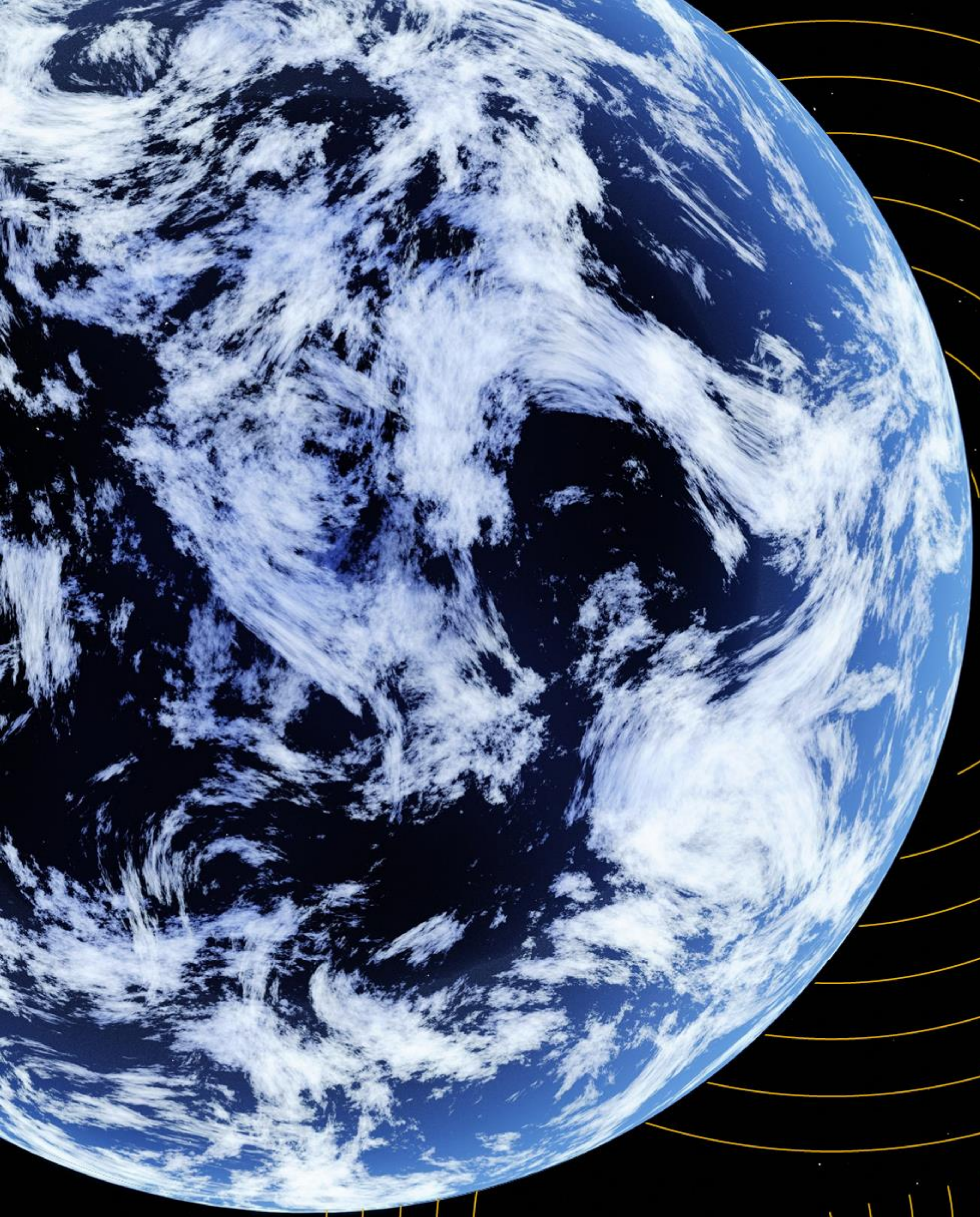


CMIP
Coupled Model
Intercomparison
Project

Streamlining model selection

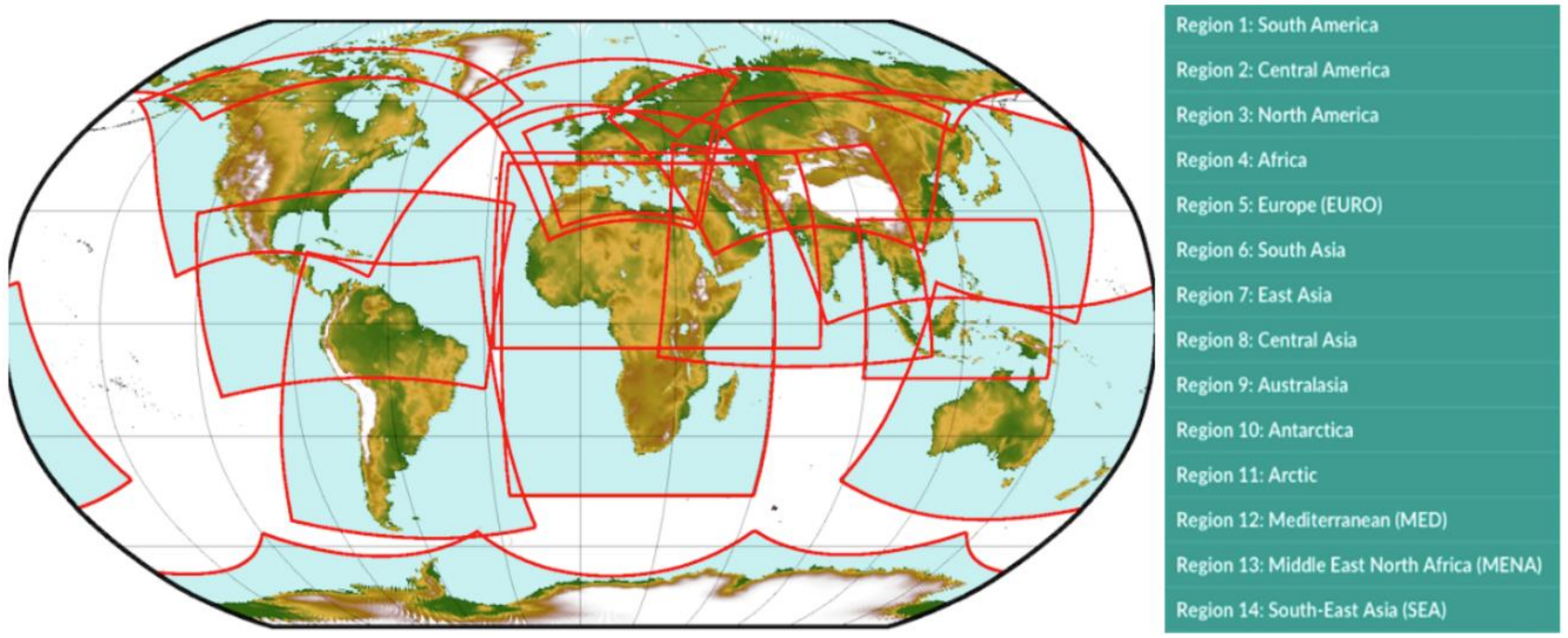
Wednesday 5th February 2025, 19:00-22:00 UTC
Virtual workshop

Introduction

Christian Steger, Deutscher Wetterdienst

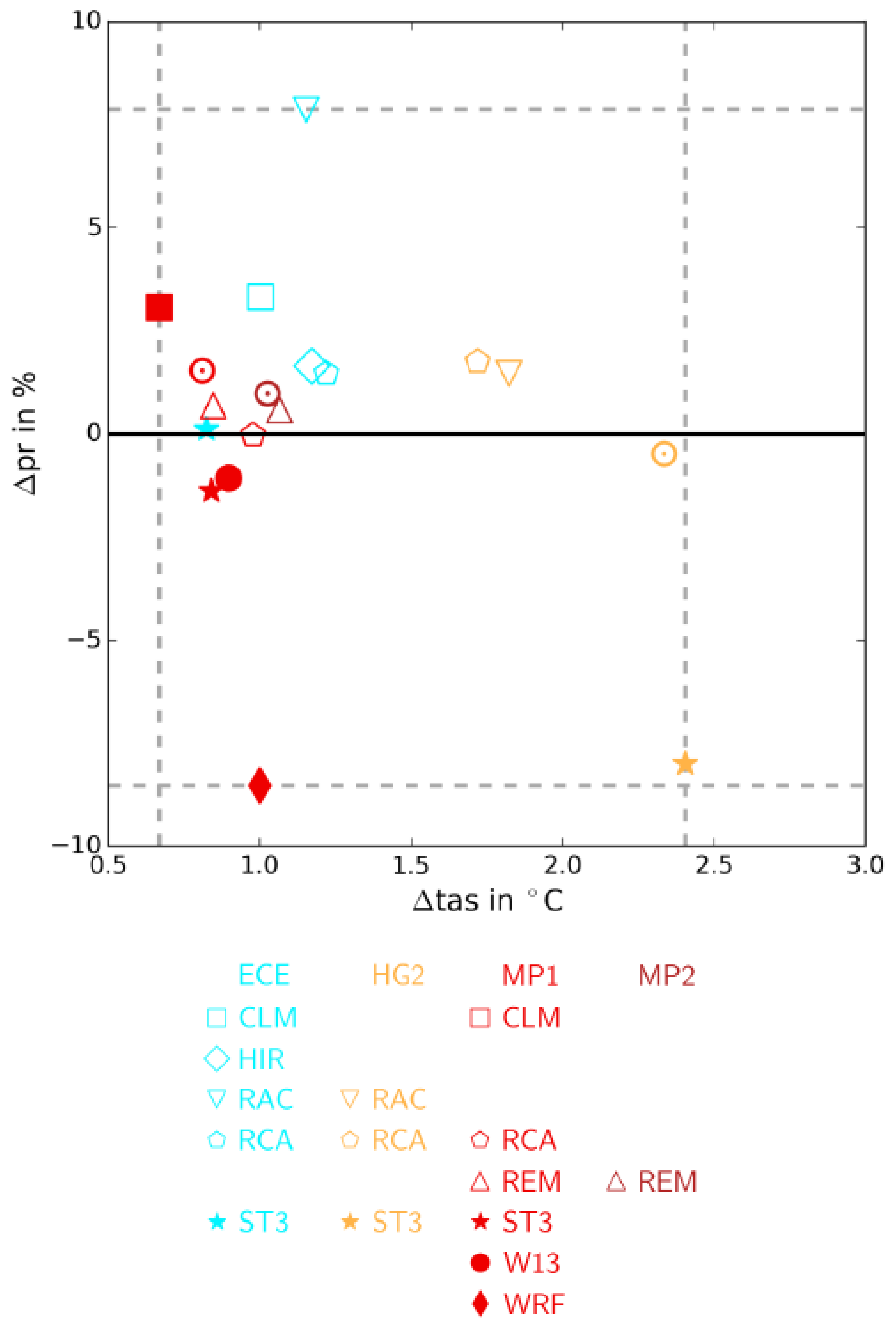
Motivation

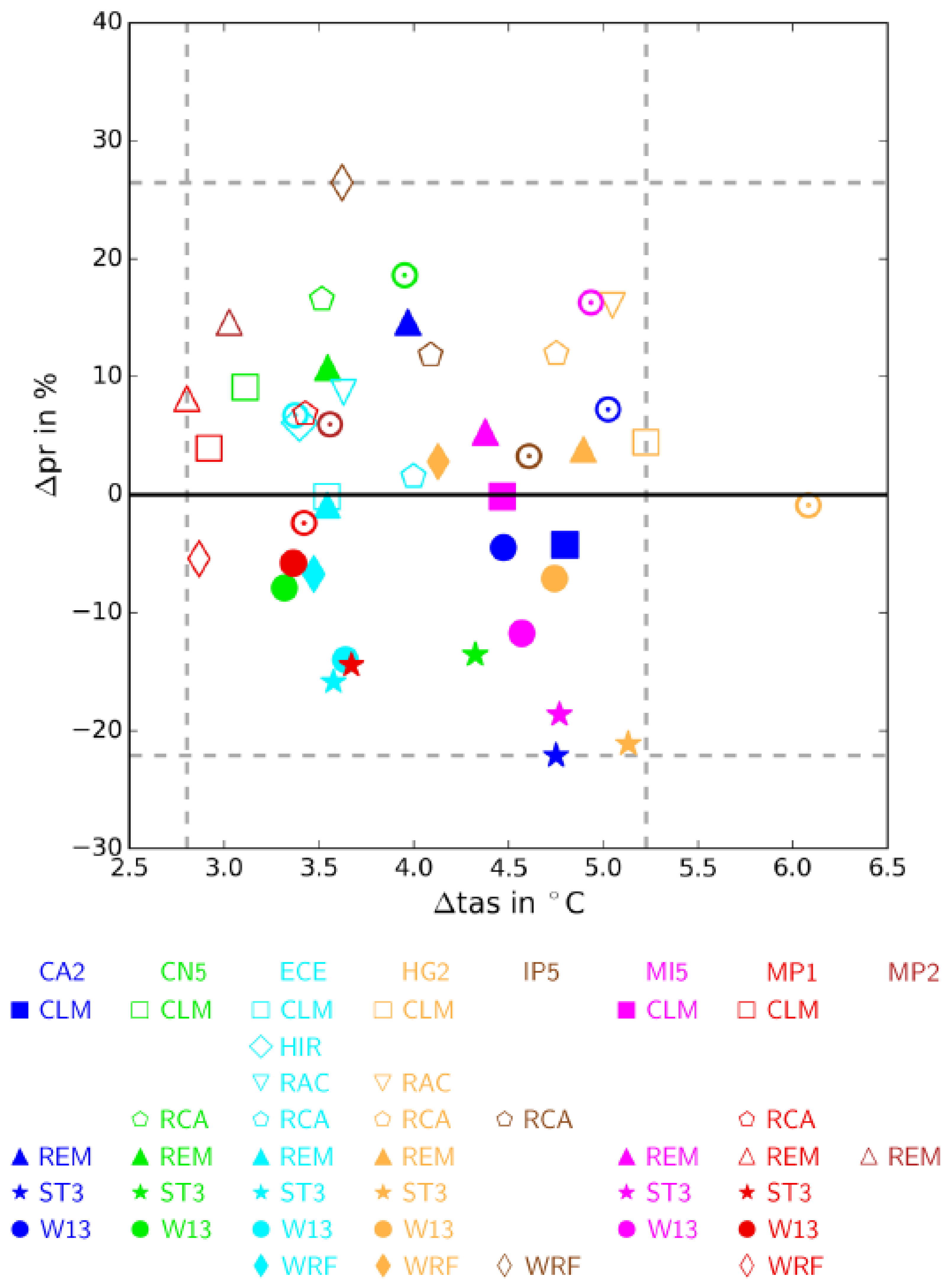
- Information for climate services, advice for decision-makers and climate adaptation are usually based on climate projection ensembles
- The composition of the ensembles is decisive for the information that is generated and passed on to users
- Various sources of information and model ensembles are used in different projects and initiatives
- This leads to different, inhomogeneous and sometimes contradictory information and, as a consequence, to uncoordinated and inconsistent decisions and measures



ReKliEs-De Ensemble

RCP2.6

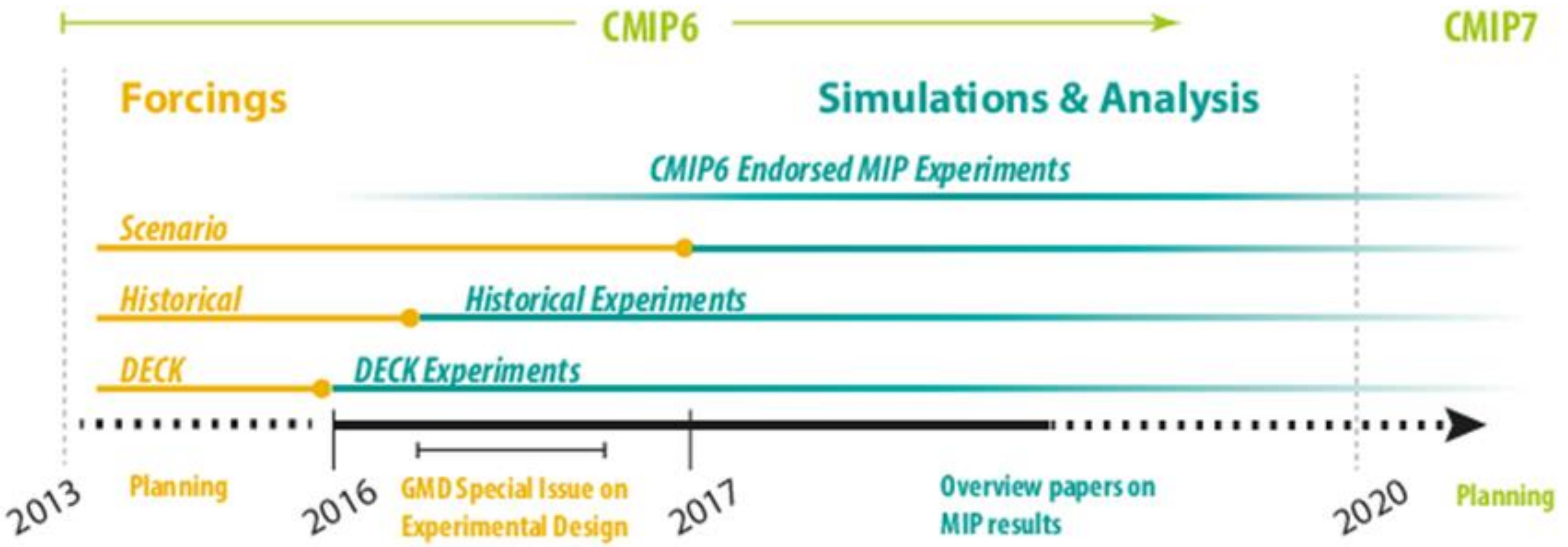




RCP8.5

Motivation

- It takes a very long time for the information from the latest generation of CMIP simulations to reach the end users



CMIP6 downscaling | Impact models | End user

2025

Key questions of the workshop

1. Is it possible to streamline, homogenize or even standardize model selection and ensemble composition across projects/activities/initiatives?

2. Is it possible to speed up the process from generating global climate projections to making the information available to end users?

Aims of the workshop:

- Identifying the issues/challenges in the current framework, both scientific and technical.
- Developing shared understanding of model selection techniques/approaches used across the communities.
- Determining if shared criteria for model selection is required and possible.
- Establishing tangible structures to support improved communication between modelling centers and downstream activities.
- Accelerate the process from the creation of global simulations to regional downscaling and impact modelling and finally, the provision of data to end users.
- Determining the needs for and identifying pathways to developing and funding the scientific and technical frameworks required to deliver to users.
- Realistically assessing whether achieving this within CMIP7 is feasible, or the focus should be on laying groundwork for a longer term plan.

Time (UTC)	Topic
19:00 -19:05	**Welcome and housekeeping** (CMIP IPO)
19:05 -19:20	**Meeting objectives** (Christian Steger, DWD)
19:20 -19:35	**Setting the context** An example of challenges faced by users in the current framework (Sven Kotlarski, Meteo Swiss)
19:35 -20:05	**Sharing examples of current best practice** EURO-CORDEX for CMIP6 (Stefan Sobolowski, University of Bergen) CORDEX-Australasia CMIP6 for Australian national projections (Michael Grose, CSIRO) Model selection for RCM downscaling in ISMIP6 (Céline Agosta, LSCE/IPSL) The ISIMIP approach (Lisa Novak, PIK Potsdam)
20:05 -20:35	**BREAKOUT GROUPS: Scientific challenges** BOG 1 & 2: Ensemble sub-selection: is a common criteria/minimum set of evaluation metrics/framework to serve all communities possible or desired? BOG 3: What do model development innovations mean for model selection e.g., the CMIP7 focus on CO_2 emission-driven simulation, increasing resolution and AI/ML? BOG 4: Understanding, quantifying and communicating uncertainty

20:35 -20:45	*BREAK: Time for a drink and stretch!*
20:45 -21:15	**BREAKOUT GROUPS: Other challenges** BOG 5: Constraints on the availability of the required GCM/ESM model data – timing, provision of all required data, including temporal frequencies and extensions BOG 6: Sustained and supported infrastructure to store, deliver and provide user friendly platforms for analysis, framework for coordinated exchange between communities. BOG 7: Speeding up the process from creation of the global simulations until data/information reaches the end user including potential role of community developed tools to support model evaluation and selection BOG 8: Balancing competing needs – funding, politics and ensuring equity, and supporting training and capacity building in a global community effort
21:15–21:40	**Feedback from breakout groups**
21:40 -21:55	**What have we learned during the workshop and what are next steps?** Immediate actions required Follow up workshop (virtual/in person) Format of outcomes – commentary paper, whitepaper etc.
21:55 -22:00	**Final comments and meeting close**

Breakout groups: Scientific challenges

1. Ensemble sub-selection: is a common criteria/minimum set of evaluation metrics/framework to serve all communities possible or desired (Christian Steger)
2. Same as 1 (Silvina Solman)
3. What do development innovations mean for model selecton e.g. the CMIP7 focus on CO2 emission-driven simulations, increasing resolution and AI/ML? (Roland Séférian)
4. Understanding, quantifying and communicating uncertainty (Michael Grose)

Breakout groups: Other challenges

1. Constraints on the availability of the required GCM/ESM model data – timing, provision of all required data, incl. temporal frequencies and extensions (Sophie Nowicki)
2. Sustained and supported infrastructure to store, deliver and provide user friendly platforms for analysis, framework for coordinated exchange between communities (Michael Grose)
3. Speeding up the process from creation of the global simulations until data/information reaches the end user including potential role of community developed tools to support model evaluation and selection (Christian Steger)
4. Balancing competing needs, funding, politics and ensuring equity, and supporting training and capacity building in a global community effort (Helene Hewitt)

Thank You

@wcrp-cmip.org

wcrp-cmip

cmip-ipo@esa.int

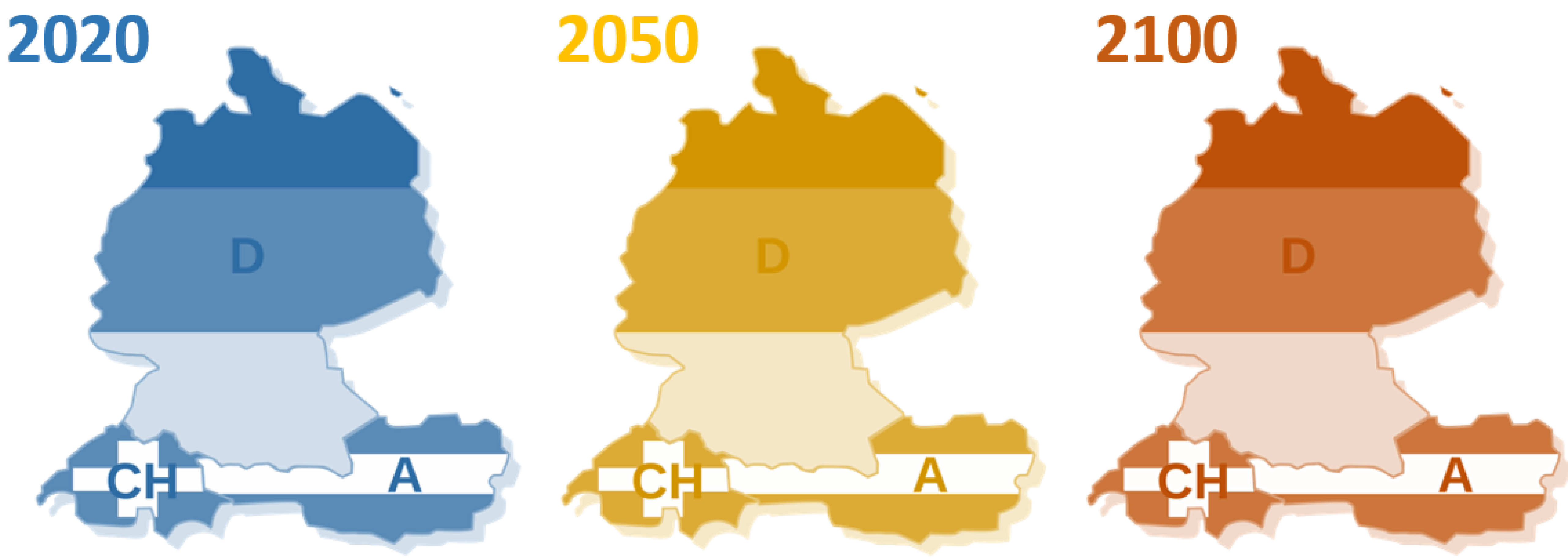
Cross-boarder climate scenarios

An example of challenges faced by users in the current framework

Sven Kotlarski (MeteoSwiss)
Harald Rybka, Nora Leps, Christian Steger (Deutscher Wetterdienst DWD)
Theresa Schellander-Gorgas (GeoSphere Austria)
Martha Kogler (University of Vienna)

D-A-CH Collaboration on climate scenarios

- **Long-lasting and intense collaboration** of the German, Austrian and Swiss Met Services on weather and climate services, including climate scenarios
- **Exchange** of models, methods, tools, know how
- **National scenarios** are mostly based on dynamically downscaled CMIP simulations (**EURO-CORDEX**)
- Mid-term aim: Provide **consistent cross-boarder scenarios**



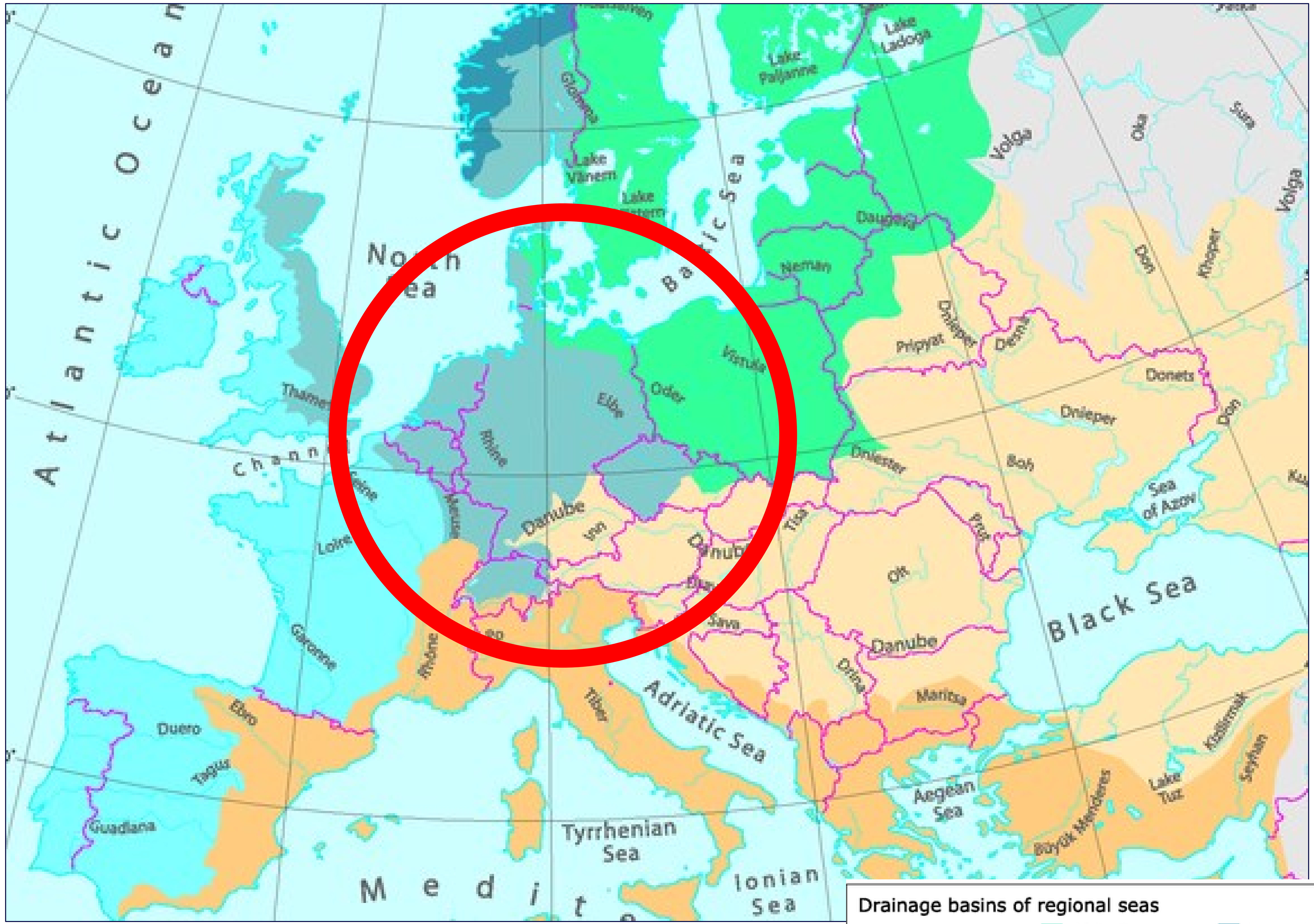
The need for consistent cross-boarder scenarios

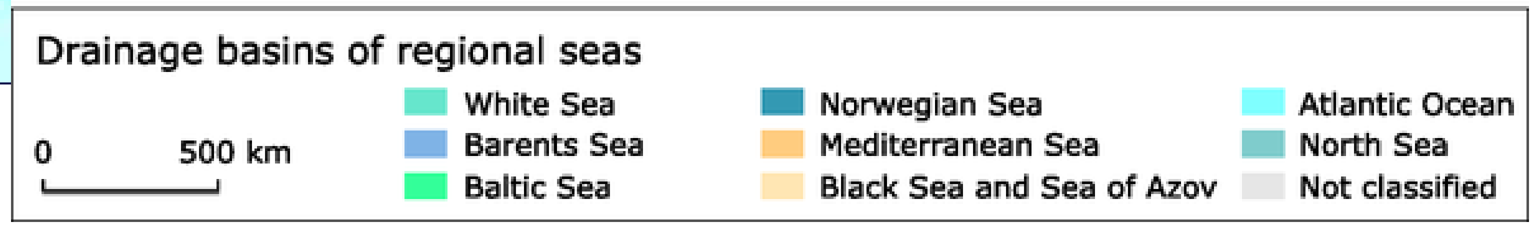
National boarders **do not align** with drainage basins

https://www.eea.europa.eu/

CMIP | WCRP | CORDEX | ISIMIP Inter-Sectoral Impact Model Intercomparison Project | ISMIP7 Ice Sheet Model Intercomparison Project

The need for consistent cross-boarder scenarios





https://www.eea.europa.eu/

National boarders **do not align** with drainage basins

Challenges (non-exhaustive):

- Differing national **time lines** (e.g., adaptation strategies)
- Differing **funding schemes**
- Differing climate **monitoring** standards and different monitoring grids
- Different **focus topics** in terms of climate change impacts (e.g. sea level rise vs. snow scarcity)
- Different "present-day" **reference periods**
- Overall: **Delayed availability** of dynamically downscaled ensembles → hard to follow IPCC cycles

CMIP | WCRP | CORDEX | ISIMIP Inter-Sectoral Impact Model Intercomparison Project | ISMIP7 Ice Sheet Model Intercomparison Project

Differences in national-scale scenario data

- *NN*
- **44** (core: **17**) EURO-CORDEX simulations
- Downscaling and bias adjustment by **quantile delta mapping and MBC**
- **Observational grid**: 5 km
- Specific set of **indicators**
- ...

- *ÖKS15*
- **26** EURO-CORDEX simulations
- Downscaling and bias adjustment by **scaled distribution mapping**
- **Observational grid**: 1 km
- Specific set of **indicators**
- ...

- *CH2018*
- **68** EURO-CORDEX simulations
- Downscaling and bias adjustment by **empirical quantile mapping**
- **Observational grid**: 2 km
- Specific set of **indicators**
- ...

Differences in national-scale scenario data

- *NN*
- **44** (core: **17**) EURO-CORDEX simulations
- **Observational grid**: 5 km
- Specific set of **indicators**
- …

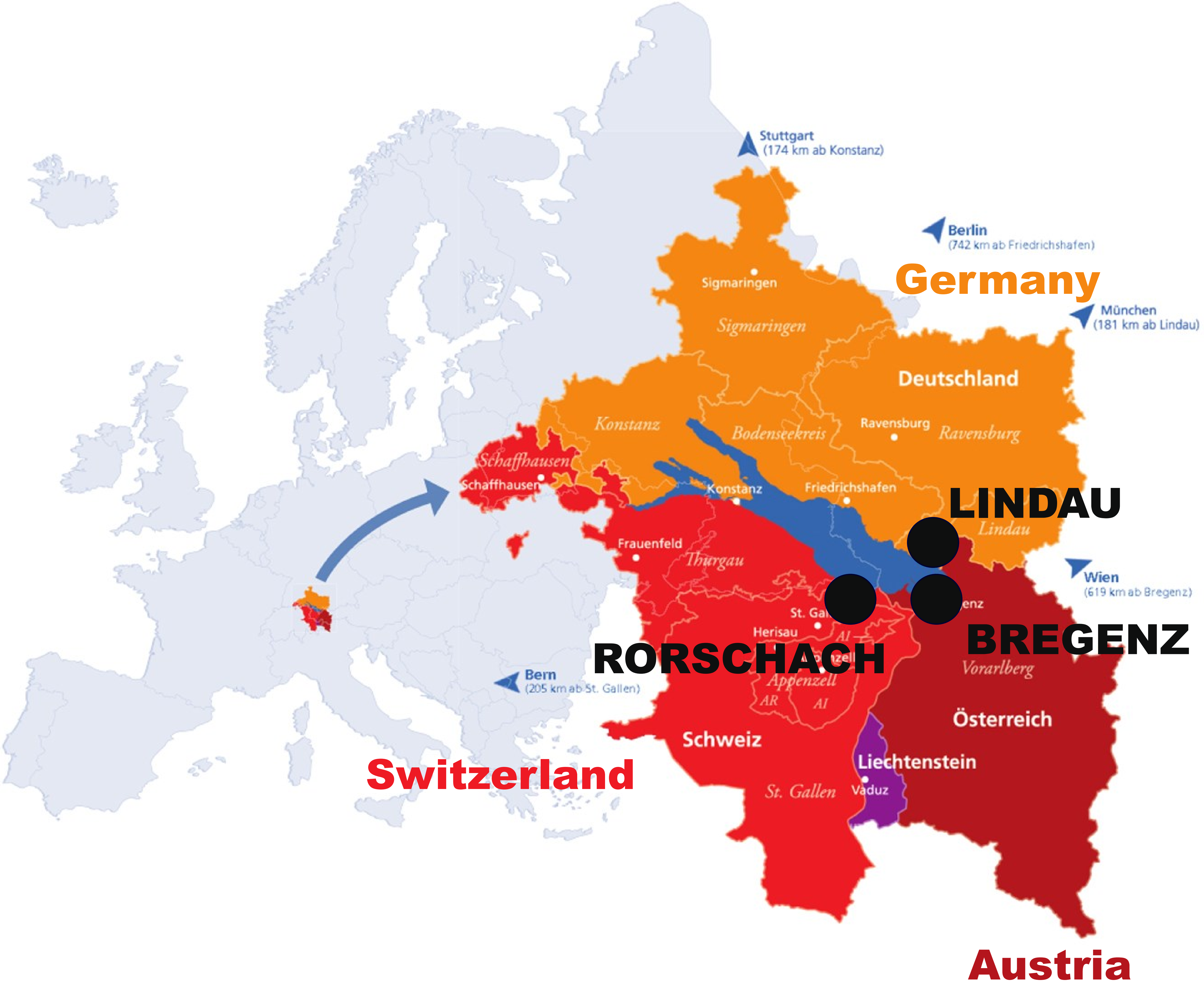
- *ÖKS15*
- **26** EURO-CORDEX simulations
- **Observational grid**: 1 km
- Specific set of **indicators**
- …

- *CH2018*
- **68** EURO-CORDEX simulations
- **Observational grid**: 2 km
- Specific set of **indicators**
- …

Selection criteria: Availability of simulations, evaluation of performance, consistent ensembles for different emission scenarios, institutional commitments, …

Consequences: Lake of Constance region

Consequences: Lake of Constance region

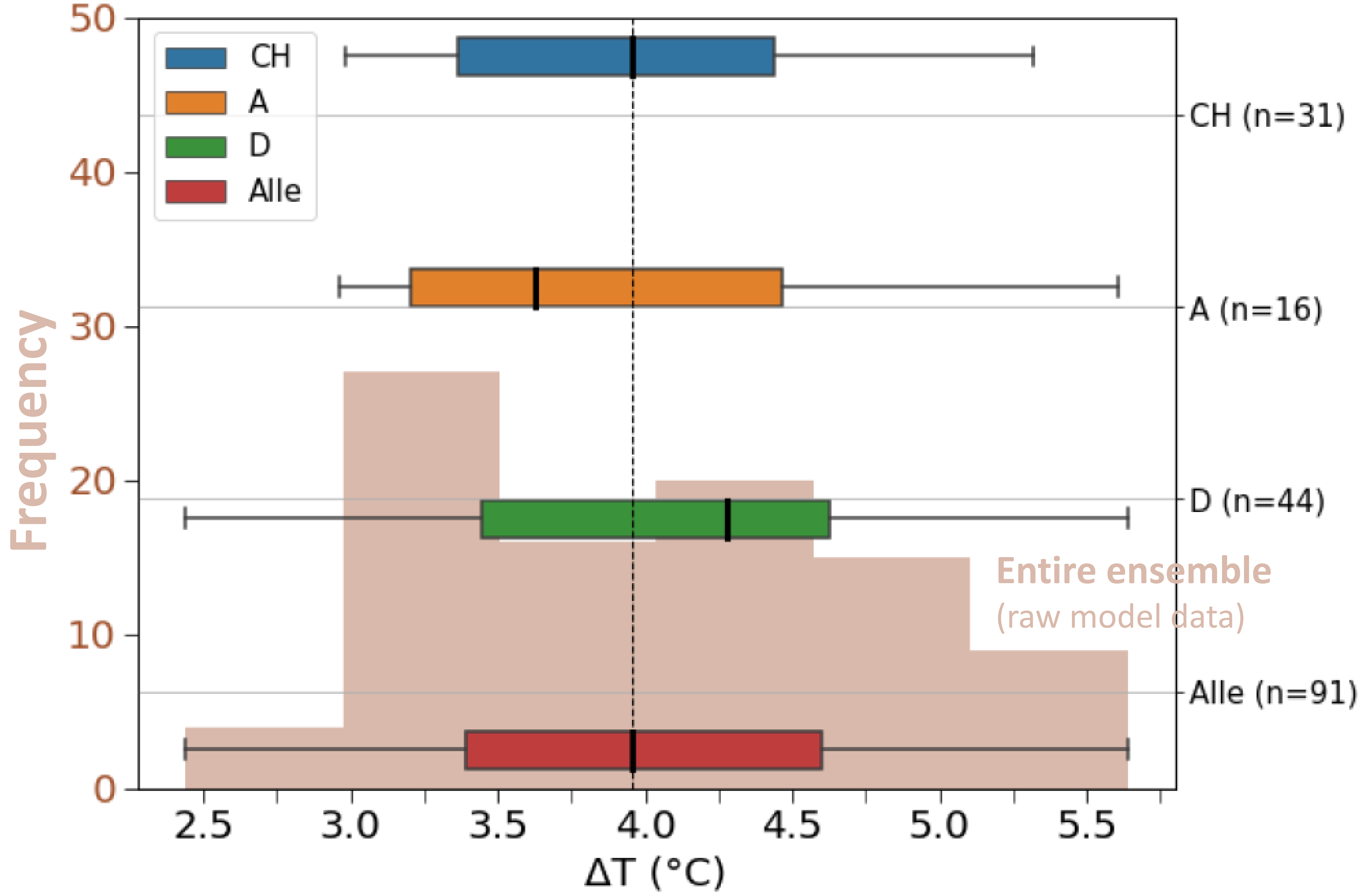


Comparison of climate scenario data obtained by each city administration from their respective national service provider

All cities located within the same 2-3 EURO-CORDEX 12 km grid cells

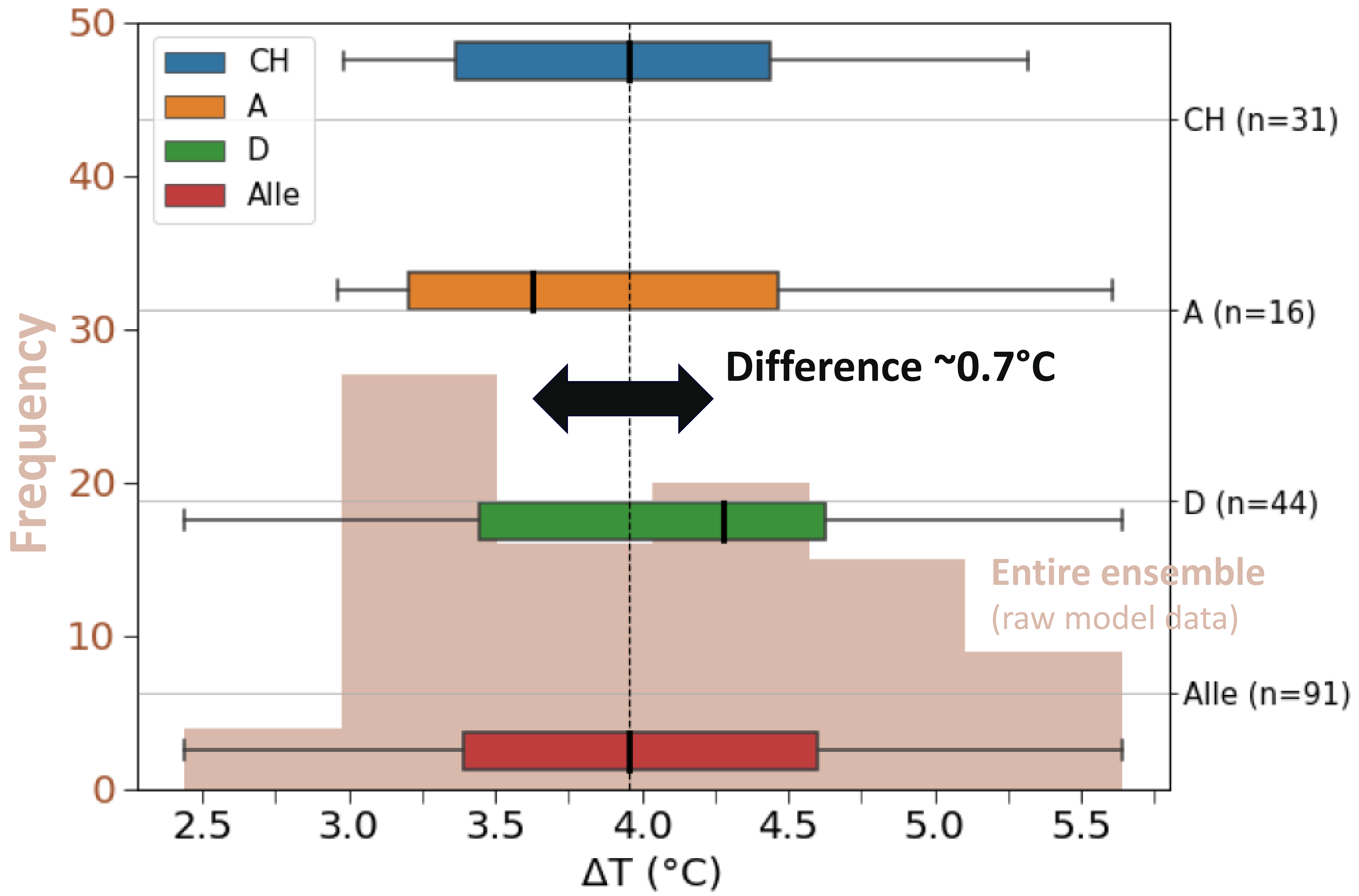
Differences in annual mean temperature change

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



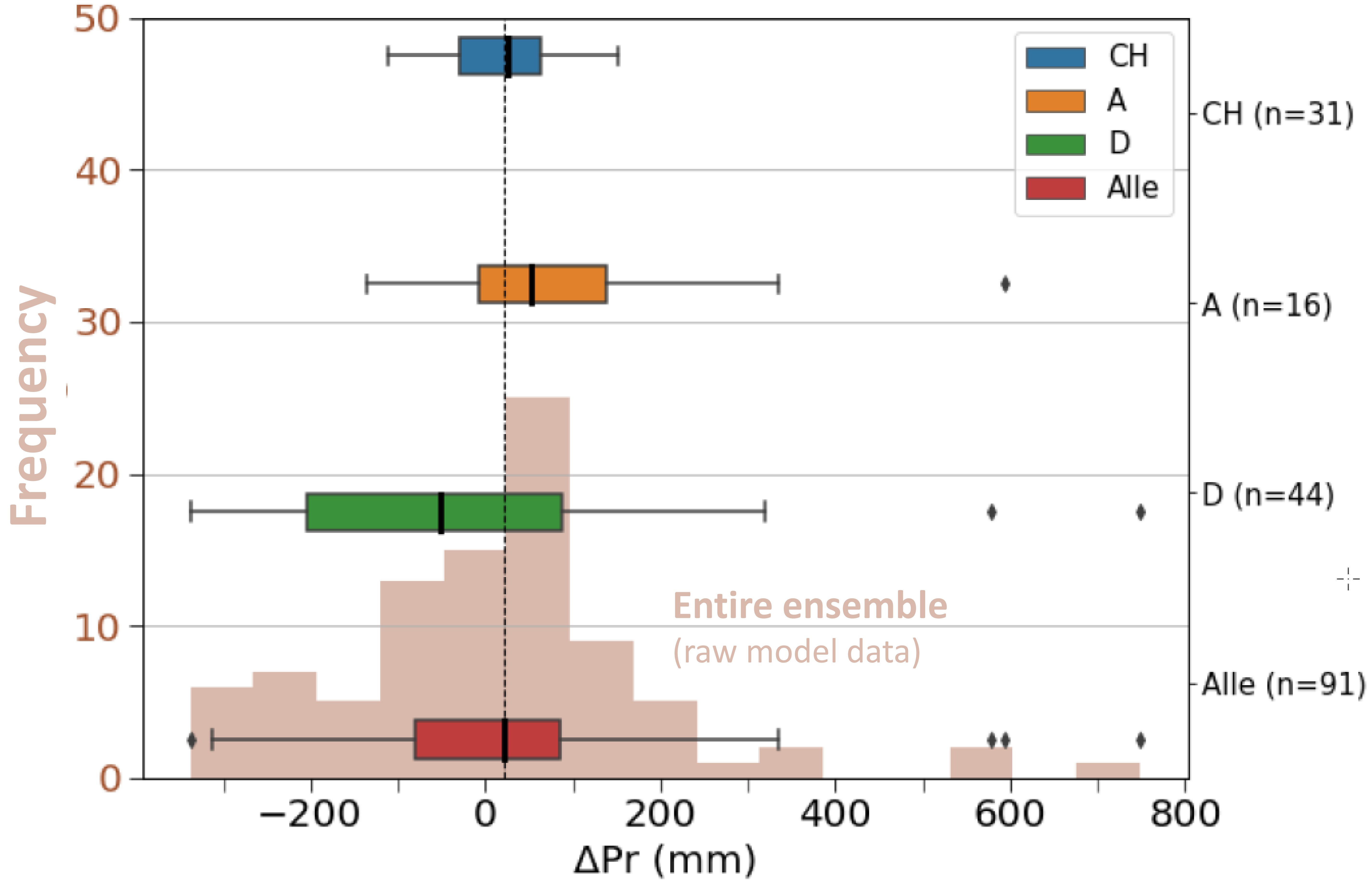
Differences in annual mean temperature change

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



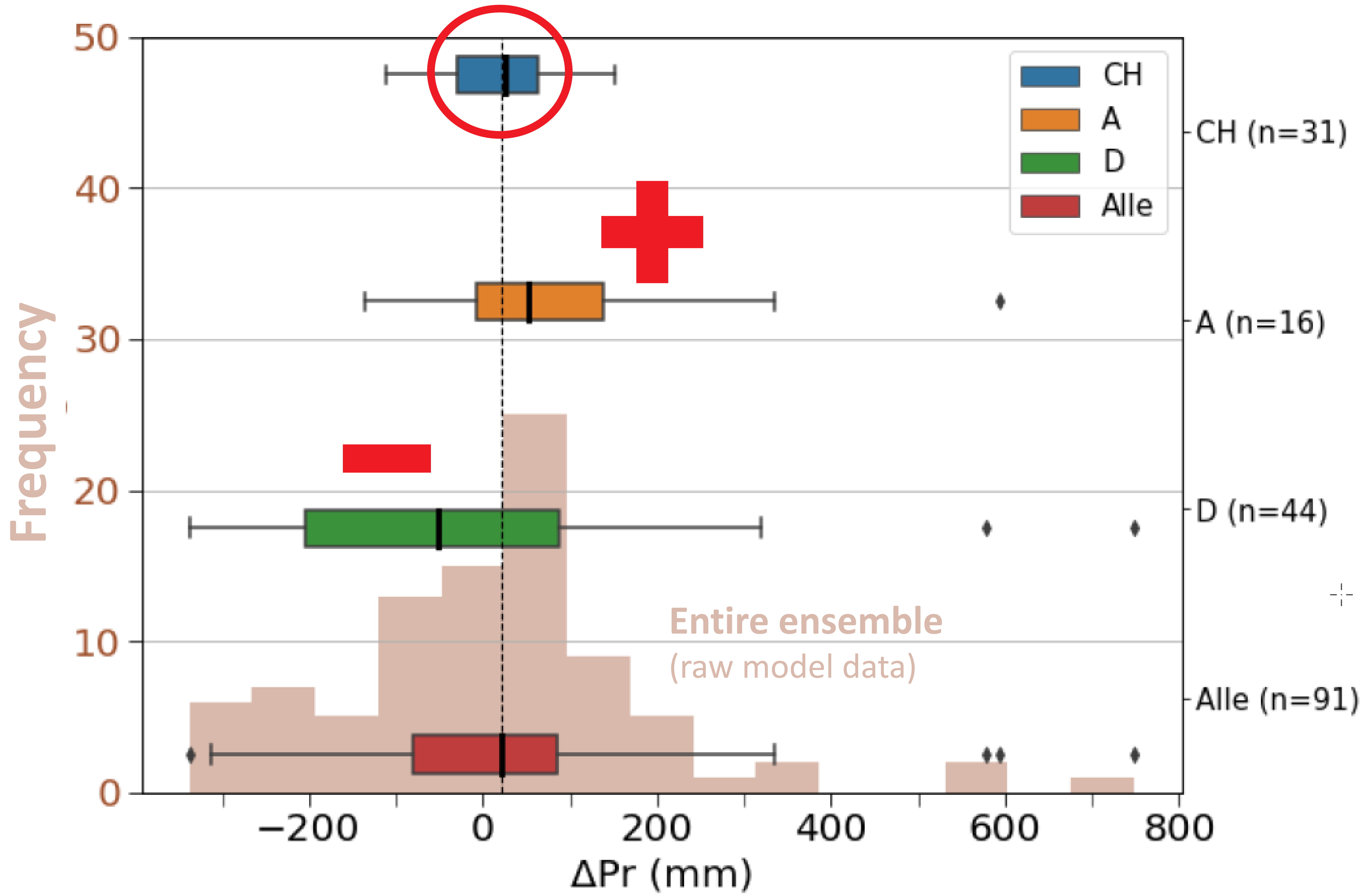
Differences in annual mean precipitation change

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



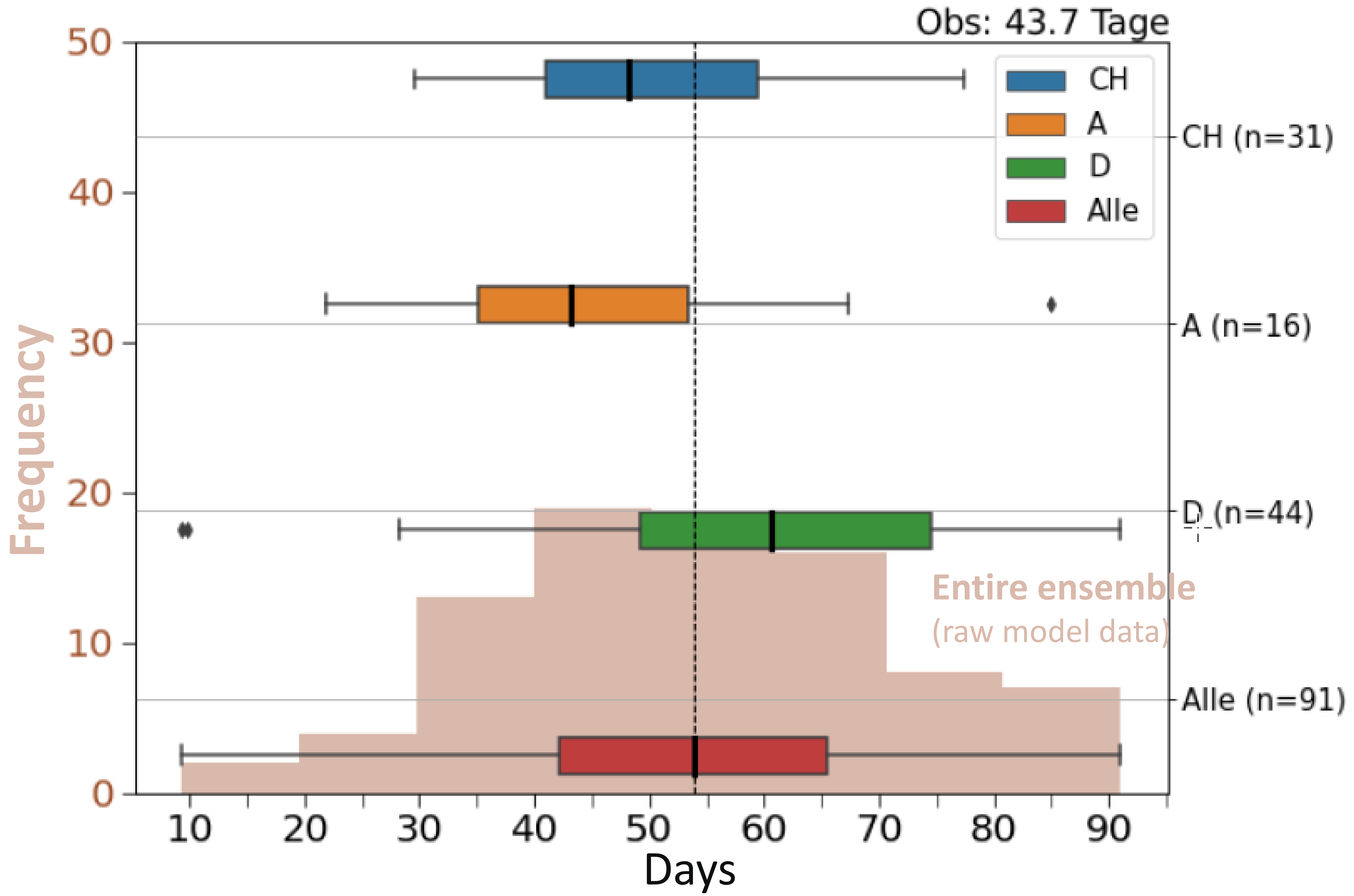
Differences in annual mean precipitation change

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



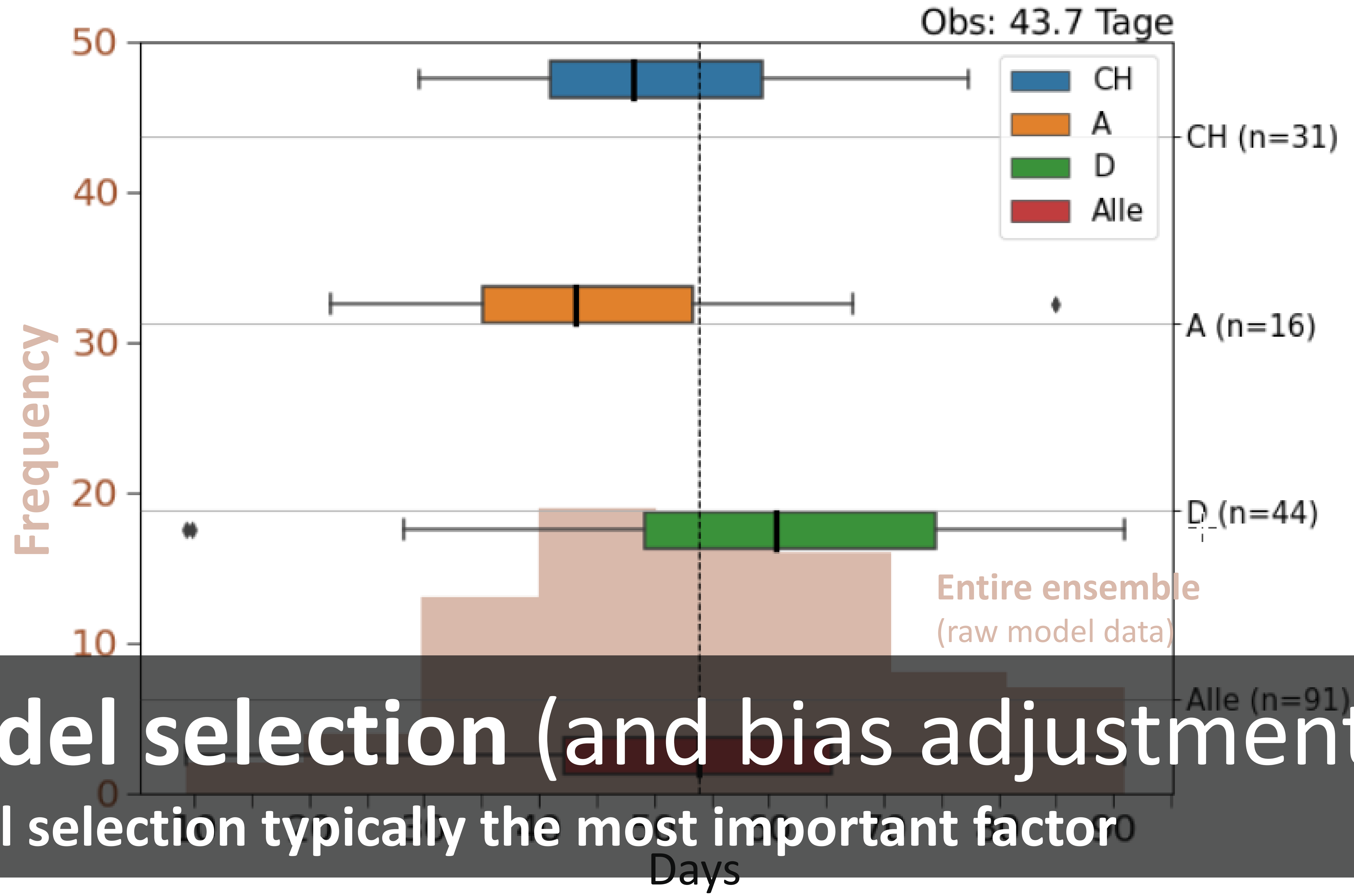
Differences change in the annual number of summer days

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



Differences change in the annual number of summer days

RCP8.5, end-of-century wrt. present-day, influence of model selection **and** bias adjustment



Model selection (and bias adjustment) matter!

Model selection typically the most important factor

Ways forward

- Harmonization of **national frameworks** (time lines, refernce periods, indicators etc.)
- **Joint evaluation and eventually model selection** on regional scale: Currently under way in EURO-CORDEX (Sobolowski et al., BAMS, in review) and D-A-CH
- **More rapid dynamical downscaling** of CMIP simulations, better integration of statistical downscaling and high-res GCMs

Ways forward

- Harmonization of **national frameworks** (time lines, refernce periods, indicators etc.)
- **Joint evaluation and eventually model selection** on regional scale: Currently under way in EURO-CORDEX (Sobolowski et al., BAMS, in review) and D-A-CH
- **More rapid dynamical downscaling** of CMIP simulations, better integration of statistical downscaling and high-res GCMs
- CMIP: Fast(er) provision of **RCM forcing data**
- CMIP: Well-informed **GCM selection** and **consistent ensembles across emission scenarios**
- …

Thank You

sven.kotlarski@meteoswiss.ch

@wcrp-cmip.org

wcrp-cmip

cmip-ipo@esa.int

Avoiding the curse of opportunity: best practices from the EURO-CORDEX community

Stefan Sobolowski and the entire CMIP6 task team
(special thanks to Jesus Fernandez, Samuel Somot)

28.09.2023 ICRC-CORDEX Session D4 Trieste, Italy
Updated 05.05.2025 For CMIP7 Model Selection Wkshp

Motivations:

- Stop the “ensemble of opportunity” approach used in the RCM community since the 90s. Improve upon GCM selection routines from CMIP5
- Make CMIP6/EURO-CORDEX ensemble more reliable to explore future climate change and therefore a better climate information source for adaptation strategies
- Avoid to run “useless” simulations (picking implausible GCMs without knowing it)
- Better explore the range of plausible futures
- Create an a ”balanced” matrix subset of simulations for practitioners and VIACS community

Goals of the “Task Team”:

- Develop a set of best-practice guidelines
- Base these on existing literature & expert judgment following internal discussions
- Execute design of RCM-GCM ensembles (i.e. “The Matrix”) in less of an ad-hoc manner

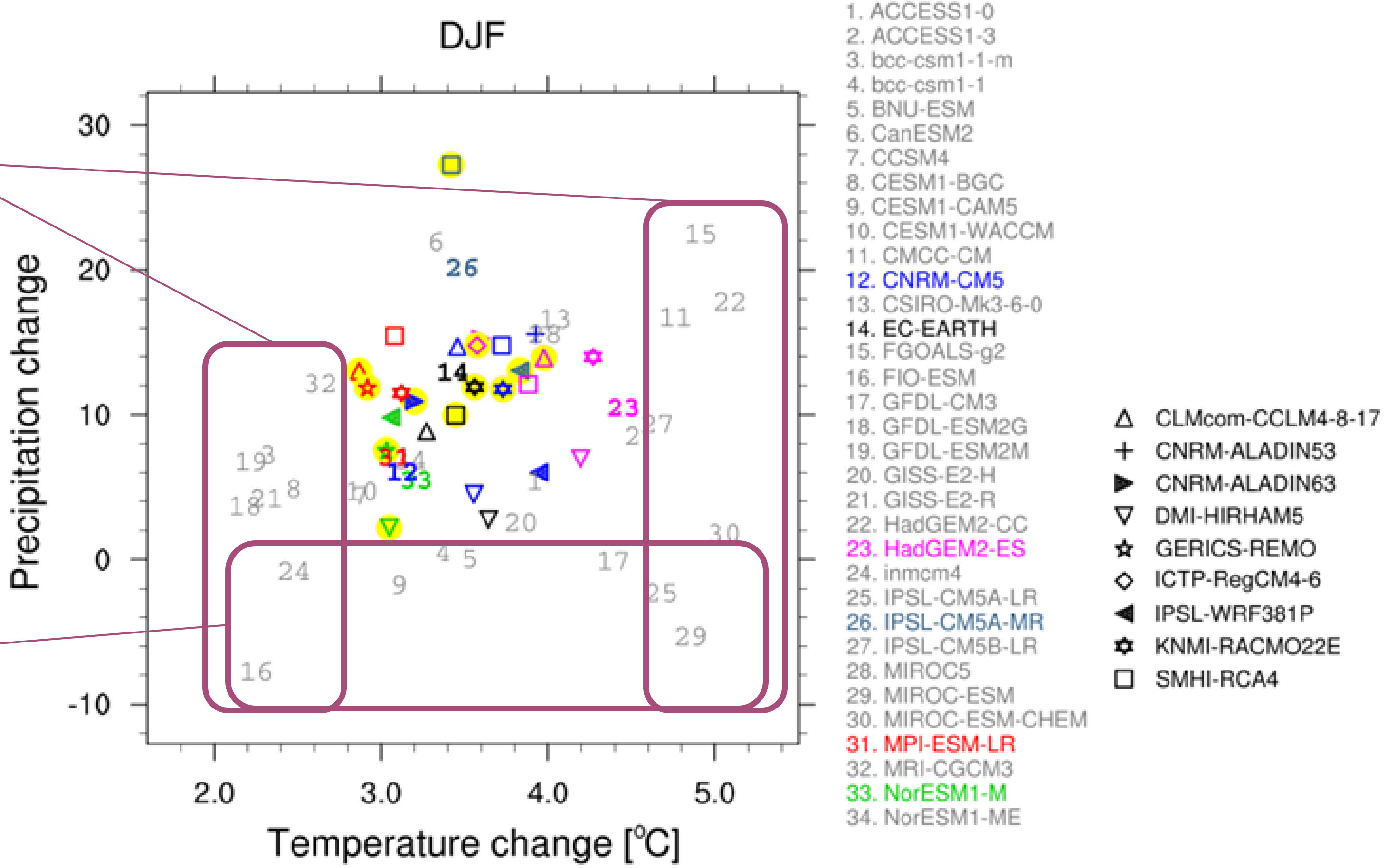
Note that the proposed protocol is strongly influenced by the spirit of McSweeney et al. 2015

We need to better explore the range of plausible futures : illustration with the current Euro-CDX ensemble for France in Winter

Future change in precipitation (%) and near-surface temperature (°C)

(France, DJF, 2071-2100 vs 1976-2005)

The GCMs with most and least warming in CMIP5 are missing in Euro-CDX forcing



The GCMs with Winter drying in CMIP5 are missing in Euro-CDX

Figures: L. Corre, Meteo-France: figure done with 6 most used driving GCMs in Euro-CORDEX among the 9 driving GCMs

Three approach: 3(4) selection criteria families

Data availability/ quality

- CORDEX-MIP
  - Availability of scenario so balanced ensembles can be produced (e.g. ssp126 and ssp585)
  - Availability of variables to (1) evaluate the GCM in step 2 and 3, (2) drive the RCMs and (3) use ESD or hybrid approaches
- Basic QA (missing values, suspect values)
- FAIR meta-data (Lars can maybe say something about this?)

Eliminate Implausible GCMs

- Global climate criteria (e.g. ECS, TCR, past trend representation of the global-mean temperature, known strange behaviour such as spurious ocean trends or breaks)
- Favour model (bio-)diversity : independence criteria (https://esd.copernicus.org/preprints/esd-2020-23/esd-2020-23.pdf), etc.
- European large scale performance criteria (.e.gg. 850hPa winds, Stormtracks, Jet stream (strength/position), trend reproducibility, low frequency variability such as NAO or weather regimes, Hadley circulation, ECS, Low-level humidity advection, African and Asian monsoon
- European forcing performance criteria: AOD, regional SST and SIC, IWV, etc. [key here is to think of factors that will influence the RCM]

Explore the range of future outcomes

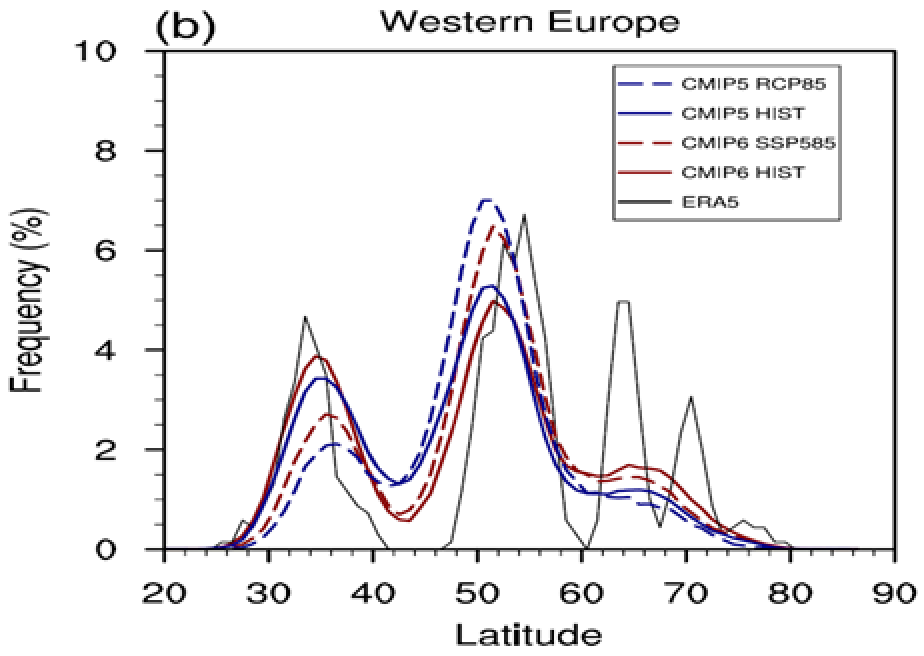
- European TAS response (delta-T spread) à la McSweeney et al. 2015, Spread in ECS or TCR (low, medium, high), Spread in GCMST for e.g. ssp585?
- Spread in future response of circulation pattern or large-scale features controlling European climate change (e.g. robust storm track shifts as in Oudar et al., 2020: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL086695), Polar vortex strength and Tropical amplification (see storyline's paper by Zappa and Shepherd), and/or other emergent constraints
- Spread in other forcings: regional SST and SIC, regional AOD trend,

Large-scale performance criteria: illustrations by the North-Atlantic storm track

Bias in the storm track north position for CMIP6 GCMs
(ONDJFM, position in °N)

CMIP6 models	Jet Bias
BCC-CSM2-MR	-1.49
CAMS-CSM1-0	-1.91
CESM2	1.42
CESM2-WACCM	0.28
CNRM-CM6-1	-2.79
CNRM-ESM2-1	-3.02
CanESM5	0.9
EC-Earth3	-0.26
EC-Earth3-Veg	-0.69
FGOALS-g3	-0.43
GFDL-CM4	-2.1
GFDL-ESM4	-3.4
INM-CM4-8	1.01
INM-CM5-0	0.03
IPSL-CM6A-LR	-1.53
MCM-UA-1-0	-0.39
MIROC6	-3.31
MIROC-ES2L	-6.97
MPI-ESM1-2-HR	-3.1
MRI-ESM2-0	-2.66
NESM3	-2.03
UKESM1-0-LL	-0.43

Maximum wind position distribution for CMIP6 GCMs
(ONDJFM)

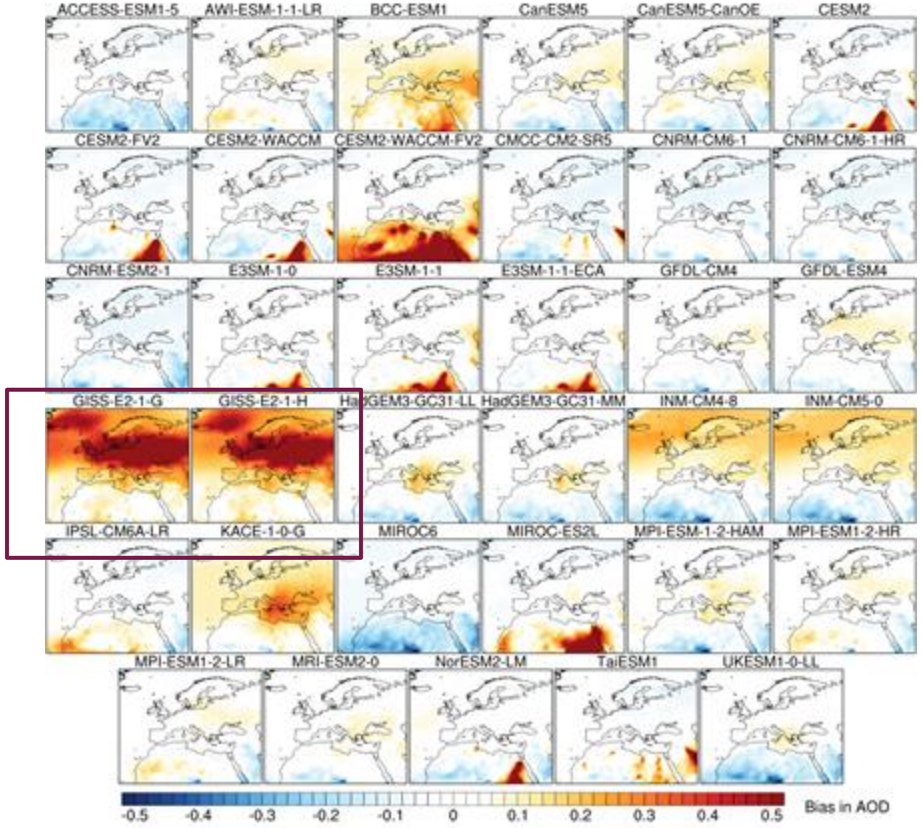


Implausible driving model
7° too far south!

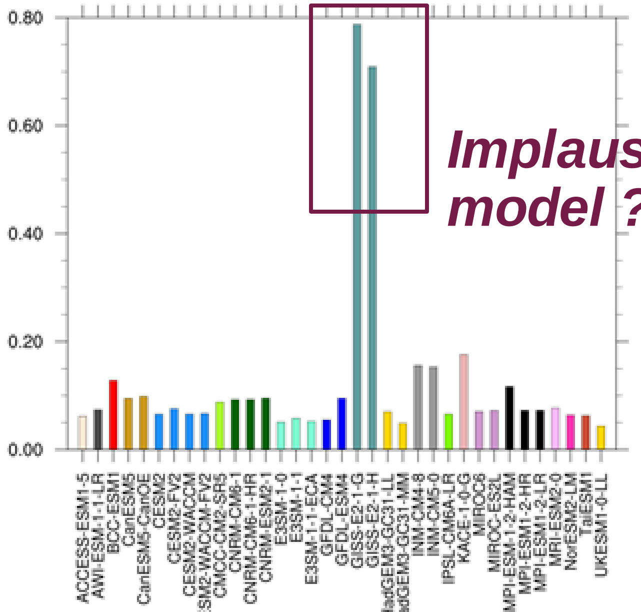
Oudar et al. 2020

Regional forcing performance criteria: illustrations by the European Aerosol Optical Depth

Yearly-mean AOD bias for CMIP6 GCMs
(yearly-mean, 2000-2014, wrt satellite data)



RMSE (yearly-mean, 2000-2014, wrt satellite data)



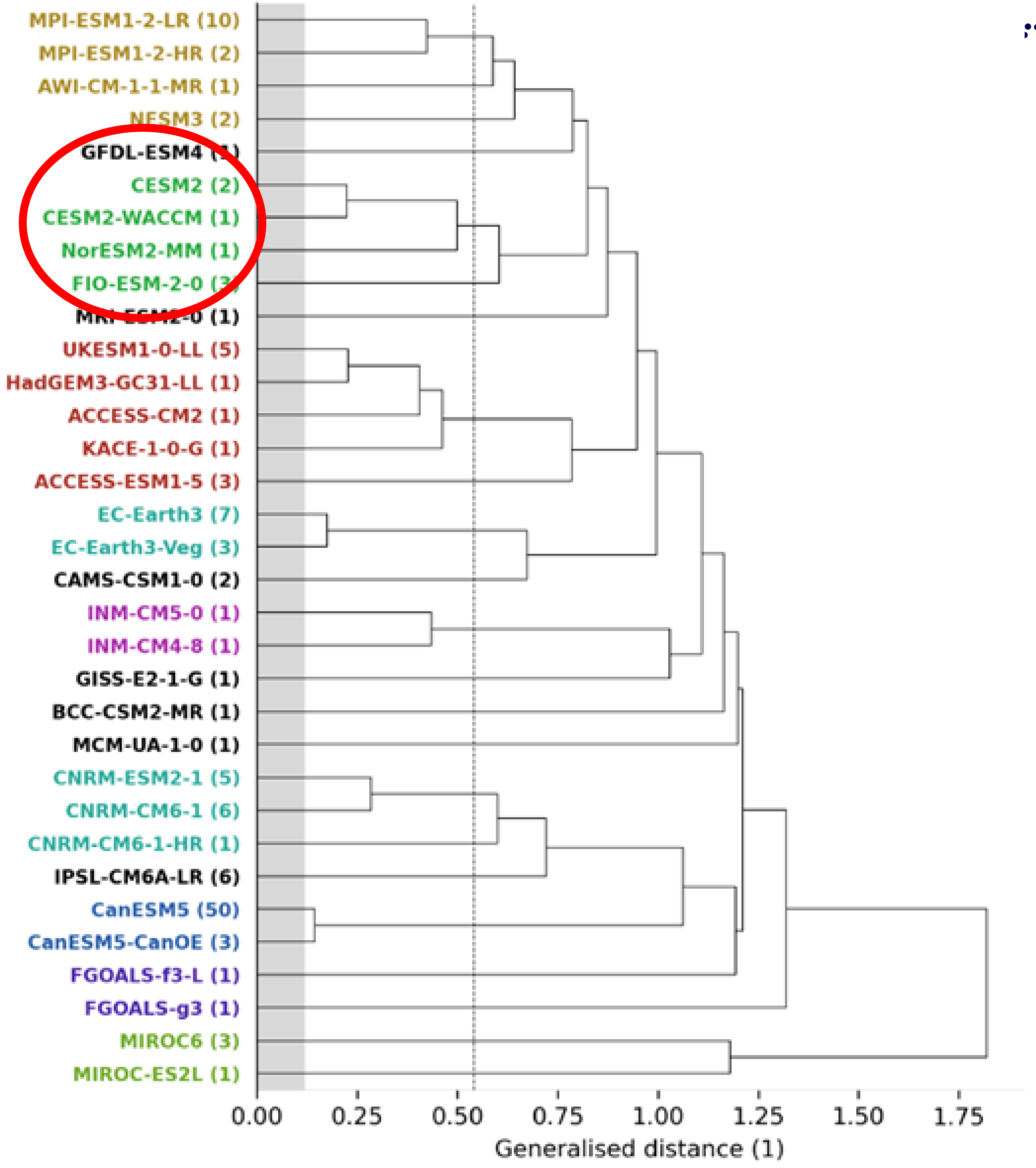
Implausible driving model ?

P. Nabat (CNRM) , pers. comm.

Global/Other criteria: model (bio-) diversity

GCM are not independent
they can be gathered by families
“End of the model democracy”. R. Knutti

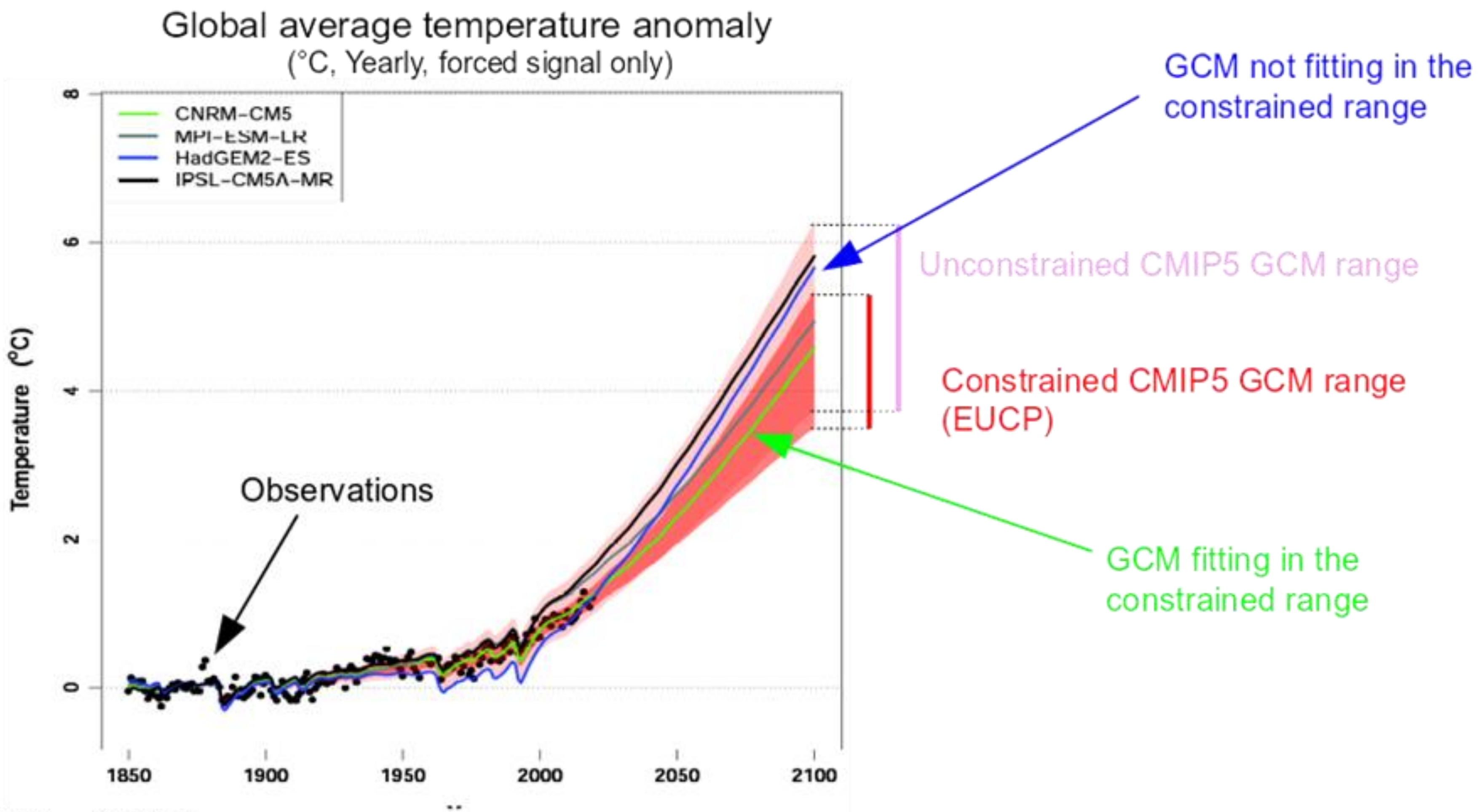
(Here similarity criteria based on global tas and psl field
1980-2014)



Brunner et al. 2020

Global criteria: constraint on the global average temperature past trend

In this example, only 2 CMIP5 GCMs out of 4 (used as drivers in Euro-CORDEX) fits the observational constraints → meaning that at least 2 Euro-CORDEX driving GCMs are implausible wrt this specific metric!



A. Ribes (CNRM), pers. comm.

How this works in practice: move to a traceable, transparent, extendable approach

New implementation to collect GCM information:

- based on published scientific literature
- extended by author contributions
- described by more than just numbers, incorporating decision thresholds
- human readable
- machine readable
- extendable (e.g. to other CORDEX domains)
- traceable, recording the decision process and alternative decisions
    - Open & collaborative
    - Version control
    - Text files
    - Programming to process the information in different ways
    - Issues to store the decision process

Outcomes: GCM recommendations for EURO-CORDEX

Table 1. *Most strict; GCMs which are available for all four scenarios (ssp126, ssp245, ssp370,ssp585) and are deemed "plausible" for each evaluated criteria. To qualify models must be evaluated for at least one criterion per score family. The third column shows the number of failed criteria over the total number of criteria for each model. Models that are also part of institutional commitments are highlighted. The fourth column shows an illustration of future spread categories for the selected GCMs (here based on TCR values).*

GCM name	Run	Marks/Criteria	TCR Plausible range (1.2K-2.4K)[12]
MPI-ESM1-2-LR	r1i1p1f1	0/18	1.84

Table 2. *Less strict; same as Table 1 except for GCMs which are "available" for all four scenarios. Scores are based on all evaluated members of a model even if only one member is "available". Only one model per family is kept in most cases and in the event of a tie criteria such as complexity and resolution may play a role as tie-breakers. Explanations appear in footnotes.*

GCM name	Run	Marks/Criteria	TCR Plausible range (1.2K-2.4K)
NorESM2-MM[13]	r1i1p1f1	1/17	1.33
MIROC6[14]	r1i1p1f1	1/20	1.55
MPI-ESM1-2-HR	r1i1p1f1	1/20	1.66
CNRM-ESM2-1	r1i1p1f2	1/19	1.86
CESM2[15]	r11i1p1f1	1/18	2.06
CMCC-CM2-SR5[16]	r1i1p1f1	1/15	2.09
IPSL-CM6A-LR[17]	r1i1p1f1	2/16	2.32
EC-Earth3-Veg[18]	r1i1p1f1	2/15	2.62
UKESM1-0-LL[19]	r1i1p1f2	2/19	2.79

Outcomes: CMIP6 - EURO-CORDEX “balanced” matrix: first final version

RCM \ GCM	EC-Earth3-Veg (*) r1i1p1f1	MPI-ESM1-2-HR r1i1p1f1	CNRM-ESM2-1 r1i1p1f2	NorESM2-MM r1i1p1f1	MIROC6 r1i1p1f1	CMCC-CM2-SR5 (*) r1i1p1f1
WRF		✗		✗		✗
ALADIN6x			✗	✗		✗
COSMO/ICON-CLM	(✗)	✗			✗	✗
HCLIM43-ALADIN	(✗)	✗	✗	(✗)	✗	
RegCM5	✗	(✗)	✗	✗		
REMO	✗	✗			✗	
RACMO23E	✗		✗	✗		

Color = TCR
Plausible range

✓ At least 3 runs by RCM and 4 runs by GCM

? GCM/RCM compatibility?

✗ planned

✗ still to be placed

(*) only total aerosol forcing available on ESGF (**od550aer). (2022.05.17 for EC-Earth3-Veg)**

CORDEX-CMIP6 downscaling plans summary tables

https://wcrp-cordex.github.io/simulation-status/CORDEX_CMIP6_status.html#EUR-11

Based on Eivin et al., 2021

Summary

- Tables to summarize the 4-step GCM selection process are ready to be used and completed with
  - Additional model runs
  - Additional studies
  - Refined decisions on thresholds, preferred metrics for a given aspect

Citable documentation link

https://wcrp-cordex.github.io/cmip6-for-cordex/CMIP6_studies_table_EUR.html

Please, explore the GitHub site and contribute

Link to GitHub pages

https://wcrp-cordex.github.io/cmip6-for-cordex/

Acknowledgements: part of this work was supported by European Union's Horizon Europe R&I programme

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101081555

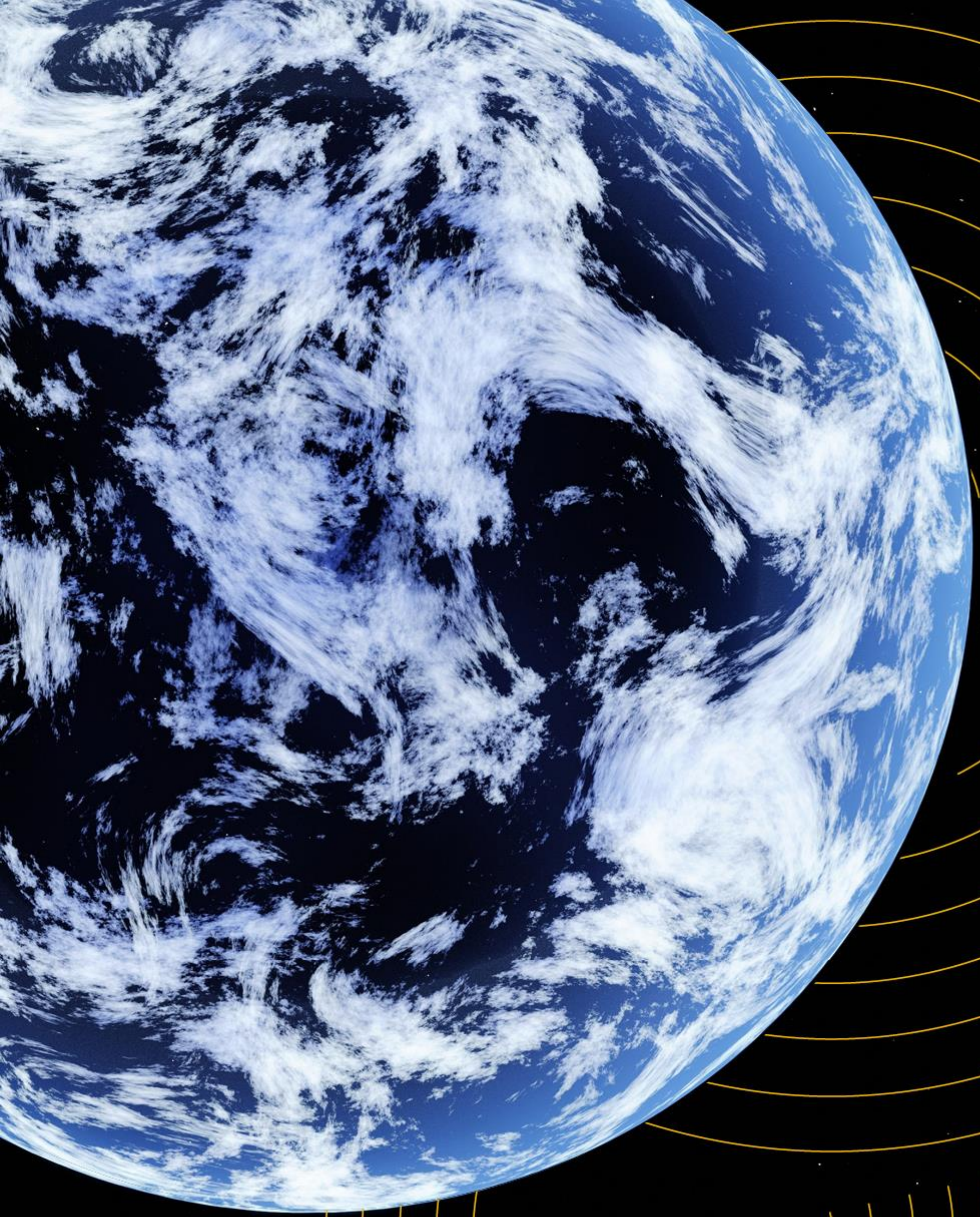
WCRP CORDEX

Thank you. Takk.

Merci. Gracias. Obrigado.

NORCE

BJERKNES CENTRE
for Climate Research



CMIP
Coupled Model Intercomparison Project

Streamlining model selection

Wednesday 5th February 2025, 19:00–22:00 UTC
Virtual workshop

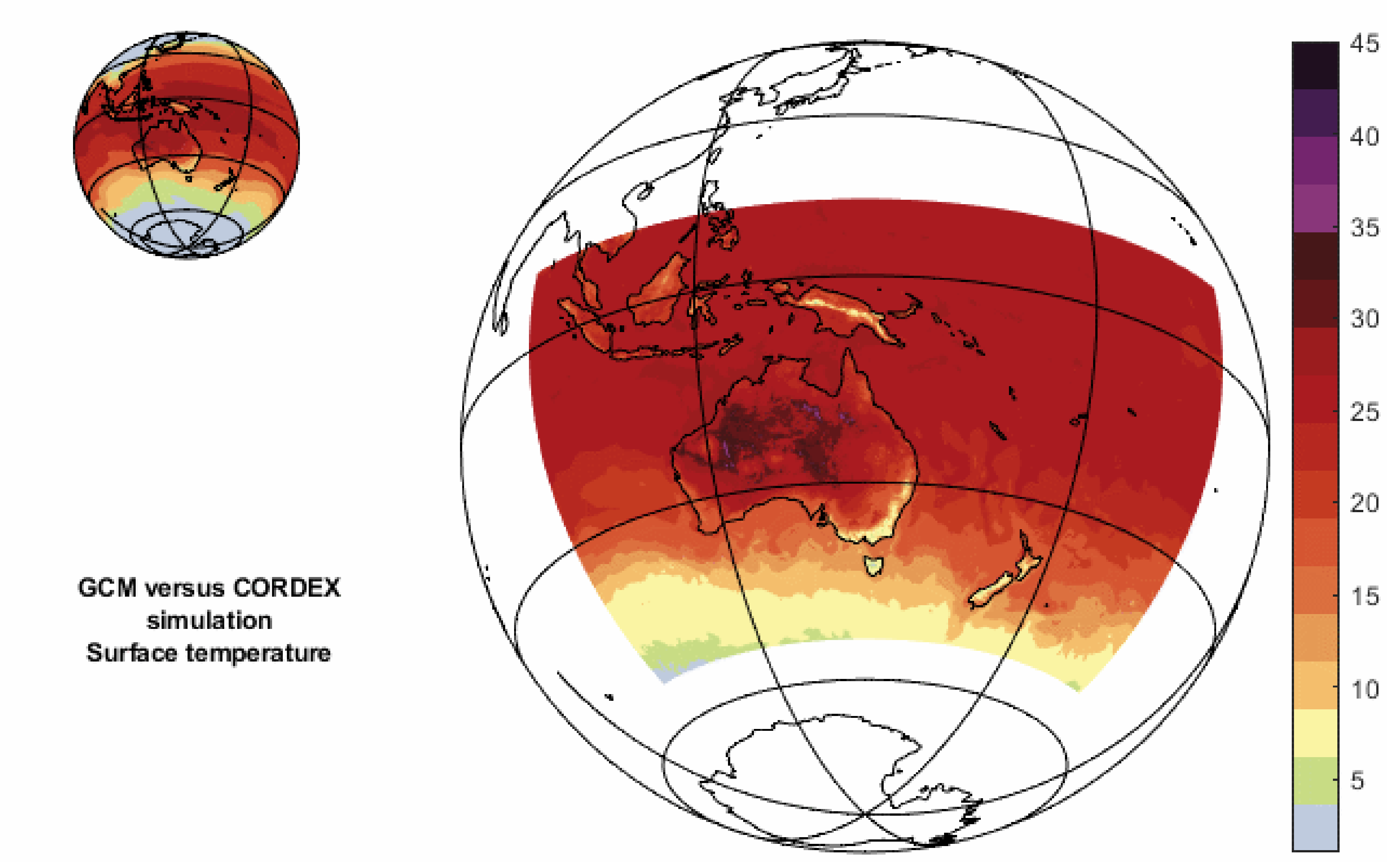
CMIP6 CORDEX-Australasia for Australian national projections

Michael Grose (CSIRO) on behalf of the National Partnership for Climate Projections (NPCP) working group

CMIP | WCRP | CORDEX | ISIMIP Inter-Sectoral Impact Model Intercomparison Project | ISMIP7 Ice Sheet Model Intercomparison Project

NextGen Projections strategy

- Interest in updated national and state-based projections
- Major new resource – a coordinated multi-model, multi-scenario RCM ensemble
- Complemented by CMIP6, large ensembles,
- CORDEX guidelines for production – international benchmarking and comparability
- Requires model selection - three studies performed – useful to compare results



- Grose et al. (2023) A CMIP6-based multi-model downscaling ensemble to underpin climate change services in Australia. Climate Services.
- DiVirgilio et al. (2022) Selecting CMIP6 GCMs for CORDEX dynamical downscaling: Model performance, independence, and climate change signals. Earth's Future.
- Syktus et al. (2022) Dynamical downscaling of CMIP6 global models with a variable resolution climate model in the Australian region. ICHSMO conference

CMIP | WCRP | CORDEX | ISIMIP Inter-Sectoral Impact Model Intercomparison Project | ISMIP7 Ice Sheet Model Intercomparison Project

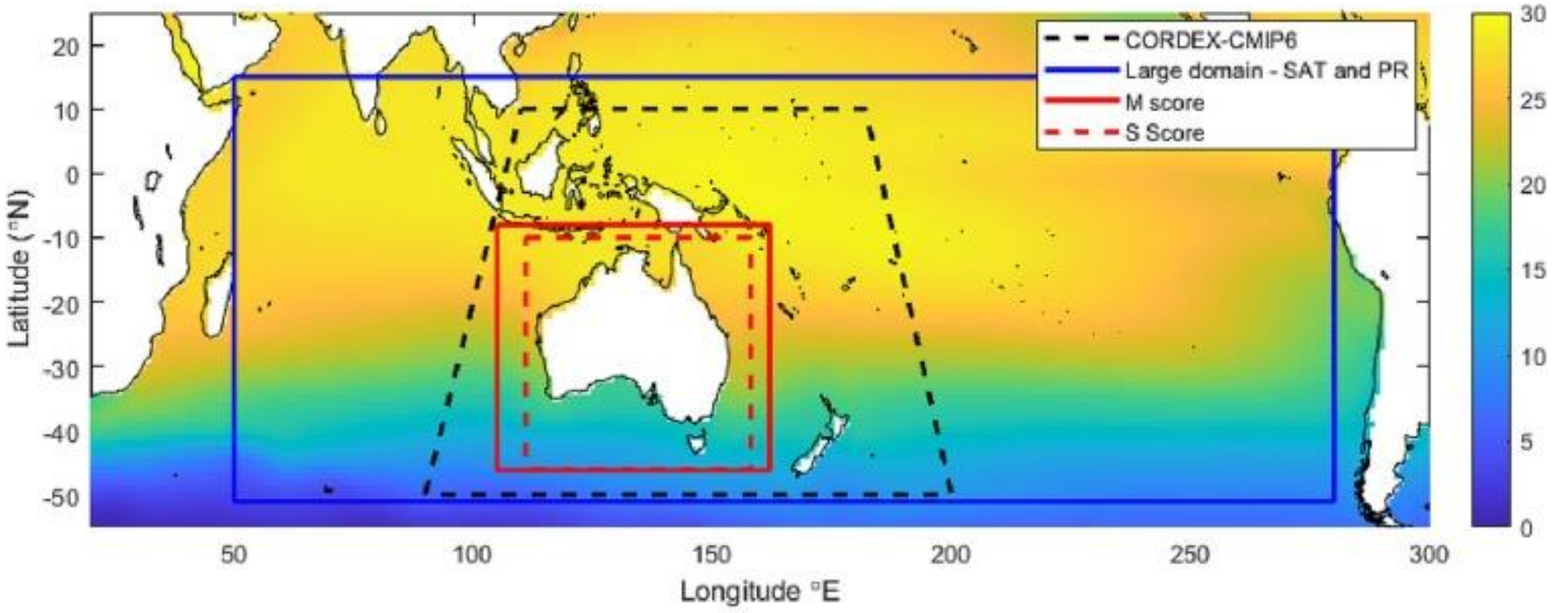
Process of selecting models

- Similarities of the three studies:
  - Standard steps ***evaluation, independence, representativeness***
  - Evaluation used to reject (not select top) – bottom category across many tests
  - Independence – generally simple approach (threshold of similarity)
  - Representativeness – spread of rainfall and temperature, some consideration of circulation, drivers
  - Consideration of 'hot model' problem
- Result – semi-coordinated 'sparse matrix' with some common selections
  - ACCESS-ESM1.5 – very dry projection
  - NorESM2–MM – cooler end
  - EC-Earth3/EC-Earth3–Veg – wet projection
  - Representative hot model

	CCAM-Qld	NARCLIM2.0 (2x WRF configurations)	CCAM	BARPA
ACCESS-CM2	r2i1p1f1oc		r4i1p1f1	r4i1p1f1
ACCESS-ESM1.5	r6i1p1f1 r20i1p1f1oc r40i1p1f1oc	r6i1p1f1	r6i1p1f1	r6i1p1f1
CESM2			r11i1p1f1	r11i1p1f1
CMCC-ESM2	r1i1p1f1		r1i1p1f1	r1i1p1f1
CNRM-CM6.1-HR	r1i1p1f2 r1i1p1f2oc			
CNRM-ESM2-1			r1i1p1f2	
EC-Earth3	r1i1p1f1		r1i1p1f1	r1i1p1f1
EC-Earth3-Veg		r1i1p1f1		
FGOALS-g3	r4i1p1f1			
GFDL-ESM4	r1i1p1f1			
GISS-E2-1-G	r2i1p1f2			
MPI-ESM1-2-HR		r1i1p1f1		
MPI-ESM1-2-LR	r9i1p1f1			
MRI-ESM2-0	r1i1p1f1			
NorESM2-MM	r1i1p1f1 r1i1p1f1oc	r1i1p1f1	r1i1p1f1	r1i1p1f1
UKESM1-0-LL		r1i1p1f1		

Process of selecting models

- Differences – open for debate/discussion
  - Evaluation - statistics of surface variables vs. some focus on drivers, processes
  - Domains of evaluation – whole Indo-Pacific region vs. Australia vs. Sub-regions
  - Different measures of relevant representative climate change signal - warming and precip, also circulation indices (e.g., subtropical ridge)

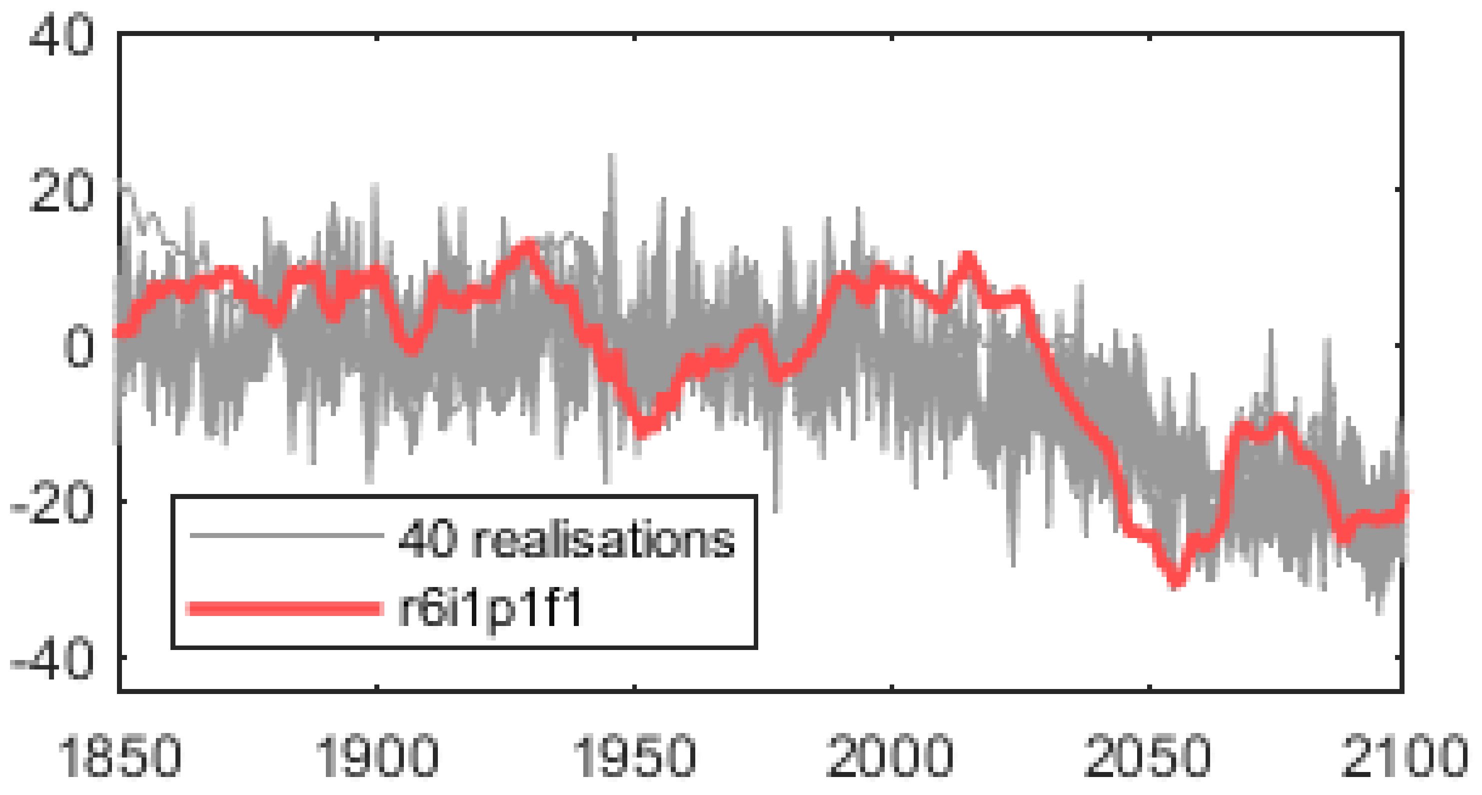


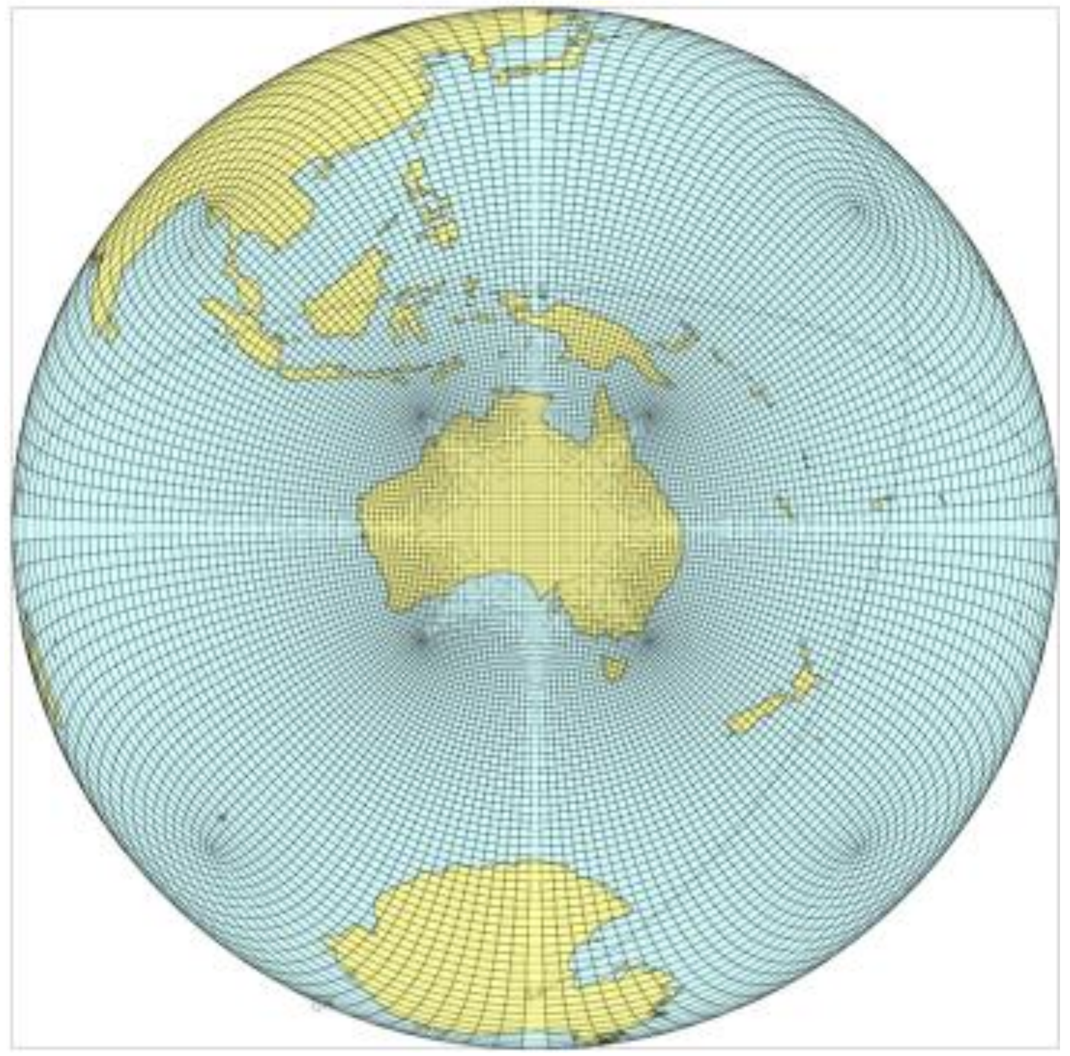
Discussion points

Innovations

- For ACCESS model, we could select realisation and request sub-daily data (selected r6) – *could this be done more widely?*
- Test and compare different RCM configurations – global variable grid vs. limited area, SST bias correction vs. not, ocean coupling

Eastern Aus rainfall – r6 chosen as a stress test through mid-century





Discussion points

Known limitations

- Not a representative sample for all applications – considers only rainfall, temperature, some broad circulation changes, not land surface, carbon cycle, etc.
- Ensemble generation is messy – CMIP6 an ensemble of opportunity, then sub-sample CMIP6, ‘sparse matrix’ not statistically balanced
- Lack of inter-comparability with other regions (different model list)

Discussion points

Barriers – can they be reduced for CMIP7?

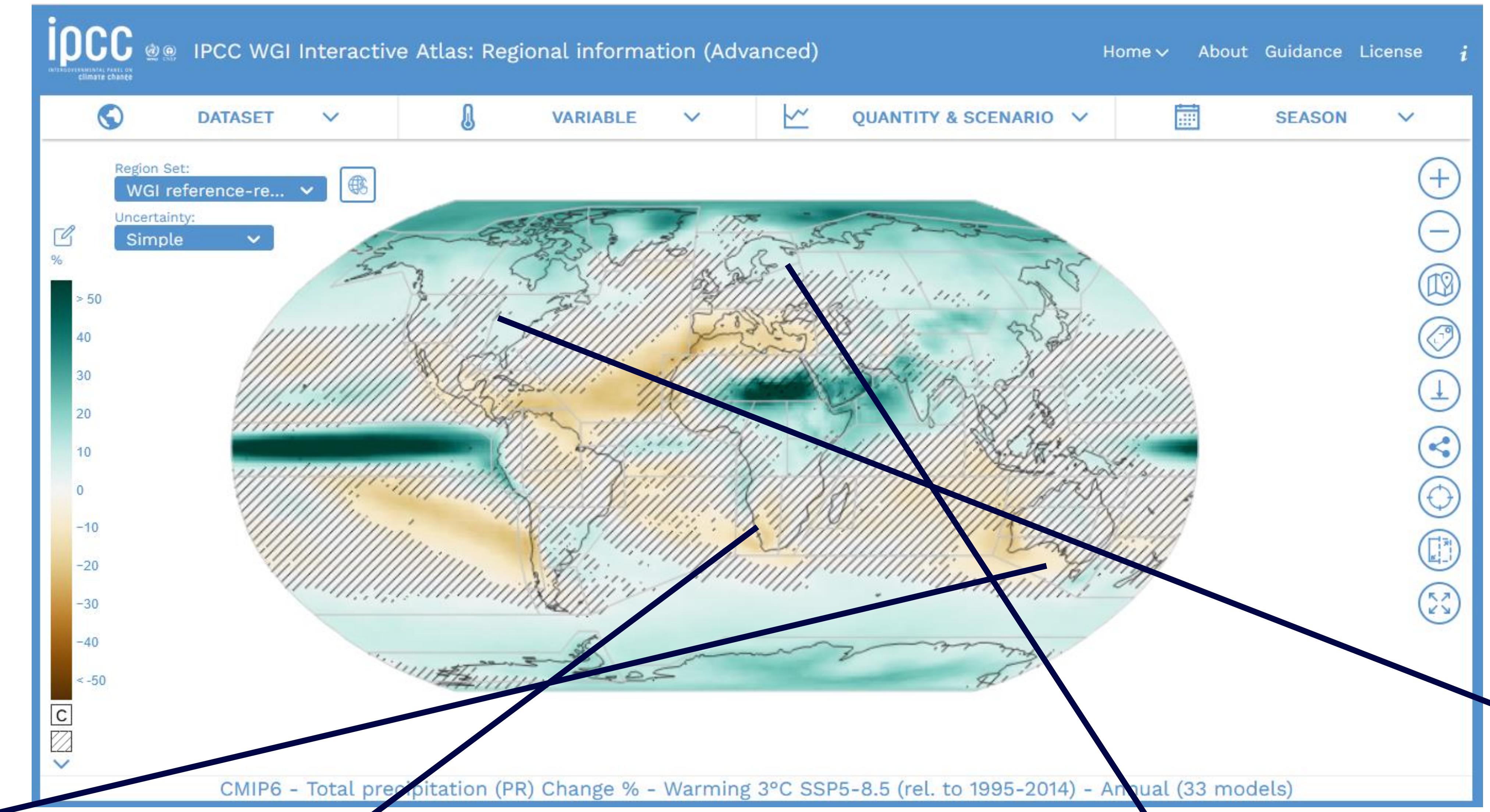
- Data availability! Only 18/50 models with sub-daily data for the SSPs
  - Overcome by the Queensland Future Climate Platrom-v2 project – uses only daily inputs
- Only one realisation available from models – can we request more?
- Lack of centralised, comprehensive lists of required diagnostics etc.
  - Climate sensitivity – found on a github page through personal connection, not all models
  - Global model evaluation – found in various studies in papers, not all models
  - Independence (family tree) – found in additional material from a paper, not all models
- No objective criteria to reject models (bottom in a set of tests not objective) – move to benchmarking?

- *Physical basis for in signal (response to forcing) and added value in the signal – central question for all modelers, especially CORDEX*

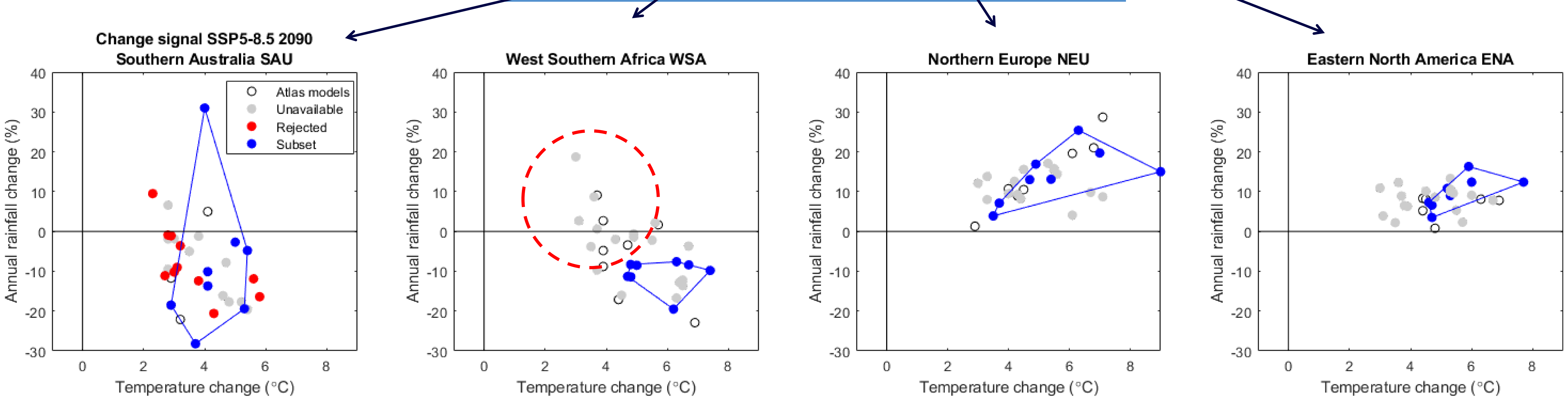
Huge data volume – is it even possible?
Provide at least training data for machine learning?

What if we applied our selection everywhere?

Simple example - spread of warming and rainfall change, SSP5-8.5 2090
IPCC Atlas results



What about extremes, ice sheets, oceans? Can a subset of models suit all purposes?



Selection for Australia (SAU example here)
Selected to cover spread (including wet outlier) after model rejection
Sub-daily data unavailability a minor problem

Applying to WSA
Not a representative sampling*
Data availability a major problem
*Will depend on model rejection

Applying to NEU
Not bad, except two outliers*
*Will depend on model rejection

Applying to ENA
Quite poor – but mainly due to data unavailability!

Thank You

@wcrp-cmip.org

wcrp-cmip

cmip-ipo@esa.int

Model selection for RCM downscaling in ISMIP6

Evaluation of CMIP5-6 global climate models in the Arctic and Antarctic regions

Cécile Agosta, Alice Barthel et al.

Ice Sheet Model Intercomparison Project for CMIP6

Ice Sheet Model Intercomparison Project

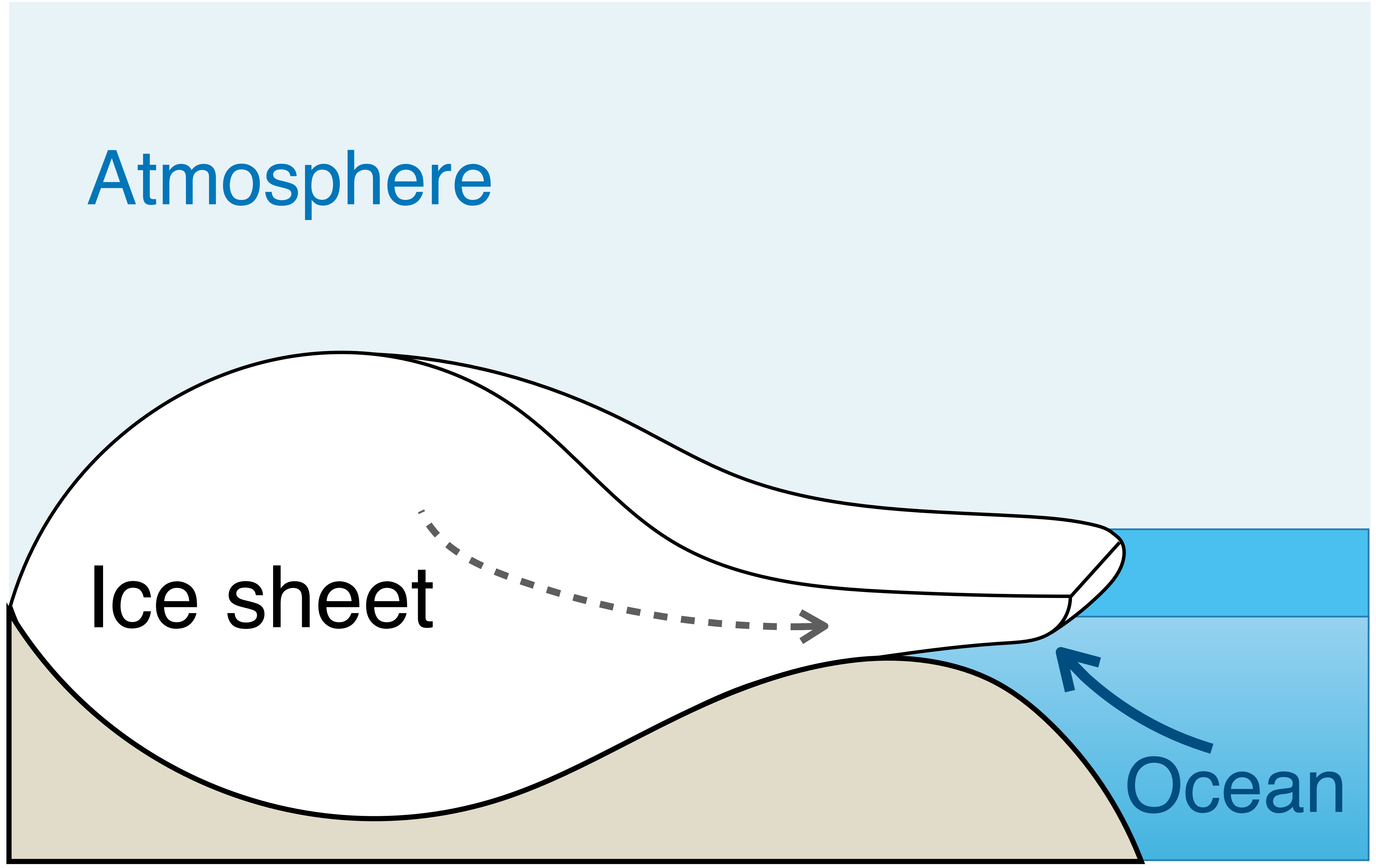
This presentation is part of the AWACA project that has received funding from the European Research Council (ERC Synergy) under the European Union's Horizon 2020 Research and Innovation Programme (Grant agreement No 951596)

This presentation is part of the PROTECT project that has received funding from the European Union's Horizon 2020 Research and Innovation Programme (Grant agreement No 869304).

ISMIP7
Ice Sheet Model Intercomparison Project

Ice Sheet models

+ Atmosphere & Ocean forcing
(from CMIP)

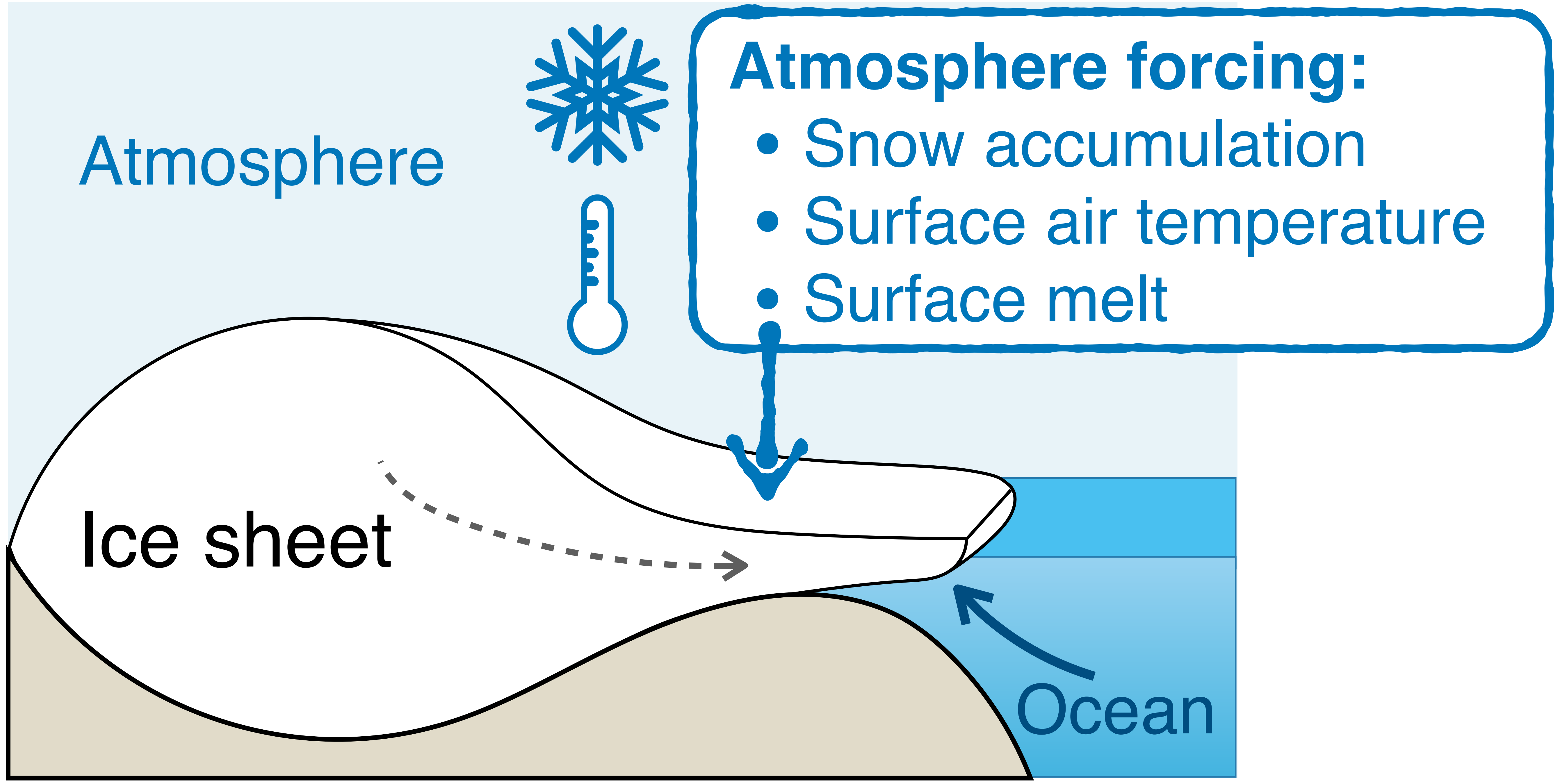


Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

Ice Sheet Model Intercomparison Project

Ice Sheet models

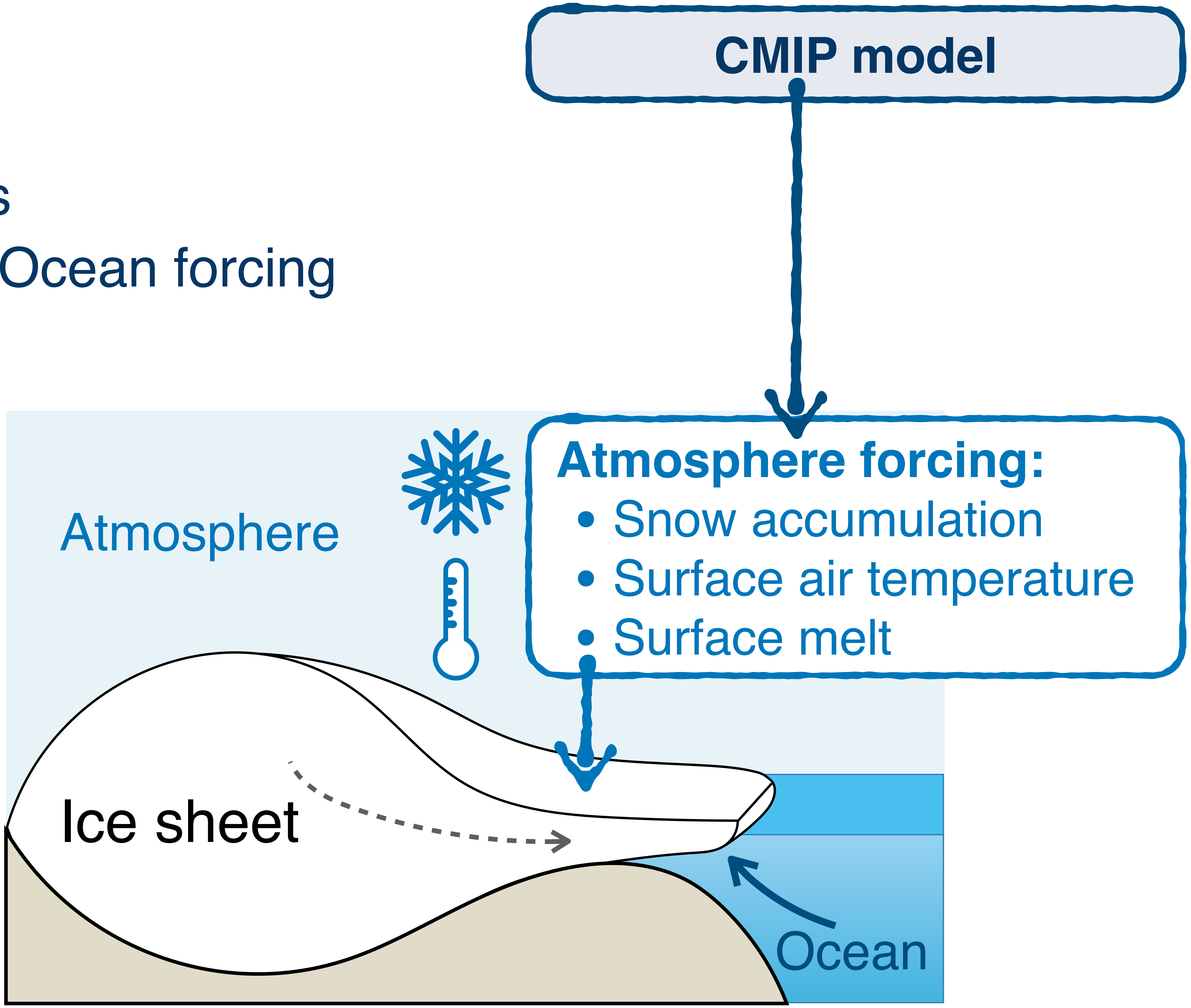
+ Atmosphere & Ocean forcing (from CMIP)



Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

ISMIP7
Ice Sheet Model Intercomparison Project

Ice Sheet models
+ Atmosphere & Ocean forcing (from CMIP)



Large scale circulation
(i.e. above boundary layer)
Sea surface conditions

+ Polar physics:
- *Clouds*
- *Snow*
- *Boundary layer*

Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

ISMIP7
Ice Sheet Model Intercomparison Project

Ice Sheet models
+ Atmosphere & Ocean forcing (from CMIP)

CMIP model

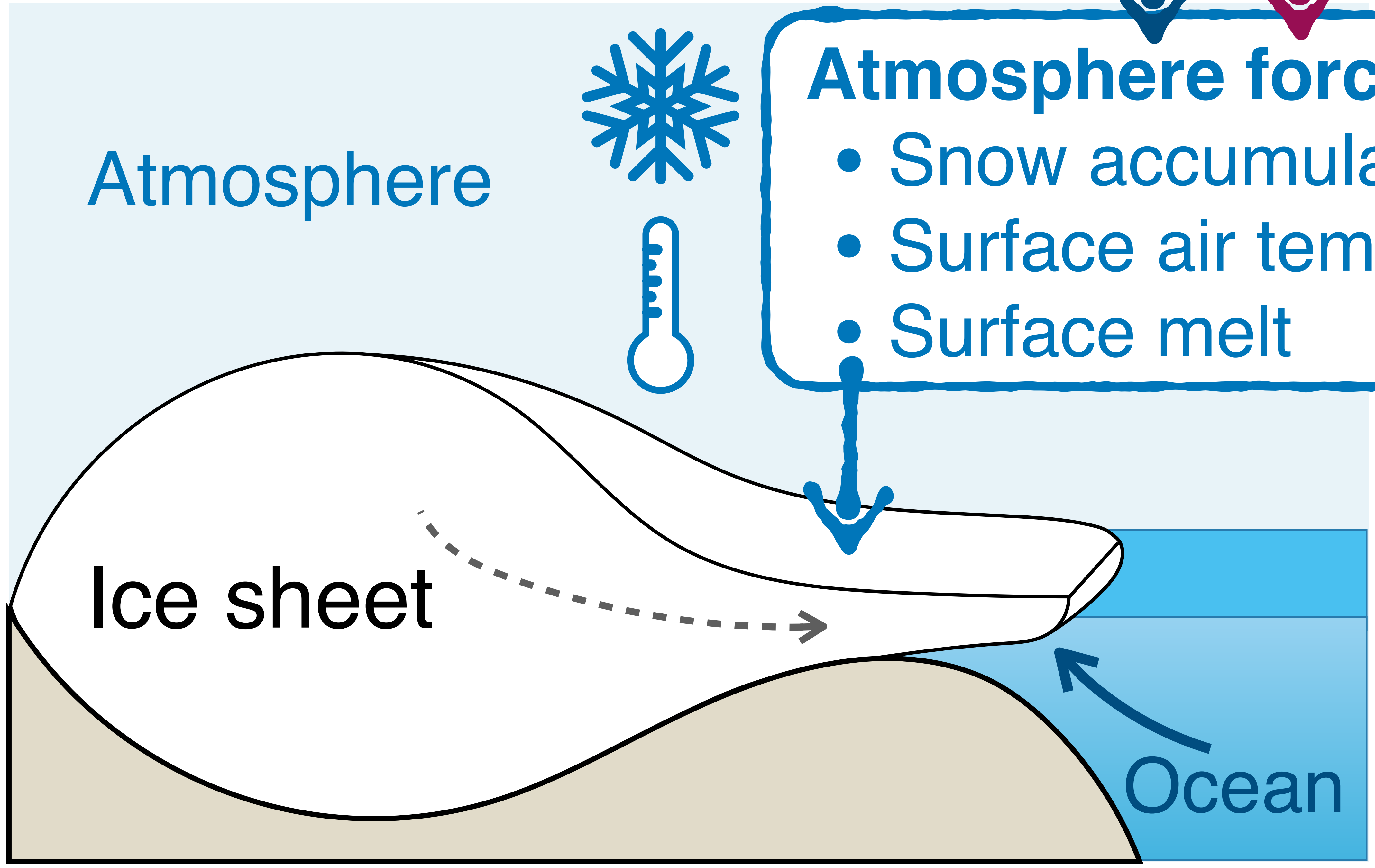
Large scale circulation
(i.e. above boundary layer)
Sea surface conditions

Downscaling (e.g. RCM)

+ Polar physics:
- *Clouds*
- *Snow*
- *Boundary layer*

+ Resolution

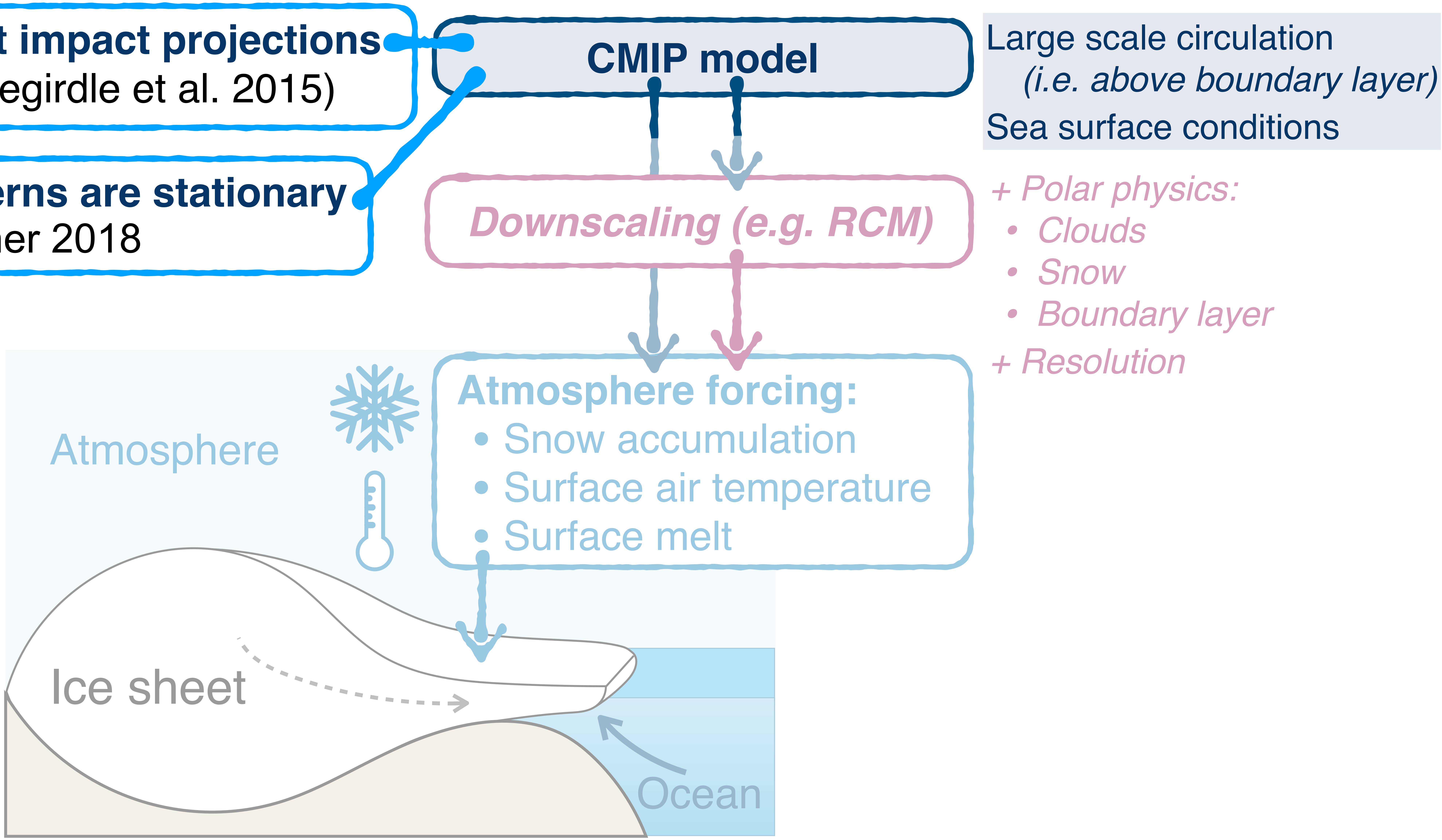
Atmosphere forcing:
- Snow accumulation
- Surface air temperature
- Surface melt



Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

Biases at present impact projections
(e.g. sea ice, Bracegirdle et al. 2015)

Spatial bias patterns are stationary
Krinner & Flanner 2018



Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

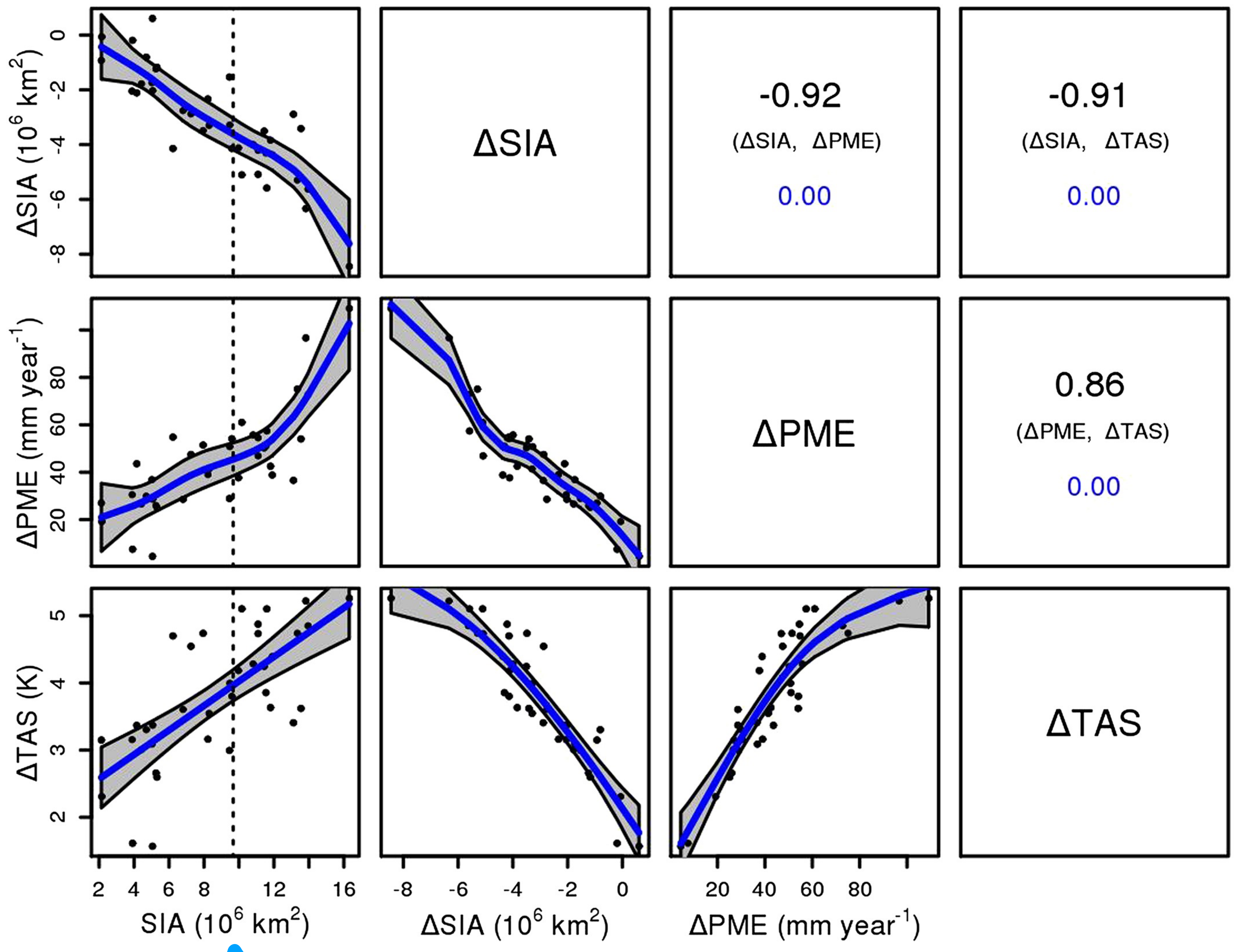
Biases at present impact projections
(e.g. sea ice, Bracegirdle et al. 2015)

CMIP model

Large scale circulation
(i.e. above boundary layer)
Sea surface conditions

ΔAccumulation Antarctica
Change during the 21st c.

ΔTAS Antarctica
Change during the 21st c.



Bracegirdle et al. 2015

Present day Antarctic Sea Ice Area

Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

⟹ Evaluation of **CMIP large-scale fields**

i.e. inputs of regional atmospheric models

Large scale circulation
(i.e. above boundary layer)
Sea surface conditions

Method :

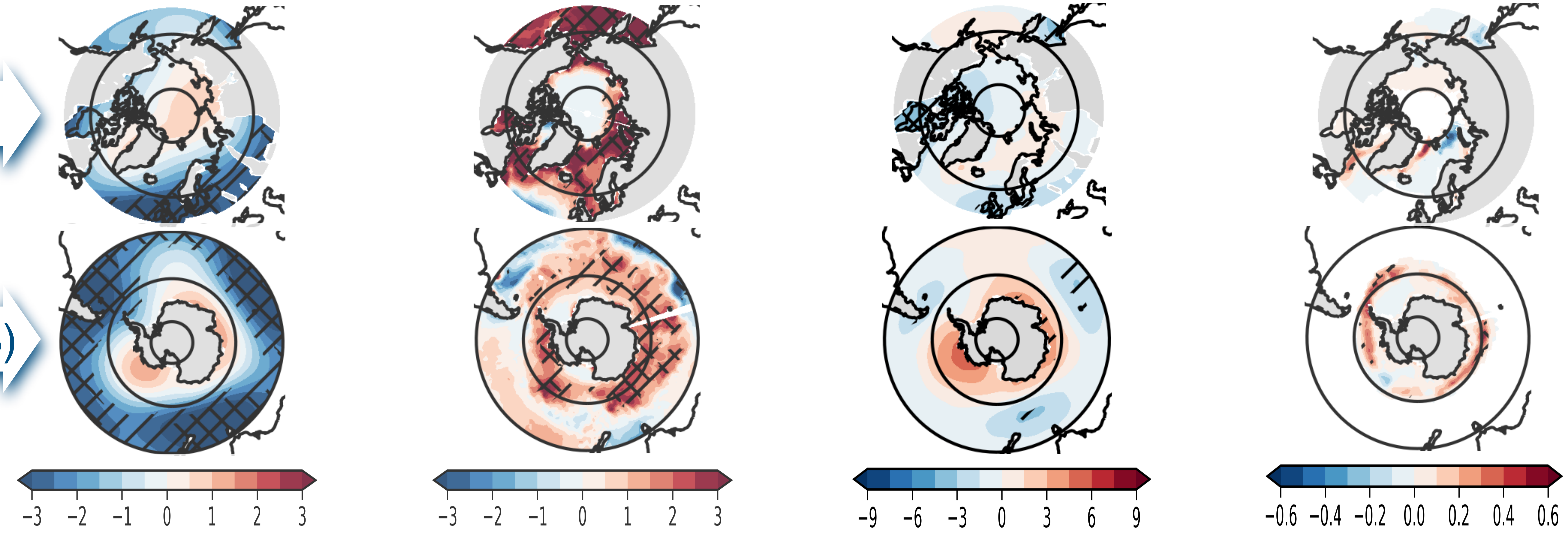
- 1979-2005 time-mean
- Annual or Seasonal
- Difference with ERA5
- 2 regions

Arctic (> 50°N)

Antarctic (< 40°S)

Exemple for one CMIP model : IPSL-CM6A-LR

Temperature | **Humidity** | **Circulation** | **Surface ocean**



Reference: ERA5, 1979-2005

9 variables
83 CMIP models
2 regions

Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

⟹ Evaluation of **CMIP large-scale fields**
i.e. inputs of regional atmospheric models

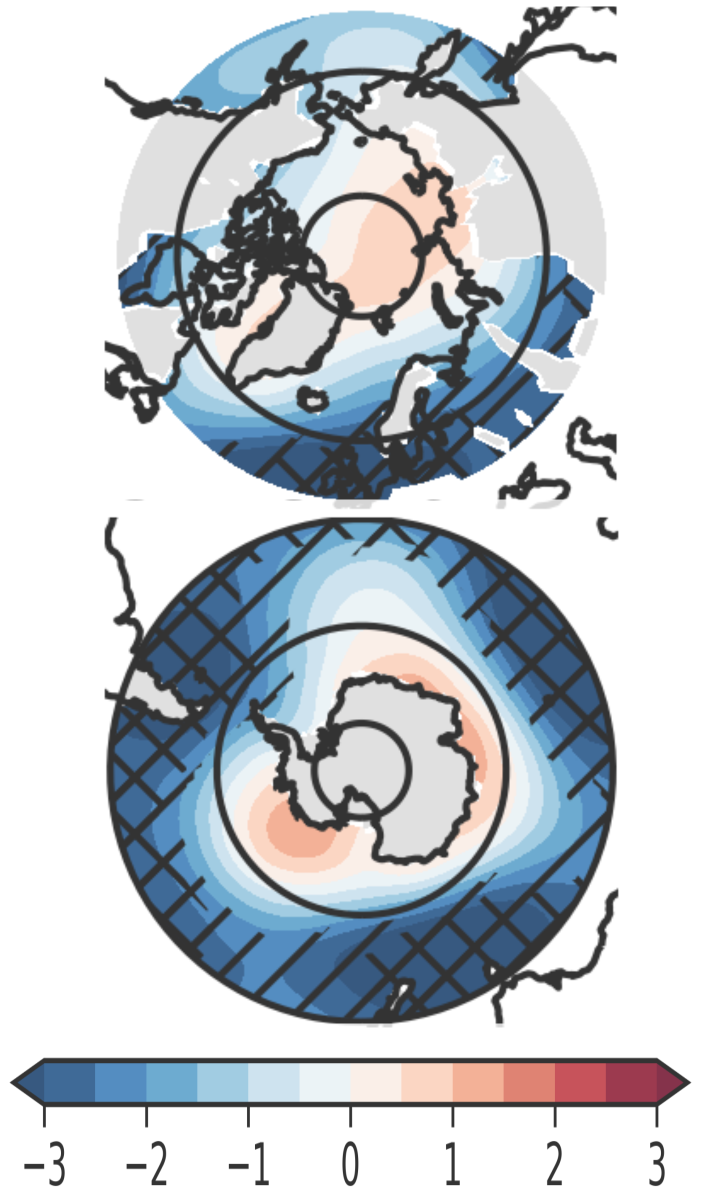
Large scale circulation
(i.e. above boundary layer)
Sea surface conditions

Defining metrics and scores

For ISMIP6 : **RMSE** for each variable + **Scaling** by median RMSE
i.e. ***Relative metrics*** *for each variable*

Agosta et al. 2015; Barthel et al. 2020 (ISMIP6), ESMValTool (e.g. Eyring et al. 2020)

IPSL-CM6A-LR
ΔT850[ann]



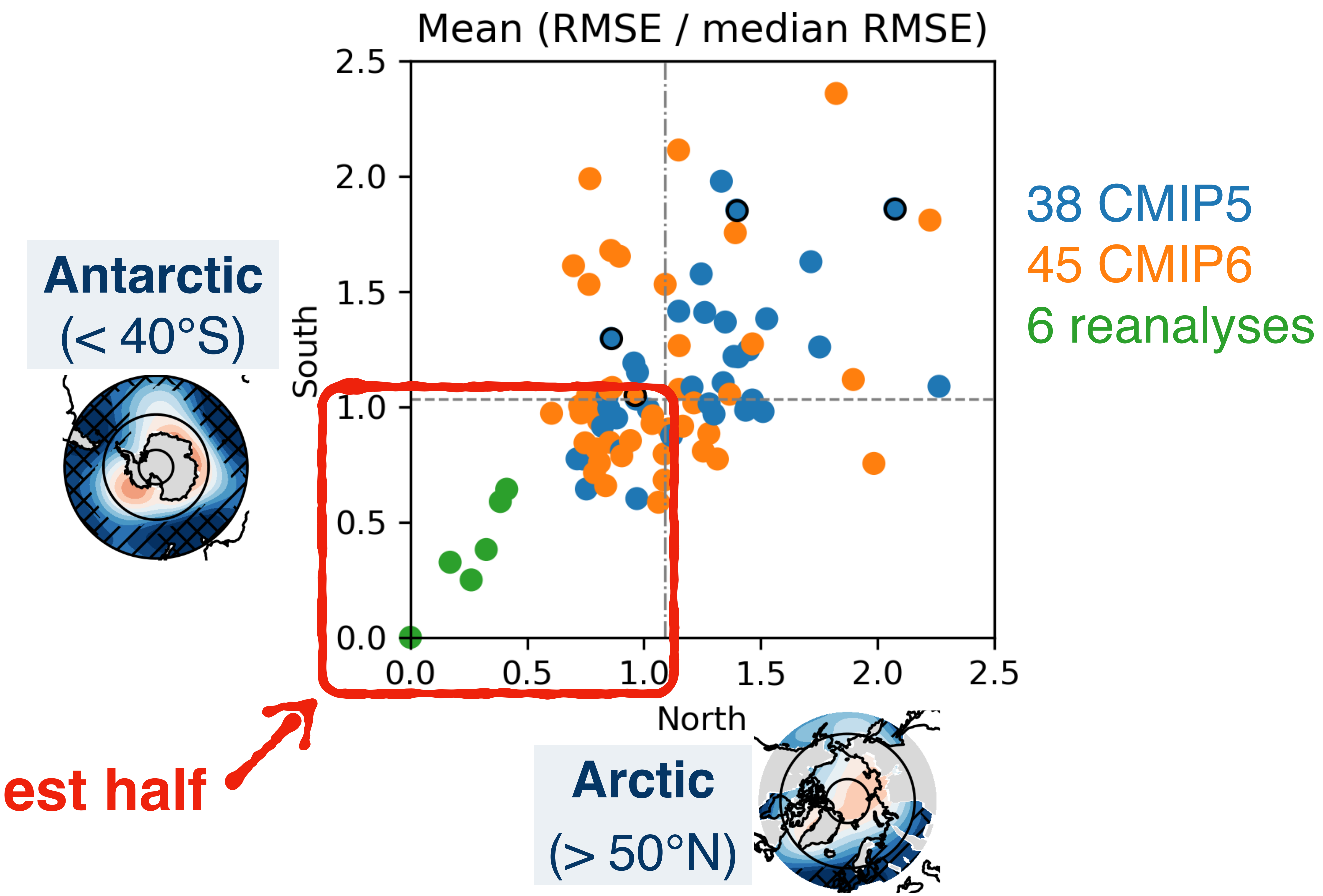
RMSE (°C) = √spatial mean(**Δ²**)

→ **Scaled-RMSE** = RMSE **/ Median RMSE among all CMIP models**

9 variables
83 CMIP models
2 regions

Score: mean Scaled-RMSE among variables
for each CMIP model and each region

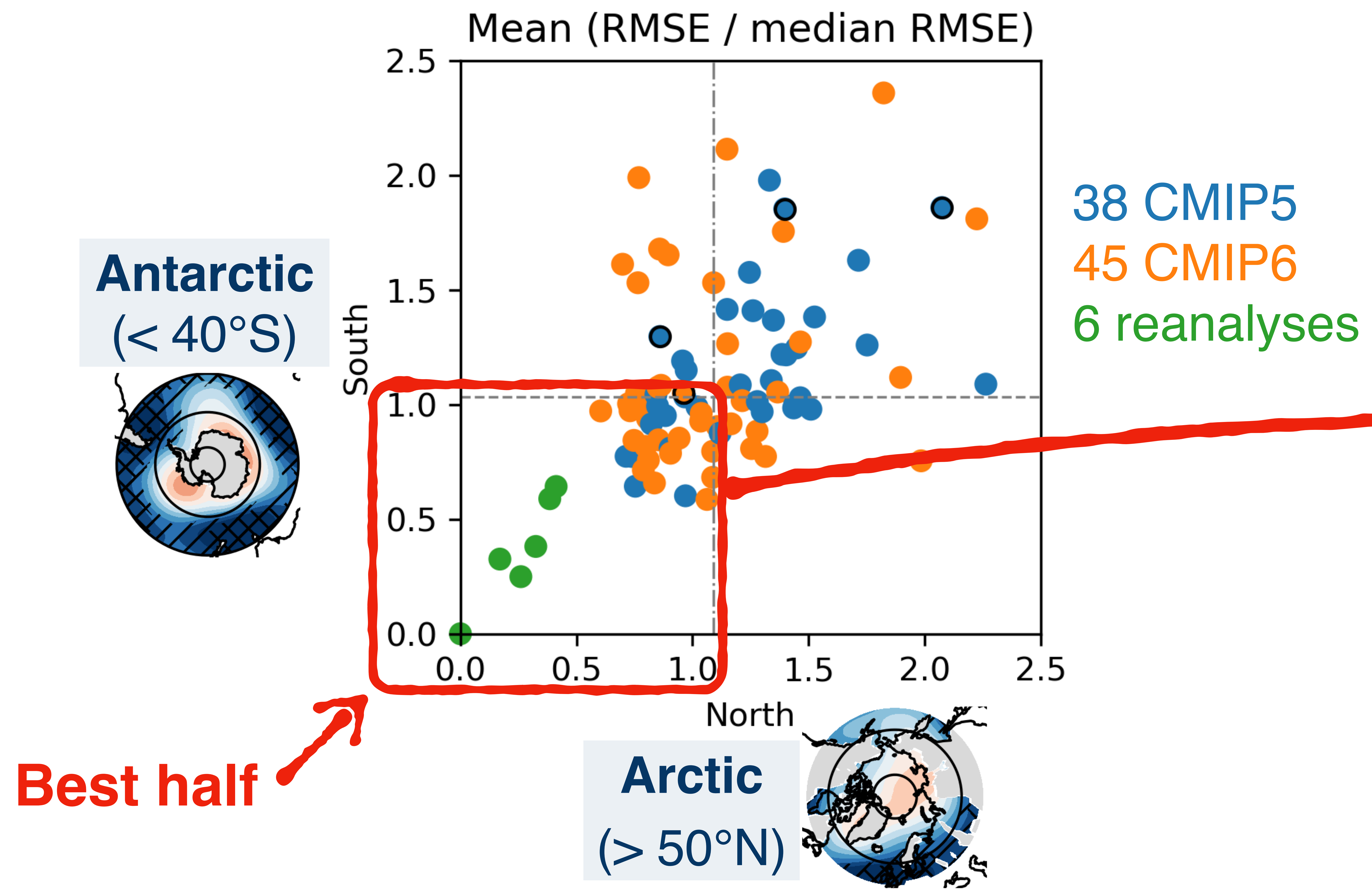
Score: mean Scaled-RMSE among variables
for each CMIP model and each region



28 (34 %) CMIP models in the « **best half** » for **both** regions
9 CMIP5, **19** CMIP6
(25% of CMIP5 vs. 40% of CMIP6)

Objective: **Assessment ('Sanity check') of CMIP models large scale fields over polar regions**

Score: mean Scaled-RMSE among variables
for each CMIP model and each region



+ Sample a diversity of climate sensitivity (local ECS)

28 (34 %) CMIP models in the « **best half** » for **both** regions

9 CMIP5, **19** CMIP6
(25% of CMIP5 vs. 40% of CMIP6)

What can we do better? (among other things…)

Replace relative metrics by absolute metrics

« **Implausibility** », absolute metric used in history matching

✓ discard only implausible models

? **require uncertainty quantification on model and observations**

Design CMIP sampling in view of emulation of the CMIP ensemble

From **selected CMIP** forcing to **all CMIP** : **need to extrapolate**

e.g. use statistical emulator (Edwards et al. 2021)

The selected CMIP models need to **sample a parameter space** impacting **ice sheet response**

⟹ **design the parameter space before selection process**

Supplement

CMIP evaluation: **Relative metrics**

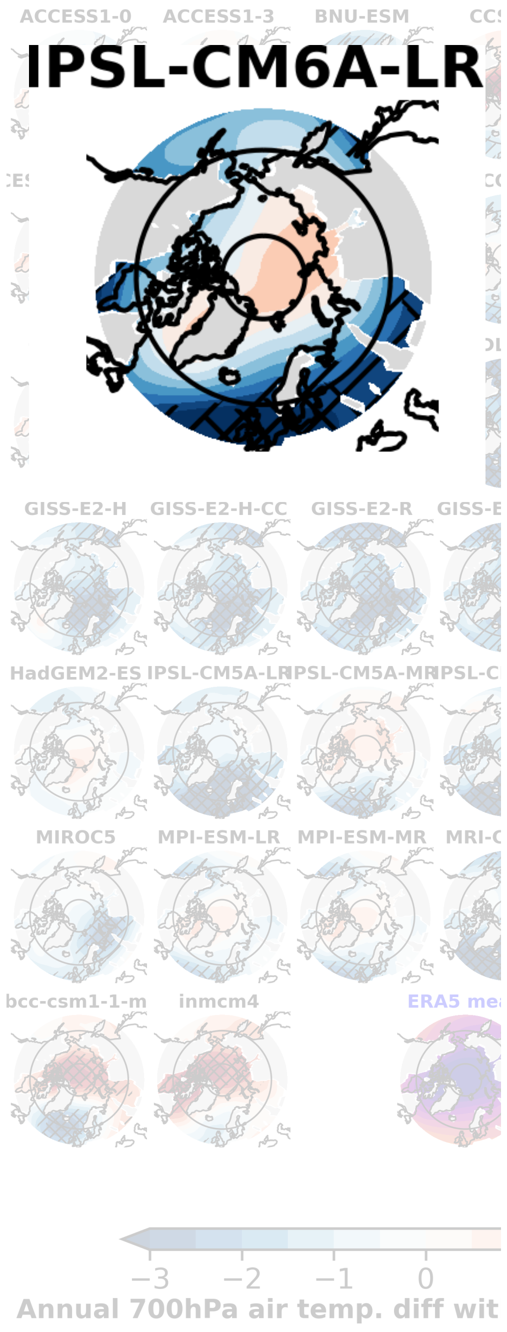
9 variables
83 CMIP models
2 regions

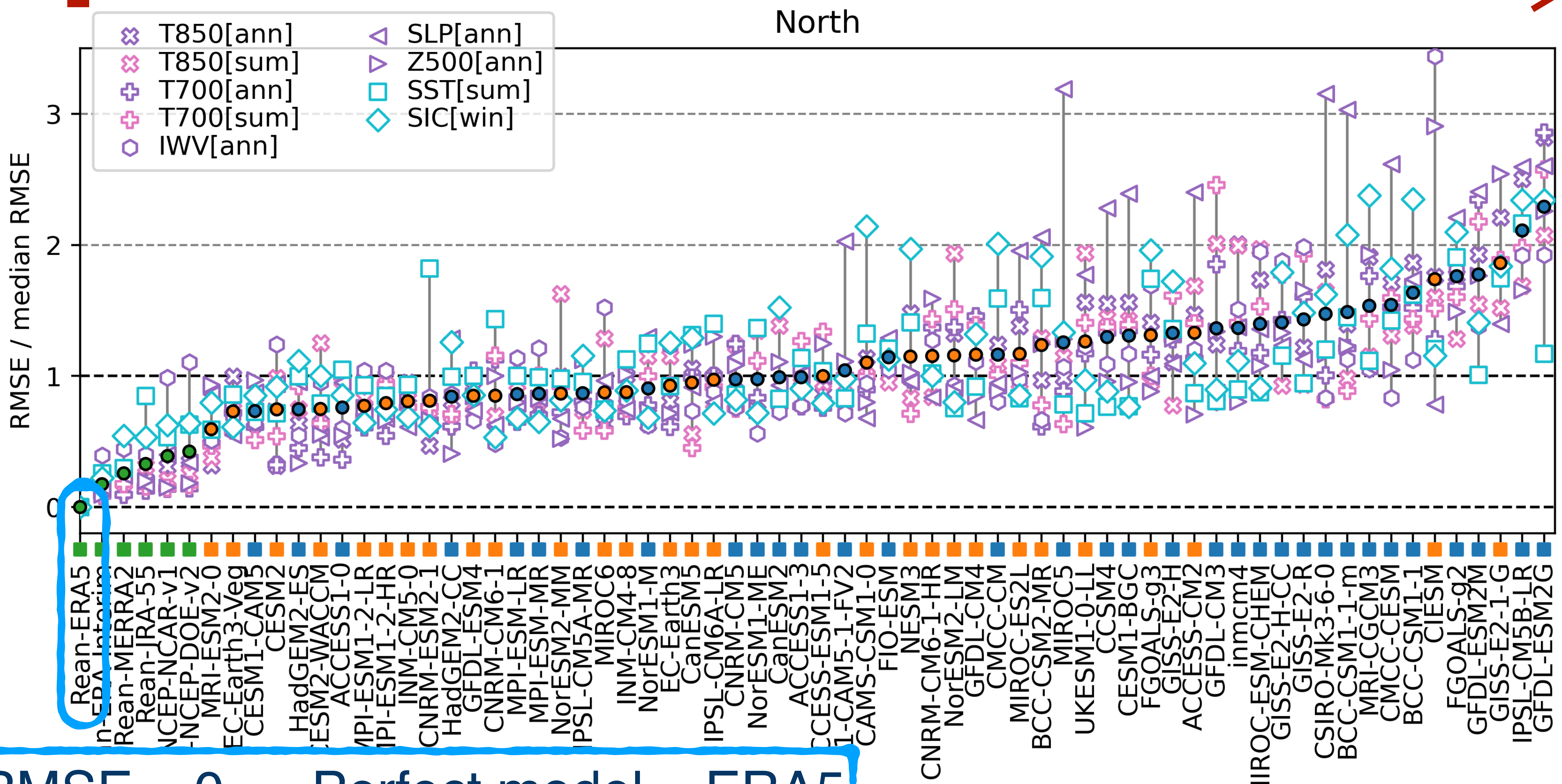
First method : **Relative metrics**

Scaling of metrics to combine them among variables

Agosta et al. 2015; Barthel et al. 2020 (ISMIP6), ESMValTool (REF)

Rank: mean Scaled-RMSE among variables





Scaled RMSE = 0 ⟹ Perfect model = ERA5

CMIP evaluation: **Relative metrics**

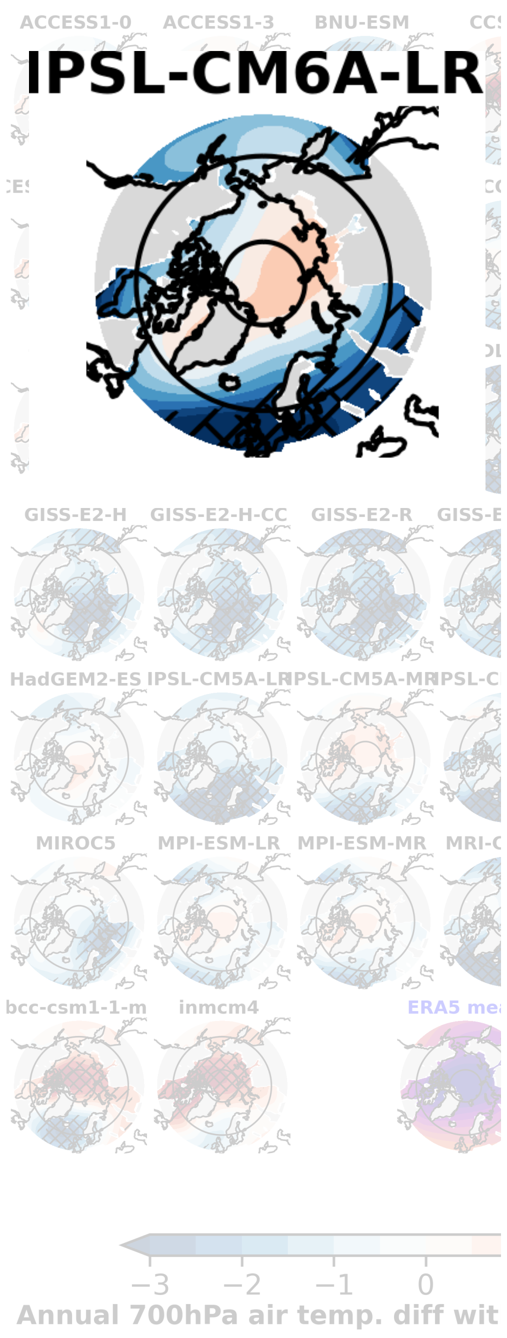
9 variables
83 CMIP models
2 regions

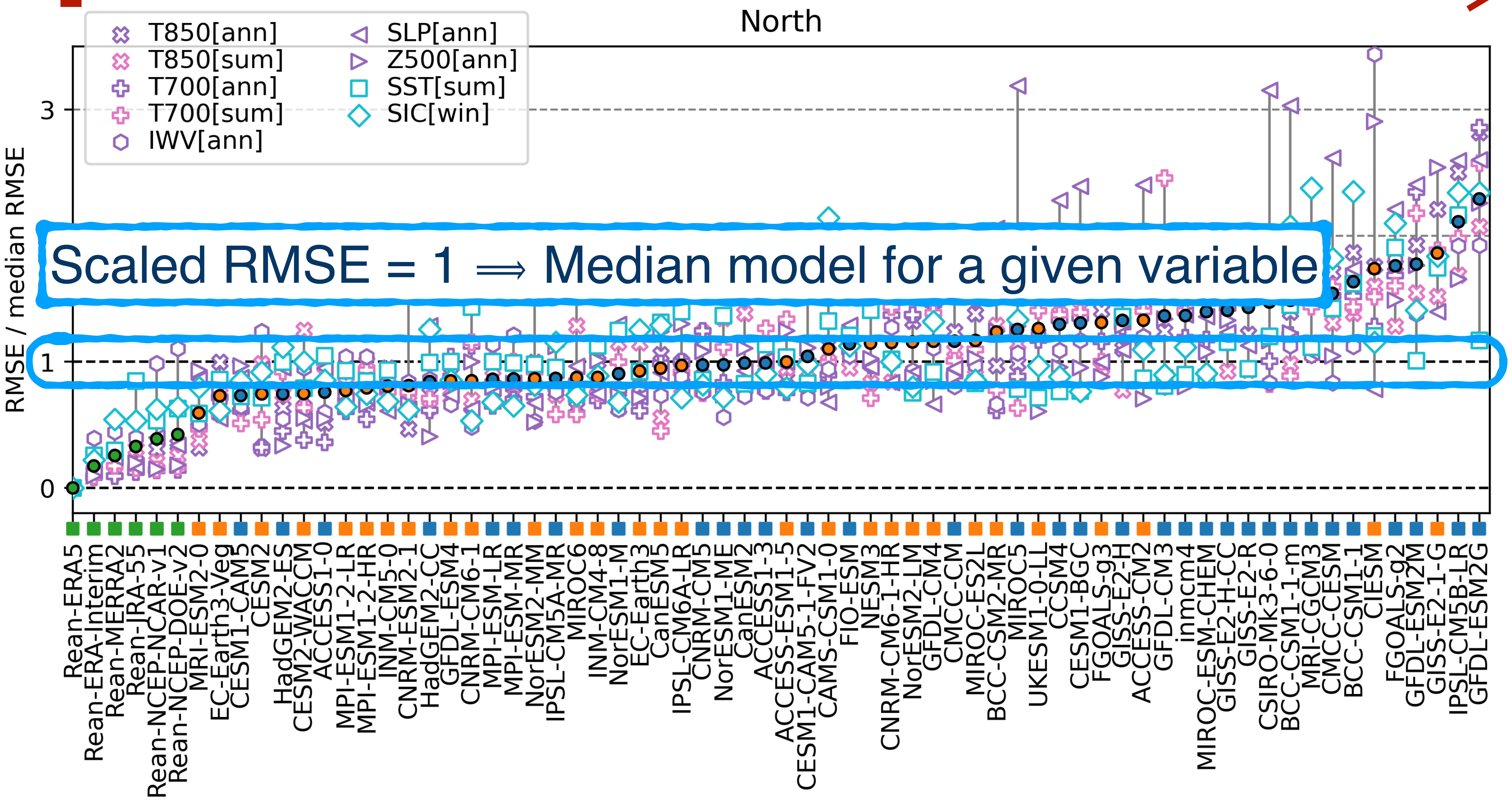
First method : **Relative metrics**

Scaling of metrics to combine them among variables

Agosta et al. 2015; Barthel et al. 2020 (ISMIP6), ESMValTool (REF)

Rank: mean Scaled-RMSE among variables





CMIP evaluation: **Relative metrics**

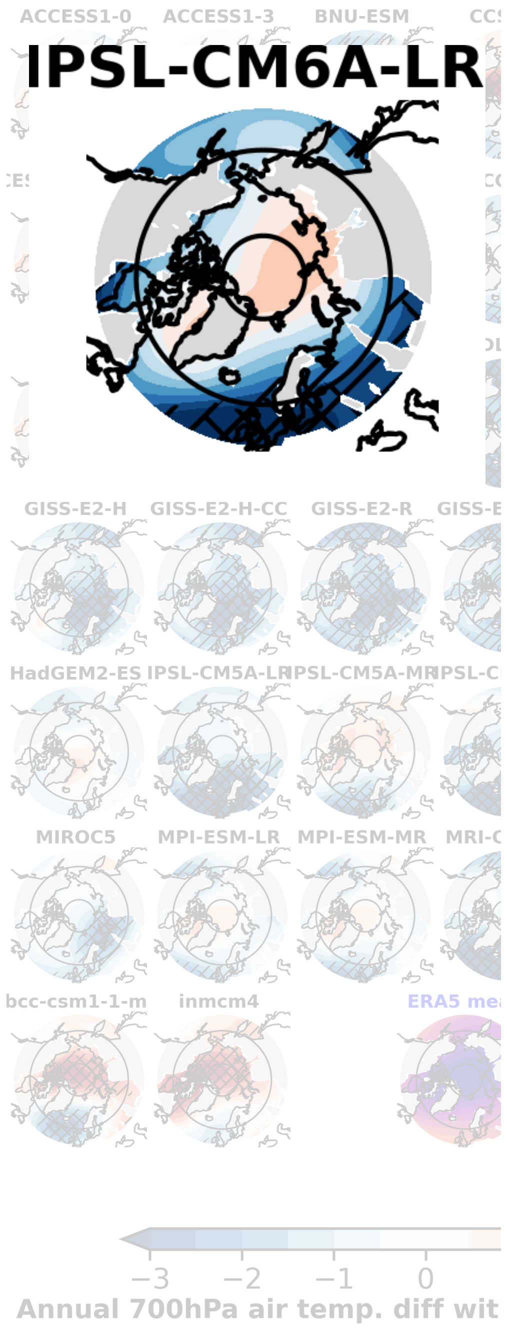
9 variables
83 CMIP models
2 regions

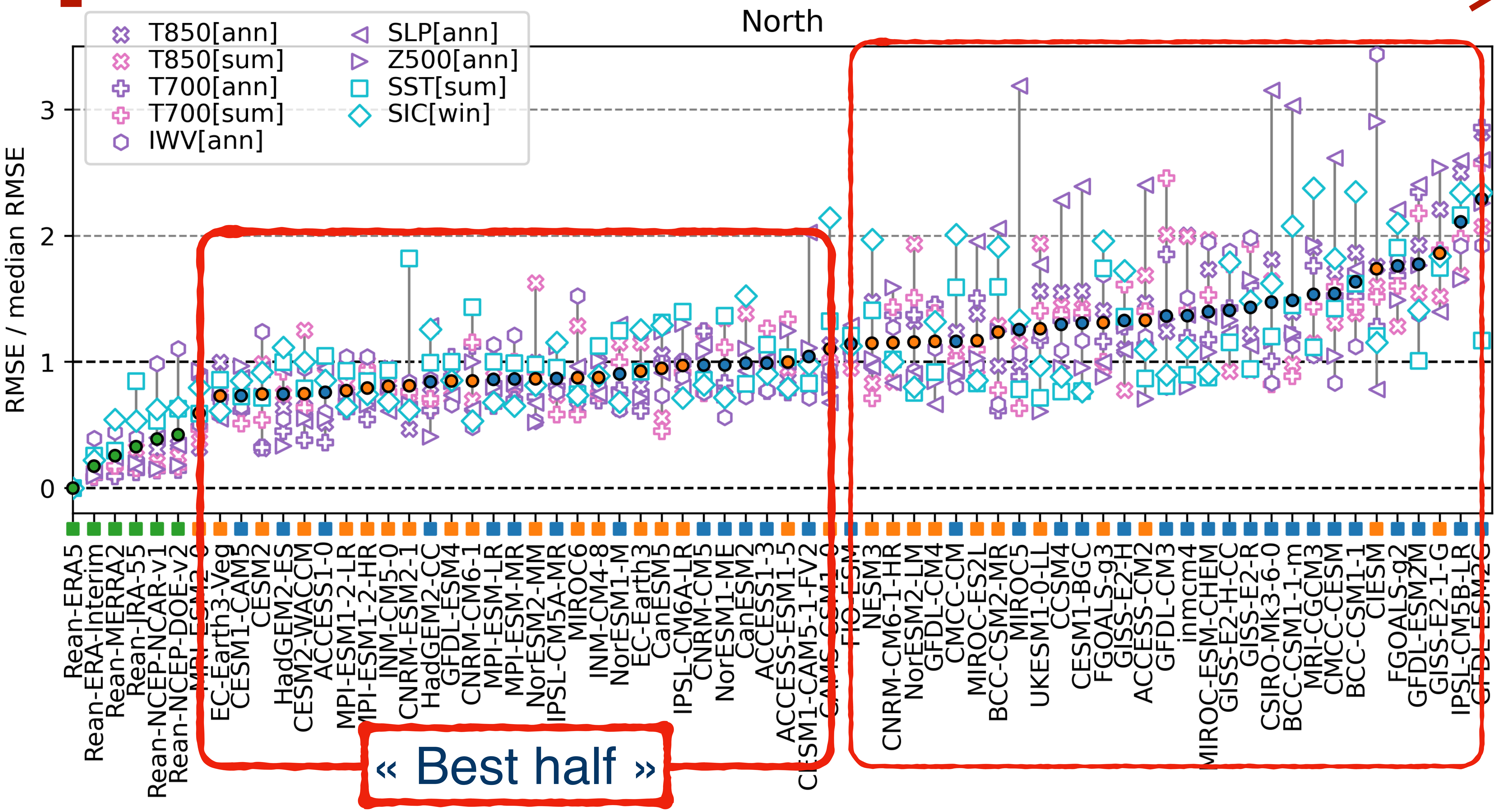
First method : **Relative metrics**

Scaling of metrics to combine them among variables

Agosta et al. 2015; Barthel et al. 2020 (ISMIP6), ESMValTool (REF)

Rank: mean Scaled-RMSE among variables





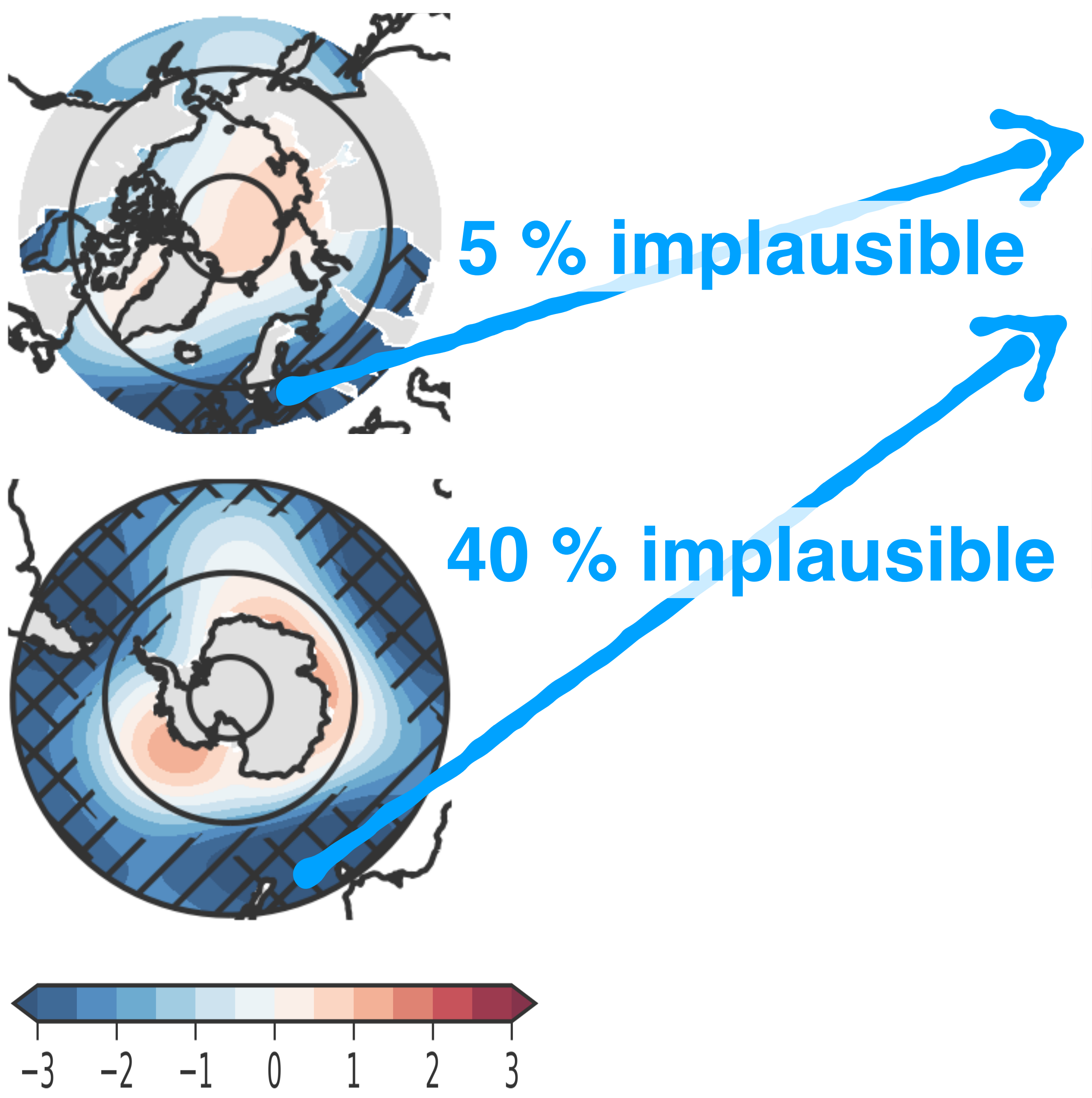
CMIP evaluation: **Relative metrics**

9 variables
83 CMIP models
2 regions

Second method : Implausible fraction
Absolute metric, **No scaling**

History matching, « Not Ruled Out Yet » method (Pukelsheim, 1994, Rougier, 2015), applied e.g. in Gladstone et al. 2012; Edwards et al., 2019

IPSL-CM6A-LR
ΔT850[ann]



Hashes: **$\Delta > 3\ std_{1yr}$(reference)**

Portion of the surface where **Δ** with ERA5
is **greater that 3 x ERA5 interannual variability**
= « **Implausible fraction** » of the surface

Score: *2nd* **max implausible fraction**
(We let 1 variable / 9 be more implausible)

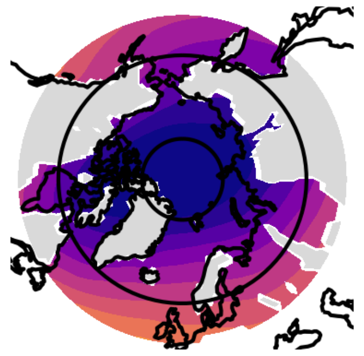
CMIP evaluation: **Absolute metrics**

9 variables
83 CMIP models
2 regions

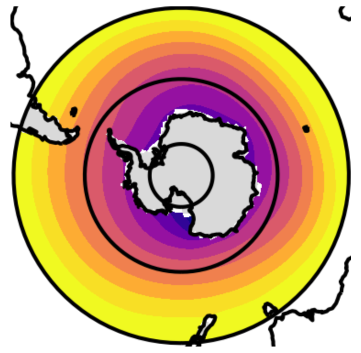
Second method : **Absolute metric**

Implausible fraction, **No scaling**

Arctic (> 50°N)

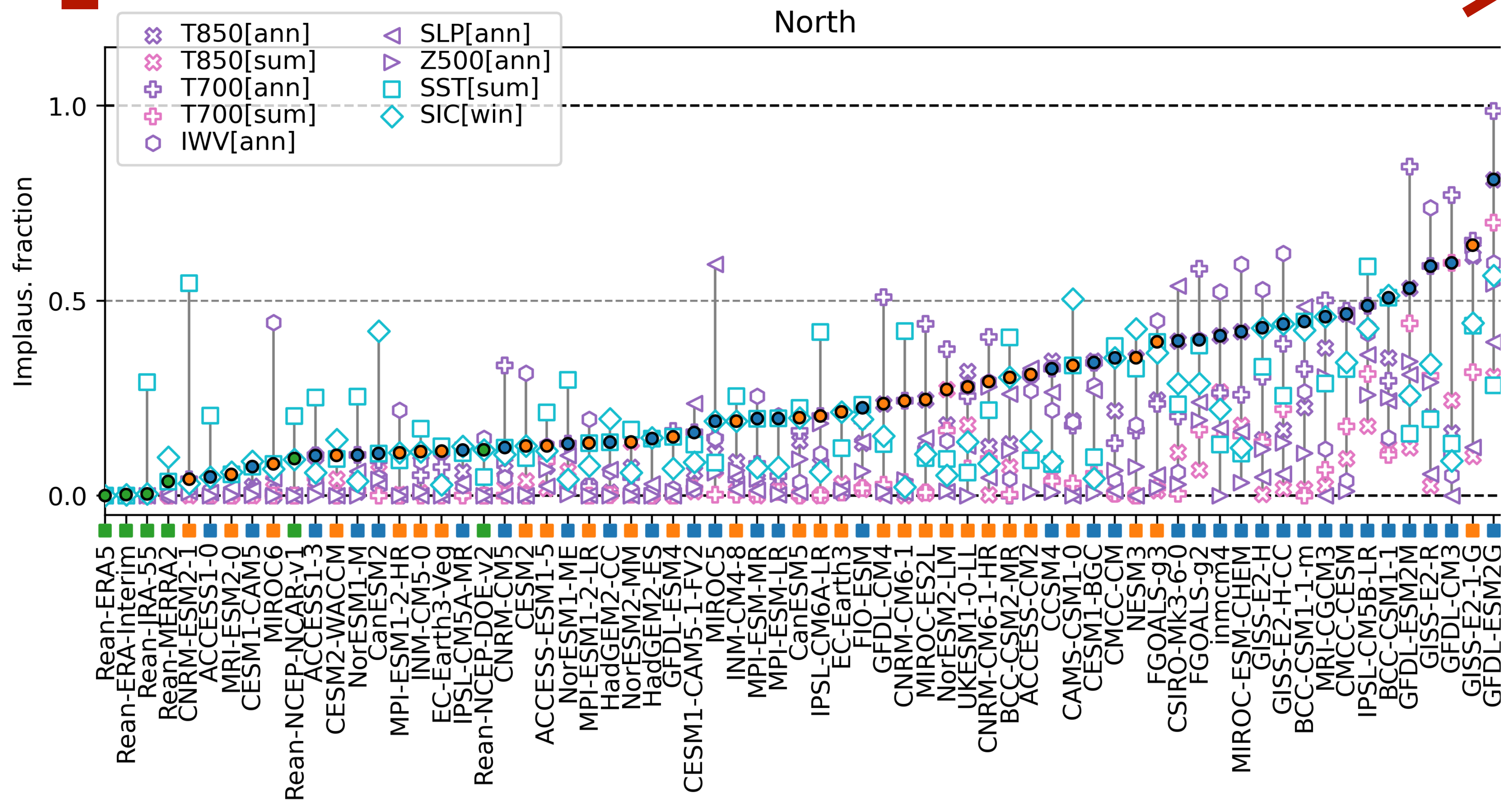


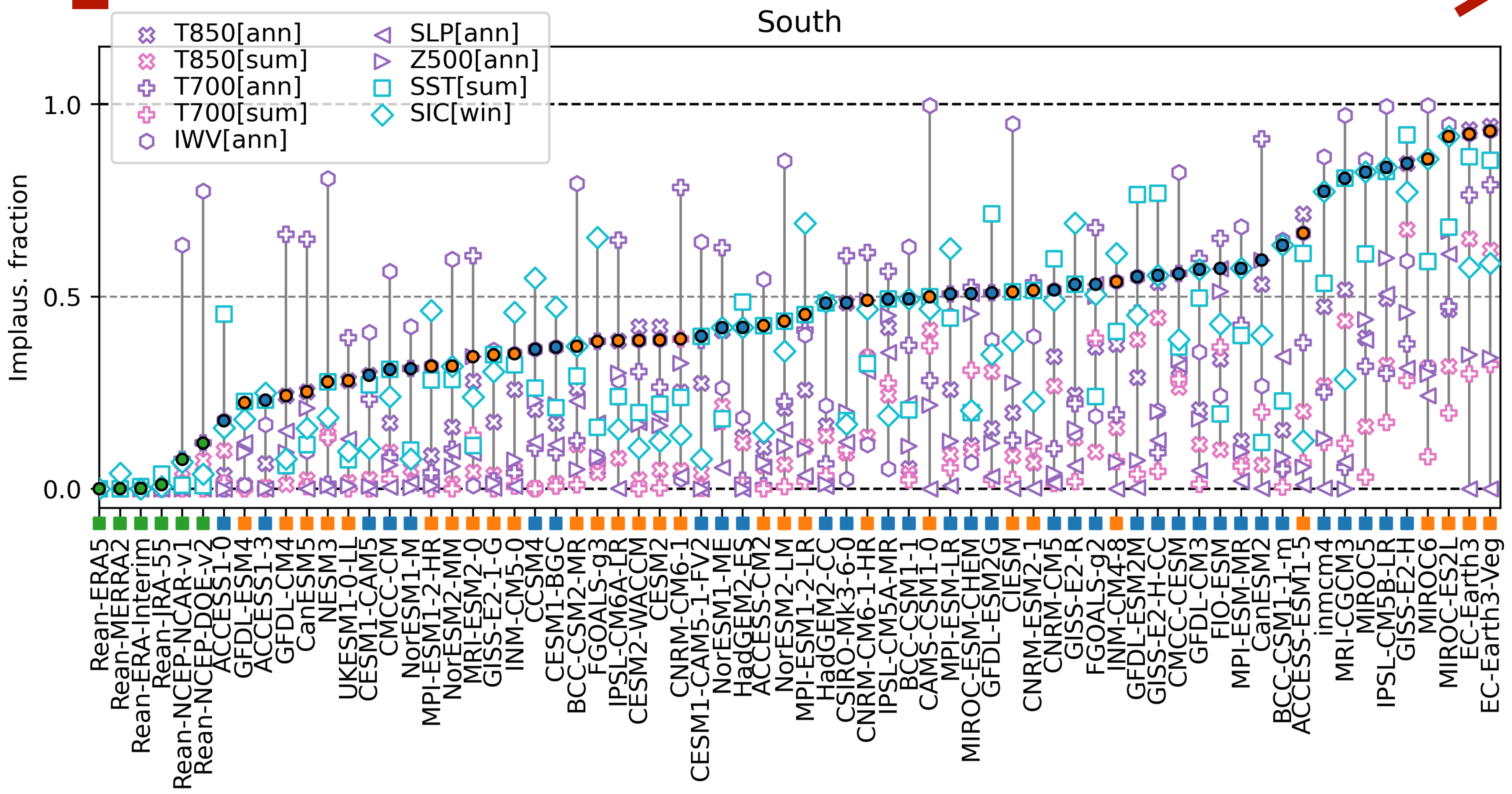
Antarctic (< 40°S)



Rank: 2nd max implausible fraction

(We let one variable be very implausible)





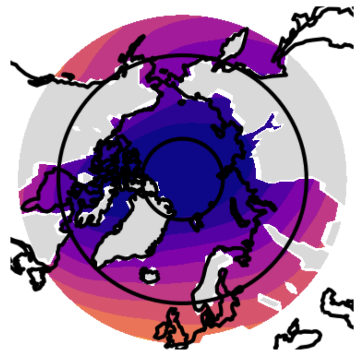
CMIP evaluation: **Absolute metrics**

9 variables
83 CMIP models
2 regions

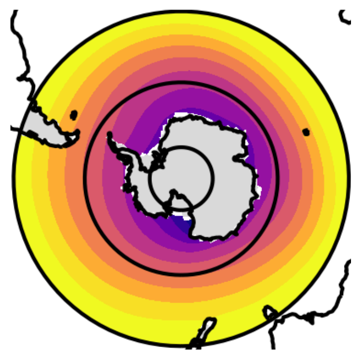
Second method : **Absolute metric**
Implausible fraction, **No scaling**

CMIP models are **more implausible in the Antarctic than in the Arctic**

Arctic (> 50°N)

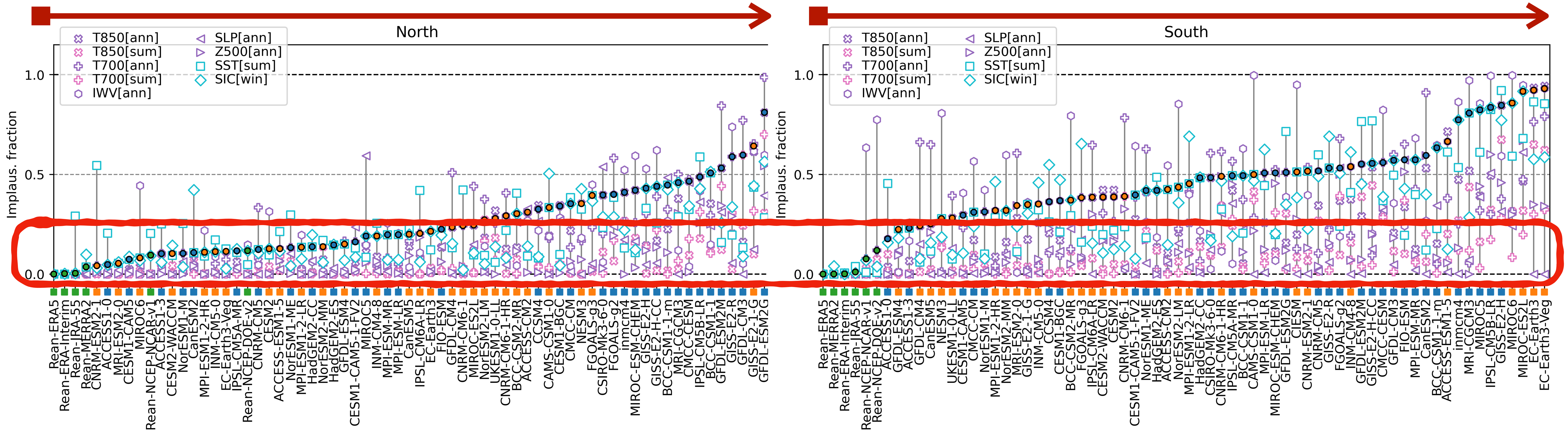


Antarctic (< 40°S)



Rank: 2nd max implausible fraction

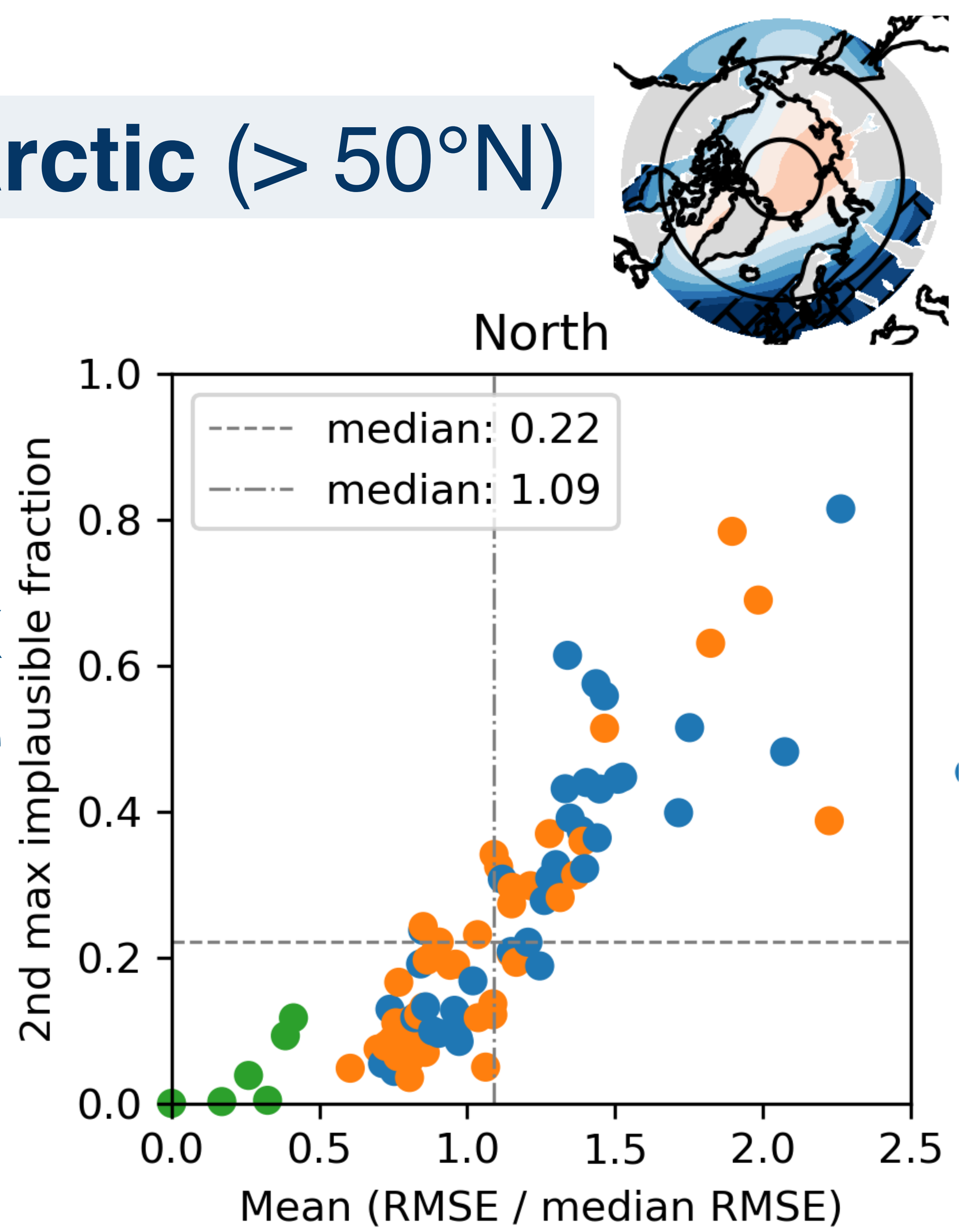
(We let one variable be very implausible)



Conclusions: CMIP models over polar regions

Contact: cecile.agosta@lsce.ipsl.fr

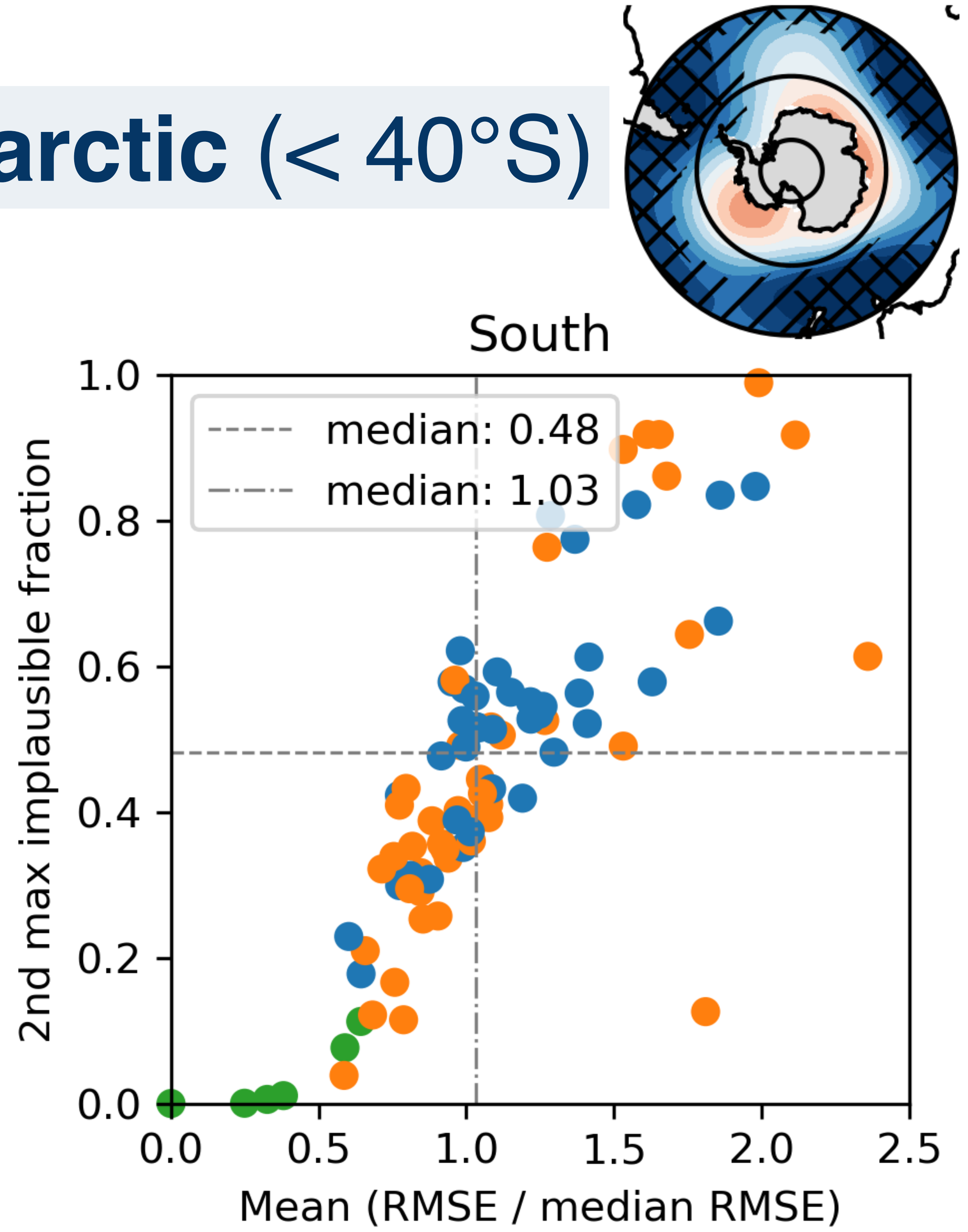
Arctic (> 50°N)



2nd max **Implausible fraction**

Mean **Scaled RMSE**

Antarctic (< 40°S)



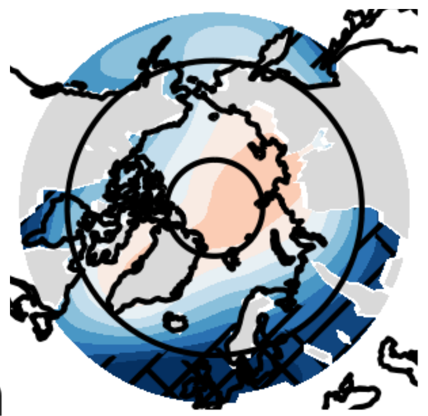
2nd max **Implausible fraction**

Mean **Scaled RMSE**

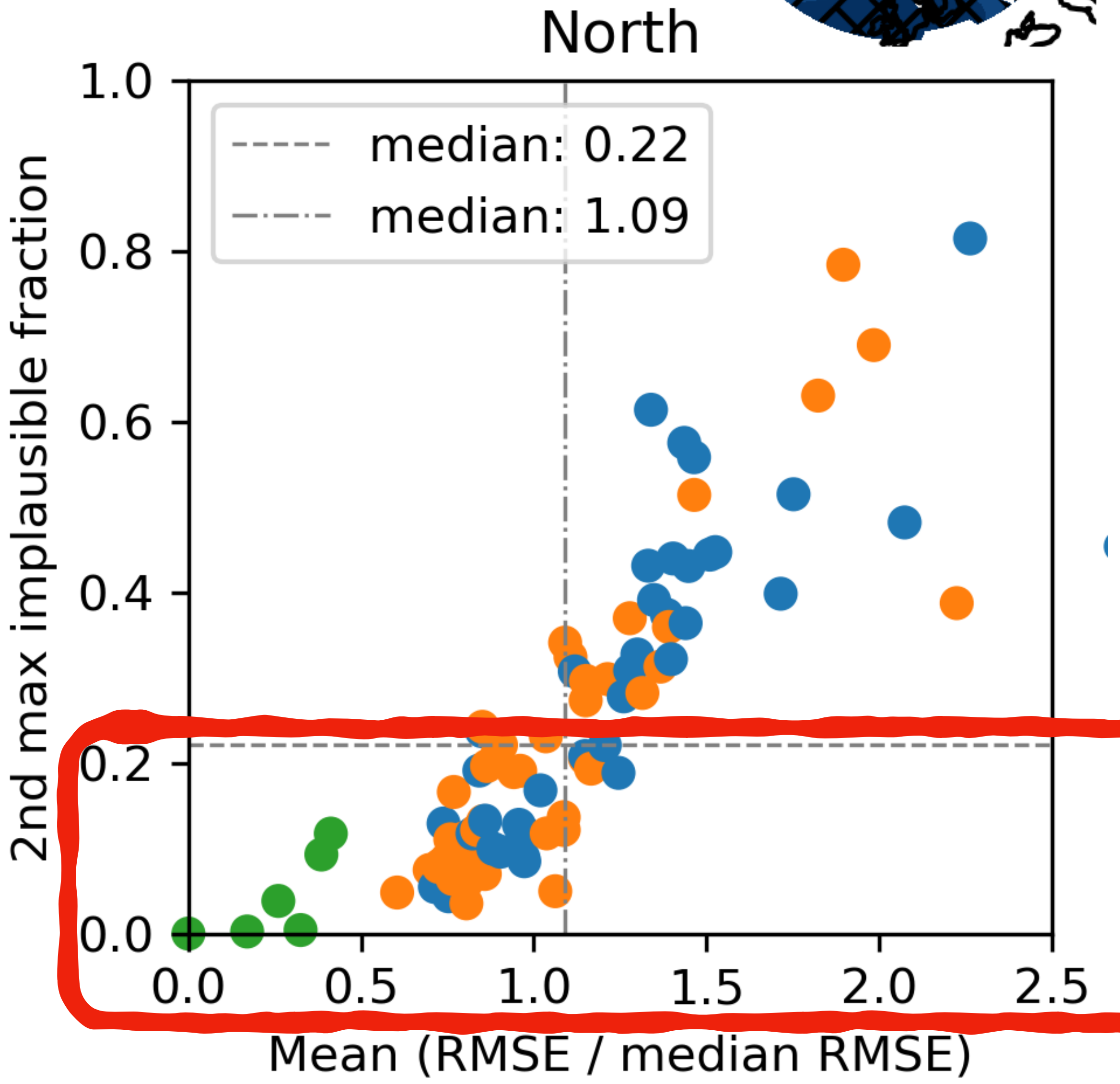
Conclusions: CMIP models over polar regions

Contact: cecile.agosta@lsce.ipsl.fr

Arctic (> 50°N)

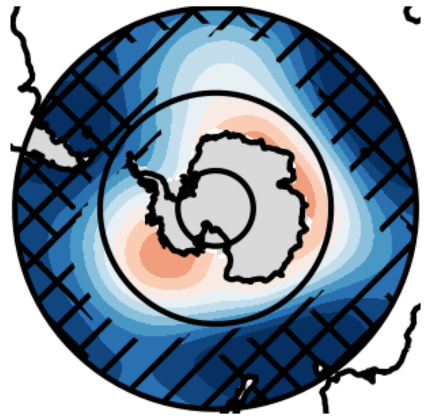


2nd max **Implausible fraction**

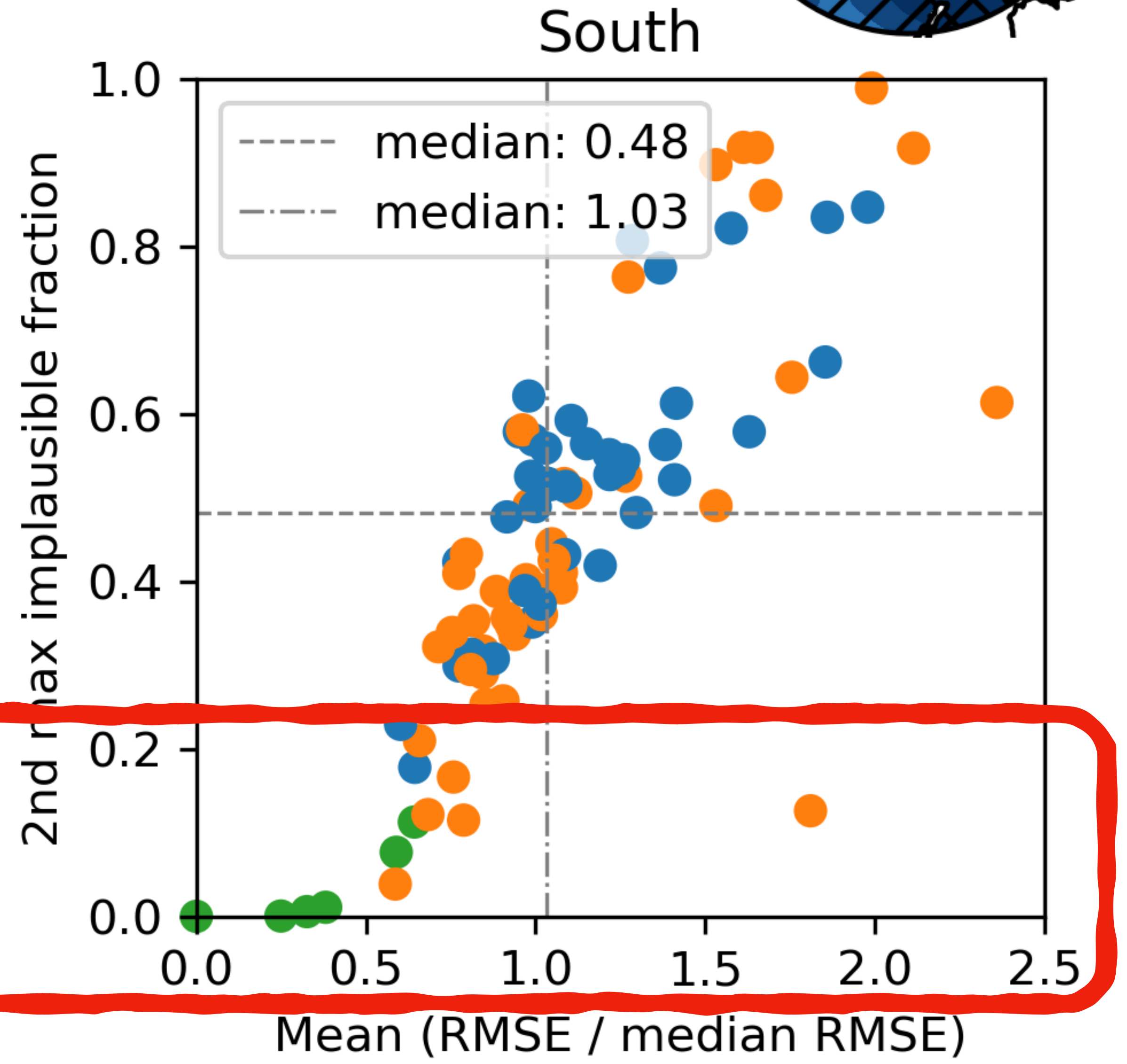


Mean **Scaled RMSE**

Antarctic (< 40°S)



2nd max **Implausible fraction**



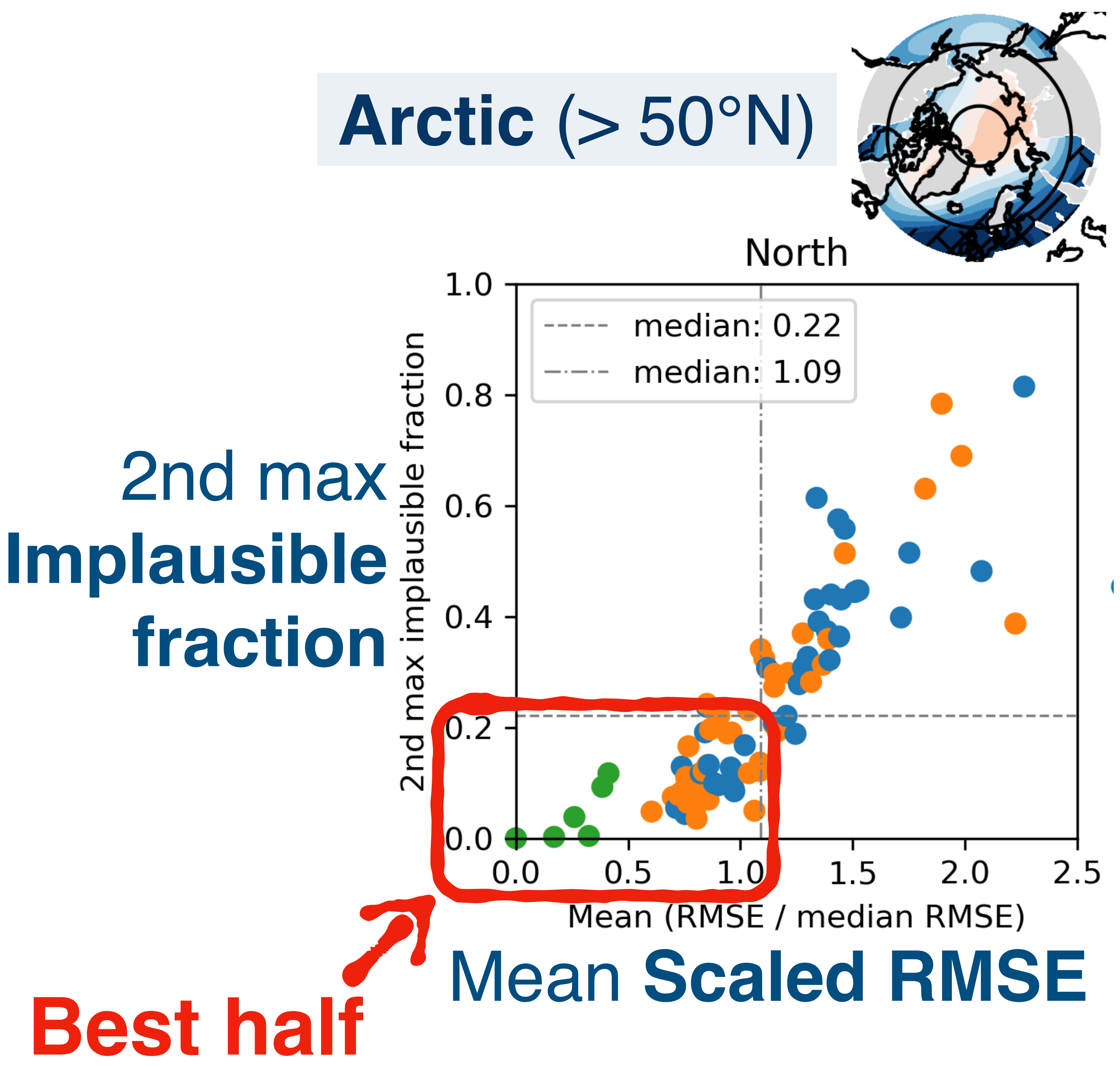
Mean **Scaled RMSE**

- CMIP models are **more implausible in the Antarctic** than in the Arctic

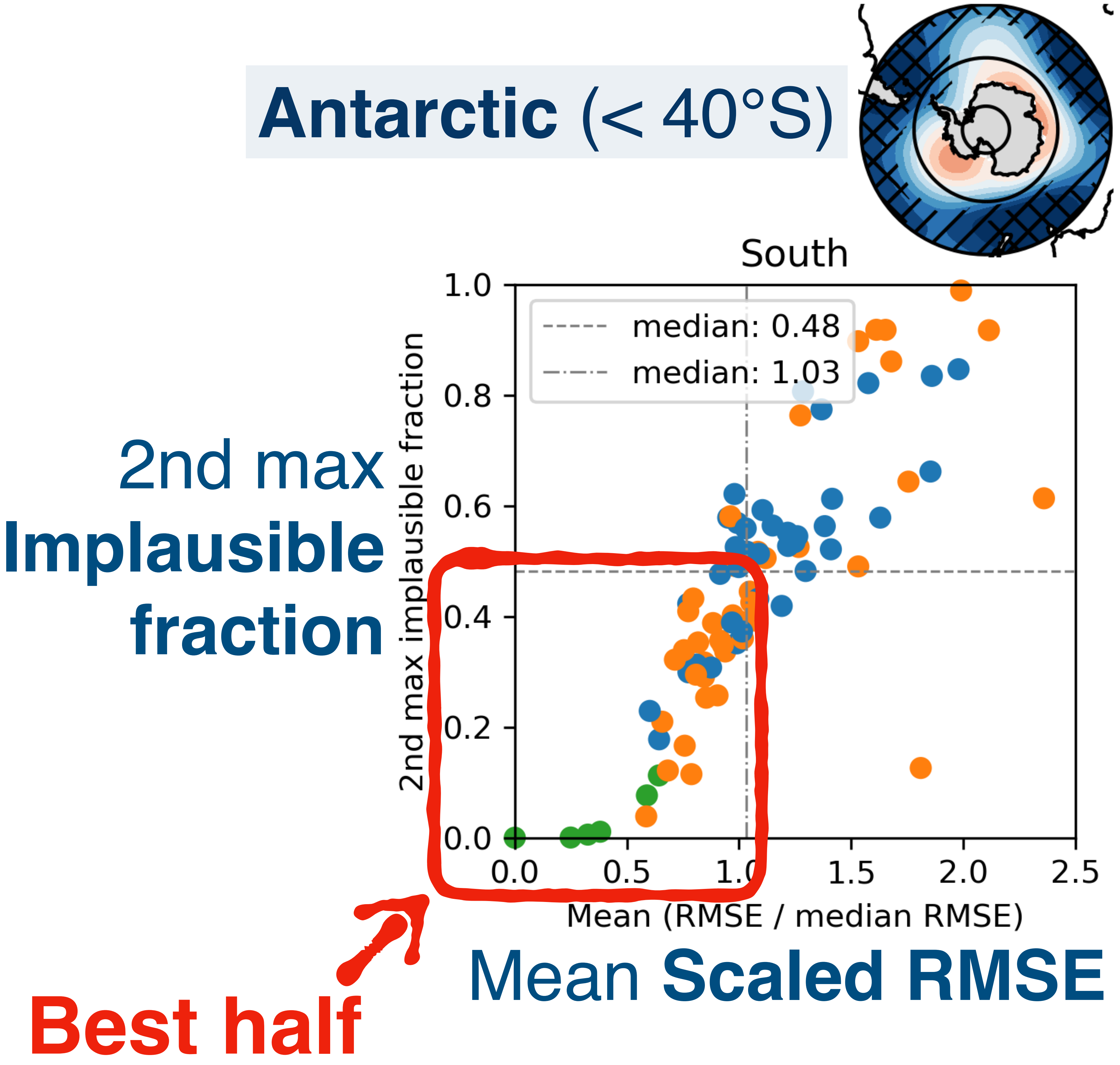
Conclusions: CMIP models over polar regions

Contact: cecile.agosta@lsce.ipsl.fr

Arctic (> 50°N)



Antarctic (< 40°S)

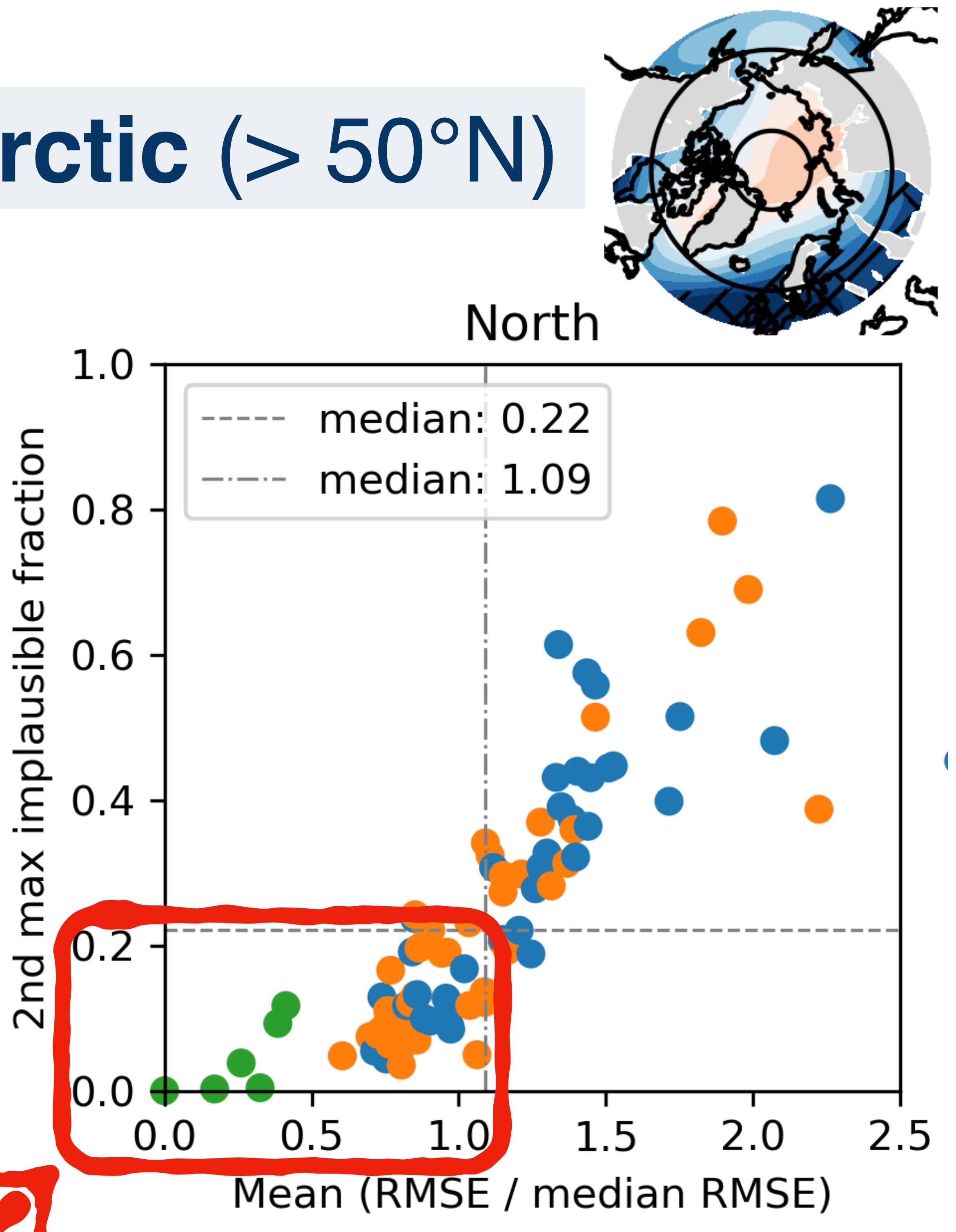


- CMIP models are **more implausible in the Antarctic** than in the Arctic
- **Same « Best half »** for Scaled RMSE and Implausibility

Conclusions: CMIP models over polar regions

Contact: cecile.agosta@lsce.ipsl.fr

Arctic (> 50°N)

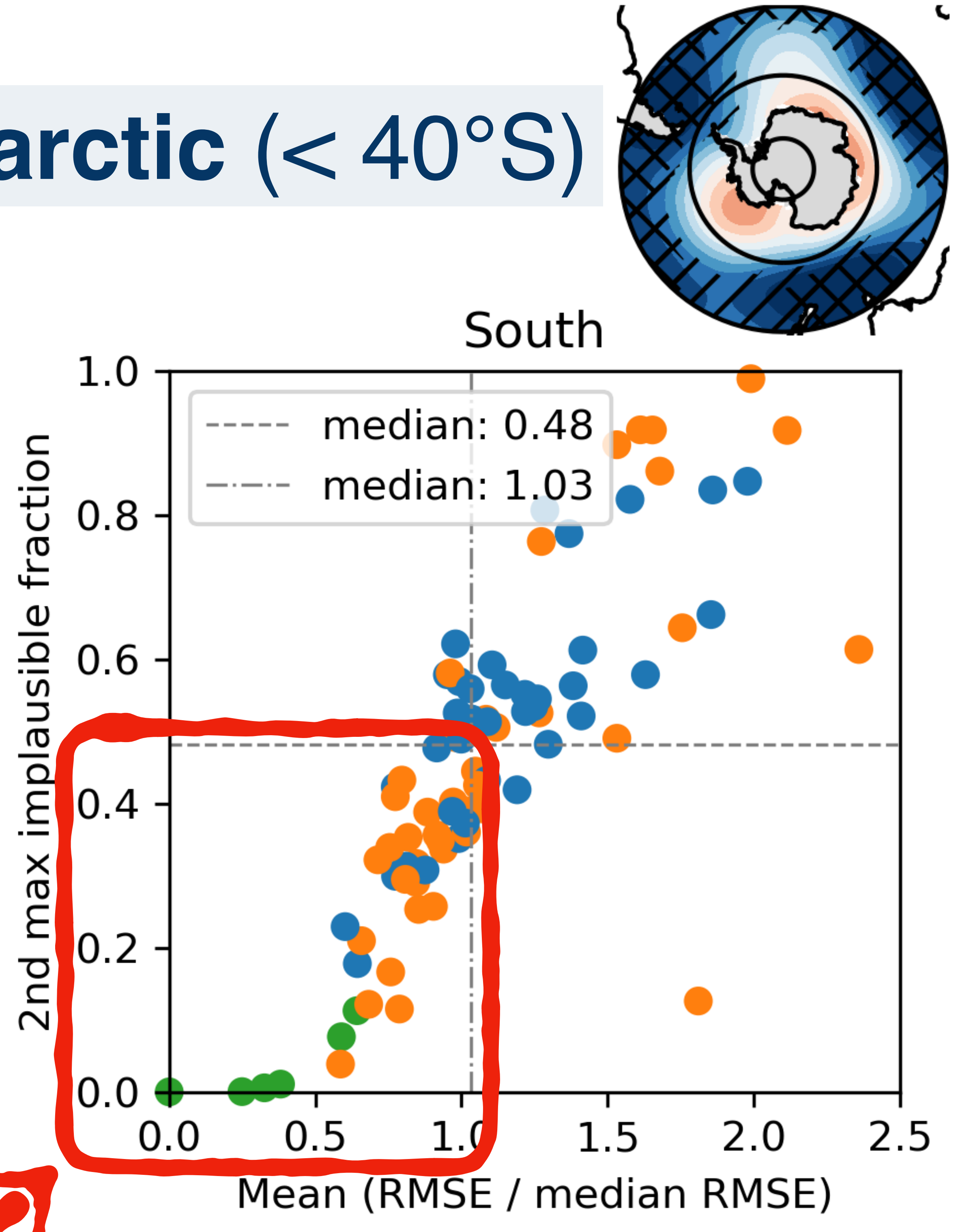


2nd max **Implausible fraction**

Mean **Scaled RMSE**

Best half

Antarctic (< 40°S)



2nd max **Implausible fraction**

Mean **Scaled RMSE**

Best half

- CMIP models are **more implausible in the Antarctic** than in the Arctic
- **Same « Best half »** for Scaled RMSE and Implausibility
- « Best half » contains **both CMIP5 and CMIP6 models**

Potsdam Institute for
Climate Impact Research

Selection of GCMs for ISIMIP -

Best(?) practice under constraints

Lisa Novak, Head of ISIMIP data team, Potsdam Institute for Climate Impact Research

Step 0: The usual situation

We are in a hurry between the availability of the next generation of CMIP simulations and the next IPCC Report….

We cannot wait for the entire CMIP ensemble if we want to keep a chance to provide CMIPX based impact simulations for the next IPCC ARX

Step 1: Data availability

Variable	Specifier	Resolution
Near-Surface Relative Humidity	**hurs**	0.5° grid, daily
Precipitation (including snowfall)	**pr**	0.5° grid, daily
Snowfall	**prsn**	0.5° grid, daily
Surface Air Pressure or sea level pressure (psl)	**ps or psl**	0.5° grid, daily
Surface Downwelling Longwave Radiation	**rlds**	0.5° grid, daily
Surface Downwelling Shortwave Radiation	**rsds**	0.5° grid, daily
Near-Surface Wind Speed or zonal wind components	**sfcwind or uas and vas**	0.5° grid, daily
Near-Surface Air Temperature	**tas**	0.5° grid, daily
Daily Maximum Near-Surface Air Temperature	**tasmax**	0.5° grid, daily
Daily Minimum Near-Surface Air Temperature	**tasmin**	0.5° grid, daily

Set of models that:

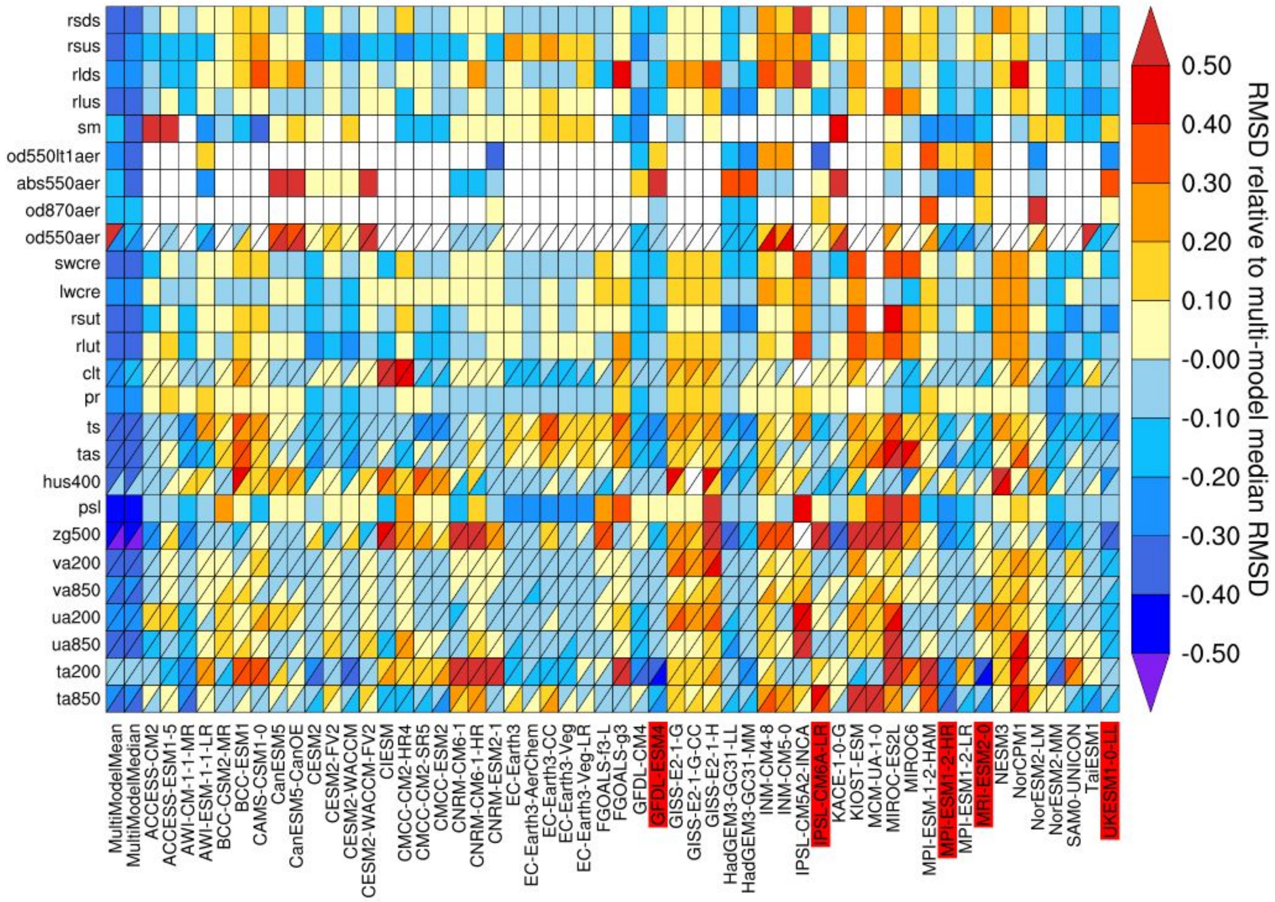
- provide **all required atmospheric variables in daily resolution**
- for selected ScenarioMIP scenarios
- at least 500 years of picontrol

Step 1.1: check for further data

- We also check for all the required oceanic forcings (fisheries and marine ecosystems sector)
  - Less than 5 models to begin with
  - → not a selection criteria

- Input data needed for the tropical cyclone modelling

Step 2: Performance in the historical period



Evaluation based on ESMValTool v2.0

From 17 GCMs reproducing the observational data well only four(!) provided the required daily data at the time of model selection (GFDL-ESM4, MPI-ESM1-2-HR, MRI-ESM2-0, UKESM1-0-LL)

→ Data availability as a severe constraint

- We want at least 5 models
  - PSL-CM6A-LR provides our data needs and has an at least average performance in the historical period

From these models

- GFDL-ESM4 does not provide all data needed for the ISIMIP tropical cyclone modelling.
- GFDL-ESM4 provides the most comprehensive oceanic bio-geochemical forcings
- Other models cover less and partly other oceanic variables.

Step 3: Structural independence, representation of feedbacks

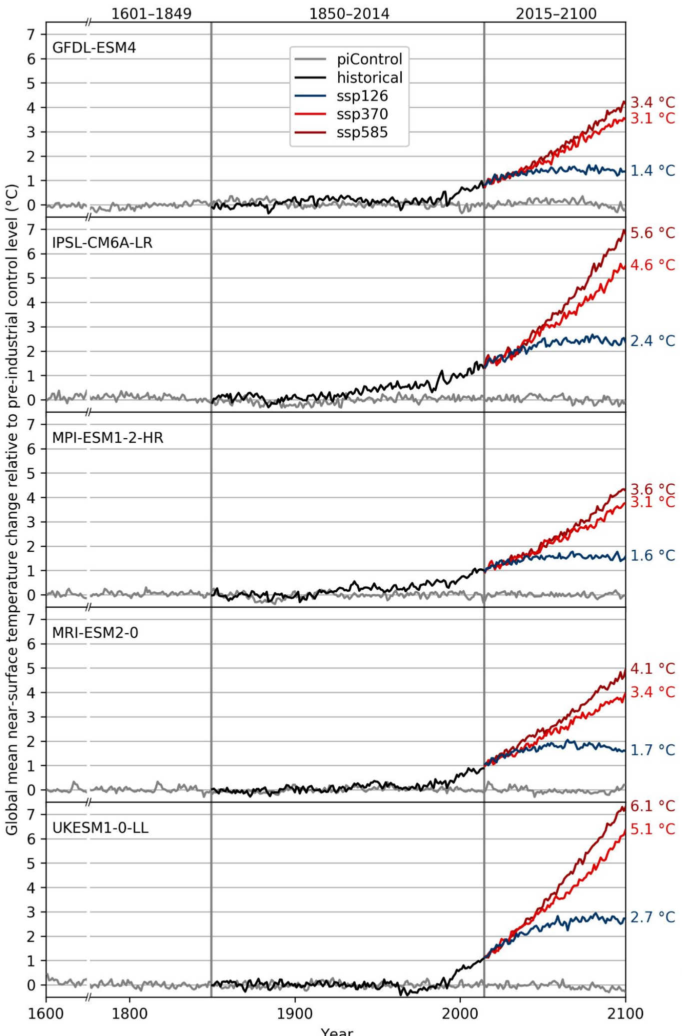
- GCMs are structurally independent in terms of their ocean and atmosphere model components.
- Coupled climate and carbon cycle
- For some: fully interactive chemistry and aerosol components.
- Prognostic couplings between processes and model domains to maximise the coverage of simulated feedbacks.

Step 4: Equilibrium Climate Sensitivity (ECS)

- The five GCMs represent both the mean and the range of the full CMIP6 multi-model ensemble ECS well.

- ISIMIP3b GCMs ESC mean matches CMIP6 multi model mean of 3.7°C.
  - Three models with below-average ECS: GFDL-ESM4, MPI-ESM1-2-HR, MRI-ESM2-0.
  - Two models with above-average ECS: IPSL-CM6A-LR, UKESM1-0-LL

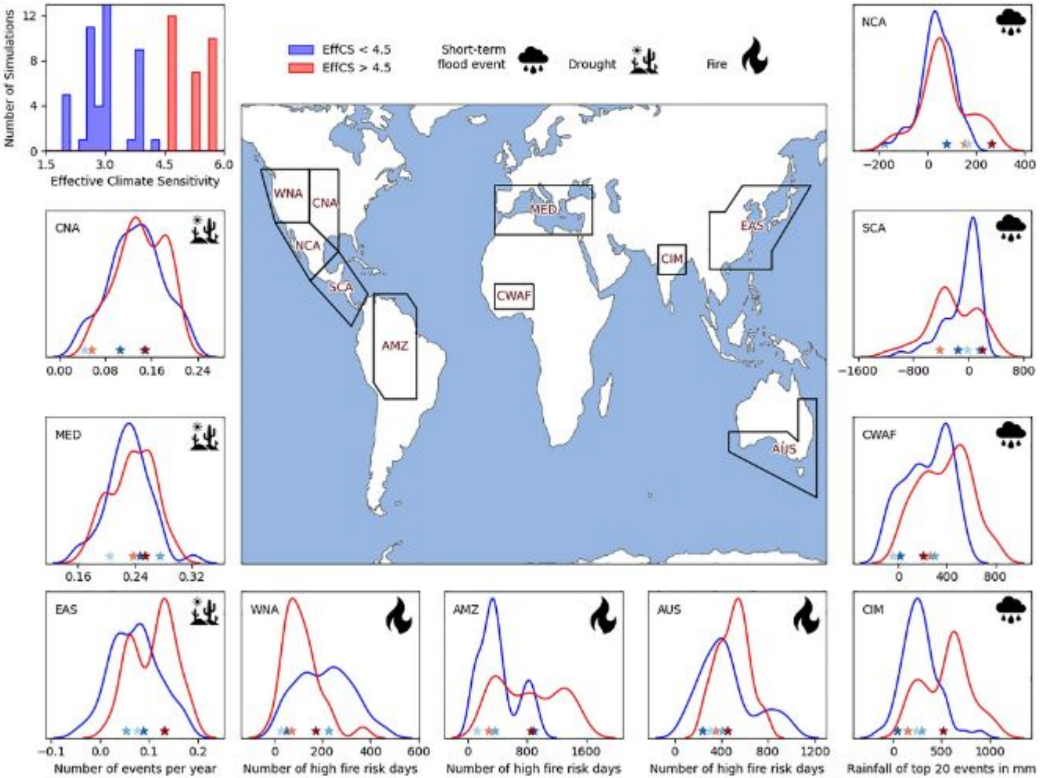
- The transient climate response (TCR) of 2.0°C is also precisely met.



Equilibrium Climate Sensitivity (ECS) as criterion for model selection?

Regional impacts poorly constrained by climate sensitivity

Swaminathan, R.**, Schewe, J.**, Walton, J., Zimmermann, K., Jones, C., Betts, R. A., Burton, C., Jones, C. D., Mengel, M., Reyer, C. P. O., Turner, A. G., & Weigel, K.



- ECS says little about regional impacts, which is what we want to study.
- Strongly recommend NOT to disregard models just based on their ECS
- We try to represent the full range of ESC of CMIP models

0. We should get started

1. Data availability (atmospheric, ocean, tropical cyclone)

2. Performance in the historical period

3. Structural independence & representation of feedbacks

4. Equilibrium climate sensitivity