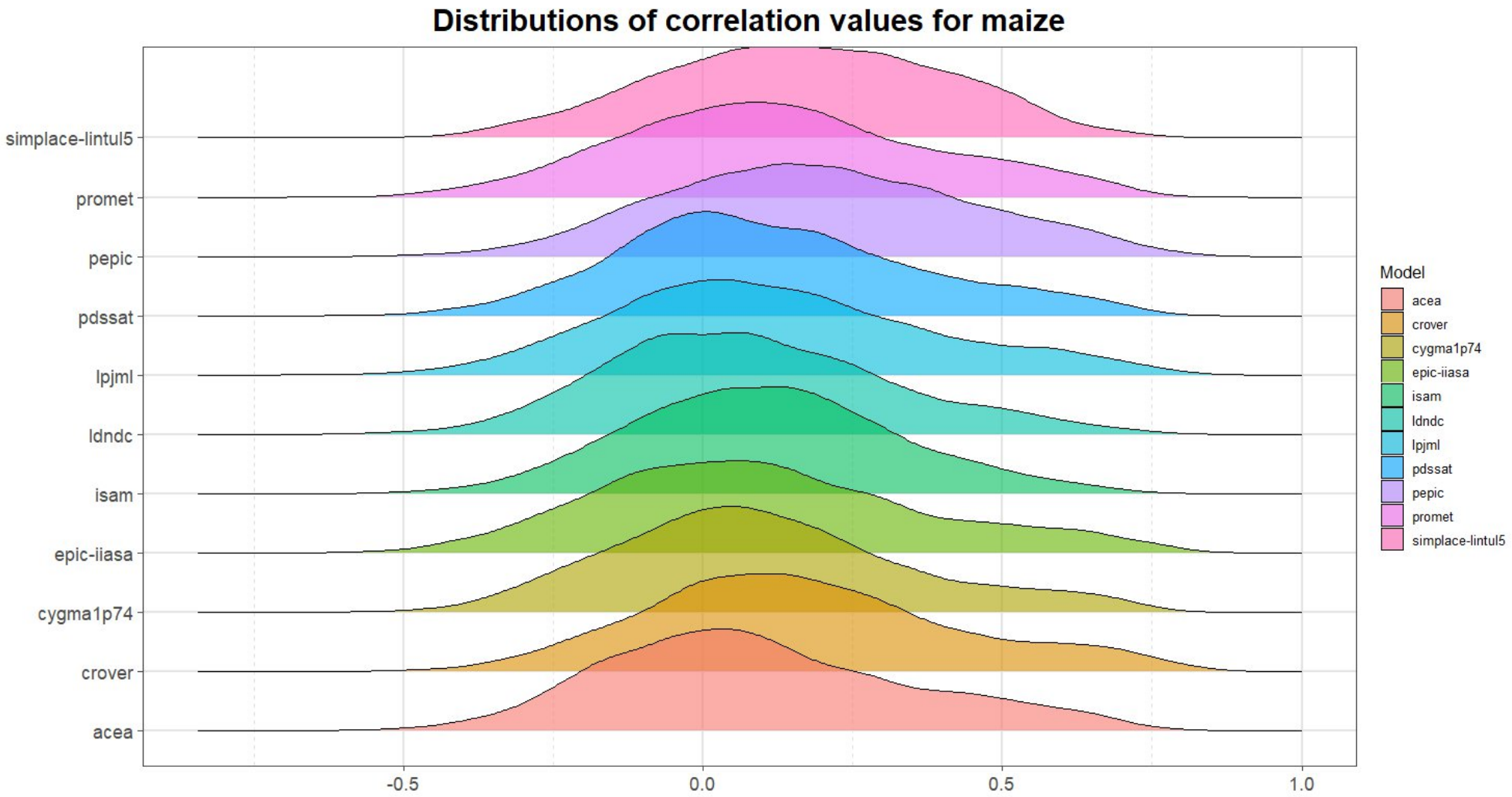
Error bars represent 89% credible intervals

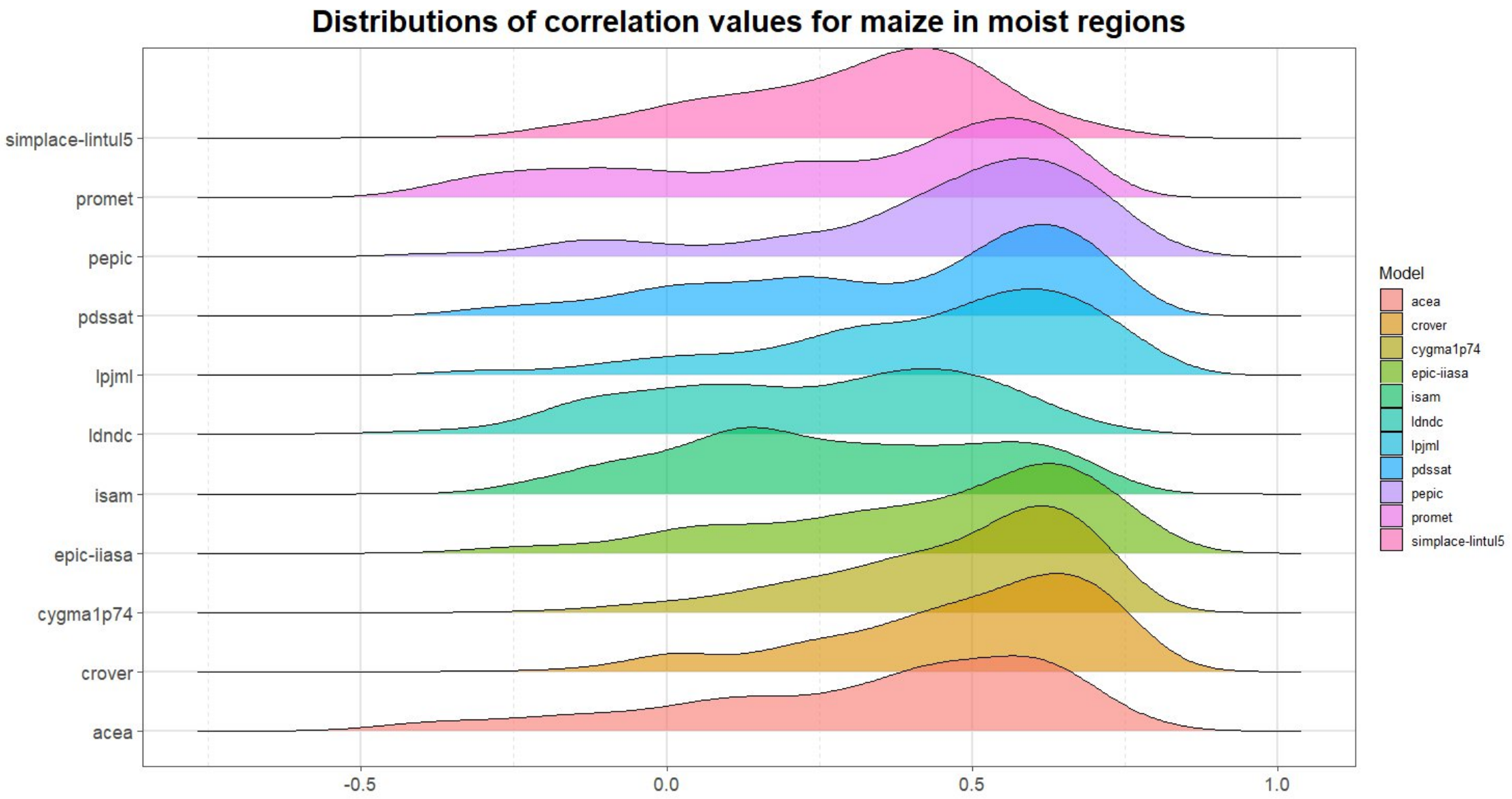
Error bars represent 89% credible intervals

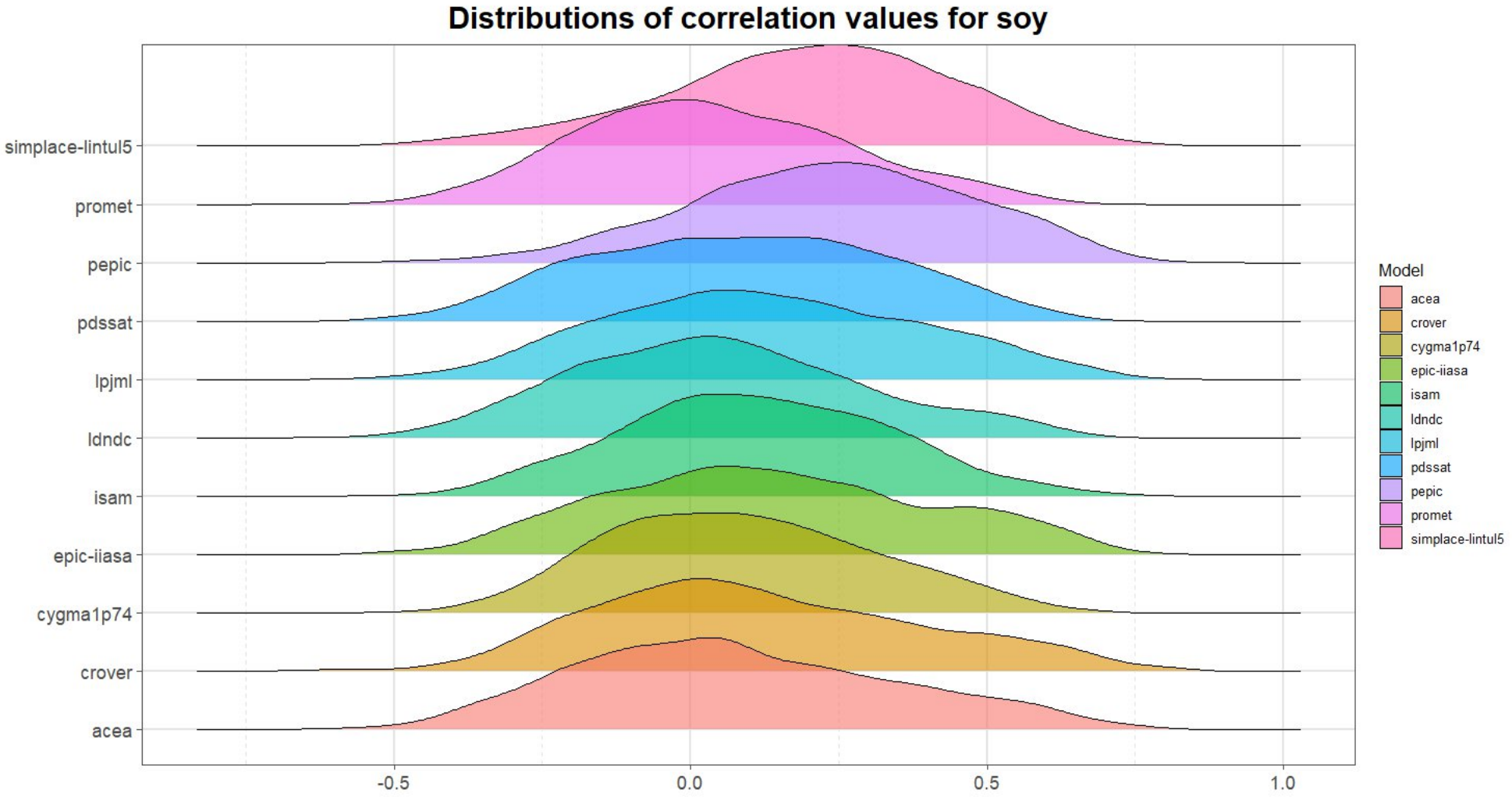
PIK

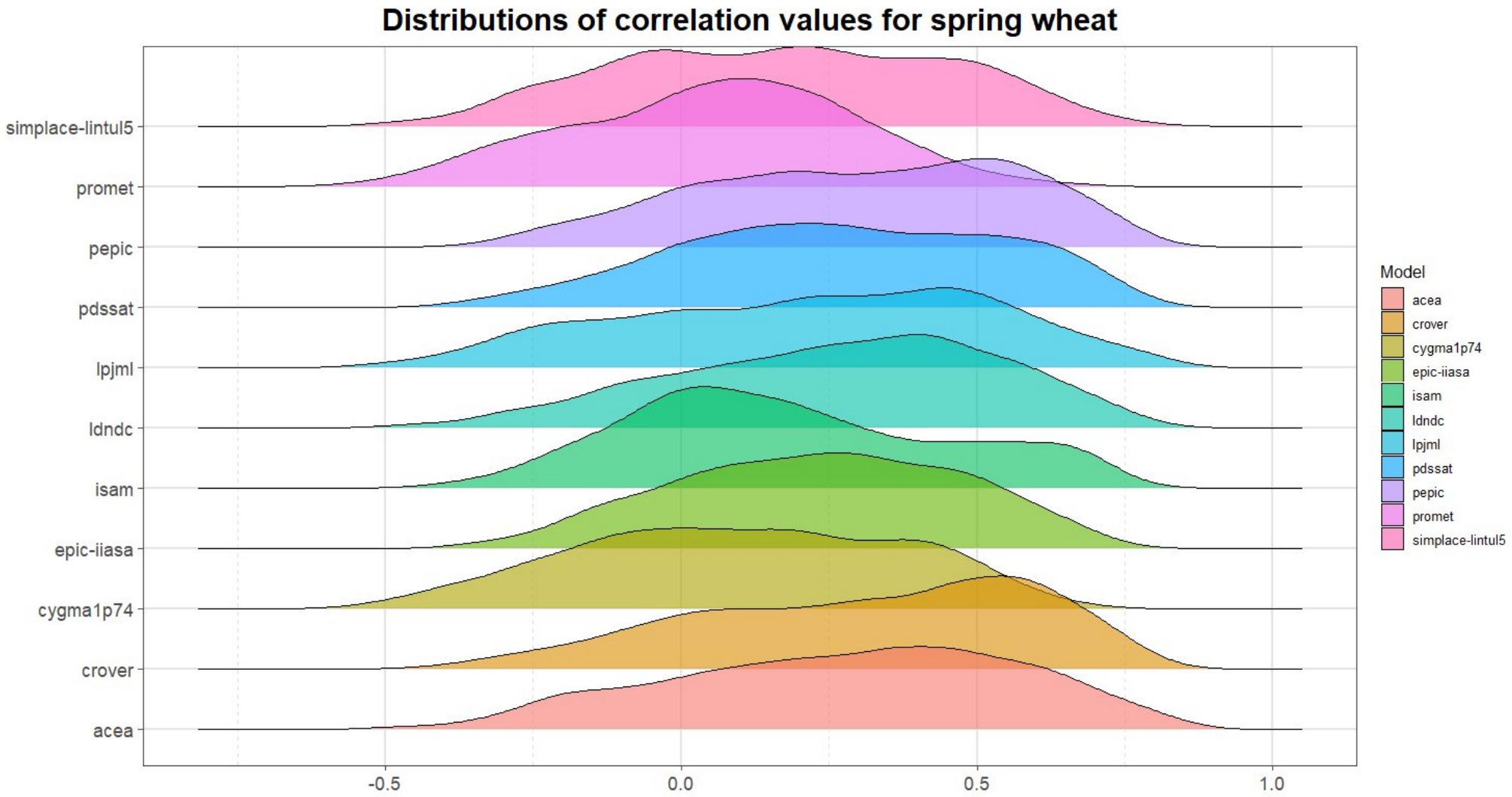
ibniz association

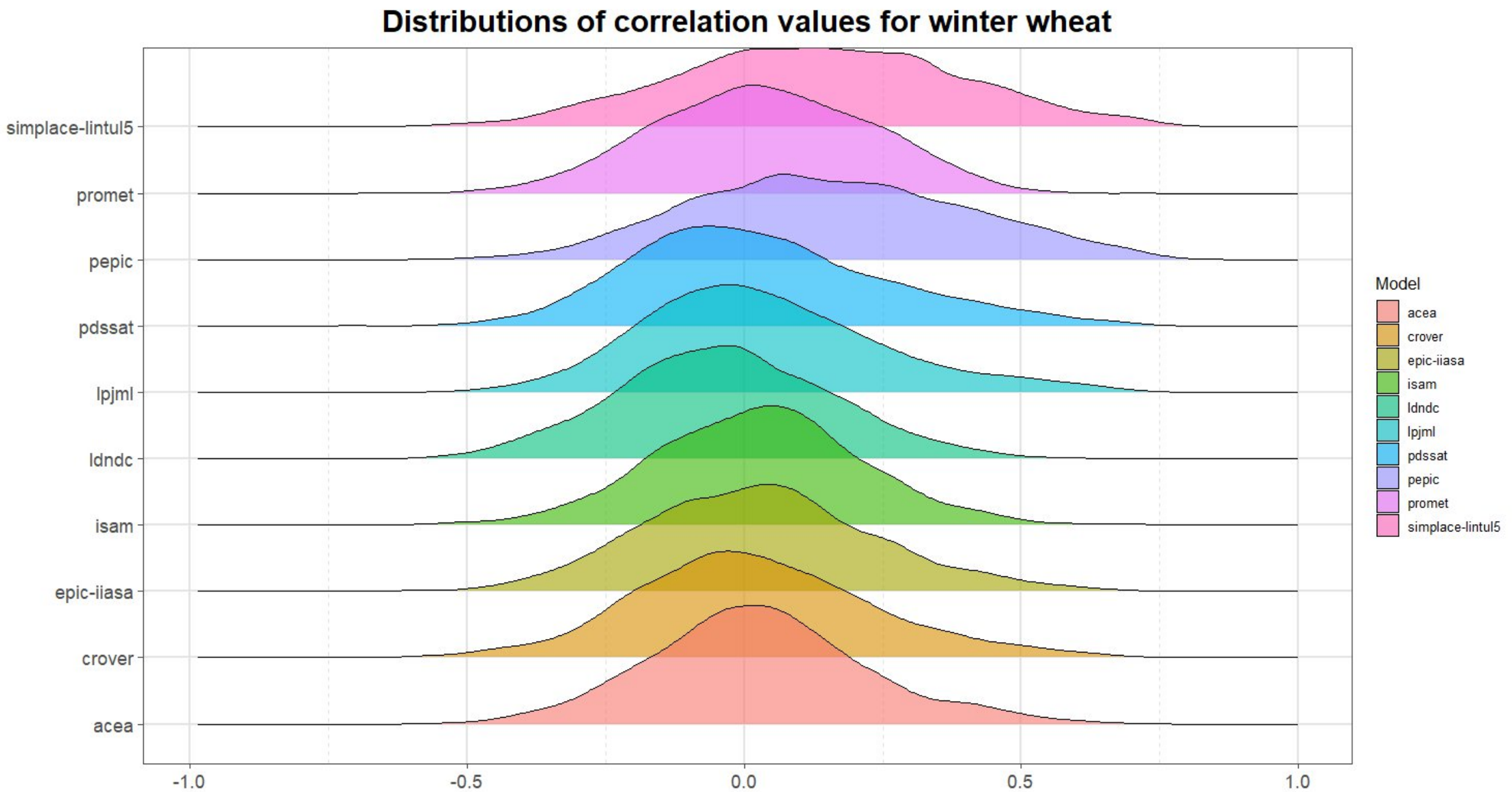
3. Raw data distributions of correlation values

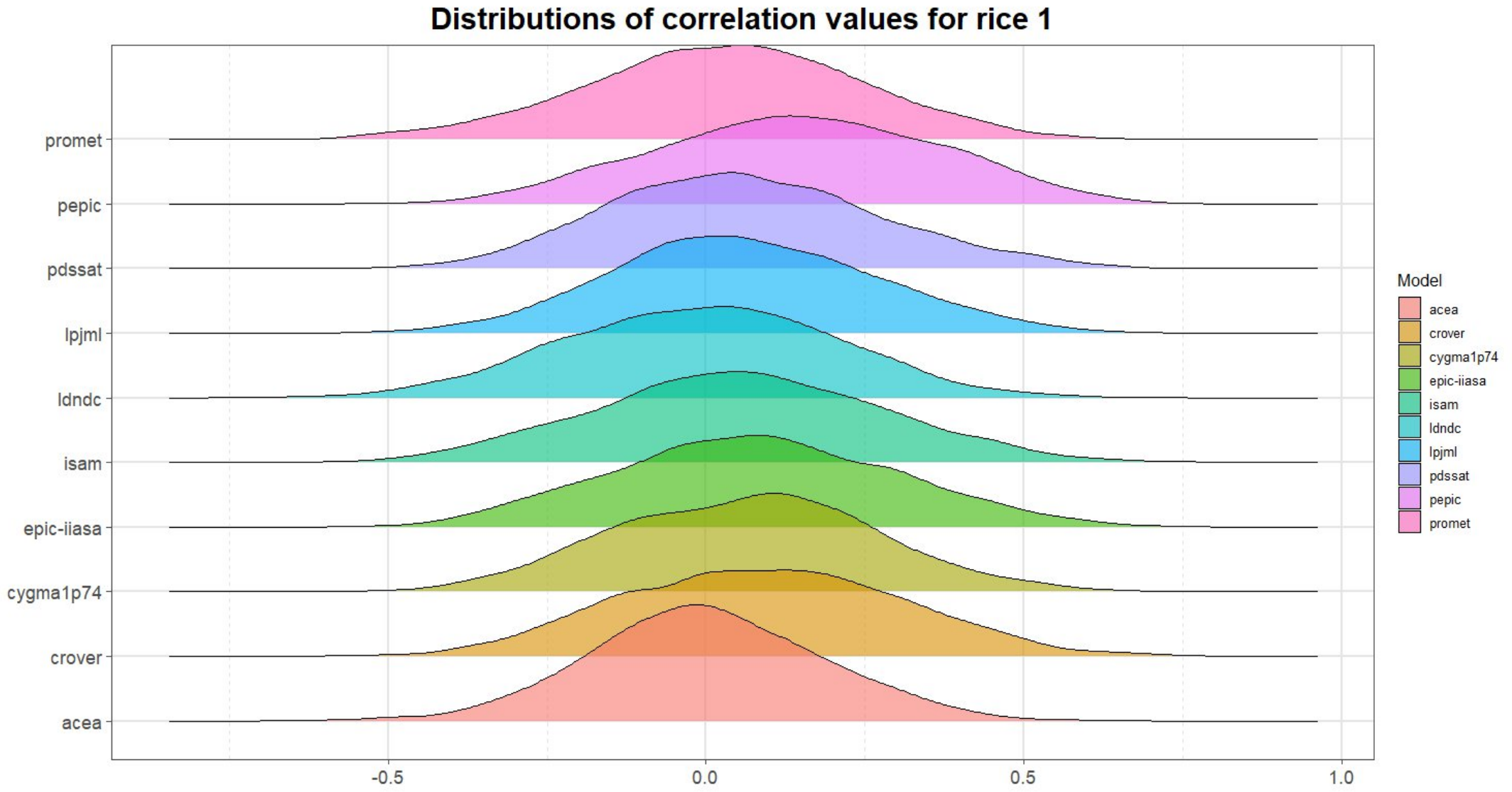


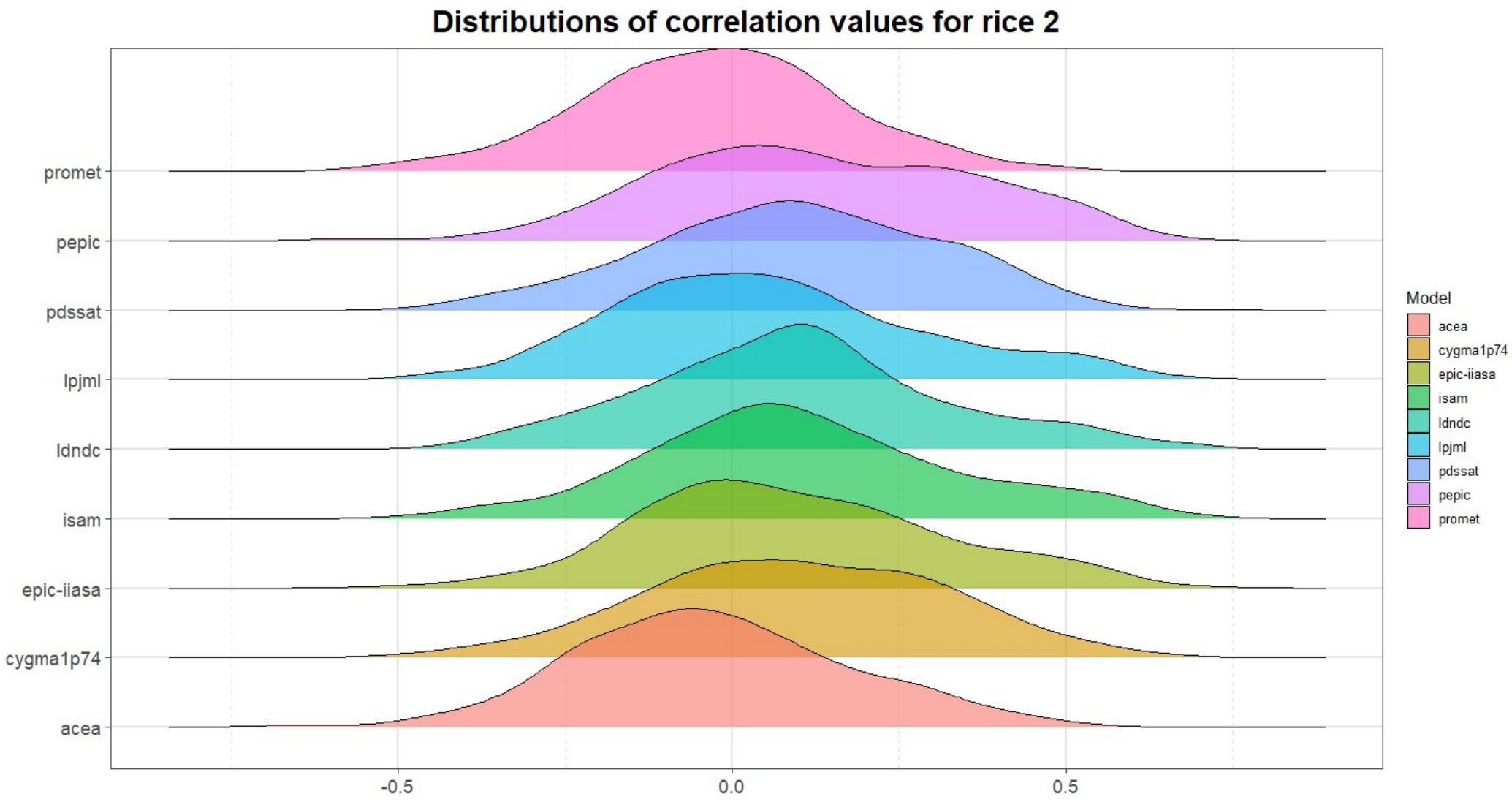







Distributions of correlation values for winter wheat
simplace-lintul5
promet
pepic
pdssat
lpjml
ldndc
isam
epic-iiasa
crover
acea
-1.0
-0.5
0.0
0.5
1.0
Model
acea
crover
epic-iiasa
isam
ldndc
lpjml
pdssat
pepic
promet
simplace-lintul5

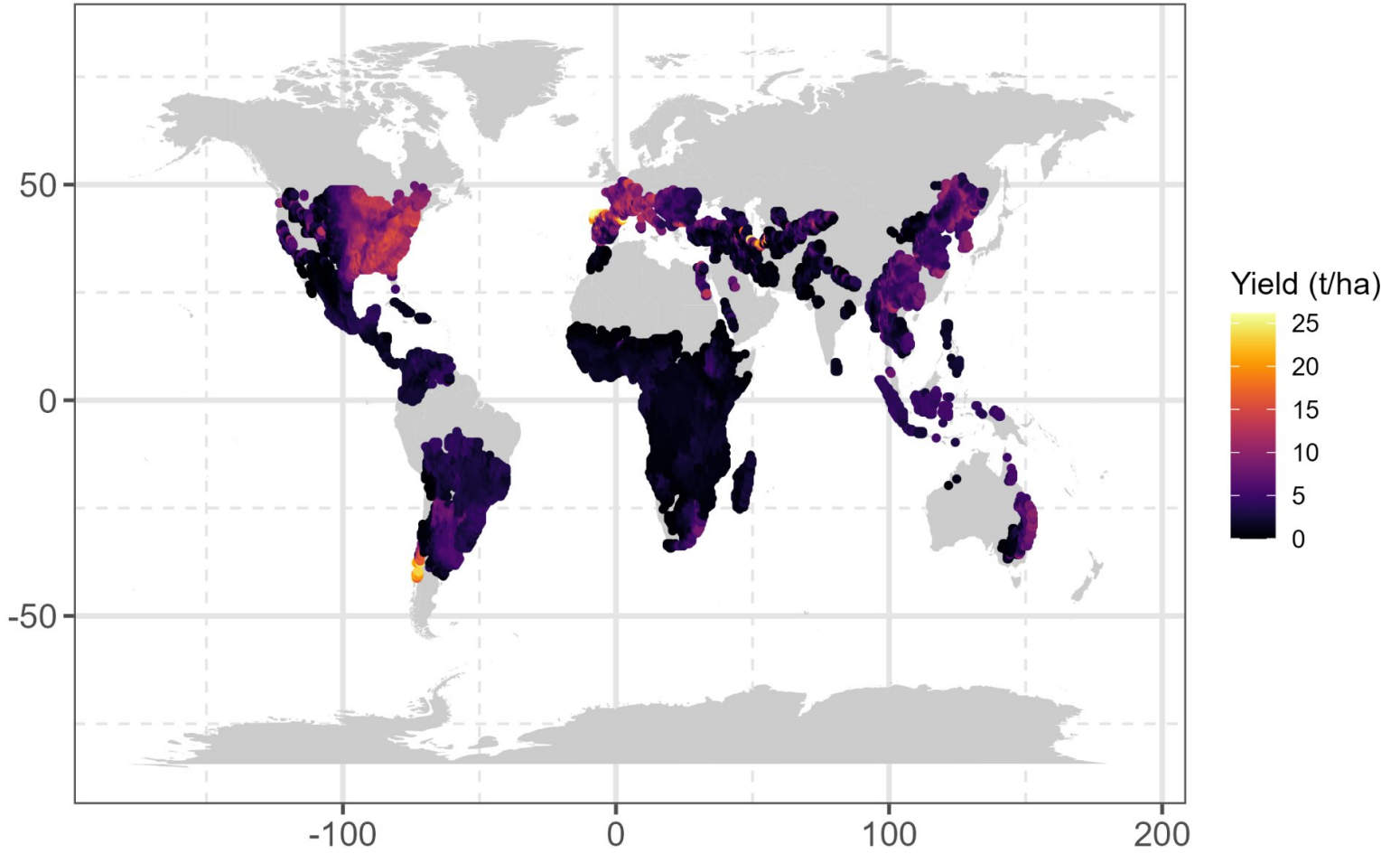




Appendix: yields benchmark data

Benchmark maize yield

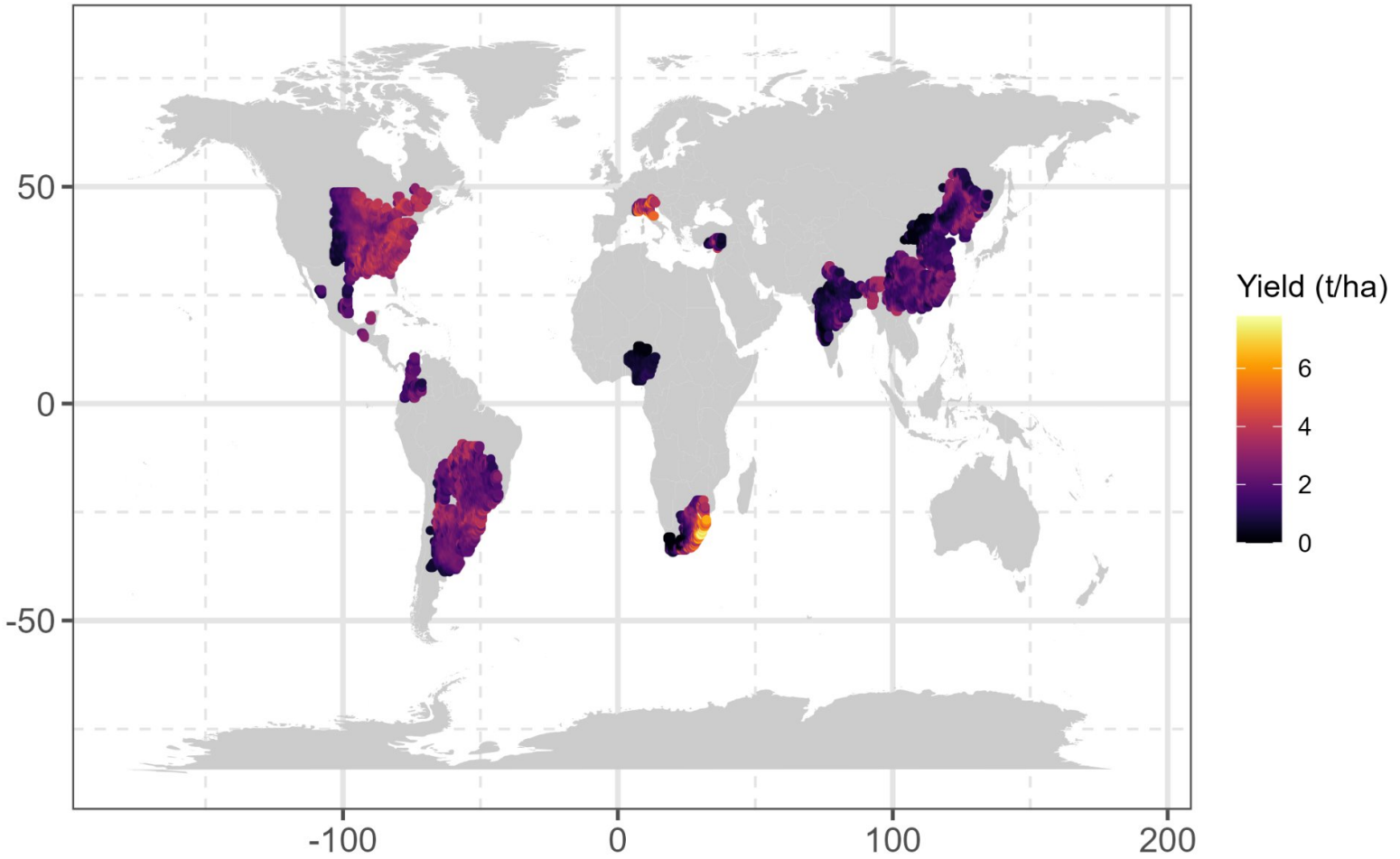
Averaged between 1981-2016



Iizumi, Toshichika (2019)

Benchmark soy yield

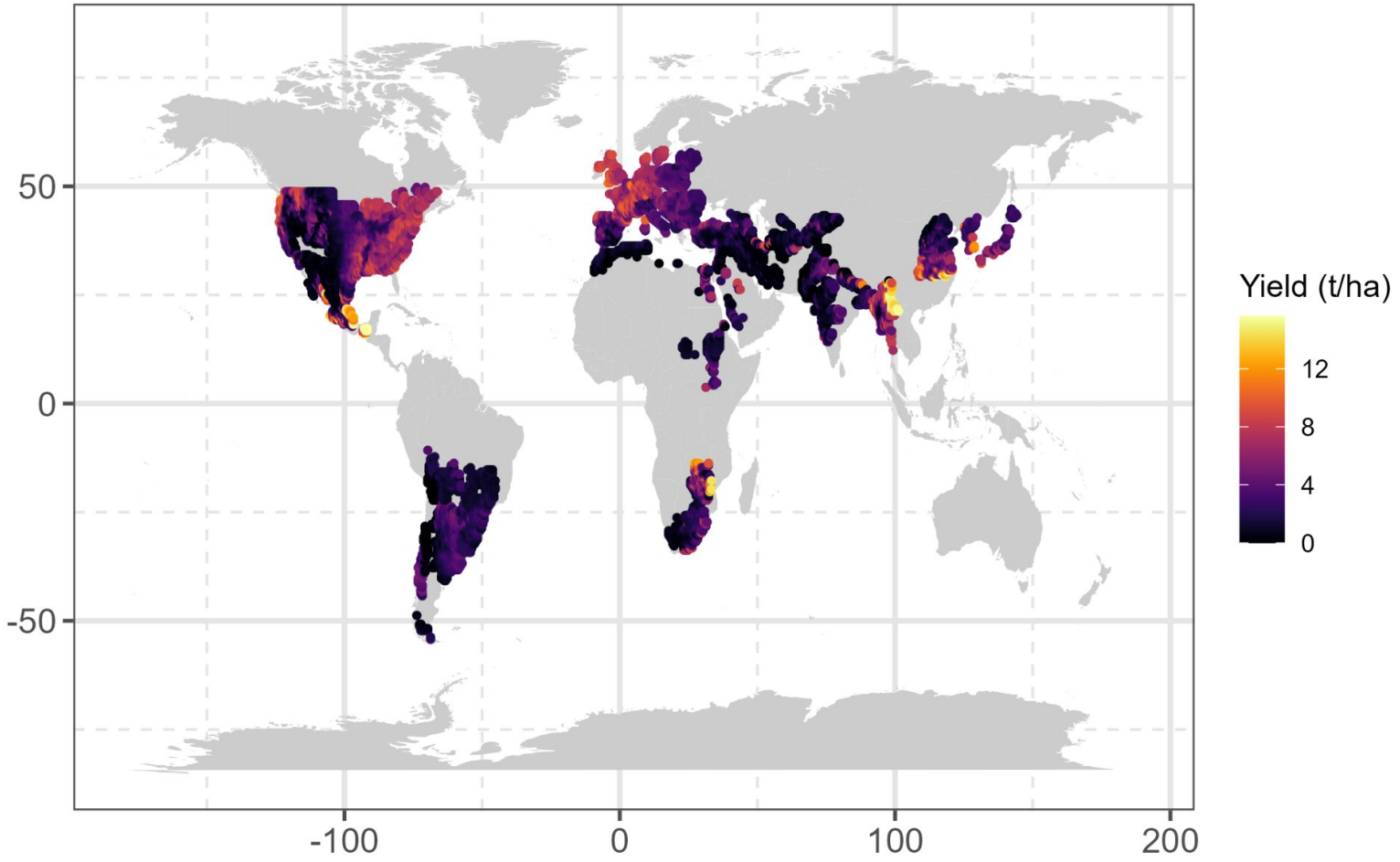
Averaged between 1981-2016



Iizumi, Toshichika (2019)

Benchmark winter wheat yield

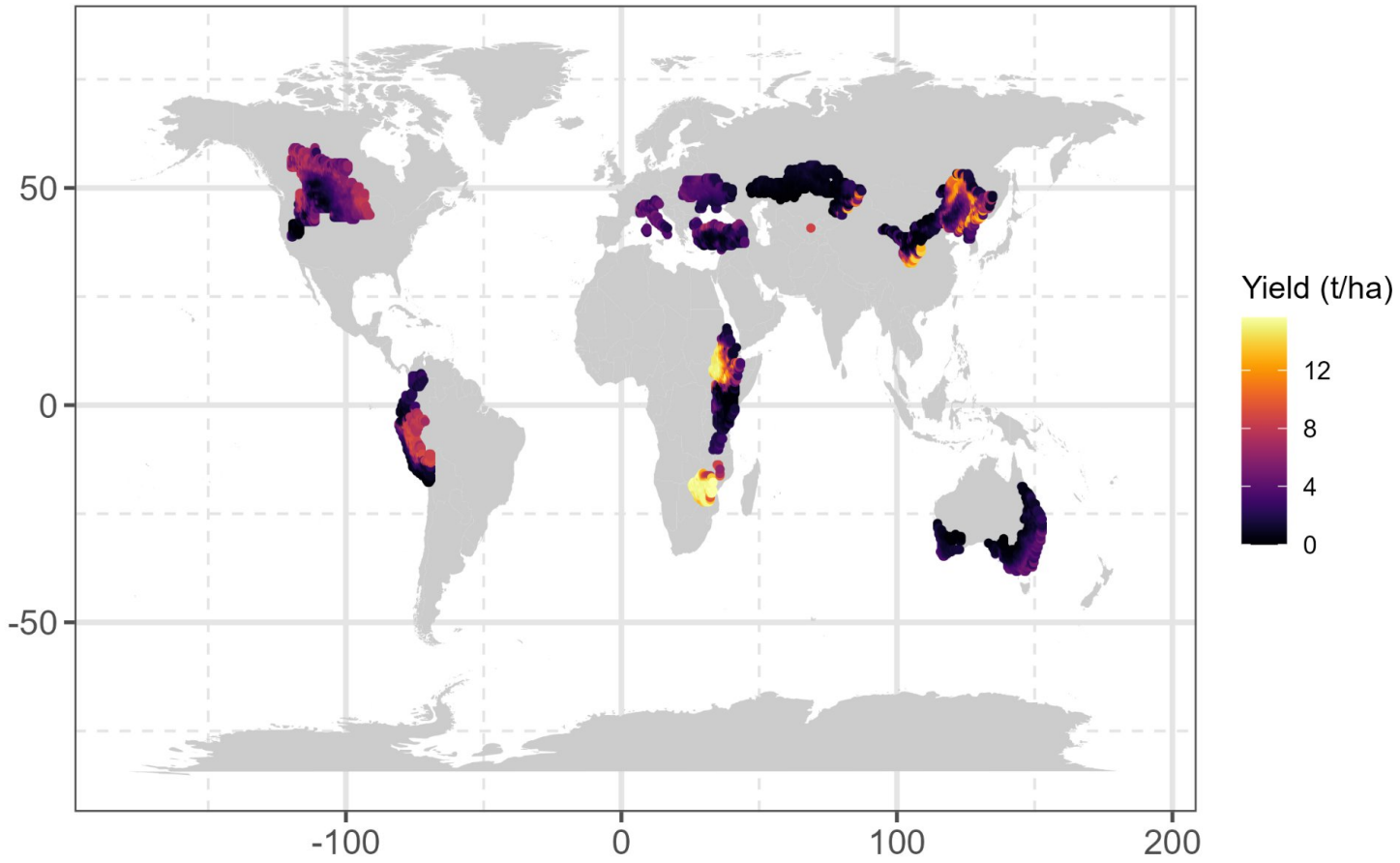
Averaged between 1981-2016



Iizumi, Toshichika (2019)

Benchmark spring wheat yield

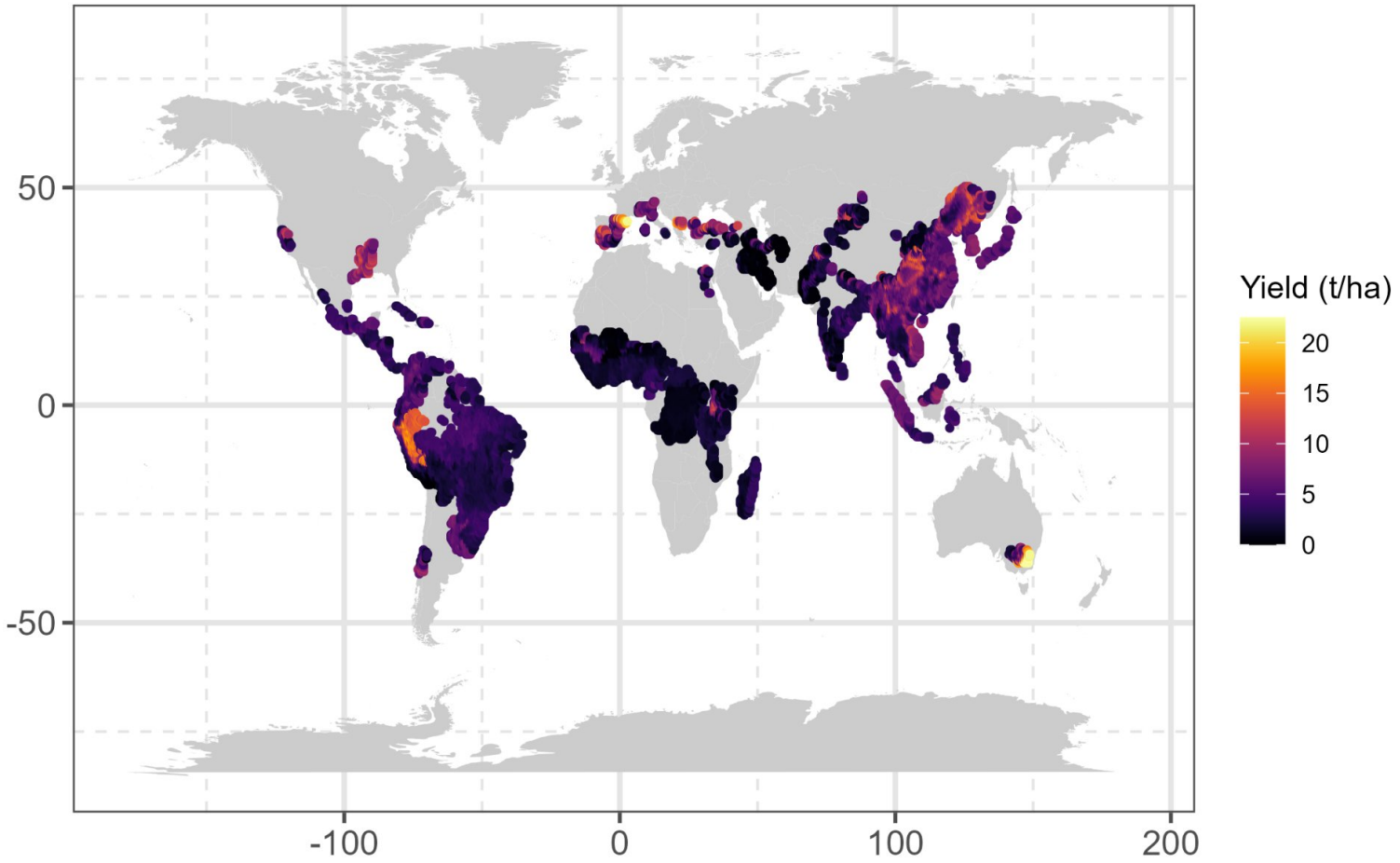
Averaged between 1981-2016



Iizumi, Toshichika (2019)

Benchmark rice 1 yield

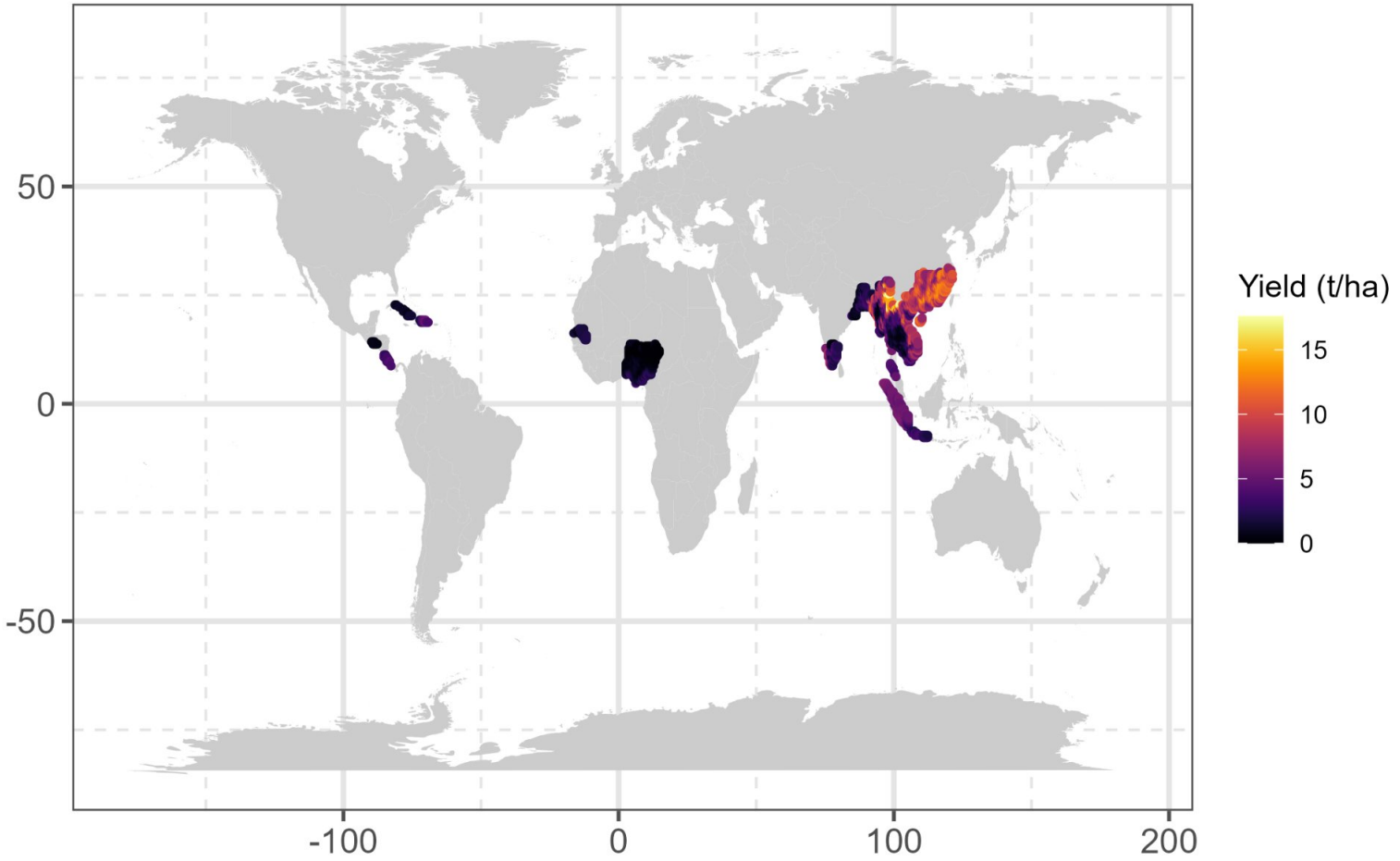
Averaged between 1981-2016



Iizumi, Toshichika (2019)

Benchmark rice 2 yield

Averaged between 1981-2016



Iizumi, Toshichika (2019)