Langley Research Center

NASA POSTDOCTORAL PROGRAM

Improving climate projections using "intelligent" ensembles

Noël C. Baker and Patrick C. Taylor

NASA Postdoctoral Program

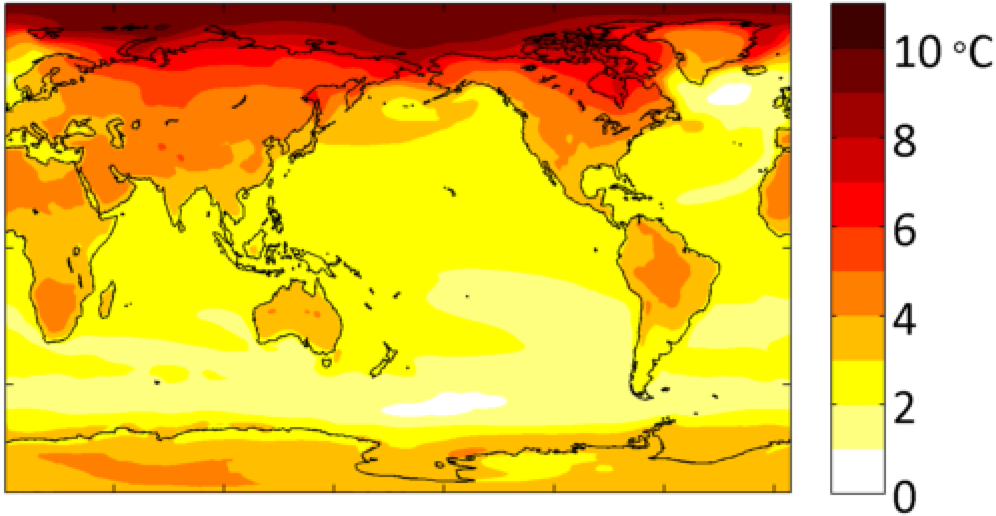
Presented at the AGU Joint Assembly in Montréal, Canada

May 5, 2015

The Intergovernmental Panel on Climate Change (IPCC) predicts that

21st-century global surface temperature change is likely to exceed 2°C

21st-century temperature trend
(RCP 8.5 multi-model ensemble mean)

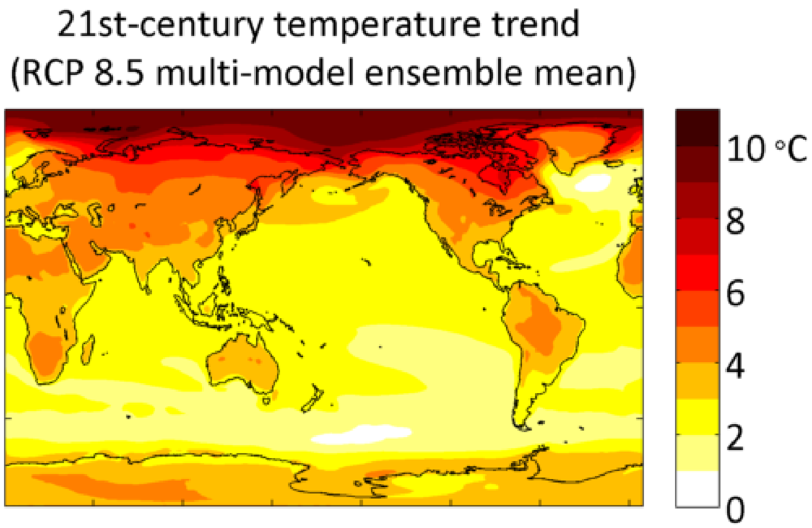


IPCC prediction comes from ensemble of global climate models: CMIP5 (Coupled Model Intercomparison Project)

CMIP5 Model

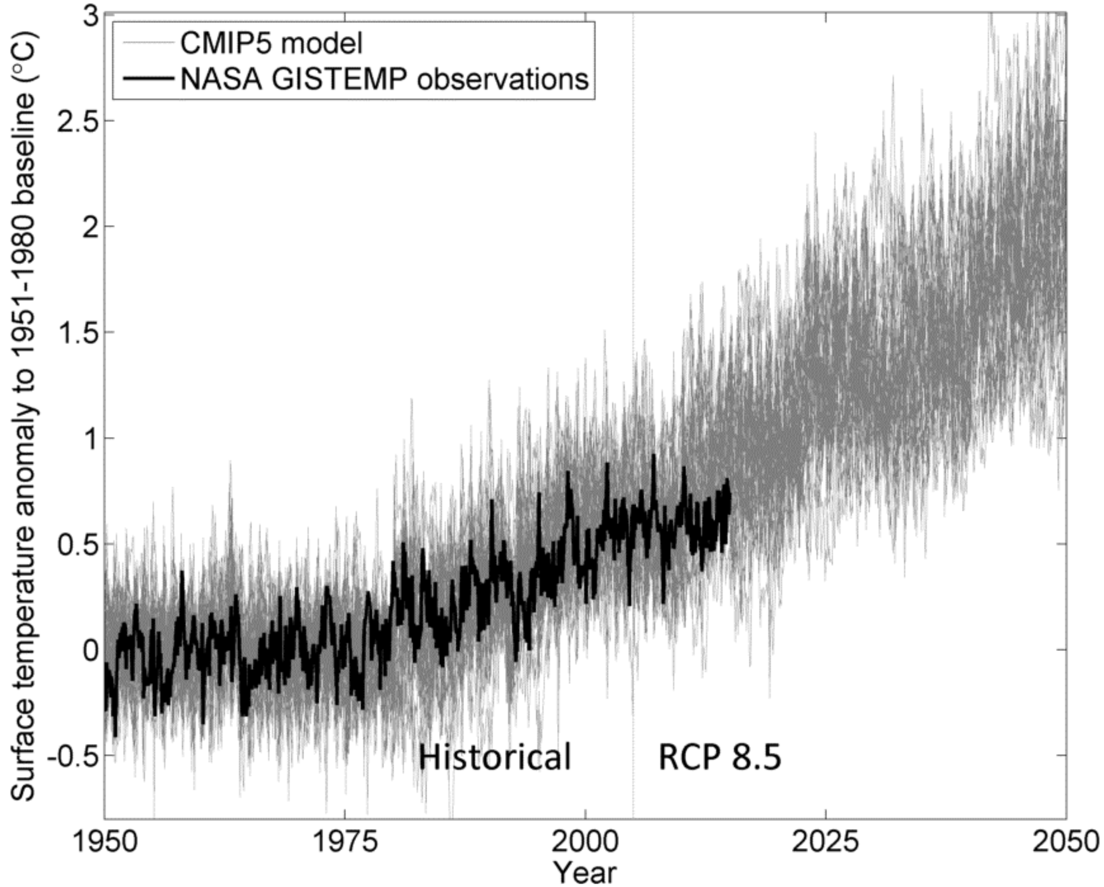
BCC-CSM1.1
BCC-CSM1.1.m
CanESM2
CCSM4
CESM1-BGC
CESM1-CAM5
CESM1-WACCM
CMCC-CESM
CMCC-CM
CMCC-CMS
CNRM-CM5
ACCESS1.0
ACCESS1.3
CSIRO-Mk3.6.0
FGOALS-g2
FIO-ESM
GFDL-CM3
GFDL-ESM2G
GFDL-ESM2M
GISS-E2-H
GISS-E2-H-CC
GISS-E2-R
GISS-E2-R-CC
HadGEM2-AO
HadGEM2-CC
HadGEM2-ES
INM-CM4
IPSL-CM5A-LR
IPSL-CM5A-MR
IPSL-CM5B-LR
MIROC5
MIROC-ESM
MPI-ESM-LR
MPI-ESM-MR
NorESM1-M
NorESM1-ME

Models are averaged together to make climate predictions

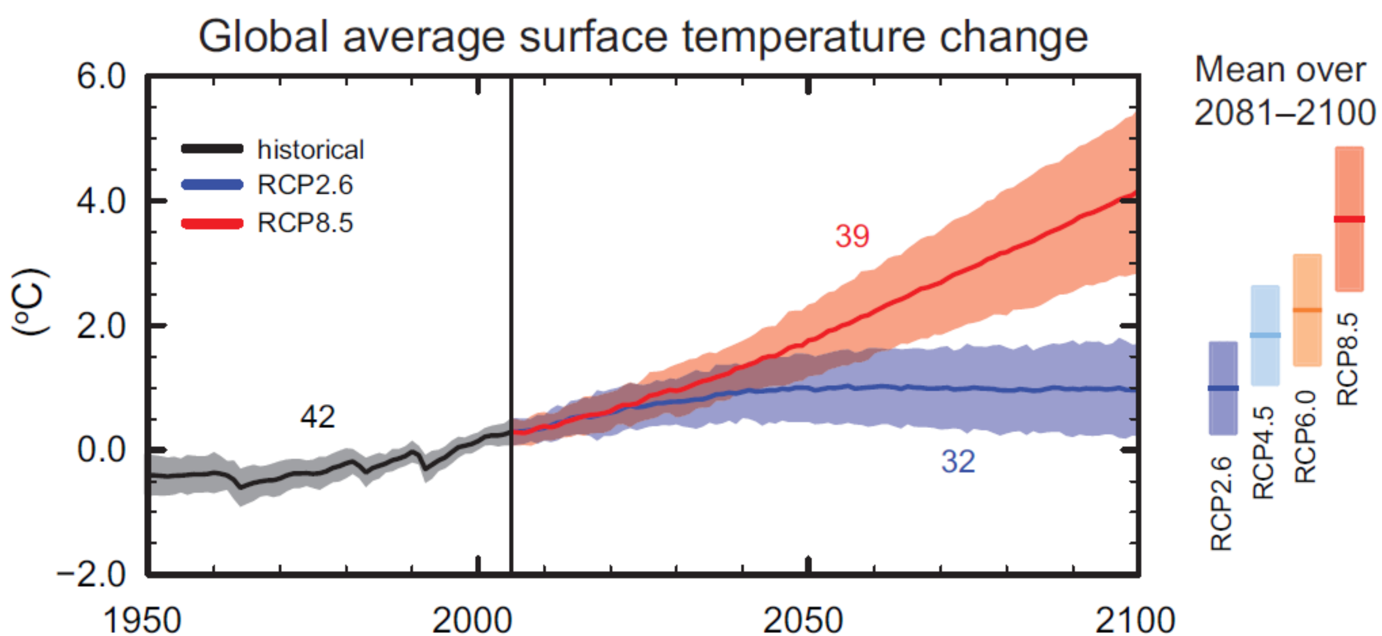


But models can have a large spread in predictions, and individual models can perform very differently from observations

Global surface temperature anomaly, from 35 CMIP5 models



The traditional **Multi-Model Ensemble (MME)** Approach uses the model mean to provide an improved "best estimate" forecast

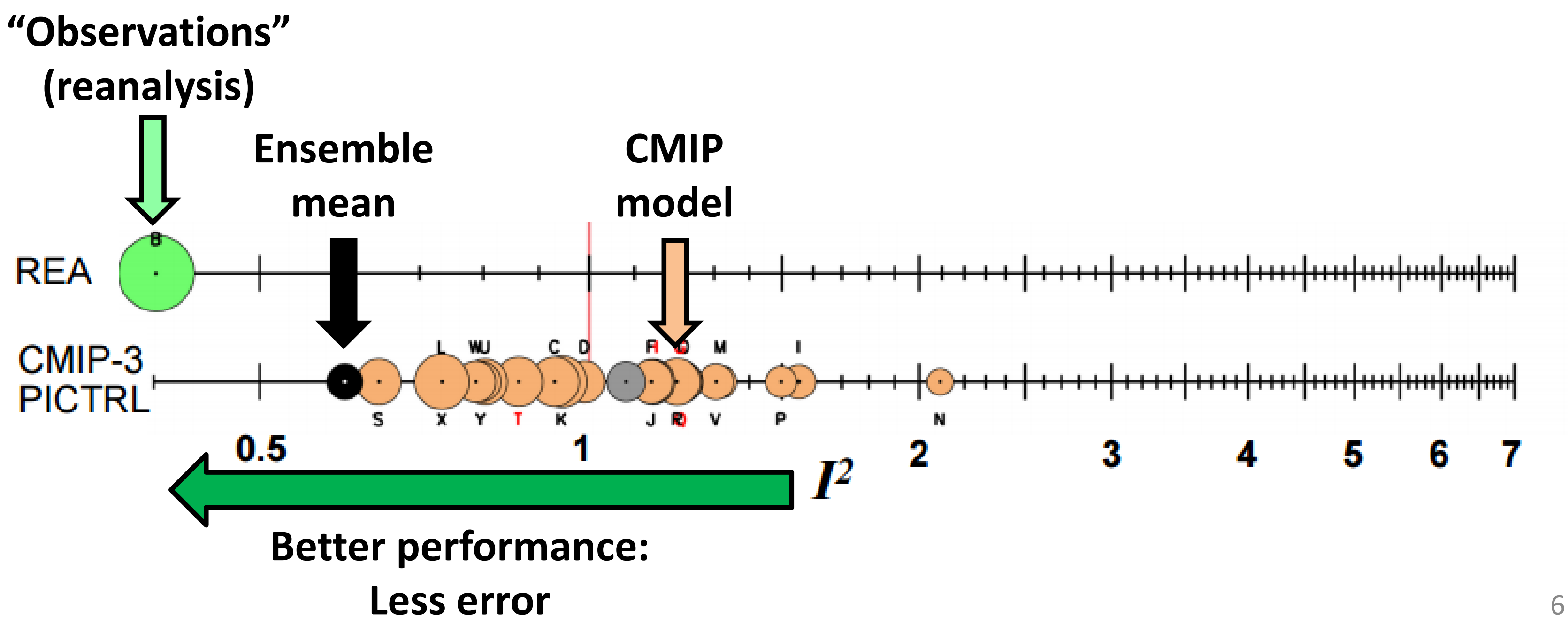


IPCC AR5 Figure SPM.7

The multi-model ensemble generally performs better than individual models

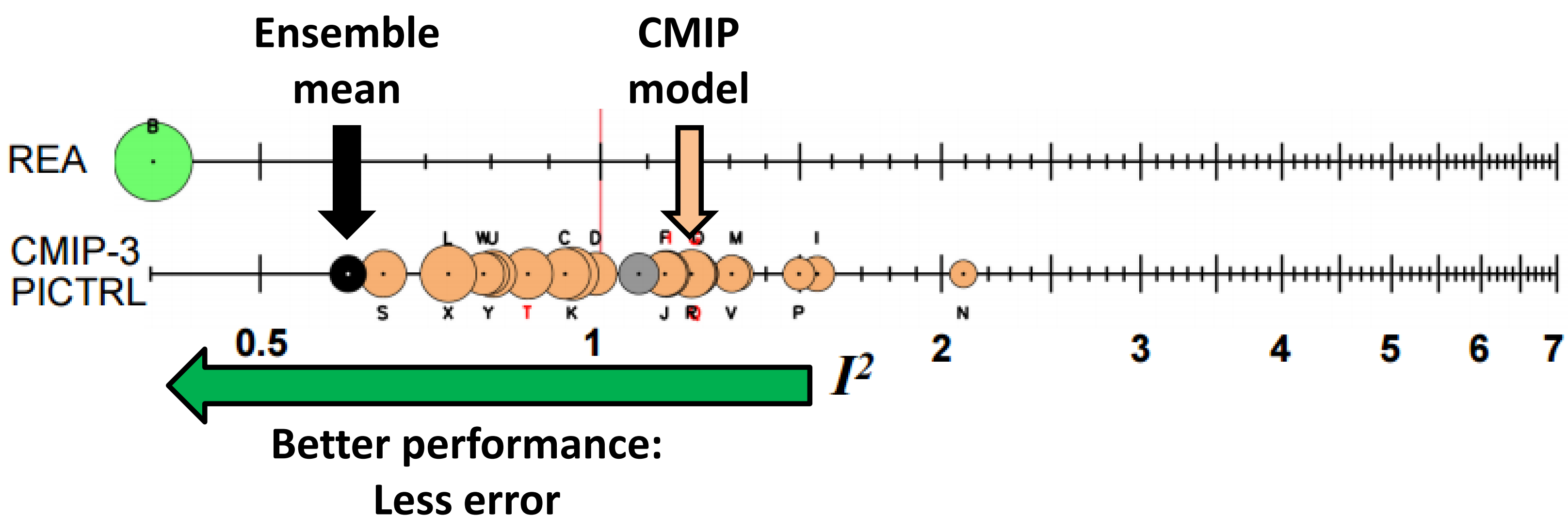
Example: I^2 *performance index* (Reichler and Kim 2008)

Calculates aggregated model errors relative to NCEP/NCAR reanalyses for multiple climate variables



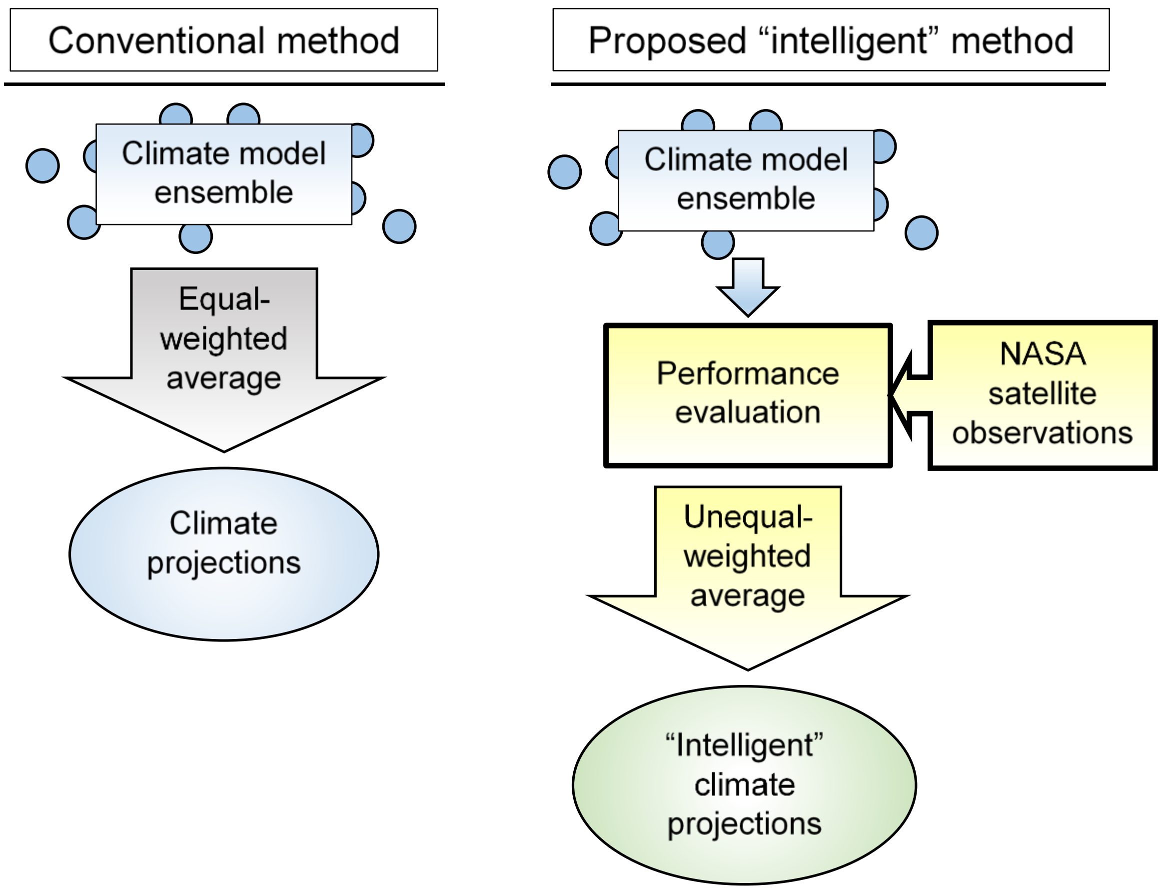
Some models perform better than others:

Can we use knowledge of model performance for a better way to combine model output?



The "intelligent ensemble" method

for creating multi-model ensemble projections



Project goal:

determine future climate state using observed current climate and an ensemble of models

$$f(x_{obs}) = \Delta x$$

Observed climate

Future climate state

Previous work has explored model performance and ensemble-weighting metrics

Several examples:

- Model subsets (USGCRP 2009)
- Performance metrics (Gleckler et al. 2008, Reichler and Kim 2008)
- Constrained projections (Tett et al. 2013; Giorgi and Mearns 2003)
- Weighted future trends (Boe et al. 2009)
- Bias correction (Baker and Huang 2012)

"The community would benefit from a larger set of proposed methods and metrics" (Knutti 2010)

New climate model performance metrics are tested:

representative of energy budget processes

Radiation budget quantities

- Top-of-atmosphere (TOA) longwave (LW) and shortwave (SW) radiation fluxes
- Surface LW and SW radiation fluxes
- Surface temperature

Statistical tests

- F-test for equal variances
- Kolmogorov-Smirnov test for distribution similarity
- Earth Mover's Distance (EMD): test for area of distribution overlap
- Local Variance: test variance of first difference time series (Baker and Taylor 2015)

New process-oriented metrics

$$\frac{\delta\, TOA\ Radiation\ flux}{\delta\, Surface\ temperature}$$: represent interannual-timescale radiative feedbacks

Model data: 32 CMIP5 models http://pcmdi9.llnl.gov/

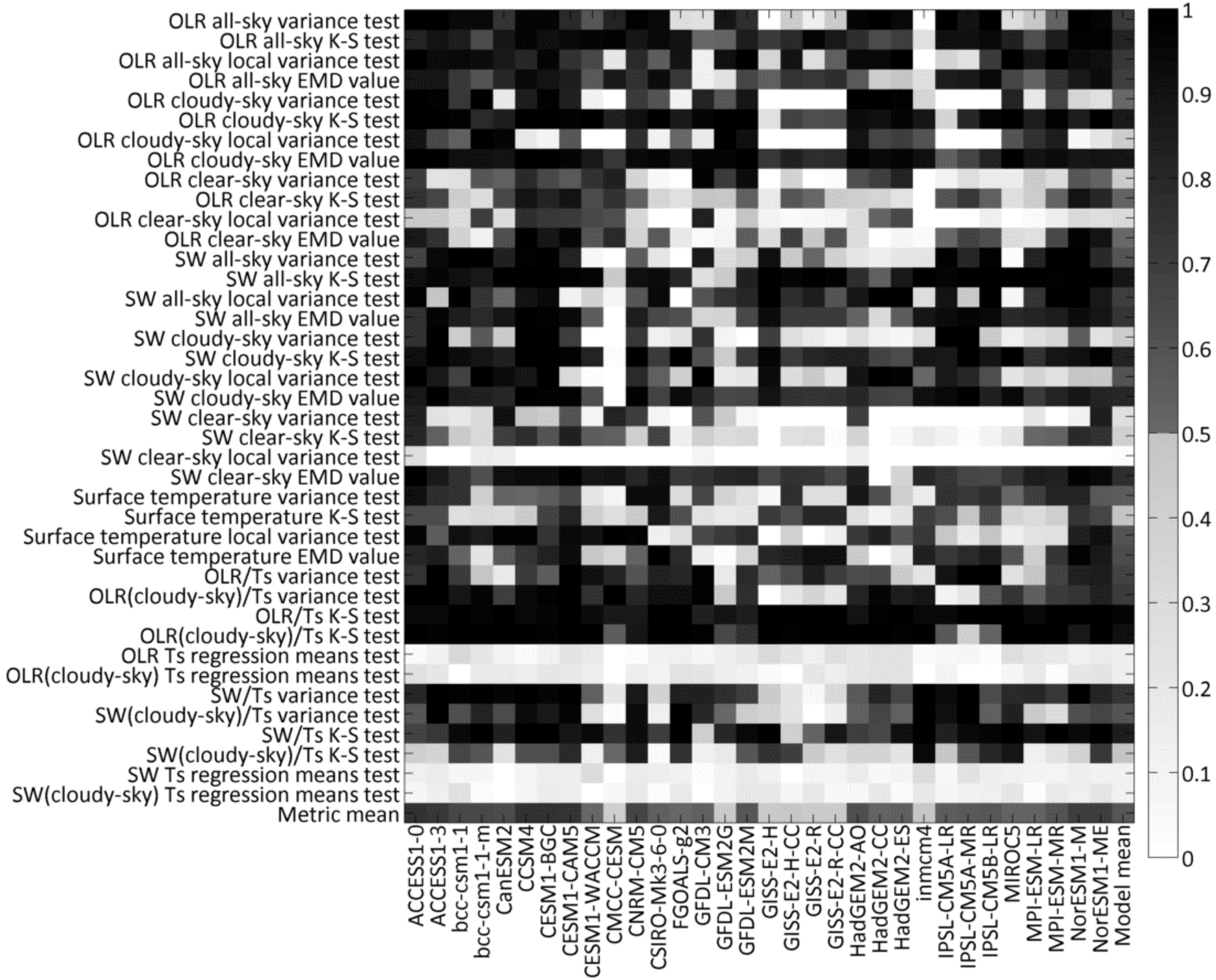
- ‘Pre-Industrial Control’ simulations (monthly mean, 100 years) to create metric weights
- ‘RCP 8.5’ future simulations (monthly mean, 2081-2100 minus 2011-2030 to produce 21st-century trends)

Observational datasets:

NASA CERES EBAF-TOA and surface monthly global-mean (full data record: 03/2000 - 05/2014)
http://ceres.larc.nasa.gov/

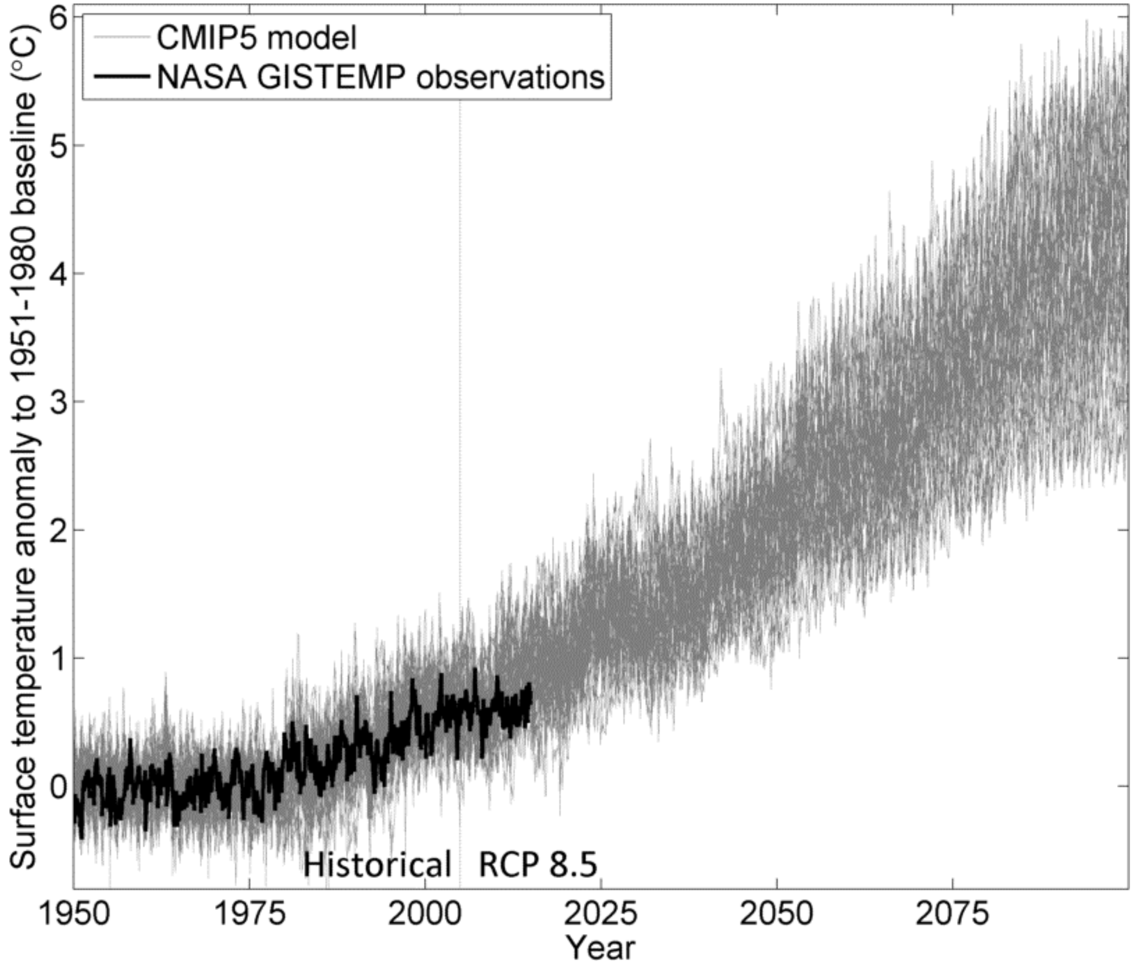
NASA GISS Surface Temperature Analysis (GISTEMP)
http://data.giss.nasa.gov/gistemp/

Step 1: Test model quality with selected metrics



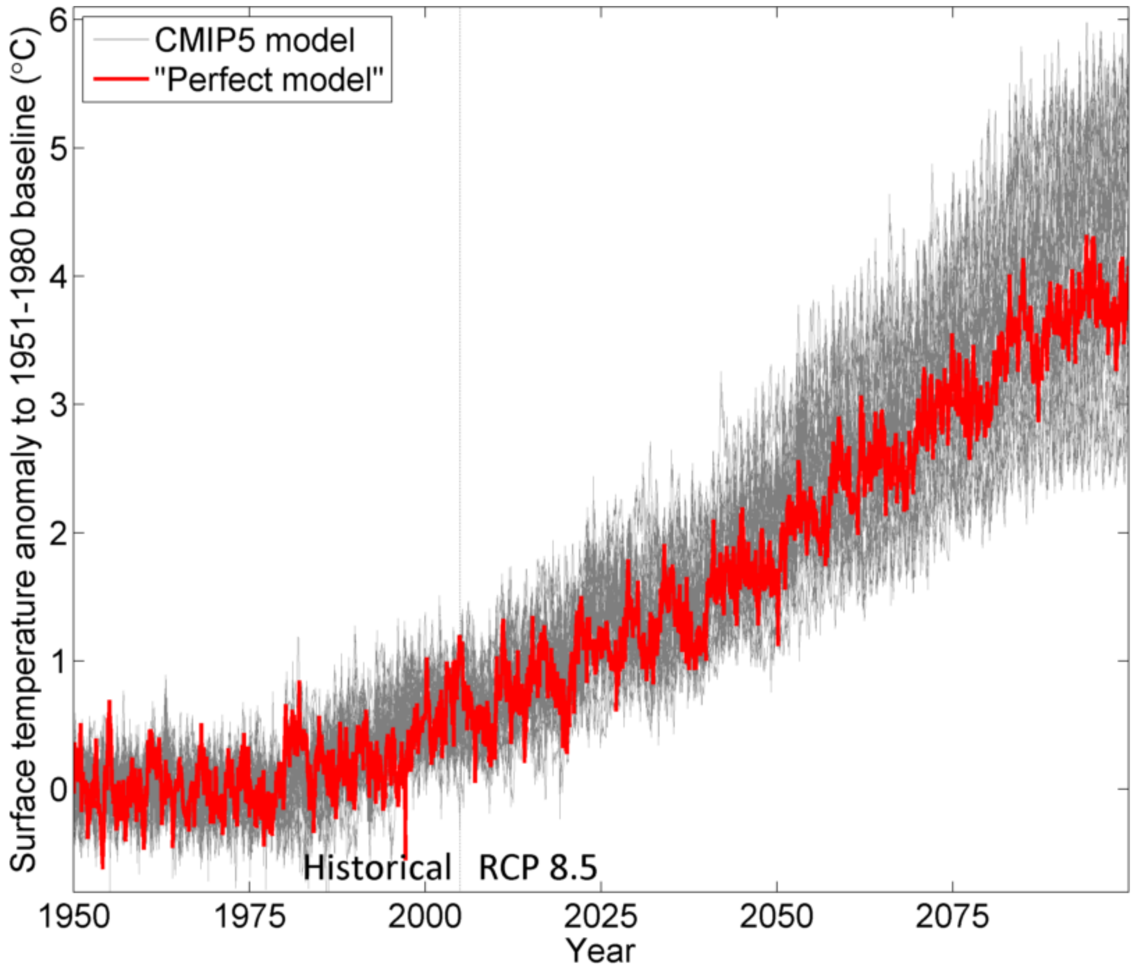
Step 2: Using skill-subset of models, apply "perfect model" approach (Räisänen and Palmer 2001)

Create set of potential "Earths" each with a continuous time series of observations

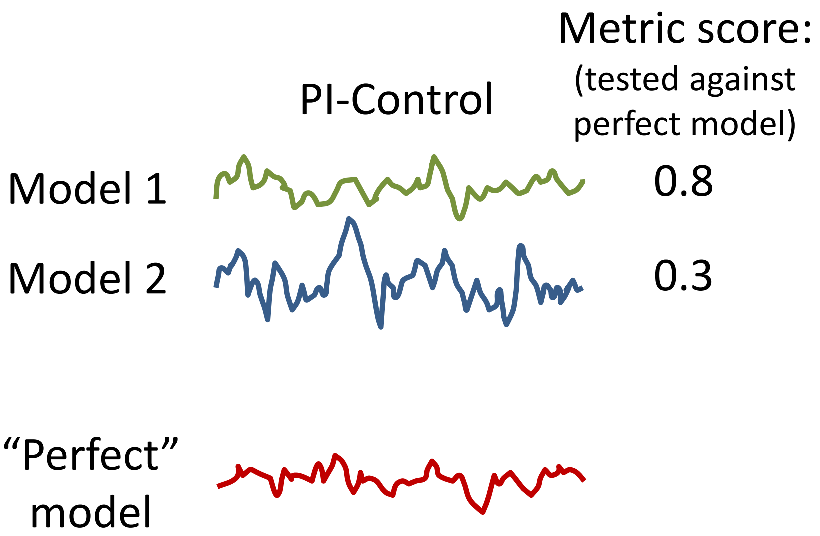


Step 2: Using skill-subset of models, apply "perfect model" approach (Räisänen and Palmer 2001)

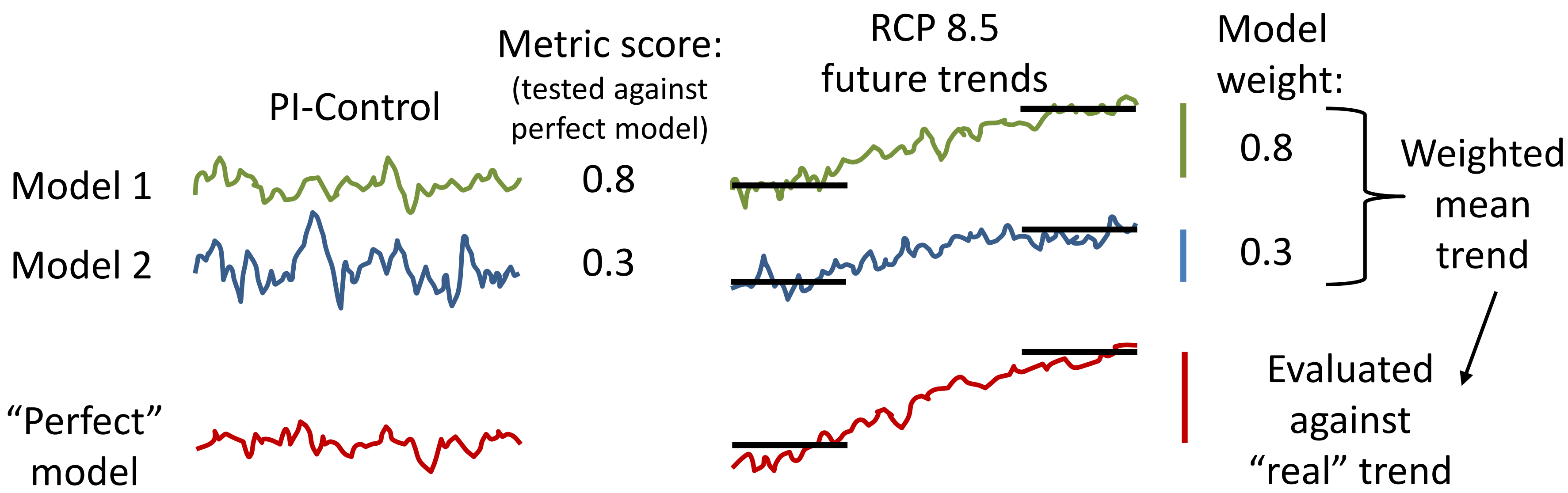
Create set of potential "Earths" each with a continuous time series of observations



- For each "perfect model" (potential Earth), the performance metrics are tested on one simulation (Pre-Industrial Control), then applied to a different simulation (RCP 8.5 future trends), <u>linking present-day quality with a future state.</u>
- Metric values are used as model weights to create unequal-weight ensemble mean trends.



- For each "perfect model" (potential Earth), the performance metrics are tested on one simulation (Pre-Industrial Control), then applied to a different simulation (RCP 8.5 future trends), linking present-day quality with a future state.
- Metric values are used as model weights to create unequal-weight ensemble mean trends.



- Metric-weighted ensemble means which have the least error compared with the "perfect model" are considered the best-performing metrics.

Reichler and Kim (2008) *I^2 performance index* is used to compare metric quality

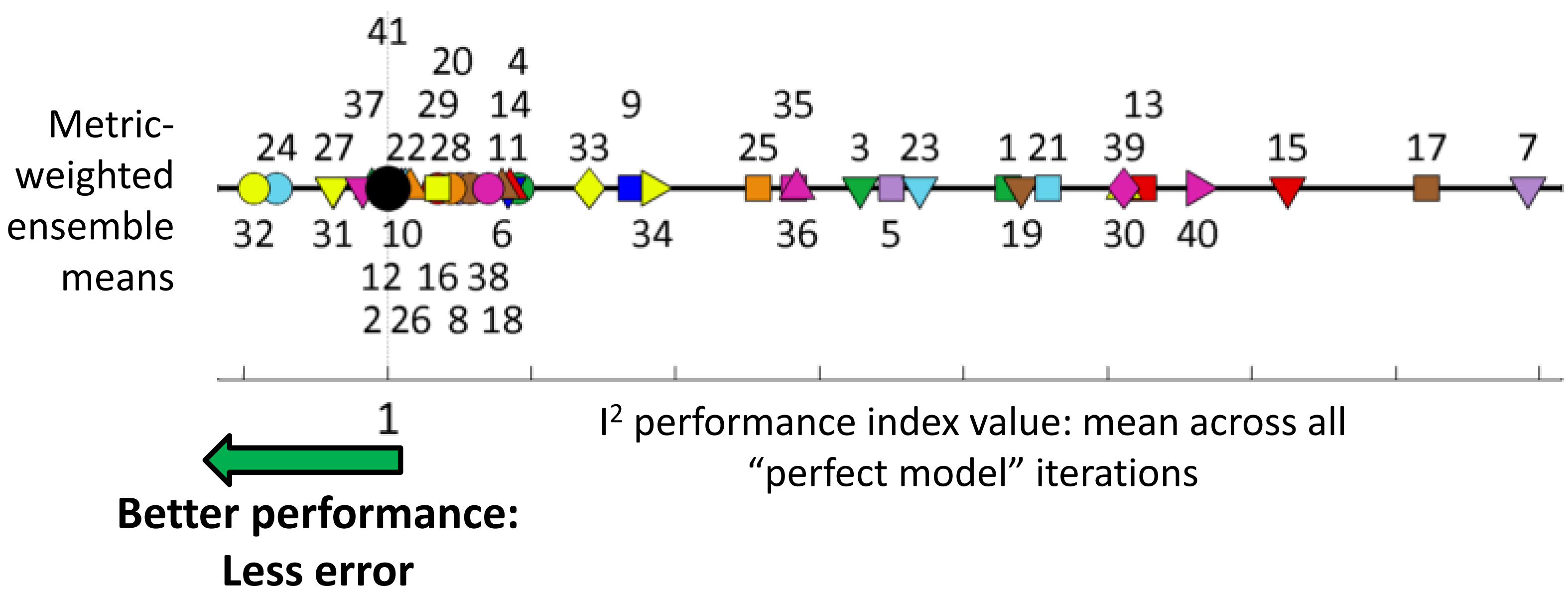
Metrics which perform well indicate a physical link between present-day model quality and reliability of projected trends

Best-performing metrics:

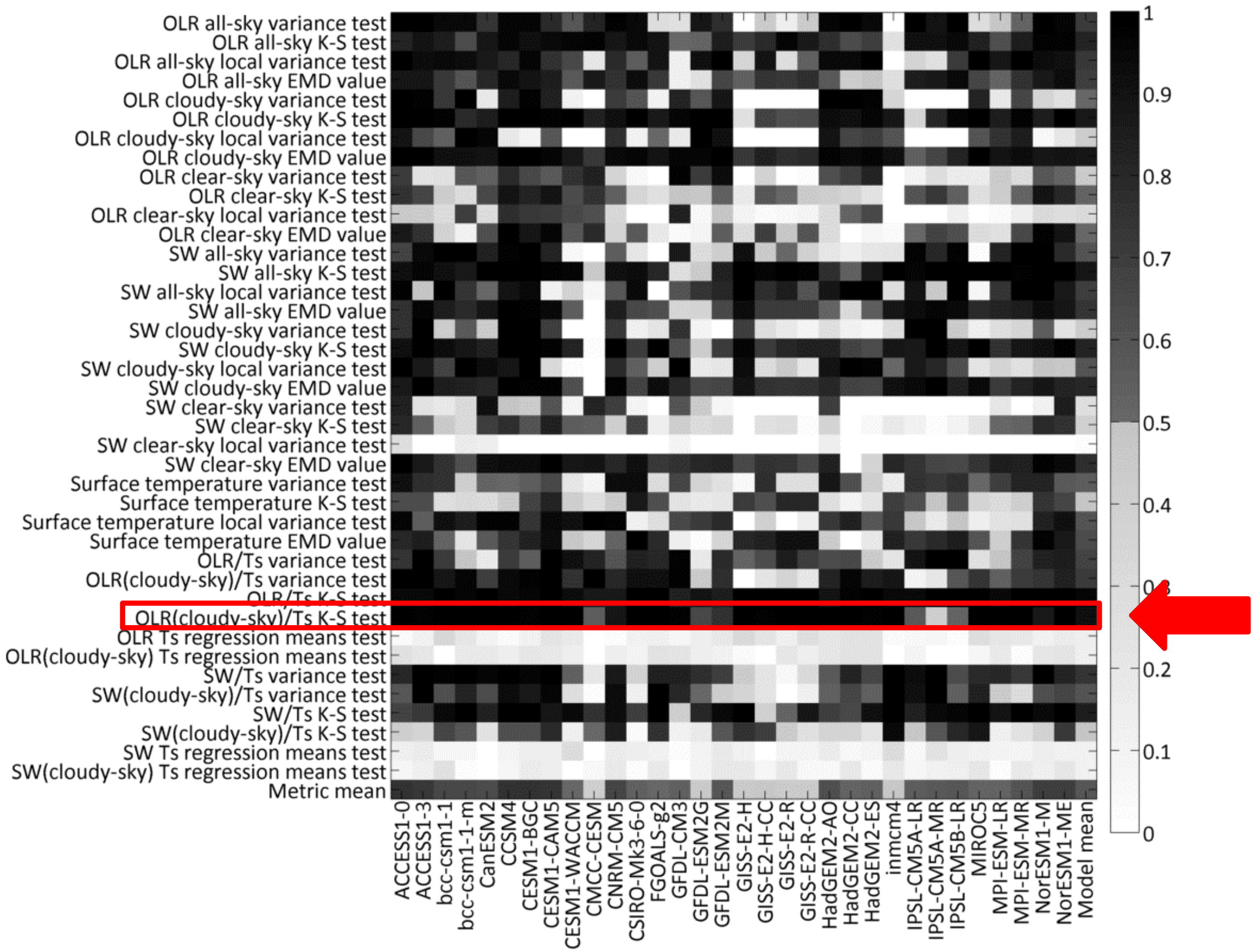
- 32 OLR(cloudy-sky)/Ts K-S test
- 24 SW clear-sky EMD value
- 31 OLR/Ts K-S test

Worst-performing metrics:

- 7 OLR cloudy-sky local variance test
- 17 SW cloudy-sky variance test
- 15 SW all-sky local variance test



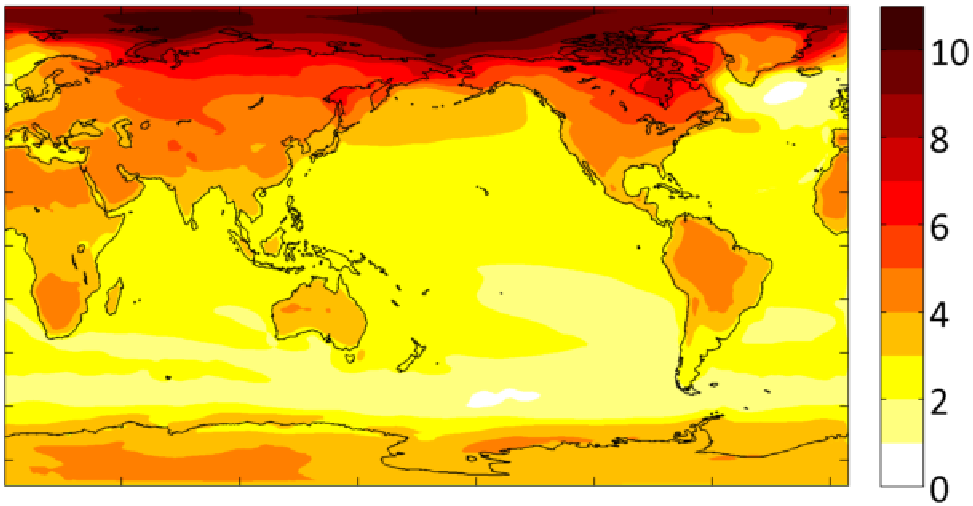
Step 3: Using best-performing metric, create new "intelligent ensemble" projections



Use metric values as model weights to create unequal-weighted mean projections

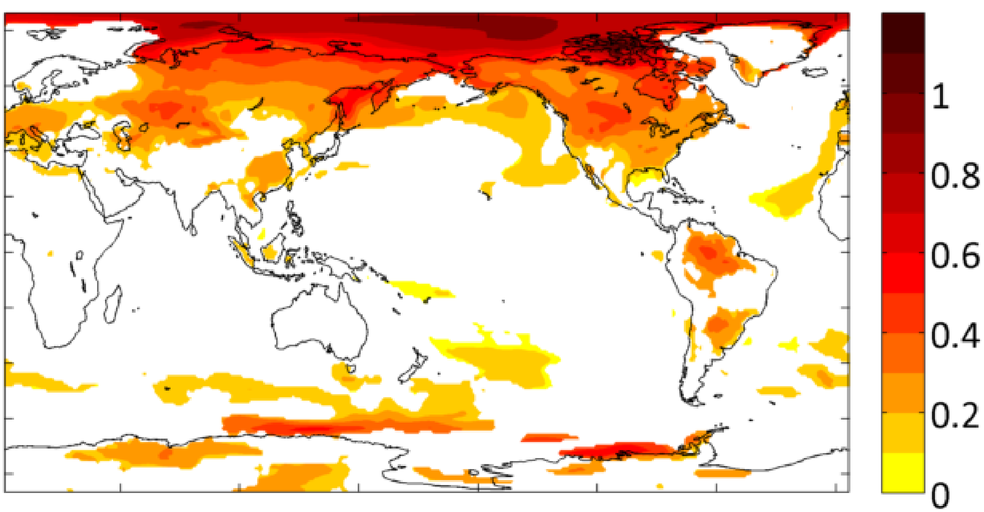
Results: new 21st-century projections (surface temperature)

"Intelligent" ensemble mean temperature trend (◦C)



Global-mean surface temperature trend: 3 °C (0.1 °C higher than the traditional equal-weight MME)

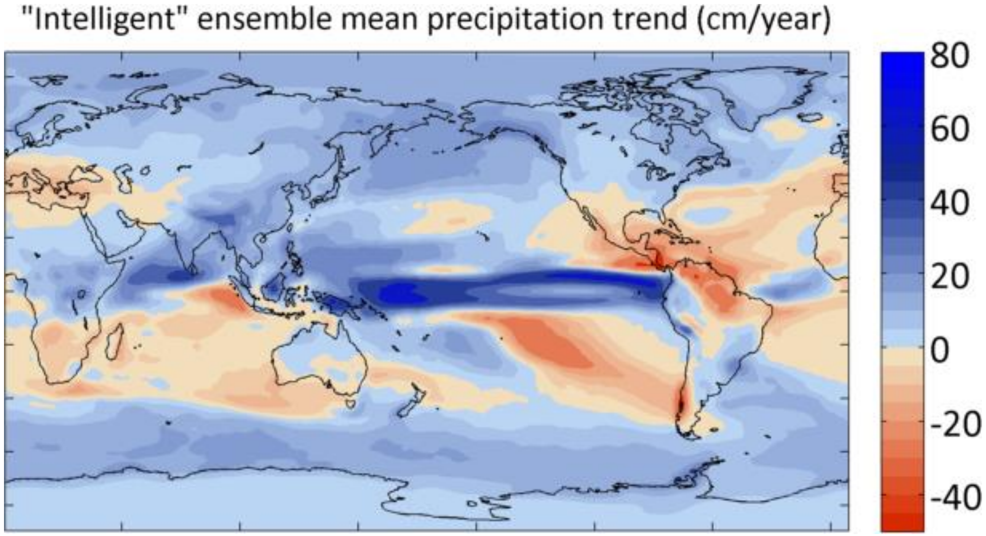
Difference between "Intelligent" and Equal-weight ensemble means (◦C)

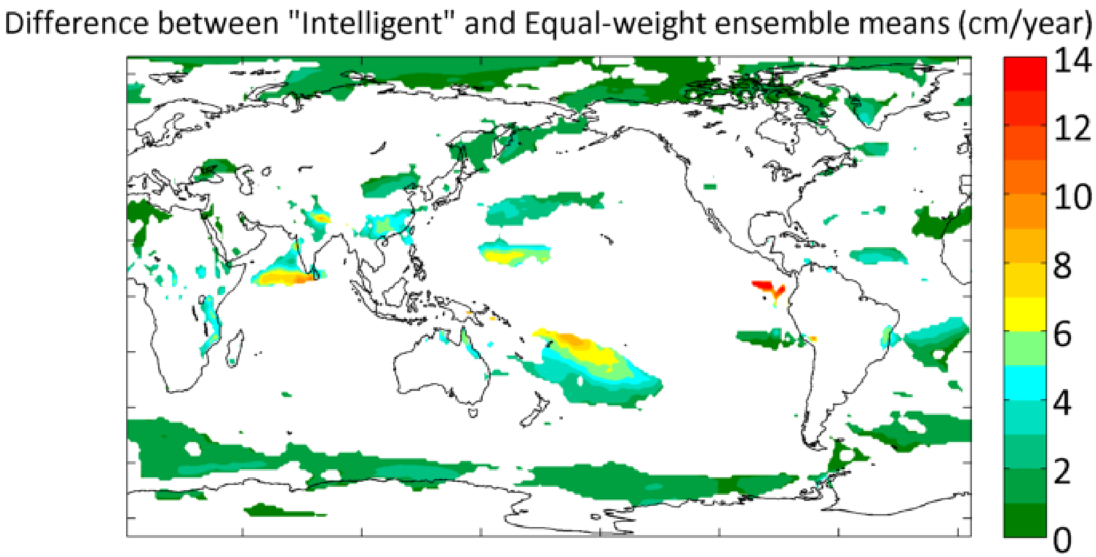


The “Intelligent Ensemble” predicts about 10% higher regional surface temperature increases than MME

Contours are shaded only where the difference is statistically significant

Results: new 21st-century projections (precipitation)



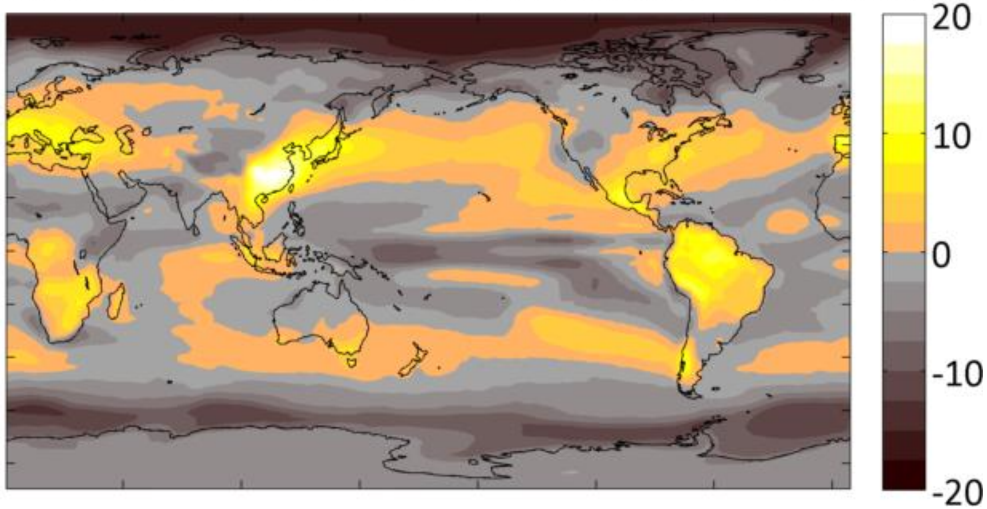


The "Intelligent Ensemble" predicts more intense precipitation increases in the tropics, especially in the South Pacific Convergence Zone (SPCZ)

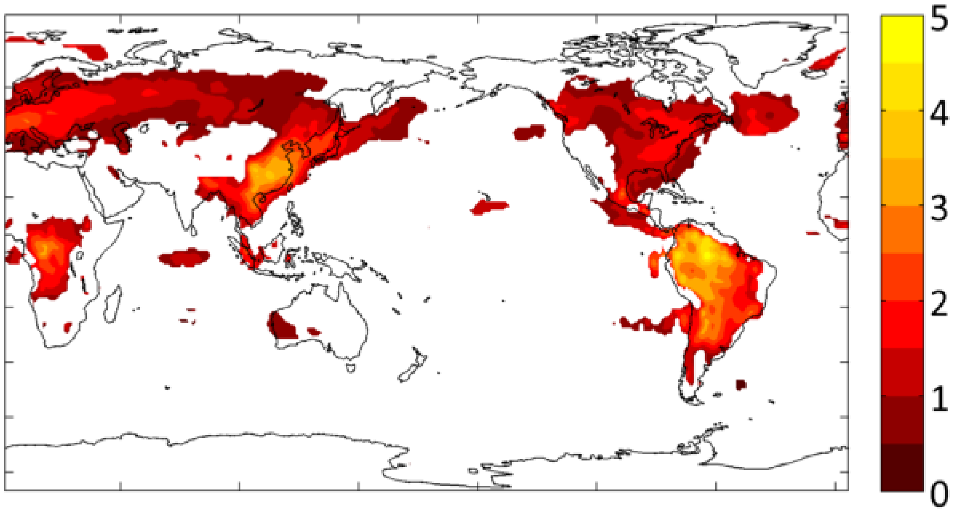
Contours are shaded only where the difference is statistically significant

Results: new 21st-century projections (surface downward SW radiation)

"Intelligent" ensemble mean surface shortwave radiation trend (W/m^2)



Difference between "Intelligent" and Equal-weight ensemble means (W/m^2)



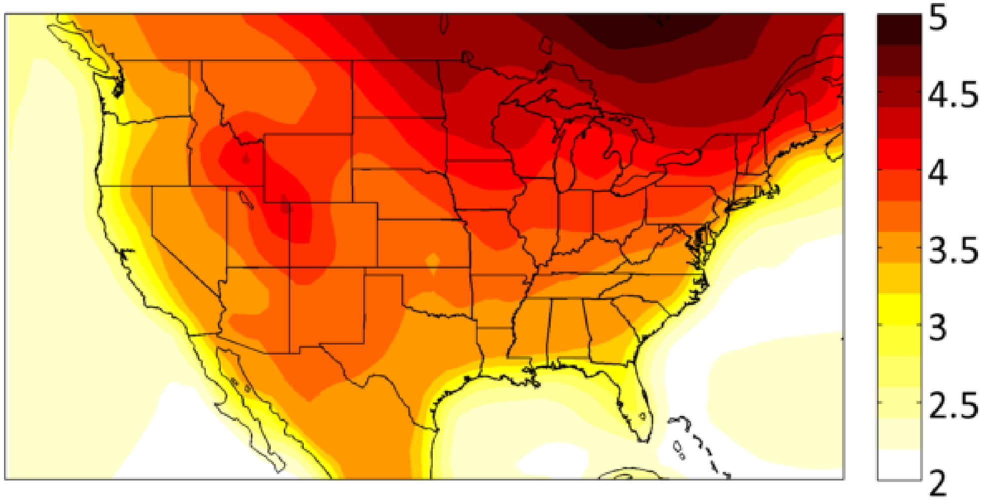
Higher surface radiation: less clouds

The "Intelligent Ensemble" predicts 10-20% less clouds than MME over certain land areas, especially in midlatitude regions

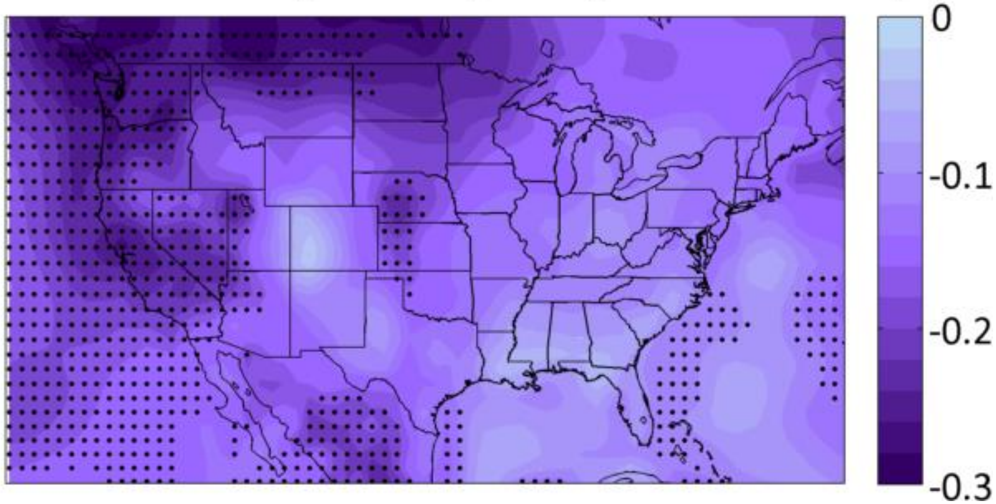
Contours are shaded only where the difference is statistically significant

Results: new 21st-century projections (regional-mean weights)

"Intelligent" ensemble mean temperature trend (◦C)



Difference between "Intelligent" and Equal-weight ensemble means (◦C)



Regional-mean weights can give very different predictions: the US-mean best-performing metric predicts less intense warming than the MME

Predicted warming: 3.9 °C (0.2 °C less than MME)

Stippling indicates where the difference is statistically significant

Conclusions

This project demonstrates:

- **New climate model performance metrics** related to radiation processes are tested on the CMIP5 archive
- **Present-day model skill is linked to quality of future projections**

The results are:

- **New "intelligent ensemble" projections** are created and compared with traditional MME projections
- For global-mean metrics, "intelligent ensemble" projections of large-scale patterns remain similar, but intensity of predicted surface temperature, precipitation, and surface radiation increase is **10-20% higher than the MME**
- Regional-mean metrics can produce very different projections: the **US-mean projected warming is 3.9 °C** (0.2 °C less than MME)