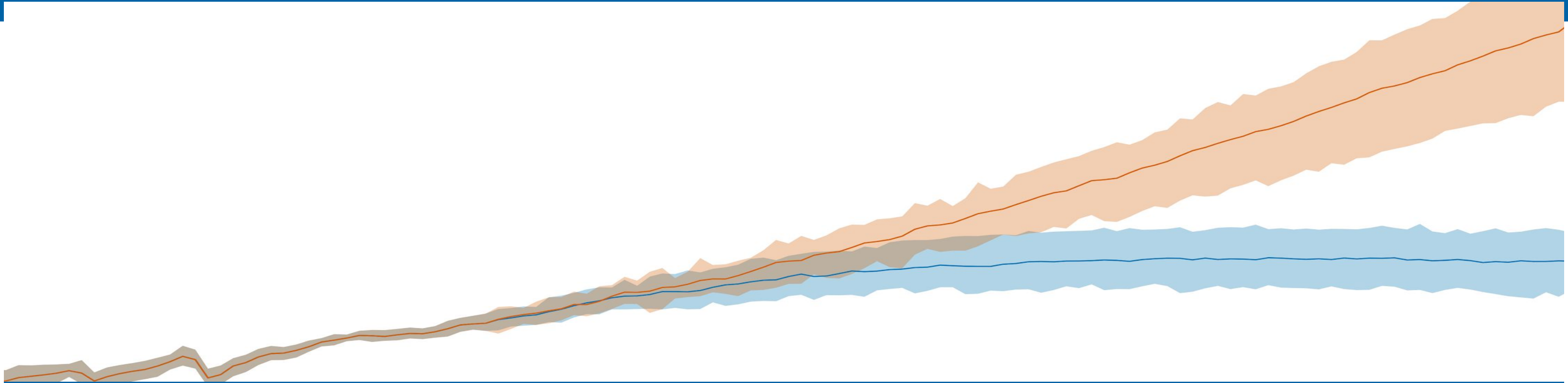
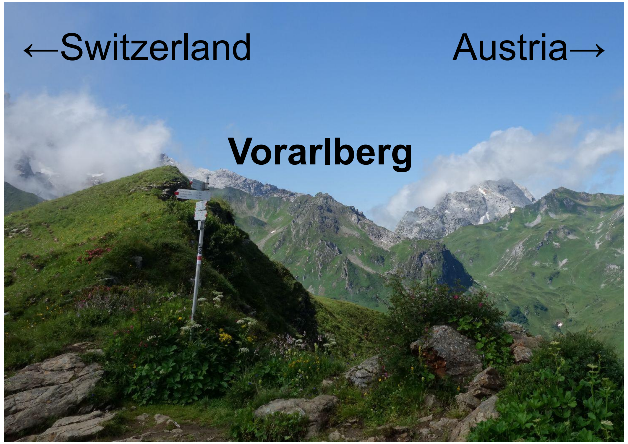


Weighting models by performance and independence Effects on projections of future climate

Lukas Brunner | Department of Meteorology and Geophysics Colloquium | January 18th 2022

With contributions from: Reto Knutti, Ruth Lorenz, Angeline G. Pendergrass, Flavio Lehner, Anna L. Merrifield and many others

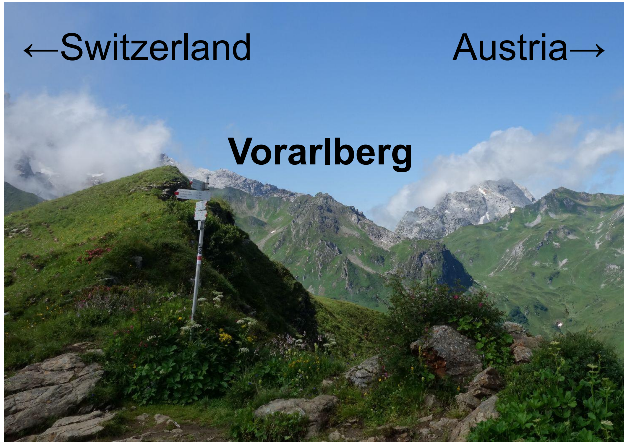
About me



- Studied Physics in Graz
- PhD in Graz, Edinburgh, Oslo
- PostDoc in Zürich
- Senior Scientist in Vienna

More:
lukasbrunner.github.io

About me

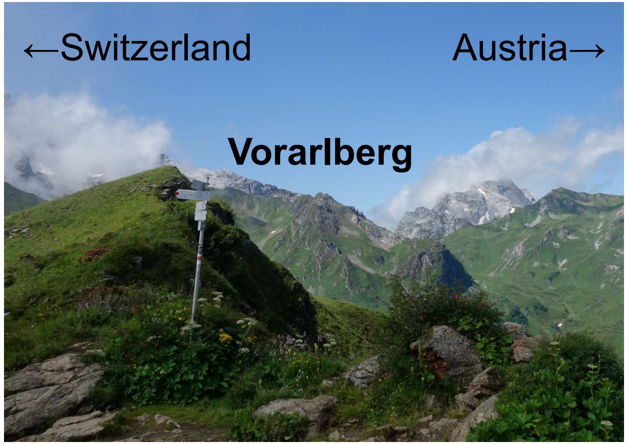


- Studied Physics in Graz
- PhD in Graz, Edinburgh, Oslo
- PostDoc in Zürich
- Senior Scientist in Vienna

More:
lukasbrunner.github.io

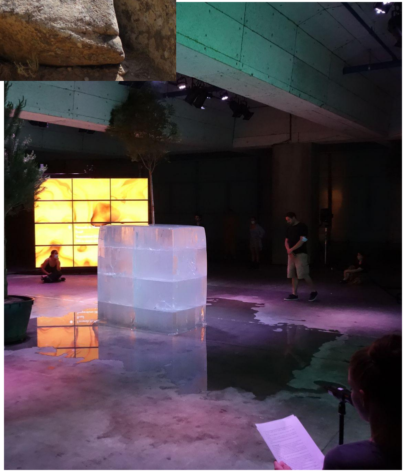
l.brunner@univie.ac.at

About me



- Studied Physics in Graz
- PhD in Graz, Edinburgh, Oslo
- PostDoc in Zürich
- Senior Scientist in Vienna

More:
lukasbrunner.github.io



Evolution of past and future climate

- ***It is unequivocal that human influence has warmed the atmosphere, ocean and land.*** (IPCC AR6 SPM)
- global temperature until today has increased by about 1°C compared to pre-industrial conditions
- estimates of future warming are based on **climate models**

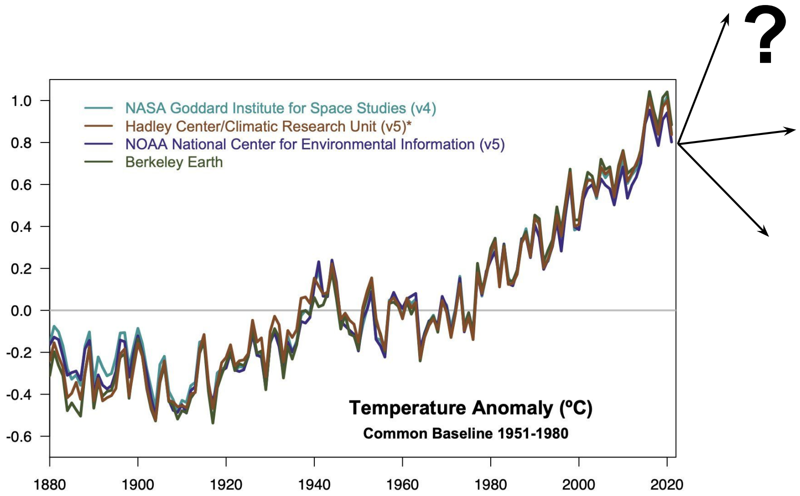


Figure: Global mean, annual mean temperature anomalies (relative to 1851-1980) based on four observational datasets. RealClimate/Gavin Schmidt, 15.1.22

*HadCRUT5: Jan-Nov mean for 2021

Climate models and climate model projections

- A model is an informative **representation** of an object, person or system. Wikipedia
- Climate models simulate the interactions of the **important** drivers of climate. Wikipedia
- Climate model are used to
  - simulate historical climate
  - understand (parts of) the climate system and interactions
  - project future climate
  - etc...

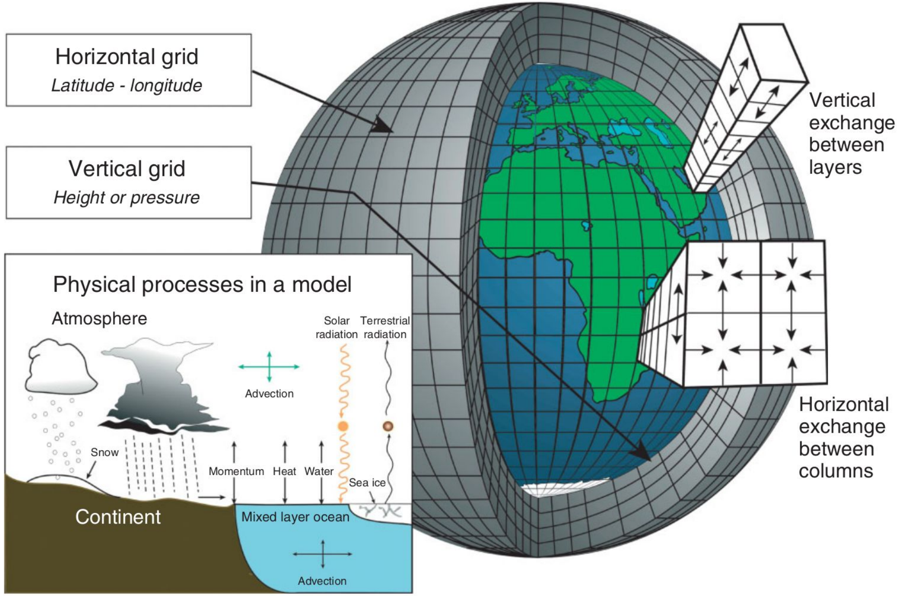
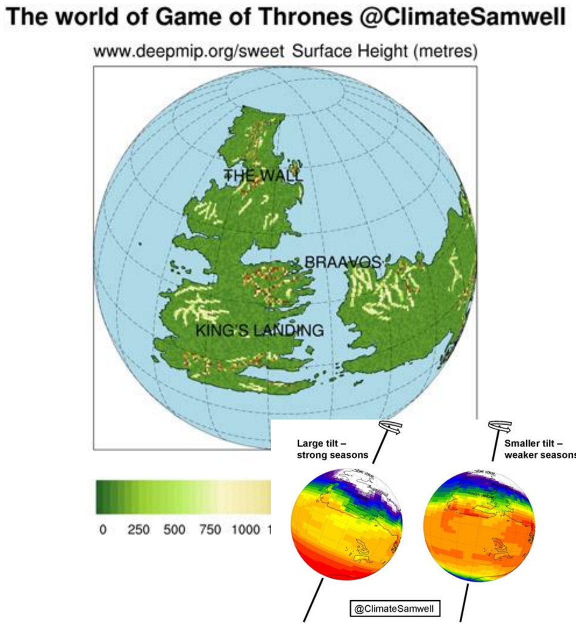
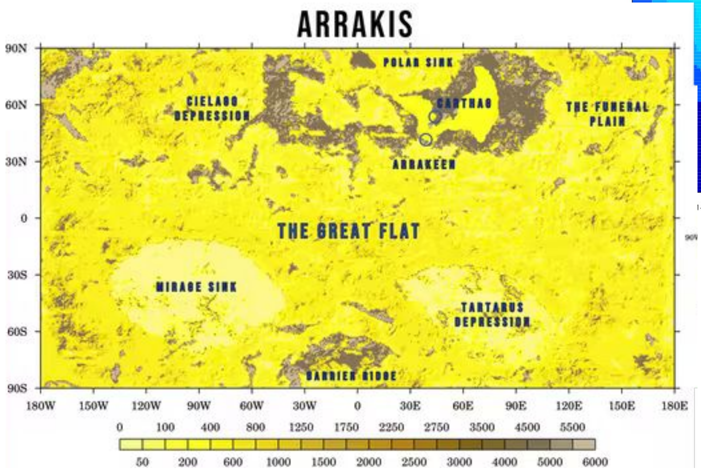
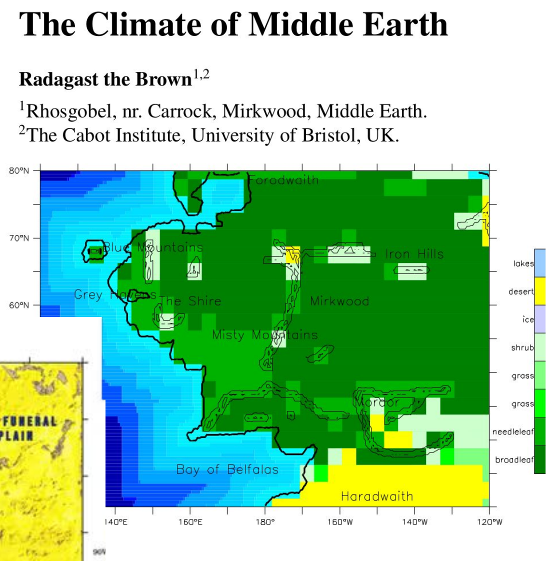


Figure: Schematic representation of a general circulation model. Edwards (2011)

What Climate models are used for







CC-BY-NC theconversation.com/Alex Farnsworth, Michael Farnsworth, Sebastian Steinig

Uncertainty in model projections of future climate

- Different socio-economic developments are represented by **scenario uncertainty**
- Multi-model assessments used to quantify **model uncertainty**
- The chaotic behavior of the climate system leads to **internal variability**

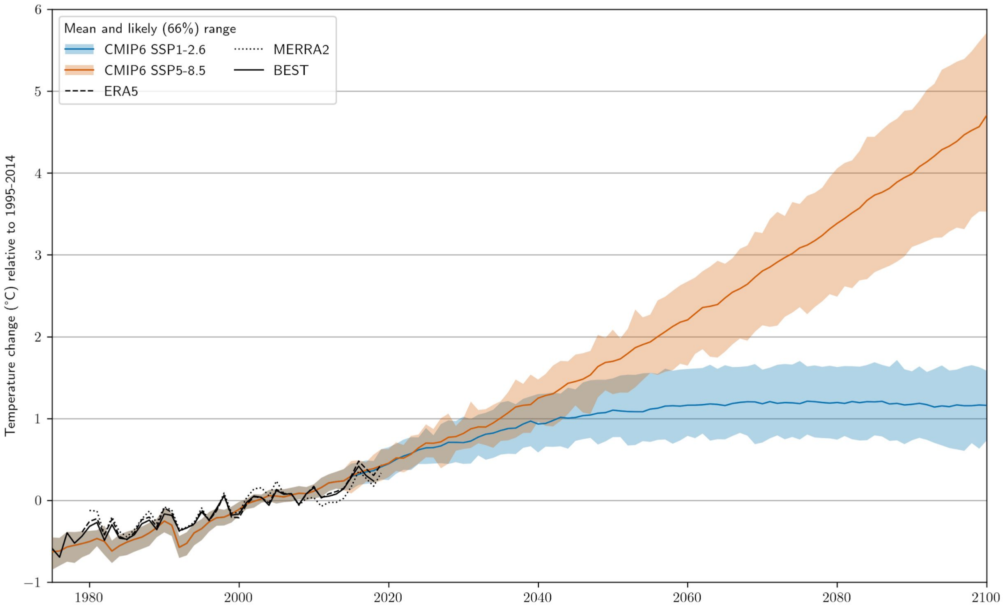


Figure: Global mean, annual mean temperature change (relative to 1995-2014) from CMIP6. Brunner et al. (2020a)

Distribution of uncertainty

- The contribution from each source is **not constant over time**
- The distribution of uncertainty also depends on a range of other parameters
- **Scenario uncertainty** can be eliminated by making projections conditional to a scenario
- **Internal variability** can, for example, be investigated using so-called SMILEs
- Leaves us with **model uncertainty**...

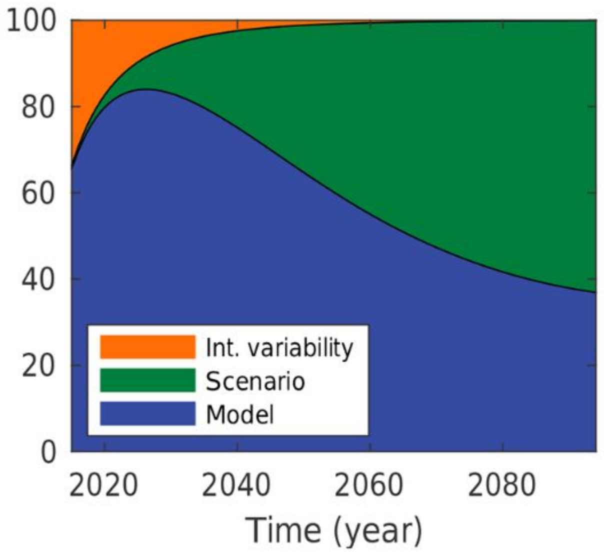


Figure: Fractional contribution to total uncertainty for 10-year running mean of global mean, annual mean temperature from CMIP6. Lehner et al. (2020)

Known and unknown model uncertainty

- Model uncertainty arises when looking at **multi-model ensembles**
- Model uncertainty ≠ actual uncertainty (e.g., IPCC AR5 & 6)
  - there are processes not covered by any model (not considered here)
  - **not all models are equally 'good'**
  - **not all model are independent**

→ Here we look at uncertainty from model spread and how to best quantify it

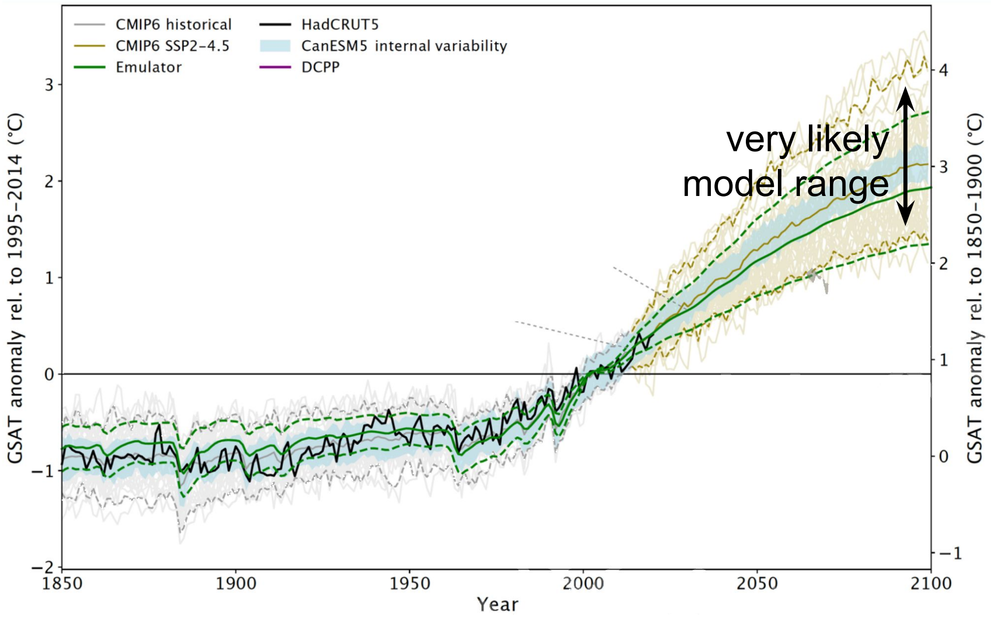


Figure: Global mean, annual mean temperature change based on 39 CMIP6 models. The dashed brown lines indicate the 90% model range. IPCC AR6

Not all models are equally 'fit for purpose'

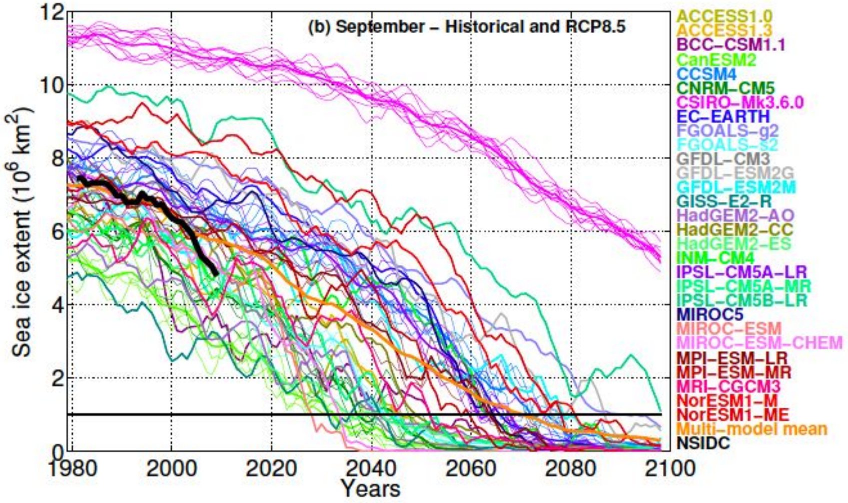


Figure: September Arctic sea ice extent in CMIP5 historical / RCP8.5 runs and observations. Massonnet et al. (2012)

Not all models are equally 'fit for purpose'

- we might want to trust models less if they are far away from observations → **weighting by performance**
- need a way to **convert** model-observation **distance into weights**
  - if we are very strict: strong weighting leaving us only with few models
  - if we are very generous: weak weighting not doing anything
- weights should be based on **metrics relevant to the target**

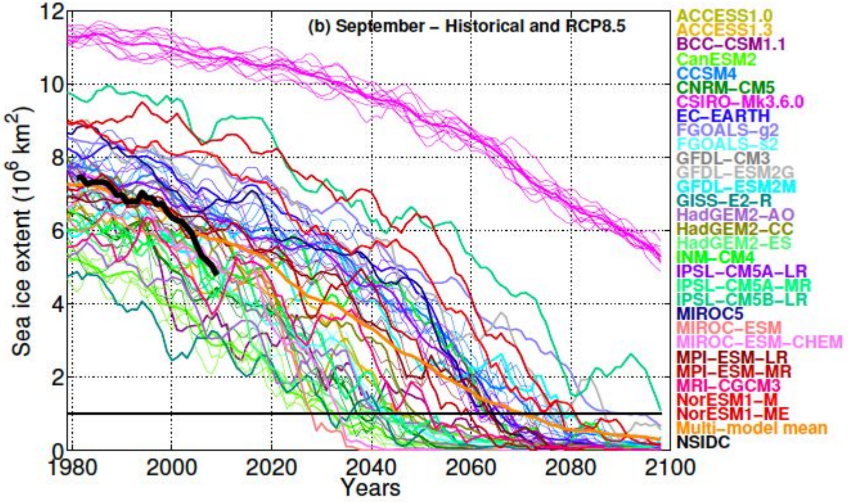


Figure: September Arctic sea ice extent in CMIP5 historical / RCP8.5 runs and observations. Massonnet et al. (2012)

Not all models are independent

- Multi-model studies often draw on **all available models**
- the CMIP multi-model ensembles are not designed to only include independent models (**‘ensembles of opportunity’**)
    - Several models are closely related (one different component, resolution)
    - Models have been branched from each other
    - Some models share components

→ **weighting by independence**

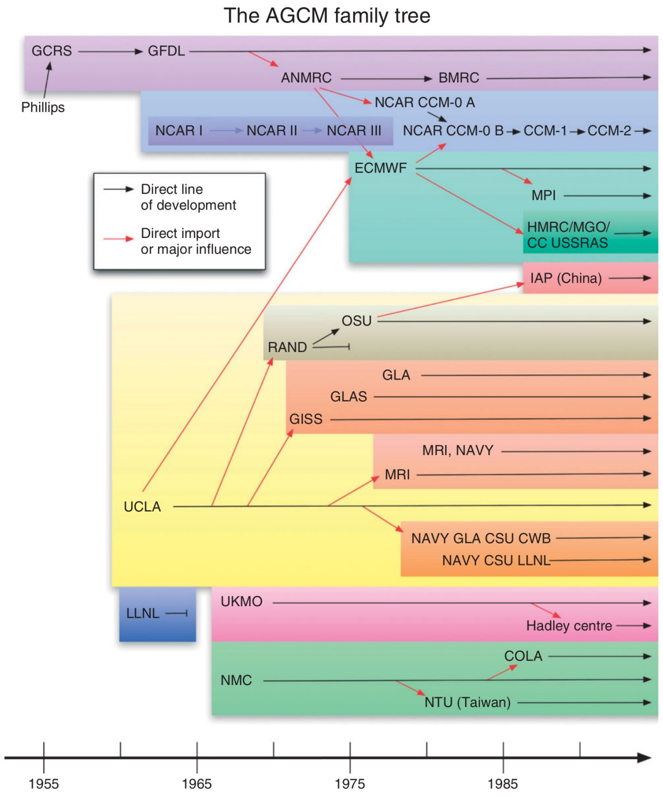


Figure: Development and dependencies for several climate models. Edwards (2010)

Putting it all together: calculation of model weights

$$w_i = \frac{e^{-\frac{D_i^2}{\sigma_D^2}}}{1 + \sum_{j \neq i}^{M} \left(e^{-\frac{S_{ij}^2}{\sigma_S^2}} \right)}$$

Knutti et al. (2017)

- w_i : weight for model i
- D_i : generalised distance of model i to observations (performance diagnostics)
- σ_D : performance shape parameter
- M: number of models
- S_{ij} : generalised distance between model pair (independence diagnostics)
- σ_S : independence shape parameter

Recap: Introduction

- Projections of future climate by climate models have three main sources of uncertainty:
  - emission scenario uncertainty
  - model uncertainty
  - internal variability
- Here I focus on **model uncertainty**
- Weighting to better quantify model uncertainty
  - accounting for model dependencies (**Part I**)
  - downweighting models which are not ‘fit for purpose’ (**Part II**)
- Finally I check if things improved (**Part III**)

Part I: Model Independence

Model independence weighting: basic assumption

Structural model similarity can be inferred from model output similarity

Model independence weighting: basic assumption

Structural model similarity can be inferred from model output similarity

- Models with multiple **shared components** have **similar output** (e.g. temperature climatologies)
- We can check this by looking at models which we know are similar
- **Two variables are enough to cluster/separate models**

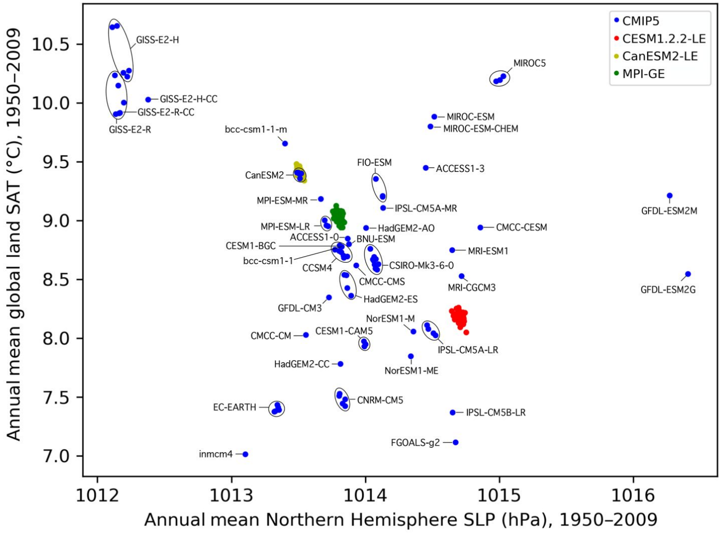


Figure: Clustering of CMIP5 models based on mean temperature and sea level pressure. Merrifield et al. (2020)

universität wien

CMIP6 model ‘family tree’

- The tree structure on the right-hand side is only based on model output
- Model branching further to the left are closer to each other in output space

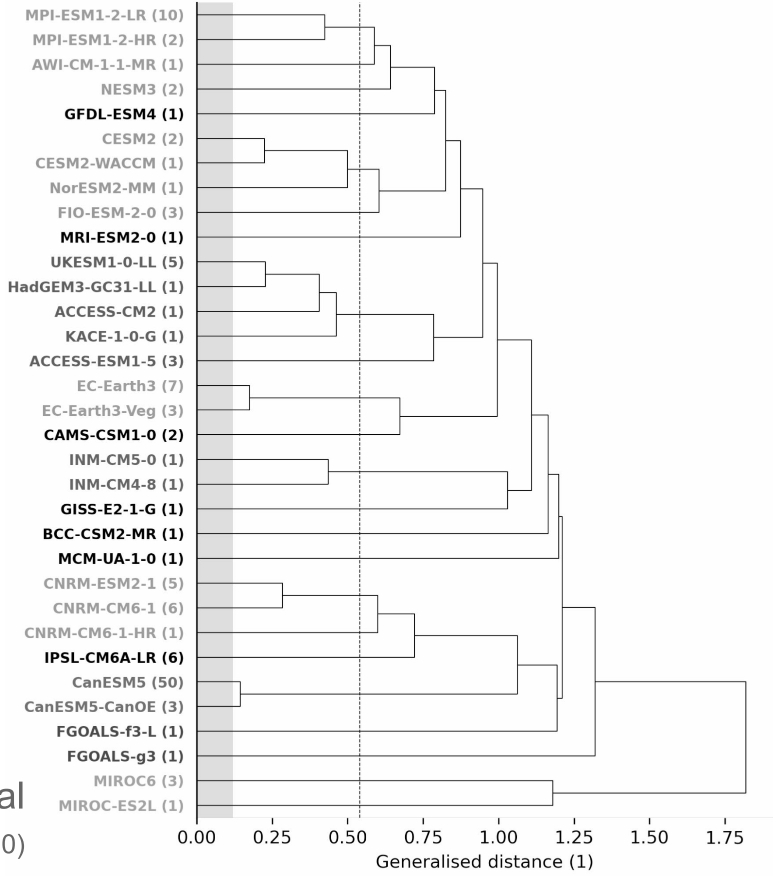


Figure: Model family tree for CMIP6, based on global temperature and sea level pressure. Brunner et al. (2020)

CMIP6 model family tree

- The tree structure on the right-hand side is only based on model output
- Model branching further to the left are closer to each other in output space
- Label colors based on expert knowledge of model components

→ **Models know to be similar are clustered together based on their output**

→ transfer generalised distance to independence weights (**shape parameter**)

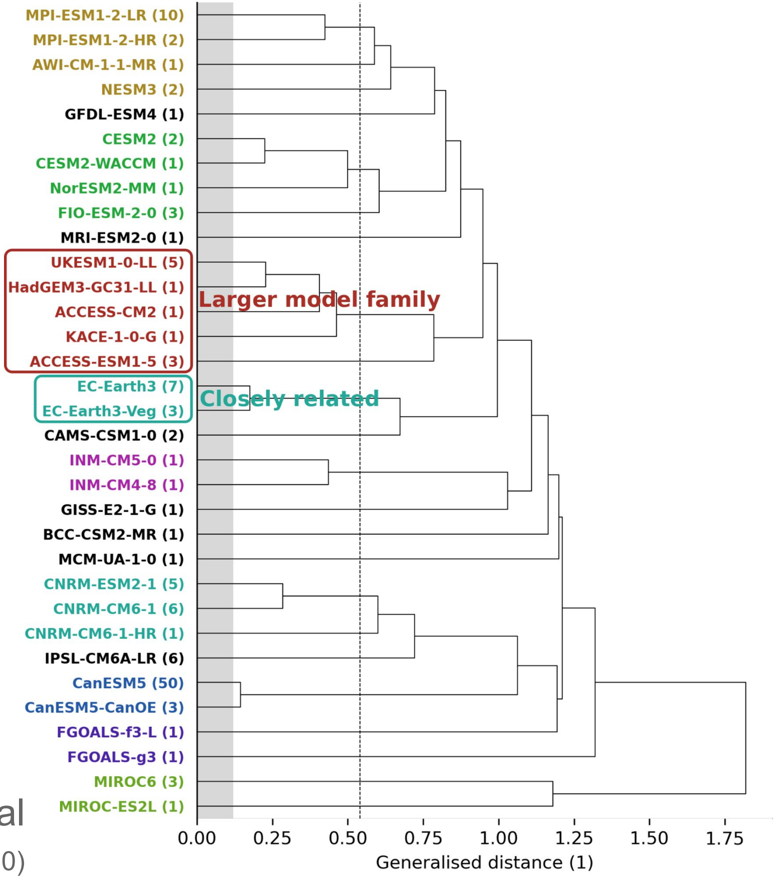


Figure: Model family tree for CMIP6, based on global temperature and sea level pressure. Brunner et al. (2020)

A look across CMIP generations

The clustering can also be used to

- track model development from CMIP5 to CMIP6 (including intermediate versions)
- investigate the importance of individual model components (atmosphere, land, etc.)
- investigate the importance of model resolution

Figure not available publicly

Figure: Model family tree for CESM, based on global temperature and sea level pressure.

Part II: Model Performance

Model-observation distances

Figure not available publicly

Figure: Generalized distance to observations (ERA5) for CMIP6 models. Based on 21-year climatology of temperature and precipitation. Brunner et al. (in prep)

universität wien

Model-observation distances

- **Model-observation distance** can be based on
  - different variables (temperature, precipitation, sea level pressure, …)
  - different time aggregations (climatology, variability, trend)
  - different geographical regions (that can differ from the target region)
  - time periods, observational datasets, resolutions, etc.
- Multiple metrics can be combined (**generalized distance**)
- Reliable observations are needed as reference

Figure not available publicly

Figure: Generalized distance to observations (ERA5) for CMIP6 models. Based on 21-year climatology of temperature and precipitation. Brunner et al. (in prep)

Model-observation distances

- **Model-observation distance** can be based on
  - different variables (temperature, precipitation, sea level pressure, …)
  - different time aggregations (climatology, variability, trend)
  - different geographical regions (that can differ from the target region)
  - time periods, observational datasets, resolutions, etc.
- Multiple metrics can be combined (**generalized distance**)
- Reliable observations are needed as reference
- Weighting: metrics should be **relevant for the target**

Figure not available publicly

Figure: Generalized distance to observations (ERA5) for CMIP6 models. Based on 21-year climatology of temperature and precipitation. Brunner et al. (in prep)

Model-observation distances across CMIP generations

Figure not available publicly

Figure: Generalized distance to observations (ERA5). Based on 21-year climatology of temperature and precipitation. Brunner et al. (in prep)

Translating distances to weights: shape parameter

The **shape parameter** σ_D needs to be carefully chosen to provide confident and meaningful weights

- small values lead to strong weighting, selecting only a few models
- large values lead to equal weighting

→ **model-as-truth test**

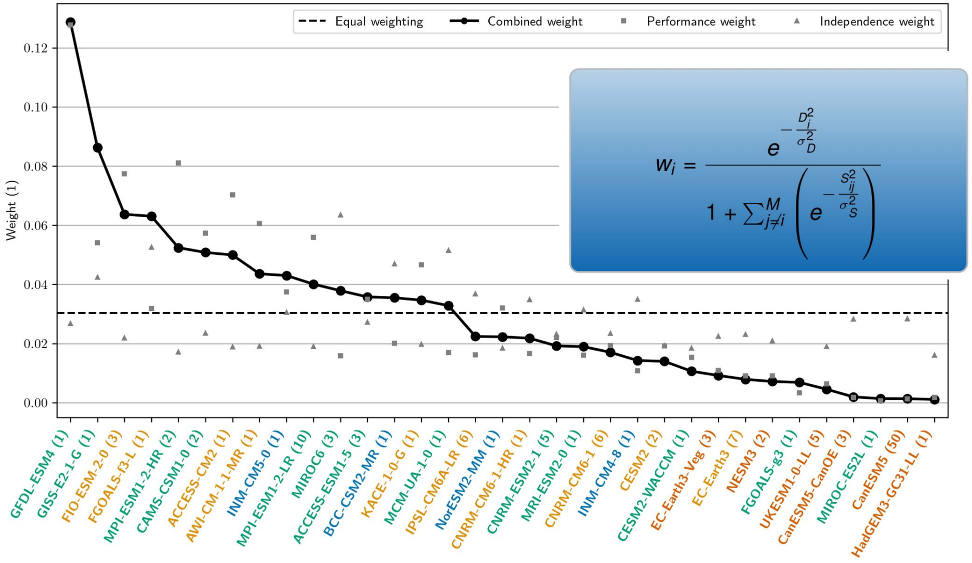


Figure: Weights for 33 CMIP6 models based on **five performance** and **two independence metrics** chosen for weighting global temperature. Brunner et al. (2020a)

Effect of weighting CMIP6 projections of future climate

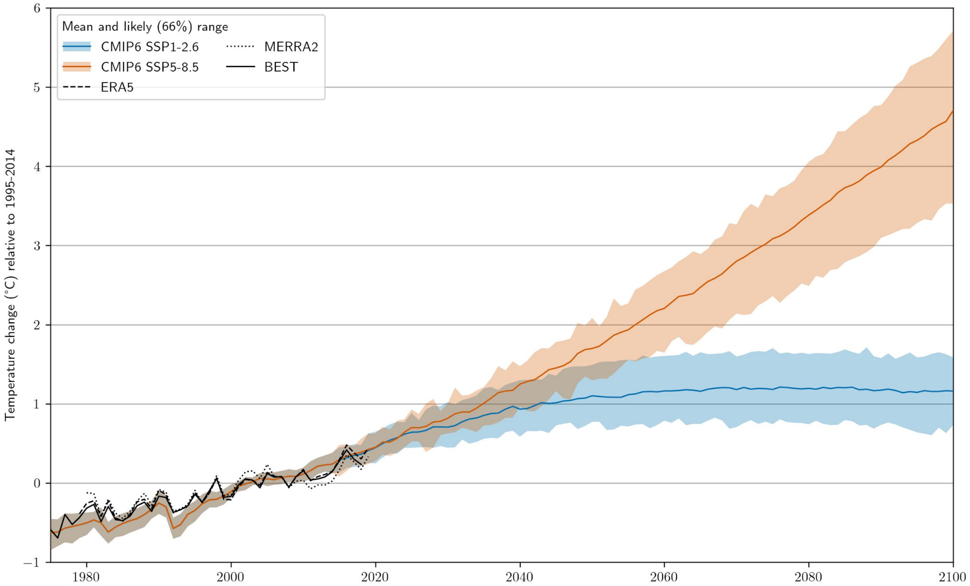


Figure: Global mean, annual mean temperature change (relative to 1995-2014) from 33 CMIP6. Brunner et al. (2020a)

Effect of weighting CMIP6 projections of future climate

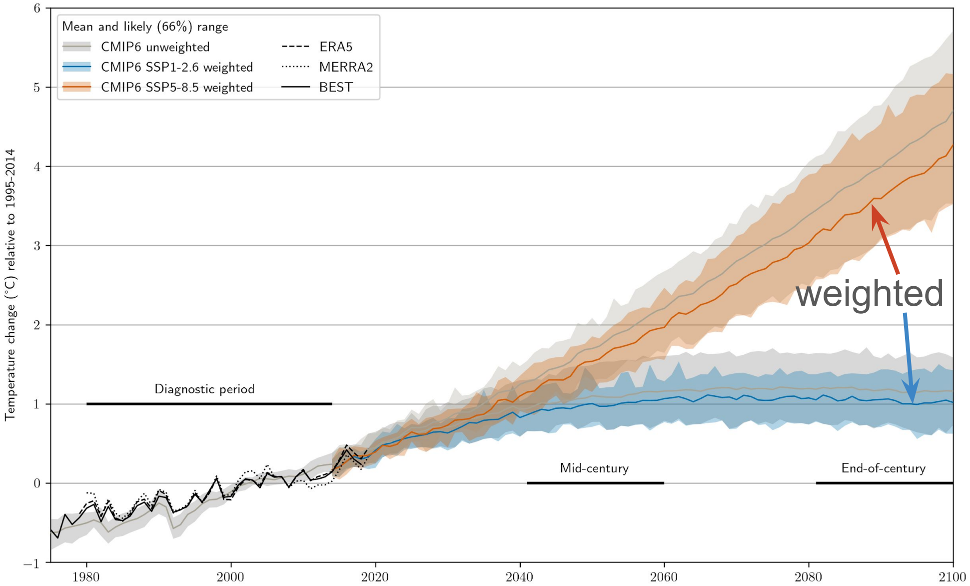


Figure: Weighted global mean, annual mean temperature change (relative to 1995-2014) from 33 CMIP6 models.
Brunner et al. (2020a)

- The weighted distribution shows **reduced mean warming from CMIP6** models broadly consistent with other studies
  - Nijsse et al. (2020)
  - Tokarska et al. (2020)
  - Ribes et al. (2021)
- **Reduction of uncertainty** by 10%-20% for the likely range due to a constraining of the upper percentiles

Recap: Performance and independence weighting

- Using the model range directly as uncertainty range disregards that
  - not all model are independent
  - not all models are equally ‘fit for purpose’
- **Model weighting** can help to account for that
- Distances are translated into weights assuming that
  - model similarity can be inferred from output similarity
  - future model performance can be inferred from past model performance
- The translation from distances to weights is done via two **shape parameters**

universität wien

Part III: Does weighting improve future projections?

Measuring the benefit of weighting climate models

From weather forecasting: “What Is a Good Forecast?” Murphy (1993)

- **Accuracy**: level of agreement between forecast and truth
- **Skill**: accuracy relative to a reference forecast
- **Reliability**: average agreement between forecasts and truth
- **Sharpness**: tendency of the forecast to predict specific values (counter-example: the climatology has no sharpness)

Quality

- **Consistency**: forecast is consistent with prior knowledge
- **Value**: degree to which the forecast helps decision makers

Measuring the benefit of weighting climate models

What Is a Good Weighting? - **we don't know the 'truth'**

- ✘ **Accuracy**: level of agreement between **weighted projection** and 'truth'
- ✘ **Skill**: accuracy relative to the **unweighted projection**
- ✘ **Reliability**: average agreement between **weighted projections** and 'truth'
- ✔ **Sharpness**: tendency of the **weighted projections** to reduce model uncertainty compared to the **unweighted projections**

- ✔ **Consistency**: is **weighting** consistent with other methods
- ✔ **Value**: degree to which the **weighted projection** helps users

Measuring the benefit of weighting climate models

What Is a Good Weighting? - we don't know the 'truth'

- ✔ **Sharpness**: determined by the performance shape parameter σ_D: smaller σ_D leads to sharper results but might no longer be **reliable**
- ✔ **Value**: determined by the users

- ✔ **Consistency**: **quantify** by comparing methods using a **common setup** (Brunner et al. 2020b, Hegerl et al. 2021, O'Reilly et al. in prep.)
- ✔ **Accuracy, Skill, Reliability**: we don't know the true climate in the future and there will be only one realisation → **model-as-truth approach**

Consistency: comparing methods to constrain projections

No **coordinated framework** to compare methods exist. They might differ for a range of reasons independent of the methods itself:

- variable (temperature vs precip)
- region (global vs Europe)
- season and time period
- models included
- uncertainties included
- ...

Figures: Comparing (top) methods and (right) apples and oranges right: CC-BY M. Johnson

A consistent framework for method comparison

We brought together **8 groups** working on constraining and developed a **level playing field for comparison**

2 conditions for participation:

1. quantify uncertainty in future projections
2. able to handle common settings

Institution name	Method acronym	Method name	References
ETH Zurich (Switzerland)	ClimWIP	Climate Model Weighting by Independence and Performance	Knutti et al. (2017b); Lorenz et al. (2018); Brunner et al. (2019)[a]
International Centre for Theoretical Physics (Italy)	REA	Reliability ensemble averaging	Giorgi and Mearns (2002, 2003)[b]
University of Edinburgh (United Kingdom)	ASK	Allen–Stott–Kettleborough	Allen et al. (2000); Stott and Kettleborough (2002); Kettleborough et al. (2007)
Centre National de Recherches Météorologiques (France)	HistC	Historically constrained probabilistic projections	Ribes et al. (2020, manuscript submitted to *Sci. Adv.*)[c]
Met Office (United Kingdom)	UKCP	U.K. Climate Projections (UKCP) Bayesian probabilistic projections method	Sexton et al. (2012); Harris et al. (2013); Sexton and Harris (2015); Murphy et al. (2018)
University of Oxford (United Kingdom)	CALL	Calibrated large ensemble projections	O'Reilly et al. (2020)
Royal Netherlands Meteorological Institute (Netherlands)	BNV*	Bootstrapped from natural variability	See the online supplemental material
Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy)	ENA*	Ensemble analysis of probability distributions	See the online supplemental material

[a] Source code available online (https://github.com/lukasbrunner/ClimWIP).
[b] Source code available online (http://doi.org/10.5281/zenodo.3890966).
[c] Method tool available online (https://saidqasmi.shinyapps.io/bayesian).

Table: Participating institutions, methods, and references. Brunner et al. (2020b)

Comparing future Central European temperature change

- Trade-off between number of methods and the **fairness of the comparison**
- **Fairest** comparison:
  4/8 methods could participate
- All methods **narrow the uncertainty** range
- All methods agree on slightly **less warming**

→ not all cases look that nice

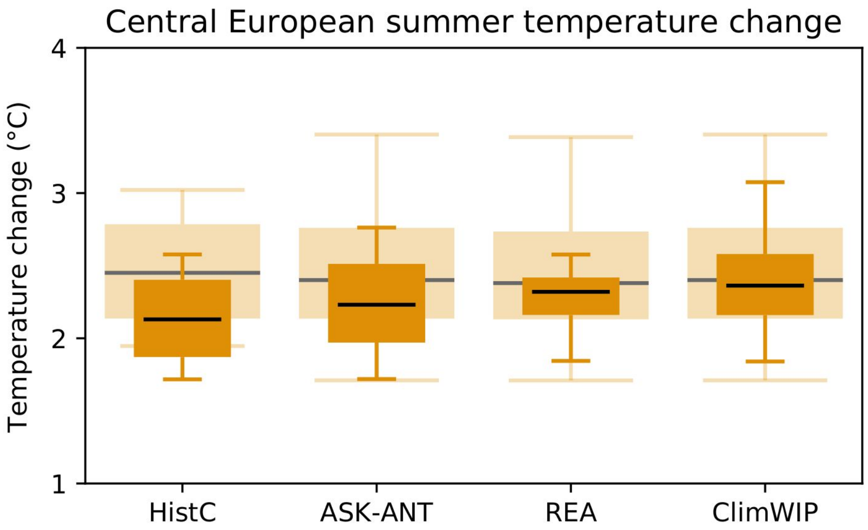


Figure: Unconstrained (light) and constrained (dark) Central European summer temperature change (2041-60 relative to 1995-2014) from CMIP5. Brunner et al. (2020b)

Take home messages

- **Uncertainty in projections of future climate** comes from
  - emission scenario uncertainty
  - climate model uncertainty
  - internal variability
- **Model spread** can be translated to **model uncertainty** but
  - **not all models are independent** estimates of the future
  - **not all models are equally 'fit for purpose'**
- **Model weighting** can help to account for this
- Model weighting is consistent with other methods

universität wien

Thank you for your attention!

Literature

- Brunner, L., Lorenz, R., Zumwald, M., & Knutti, R. (2019). Quantifying uncertainty in European climate projections using combined performance-independence weighting. Environmental Research Letters, 14(12), 124010. https://doi.org/10.1088/1748-9326/ab492f
- Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020a). Reduced global warming from CMIP6 projections when weighting models by performance and independence. Earth System Dynamics, 11(4), 995–1012. https://doi.org/10.5194/esd-11-995-2020
- Brunner, L., McSweeney, C., Ballinger, A. P., Befort, D. J., Benassi, M., Booth, B., … Undorf, S. (2020b). Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework. Journal of Climate, 33(20), 8671–8692. https://doi.org/10.1175/JCLI-D-19-0953.1
- Brunner et al. (in preparation): Evolution of climate models in past, present, and future
- Edwards, P. N. (2011). History of climate modeling. Wiley Interdisciplinary Reviews: Climate Change, 2(1), 128–139. https://doi.org/10.1002/wcc.95
- Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., … Hawkins, E. (2020). Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. Earth System Dynamics, 11(2), 491–508. https://doi.org/10.5194/esd-11-491-2020
- Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., & Eyring, V. (2017). A climate model projection weighting scheme accounting for performance and interdependence. Geophysical Research Letters, 44(4), 1909–1918. https://doi.org/10.1002/2016GL072012
- Massonnet, F., Fichefet, T., Goosse, H., Bitz, C. M., Philippon-Berthier, G., Holland, M. M., & Barriat, P. Y. (2012). Constraining projections of summer Arctic sea ice. Cryosphere, 6(6), 1383–1394. https://doi.org/10.5194/tc-6-1383-2012
- Merrifield, A. L., Brunner, L., Lorenz, R., Medhaug, I., & Knutti, R. (2020). An investigation of weighting schemes suitable for incorporating large ensembles into multi-model ensembles. Earth System Dynamics, 11(3), 807–834. https://doi.org/10.5194/esd-11-807-2020