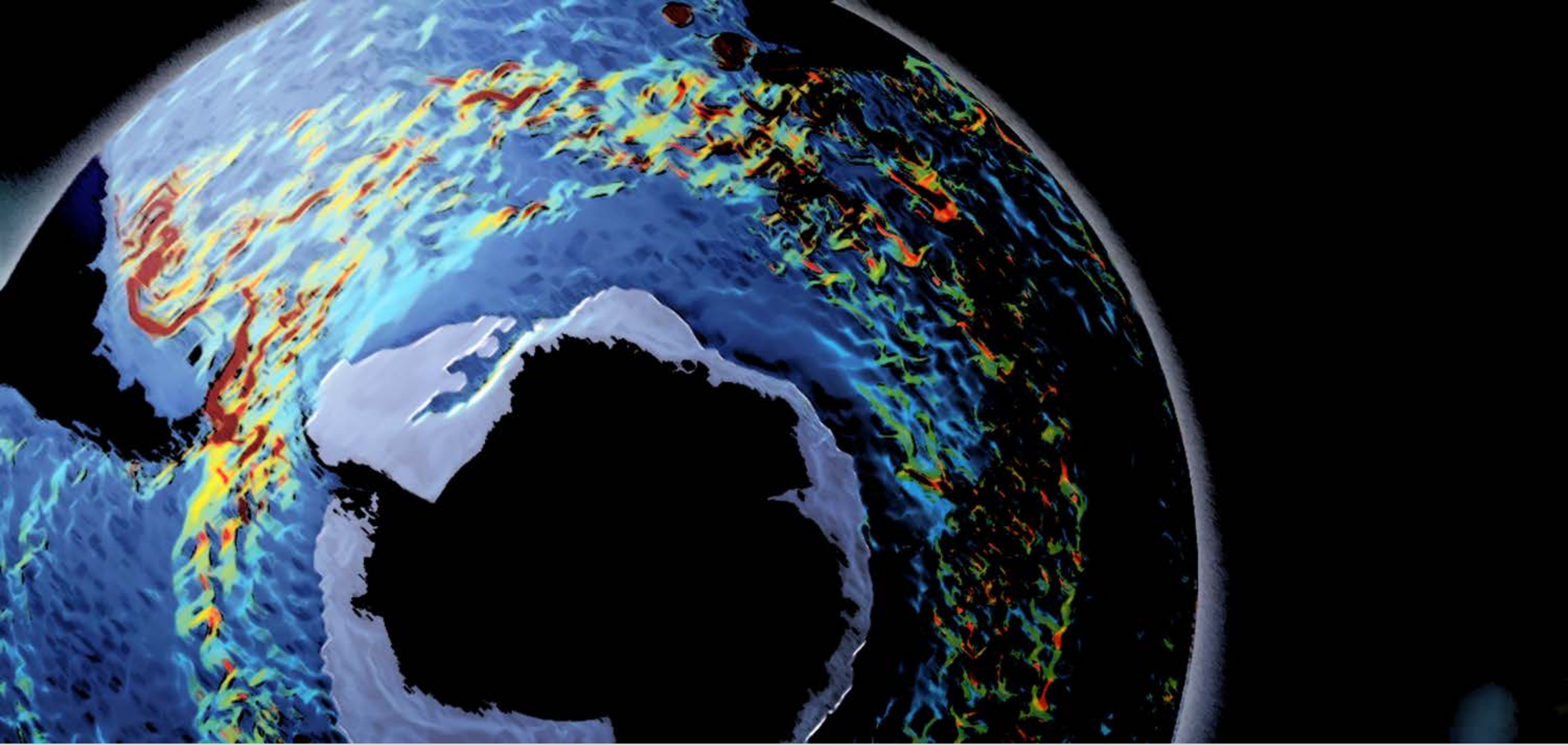
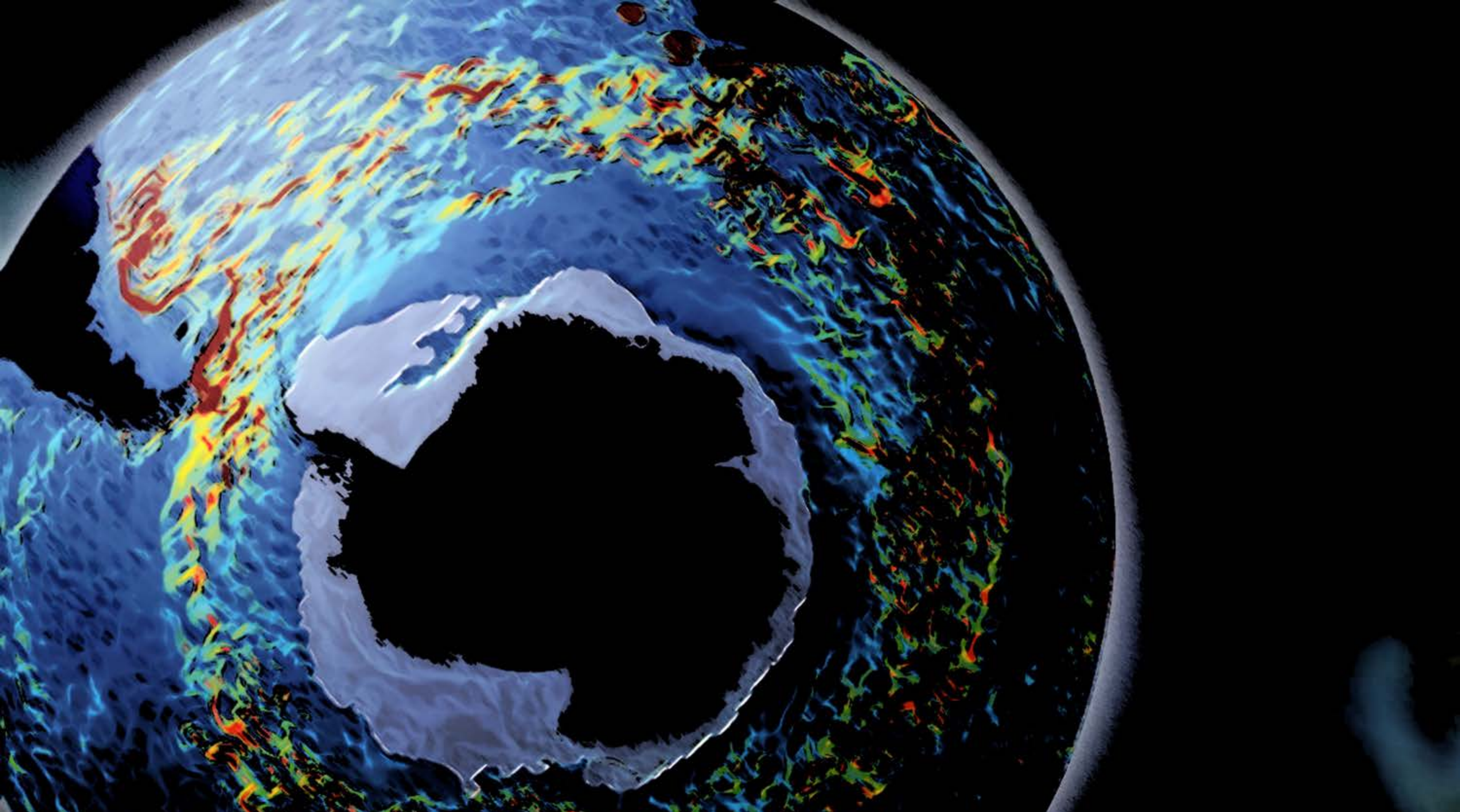


Process-oriented diagnostics of Earth system models

advances & & challenges for understanding the ocean's role in the climate system.

Becki Beadling, PhD

2023 US CLIVAR SUMM



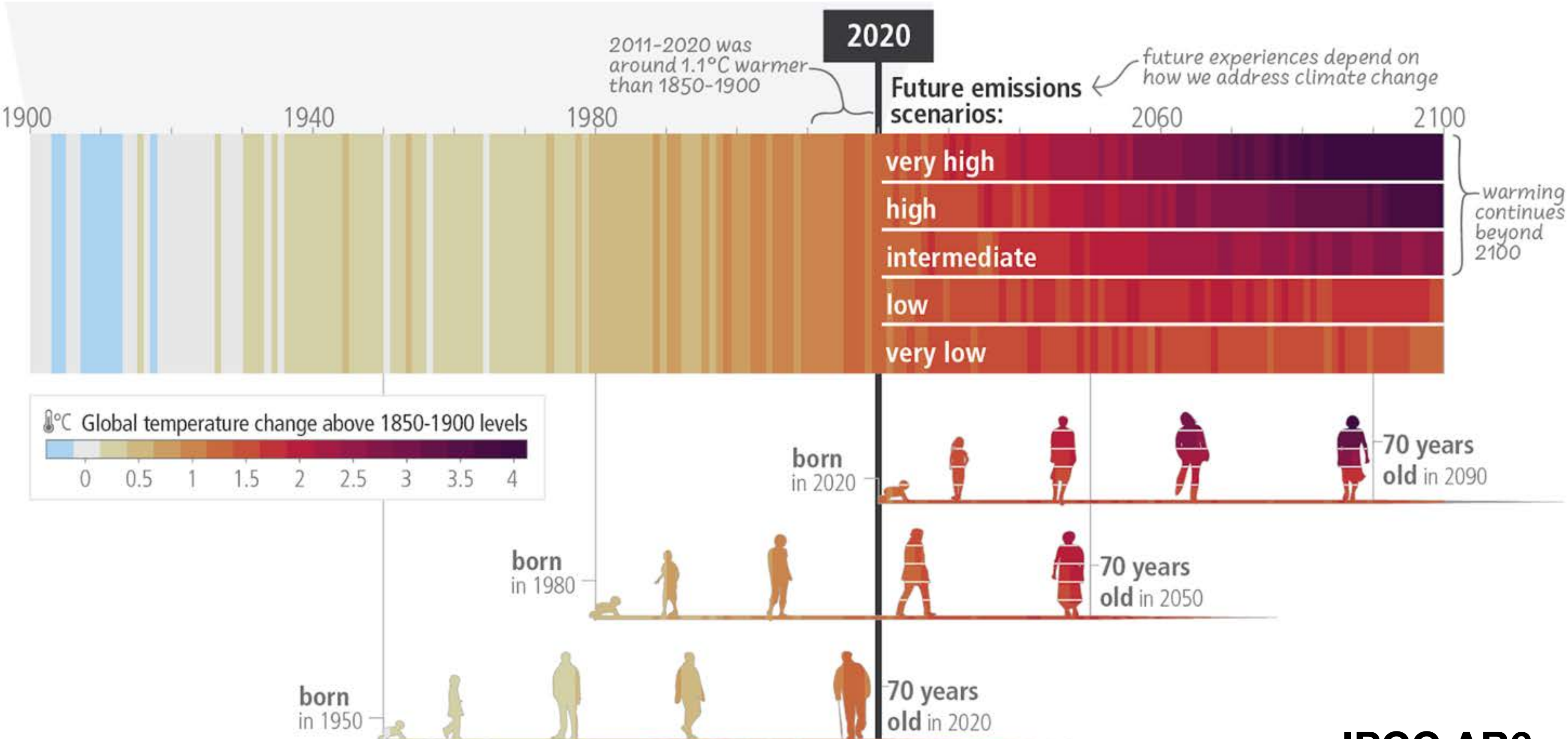
Part 1

Motivation for model evaluation, current capabilities, emphasizing need for process-oriented ocean diagnostics

Part 2

An example of development of process-oriented diagnostics for circulation in the Southern Ocean

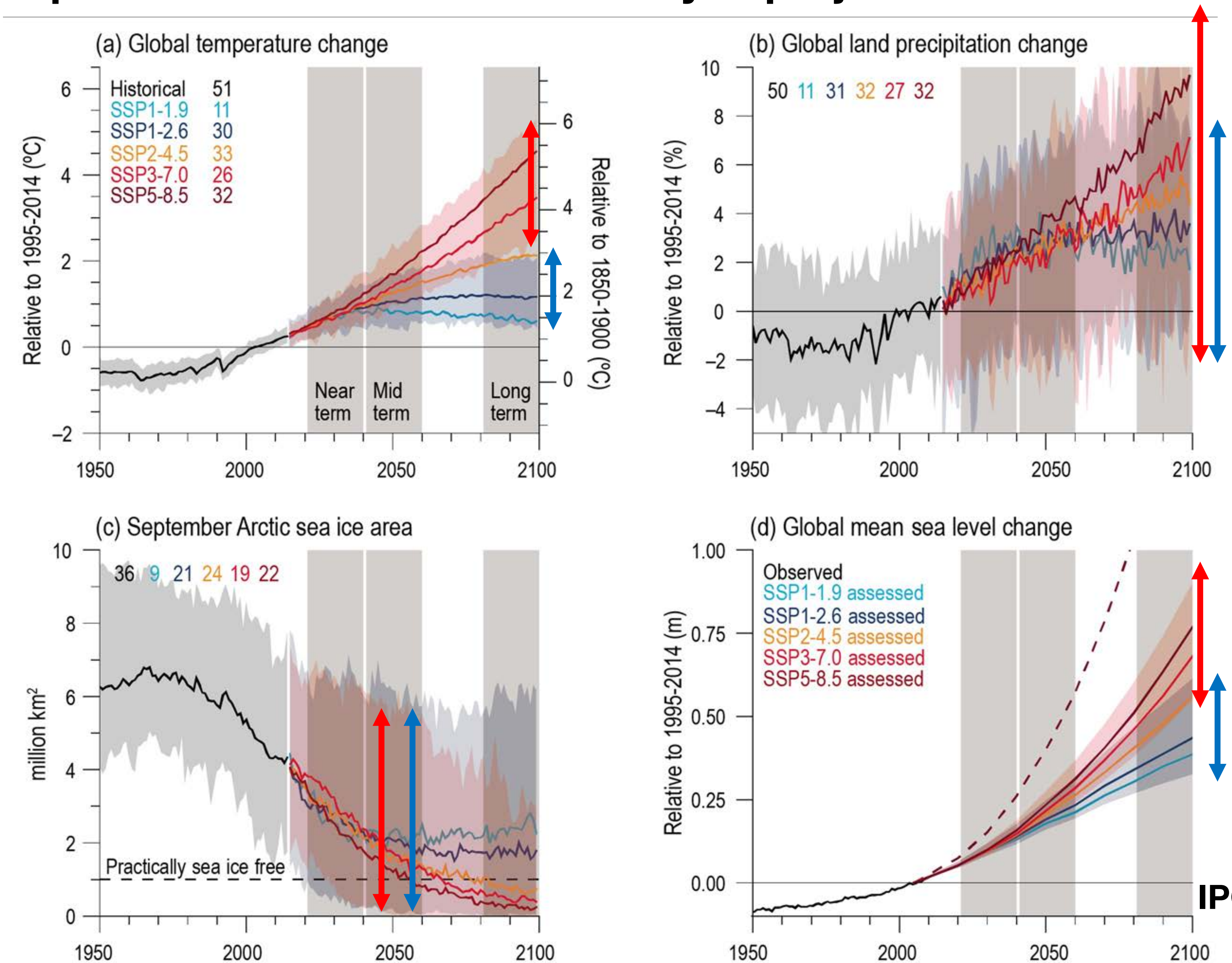
The planet is warming and will continue to do so in the face of unabated climate change.



IPCC AR6

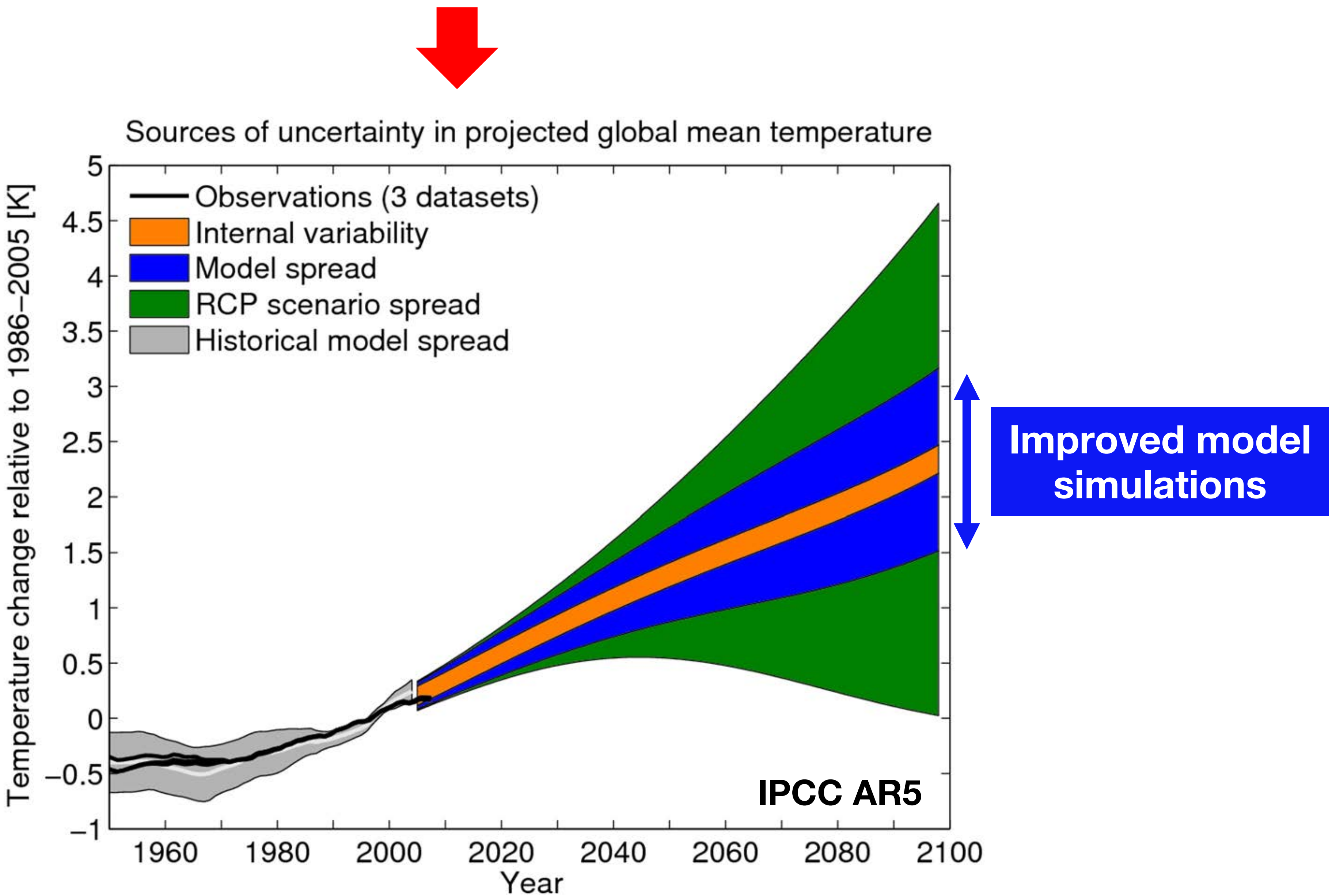
Such a potentially large perturbation to the climate system is and will have adverse impacts on humans & natural ecosystems.

Developing appropriate climate adaptation & mitigation strategies requires a reduction in uncertainty in projected climate change.

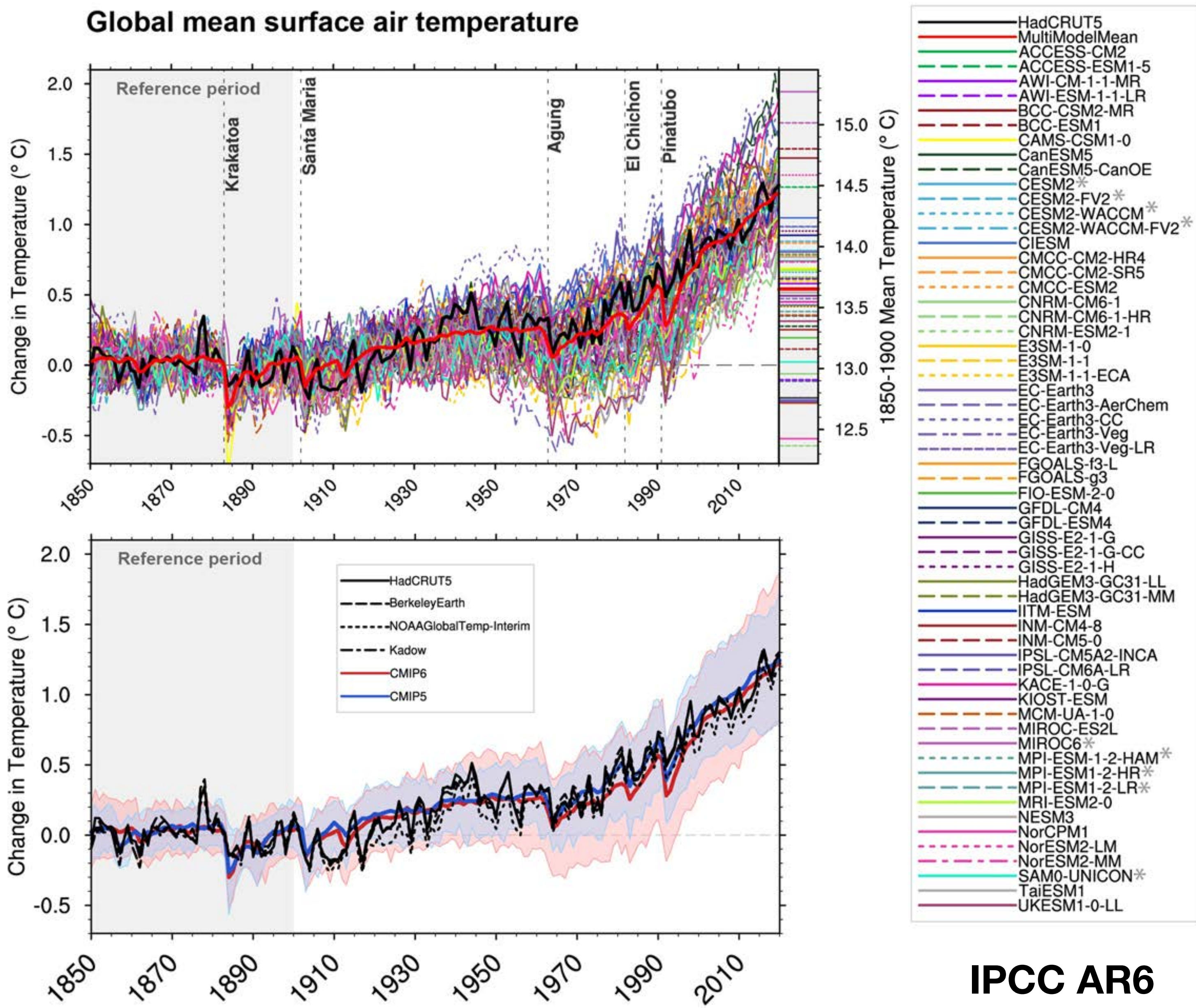


IPCC AR6

Developing appropriate climate adaptation & mitigation strategies requires a reduction in uncertainty in projected climate change.



Climate models do a pretty good job at capturing the observed increase in global average surface air temperature.

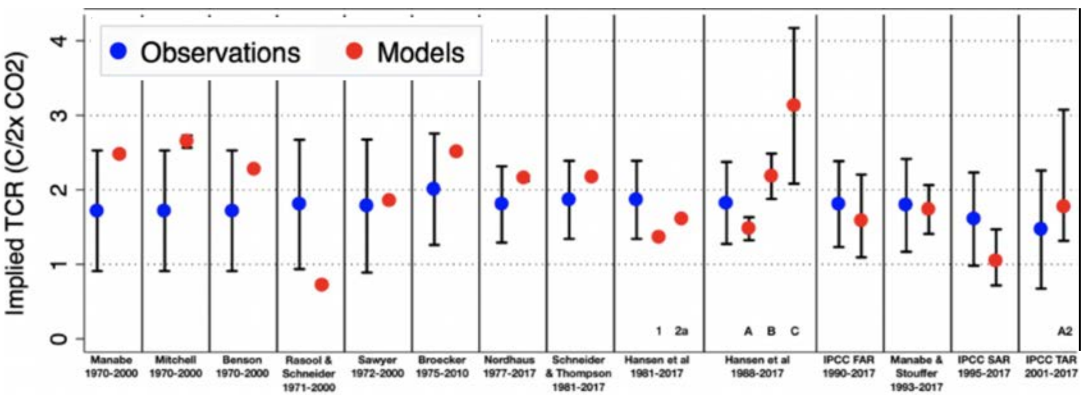


IPCC AR6

Climate models do a pretty good job at capturing the observed increase in global average surface air temperature.

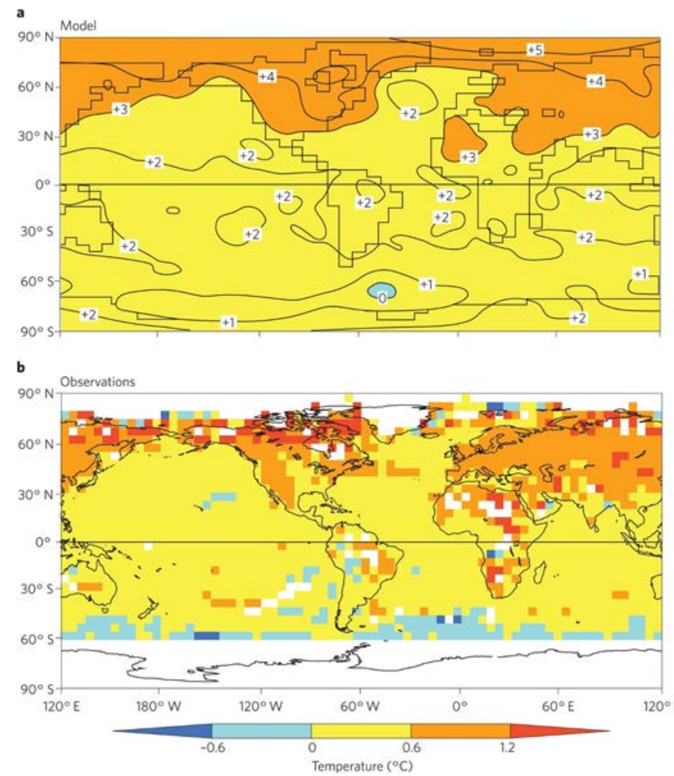
Evaluating the Performance of Past Climate Model Projections

Zeke Hausfather[1], Henri F. Drake[2,3], Tristan Abbott[3], and Gavin A. Schmidt[4]



Assessing temperature pattern projections made in 1989

Ronald J. Stouffer[1*] and Syukuro Manabe[2]



Climate models published over the past five decades were generally accurate in predicting global warming and the spatial distribution of it in the years after publication.

At the regional scale & considering other climate variables outside of the temperature response, **confidence lowers, uncertainties rise, and inter-model spread increases ….**

Need for advanced and coordinated model evaluation capabilities for improved model development and for better interpretation of future projections.

Need for advanced and coordinated model evaluation capabilities for improved model development and for better interpretation of future projections.

Climate Model Evaluation:

How well do climate models simulate aspects of the current mean climate for which we have an observational constraint?

- Differences between modeled and observed indicate systematic errors.
- **Show where** and **in what ways** models are succeeding or failing at reproducing the climate under current / past conditions.

Need for advanced and coordinated model evaluation capabilities for improved model development and for better interpretation of future projections.

Climate Model Evaluation:

How well do climate models simulate aspects of the current mean climate for which we have an observational constraint?

- Differences between modeled and observed indicate systematic errors.
- **Show where** and **in what ways** models are succeeding or failing at reproducing the climate under current / past conditions.

Mean-state
Observed change
Observed spatial and temporal variability
Modes of variability (ENSO, SAM, MJO … etc.)

Climate Model Evaluation:
How well do climate models simulate aspects of the current mean climate for which we have an observational constraint?

- Differences between modeled and observed indicate systematic errors.
- **Show where** and **in what ways** models are succeeding or failing at reproducing the climate under current conditions

Intercomparison Makes for a Better Climate Model

1997

PAGES 445–446, 451

Gerald A. Meehl, George J. Boer, Curt Covey, Mojib Latif, and Ronald J. Stouffer

“Since simulation results are widely used to identify vulnerabilities and study societal impacts that have policy implications, **the simulation capabilities of these models must be systematically assessed. CMIP fills this role**. ”

Climate Model Evaluation:
How well do climate models simulate aspects of the current mean climate for which we have an observational constraint?

- Differences between modeled and observed indicate systematic errors.
- **Show where** and **in what ways** models are succeeding or failing at reproducing the climate under current conditions

Intercomparison Makes for a Better Climate Model

1997

PAGES 445–446, 451

Gerald A. Meehl, George J. B
and Ronald J. Stouffer

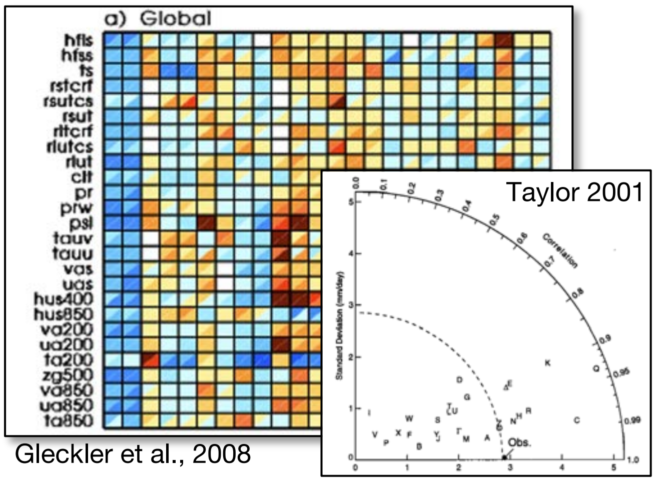
Expansion to multi-model evaluation:
Evaluate **different models**
Evaluate **different model versions**
Evaluate performance **across model generations**
Evaluate **response to forcings**

"Since simulation results are widely used to identify vulnerabilities and study societal impacts that have policy implications, **the simulation capabilities of these models must be systematically assessed. CMIP fills this role**. "

Climate Model Evaluation

How well do climate models simulate aspects of the current mean climate for which we have an observational constraint ?

"Benchmarking": Performance Metrics (Scalar) [The symptom, the *what*]



Gleckler et al., 2008

"Benchmarking": to compare a model simulation to a standard (observational constraint)

"Benchmark *Experiment*": a critical test that a model should pass to demonstrate its viability as a tool to probe the climate system (i.e., the CMIP historical experiment)

Climate Model Evaluation

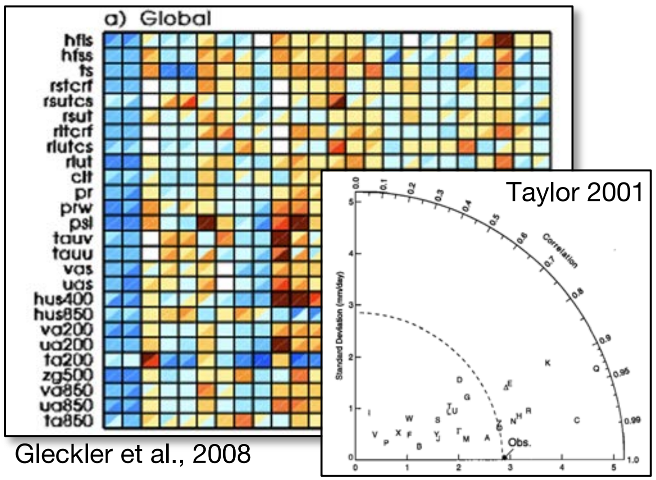
How well do climate models simulate aspects of the current mean climate for which we have an <u>observational constraint</u> ?

"Benchmarking": Performance Metrics (Scalar) [The symptom, the *what*]

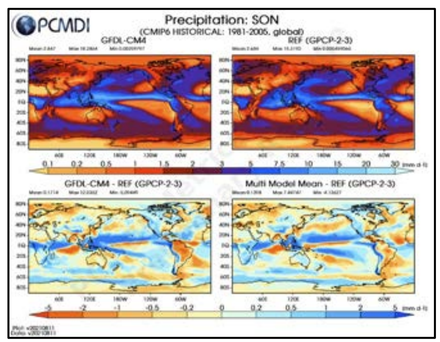
Diagnostics

Maps, timeseries, zonal-averaged fields, power spectra, etc ....

Helps "diagnose" the scalar quantity.



Gleckler et al., 2008



ESMValTool
(recipe ocean_ice_extent)

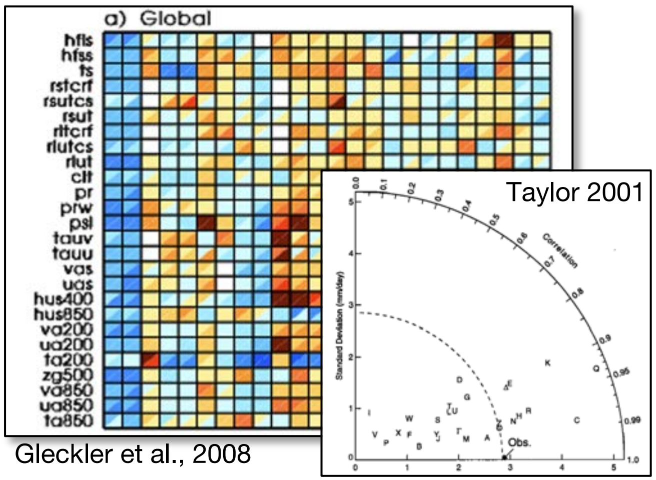
"Benchmarking": to compare a model simulation to a standard (observational constraint)

"Benchmark *Experiment*": a critical test that a model should pass to demonstrate its viability as a tool to probe the climate system (i.e., the CMIP historical experiment)

Climate Model Evaluation

How well do climate models simulate aspects of the current mean climate for which we have an <u>observational constraint</u> ?

"Benchmarking": Performance Metrics (Scalar) [The symptom, the *what*]



Gleckler et al., 2008

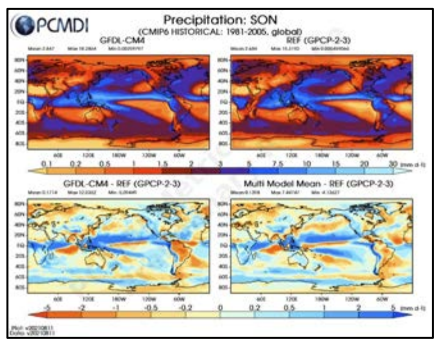
"Benchmarking": to compare a model simulation to a standard (observational constraint)

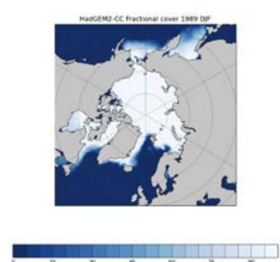
"Benchmark *Experiment*": a critical test that a model should pass to demonstrate its viability as a tool to probe the climate system (i.e., the CMIP historical experiment)

Diagnostics

Maps, timeseries, zonal-averaged fields, power spectra, etc ....

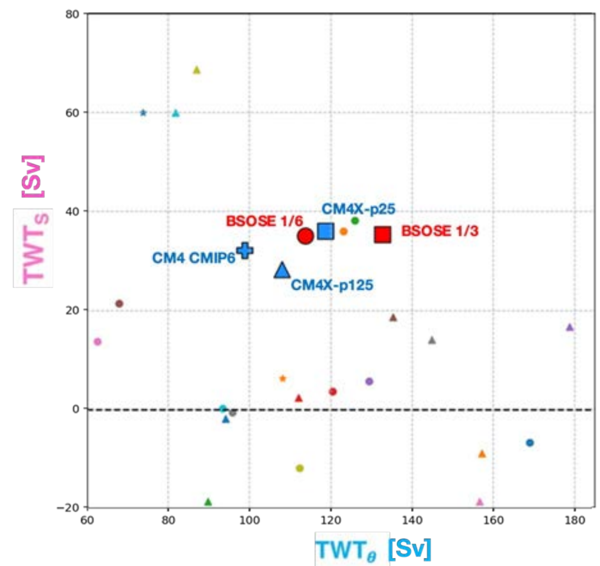
Helps "diagnose" the scalar quantity.





ESMValTool
(recipe ocean_ice_extent)

"Process Oriented Diagnostics": [The *why*]



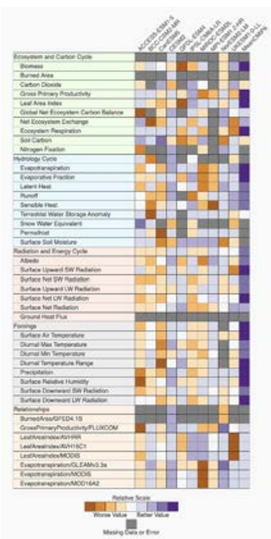
What's out there already for model diagnostics?

Coordinated Model Evaluation Capabilities (CMEC)

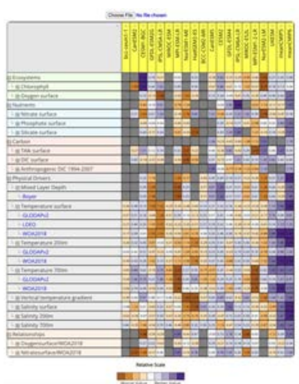
U.S. DEPARTMENT OF ENERGY | Office of Science

ILAMB

The International Land Model Benchmarking (ILAMB) Package

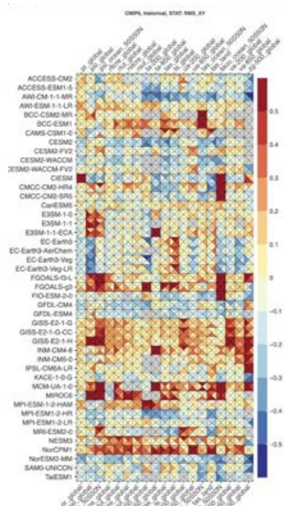


The International Ocean Model Benchmarking (OLAMB) Package



PCMDI Earth System Model Evaluation Project

PCMDI Metrics Package (PMP)



- Mean Climate
- Benchmarking Simulated Precipitation
- El Niño–Southern Oscillation (ENSO)
- Extratropical Modes of Variability
- Madden-Julian Oscillation (MJO)
- Monsoon Characteristics (example)

Toolkit for Extreme Climate Analysis (TECA)

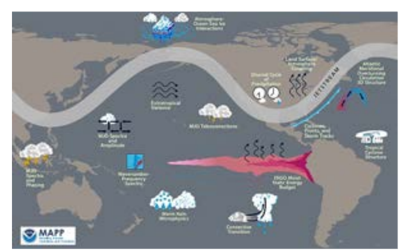
Analyzing Scales of Precipitation (ASoP)

Cyclone Metrics Package (CyMeP)

Drought Metrics Package

The MDTF-Diagnostics Package

A Portable Framework for Weather and Climate Model Data Analysis



NOAA

Specifically, Process-Oriented Diagnostics (PODs)

What's out there already for model diagnostics?

Coordinated Model Evaluation Capabilities (CMEC)

U.S. DEPARTMENT OF ENERGY | Office of Science

- PCMDI Metrics Package (PMP)
- The International Land Model Benchmarking (ILAMB) Package
- The International Ocean Model Benchmarking (OLAMB) Package
- Toolkit for Extreme Climate Analysis (TECA)
- Analyzing Scales of Precipitation (ASoP)
- Cyclone Metrics Package (CyMeP)
- Drought Metrics Package

The MDTF-Diagnostics Package

A Portable Framework for Weather and Climate Model Data Analysis

Emphasis on software development : community-based framework – emphasizing code / diagnostics that are maintainable, interoperable, and portable.

PROCESS-ORIENTED EVALUATION OF CLIMATE AND WEATHER FORECASTING MODELS

Eric D. Maloney, Andrew Gettelman, Yi Ming, J. David Neelin, Daniel Barrie, Annarita Mariotti, C.-C. Chen, Danielle R. B. Coleman, Yi-Hung Kuo, Bohar Singh, H. Annamalai, Alexis Berg, James F. Booth, Suzana J. Camargo, Aiguo Dai, Alex Gonzalez, Jan Hafner, Xianan Jiang, Xianwen Jing, Daehyun Kim, Arun Kumar, Yumin Moon, Catherine M. Naud, Adam H. Sobel, Kentaroh Suzuki, Fuchang Wang, Junhong Wang, Allison A. Wing, Xiaobiao Xu, and Ming Zhao

Specifically, Process-Oriented Diagnostics (PODs)

Targets a specific physical process or emergent behavior, with the goals of determining how accurately the model represents that process, ensuring that models produce **the right answers for the right reasons**, and identifying gaps in the understanding of phenomena.

What's out there already for model diagnostics?

CMEC

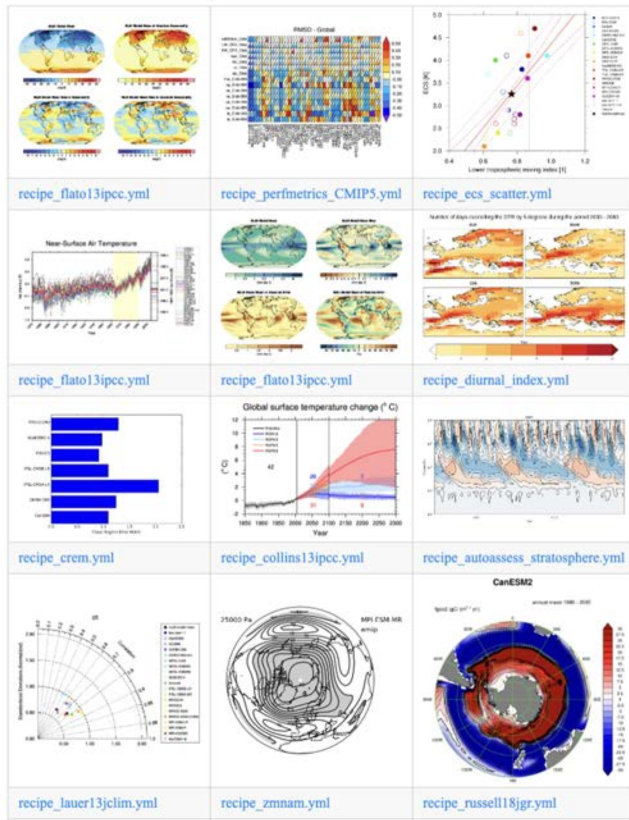
U.S. DEPARTMENT OF ENERGY | Office of Science

- PCMDI Metrics Package (PMP)
- The International Land Model Benchmarking (ILAMB) Package
- The International Ocean Model Benchmarking (OLAMB) Package
- Toolkit for Extreme Climate Analysis (TECA)
- Analyzing Scales of Precipitation (ASoP)
- Cyclone Metrics Package (CyMeP)
- Drought Metrics Package
- NOAA's MDTF

ESMValTool — Earth System Model Evaluation Tool

A community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP

ESMValTool examples



What's out there already for model diagnostics?

CMEC

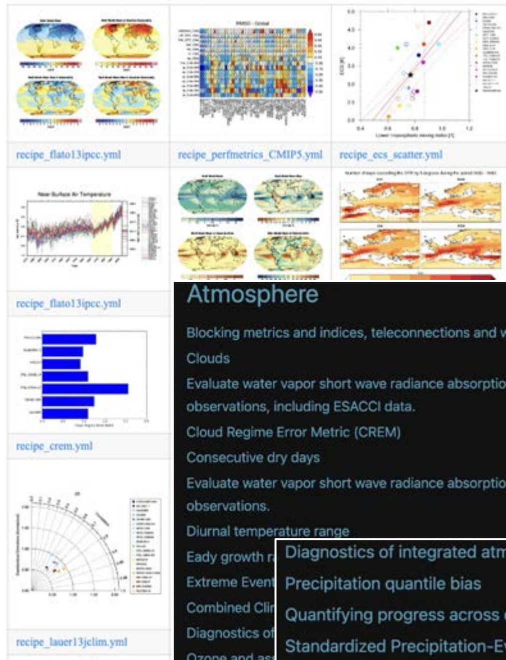
U.S. DEPARTMENT OF ENERGY | Office of Science

- PCMDI Metrics Package (PMP)
- The International Land Model Benchmarking (ILAMB) Package
- The International Ocean Model Benchmarking (OLAMB) Package
- Toolkit for Extreme Climate Analysis (TECA)
- Analyzing Scales of Precipitation (ASoP)
- Cyclone Metrics Package (CyMeP)
- Drought Metrics Package
- NOAA's MDTF

ESMValTool
Earth System Model Evaluation Tool

A community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP

ESMValTool examples



Land

Landcover - Albedo
Turnover time of carbon over land ecosystems
Hydrological models - data pre-processing
Hydro forcing comparison
Landcov
Land an
Runoff,

Ocean

Recipe for evaluating Arctic Ocean
Climate Variability Diagnostics Package (CVDP)
Nino indices, North Atlantic Oscillation (NAO), Souther Oscillation Index (SOI)
Ocean chlorophyll in ESMs compared to ESA-CCI observations.
Ocean diagnostics

Atmosphere

Blocking metrics and indices, teleconnections and weather regimes (MiLES)
Clouds
Evaluate water vapor short wave radiance absorption schemes of ESMs with the observations, including ESACCI data.
Cloud Regime Error Metric (CREM)
Consecutive dry days
Evaluate water vapor short wave radiance absorption schemes of ESMs with the observations.
Diurnal temperature range
Eady growth r
Extreme Event
Combined Cli
Diagnostics of
Ozone and as
Spatially resol
Heat wave an
Hydroclimatic
Quick insights
Modes of vari
Diagnostics of
Precipitation q
Quantifying pr

Diagnostics of integrated atmospheric methane (XCH4)
Precipitation quantile bias
Quantifying progress across different CMIP phases
Standardized Precipitation-Evapotranspiration Index (SPE
Drought characteristics following Martin (2018)
Stratosphere - Autoassess diagnostics
Land-surface Permafrost - Autoassess diagnostics
Land-surface Surface Radiation - Autoassess diagnostics
Land-surface Soil Moisture - Autoassess diagnostics
Stratosphere-troposphere coupling and annular modes indices (ZMNAM)
Thermodynamics of the Climate System - The Diagnostic Tool TheDiaTo v1.0
Zonal and Meridional Means
Radiation Budget

Climate metrics

Performance metrics for essential climate parameters
Single Model Performance Index (SMPI)

Future projections

Climate model Weighting by Independence and Performance (ClimWIP)
Constraining future Indian Summer Monsoon projections with the present-day precipitation over the tropical western Pacific
Constraining uncertainty in projected gross primary production (GPP) with machine learning
Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models
Emergent constraints for equilibrium climate sensitivity
Emergent constraints on carbon cycle feedbacks
Emergent constraints on equilibrium climate sensitivity in CMIP5: do they hold for CMIP6?
Emergent constraint on equilibrium climate sensitivity from global temperature variability
Emergent constraint on snow-albedo effect
Equilibrium climate sensitivity
KNMI Climate Scenarios 2014
Multiple ensemble diagnostic regressi

Equilibrium climate sensitivity
KNMI Climate Scenarios 2014
Multiple ensemble diagnostic regression (MDER) for constraining future austral jet position
Projected land photosynthesis constrained by changes in the seasonal cycle of atmospheric CO_2
Transient Climate Response
Climate model projections from the ScenarioMIP of CMIP6
Climate Change Hotspot

What's out there already for model diagnostics?

CMEC

U.S. DEPARTMENT OF ENERGY | Office of Science

- **PCMDI Metrics Package (PMP)**
- **The International Land Model Benchmarking (ILAMB) Package**
- **The International Ocean Model Benchmarking (OLAMB) Package**
- **Toolkit for Extreme Climate Analysis (TECA)**
- **Analyzing Scales of Precipitation (ASoP)**
- **Cyclone Metrics Package (CyMeP)**
- **Drought Metrics Package**
- **NOAA's MDTF**

ESMValTool – Earth System Model Evaluation Tool

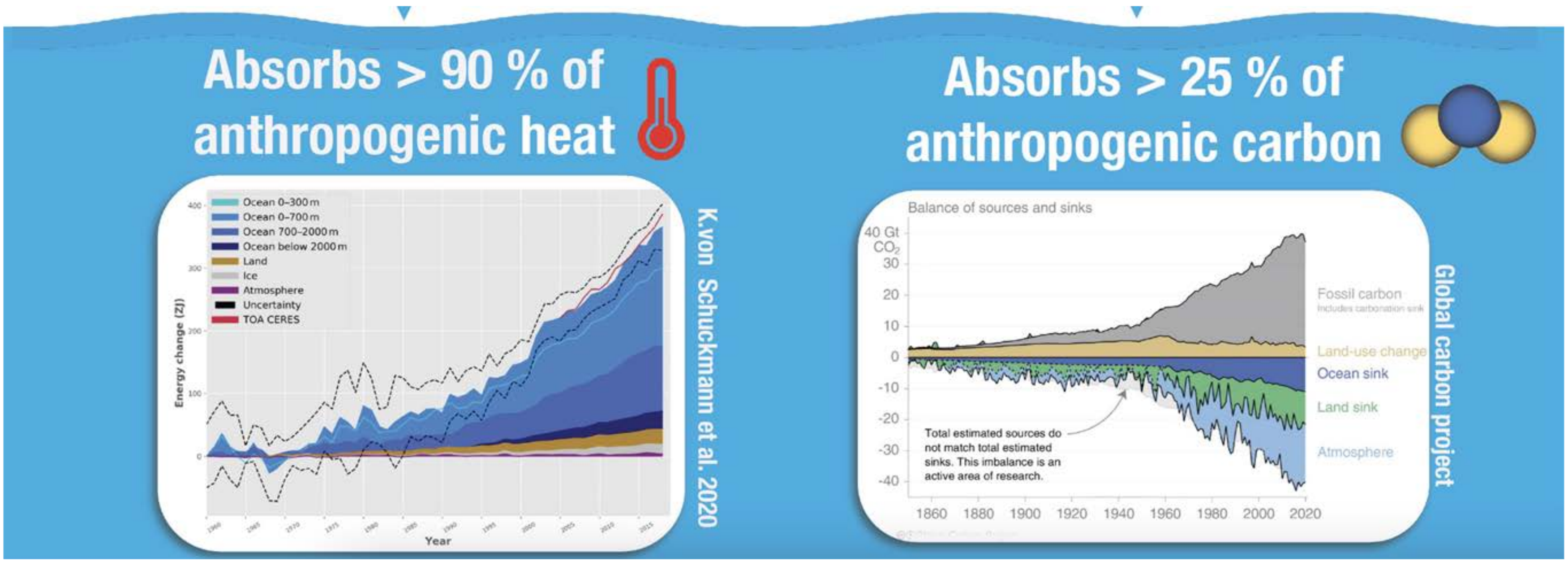
A community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP

And many other on-going efforts on **package development**, **coordination**, and **standardization / best practices** ….
(hoping to learn more from this meeting!)

And many other on-going efforts on **package development**, **coordination**, and **standardization / best practices** ….
(hoping to learn more from this meeting!)

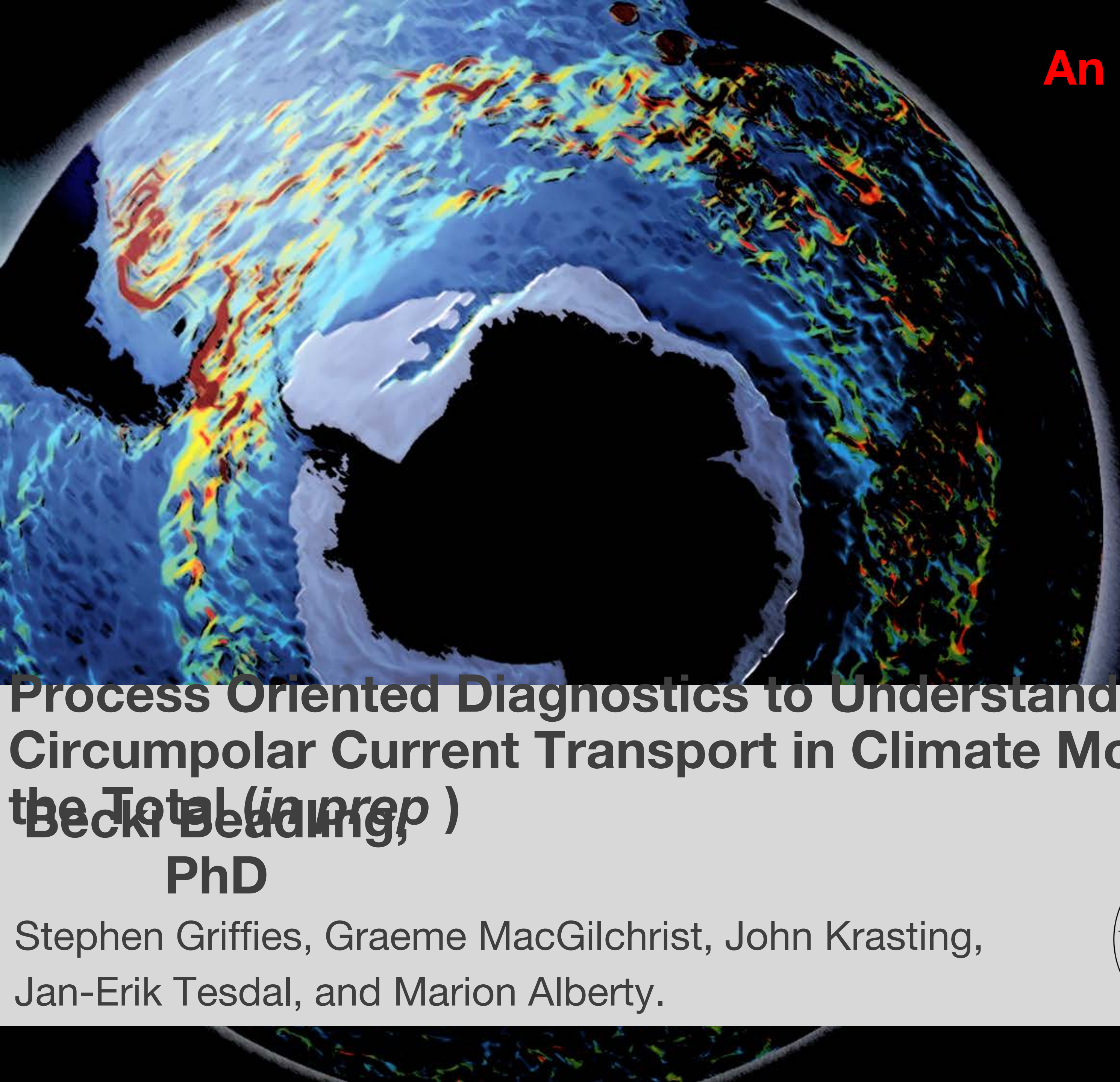
The majority of existing performance metrics & process oriented diagnostic focus on atmosphere, land, & the *surface* ocean

The majority of existing performance metrics & process oriented diagnostic focus on atmosphere, land, & the *surface* ocean



Given the critical role of the ocean in the climate system and the transient climate response, it is vital to accurately model 3D ocean processes from the air-sea interface to the ocean interior.

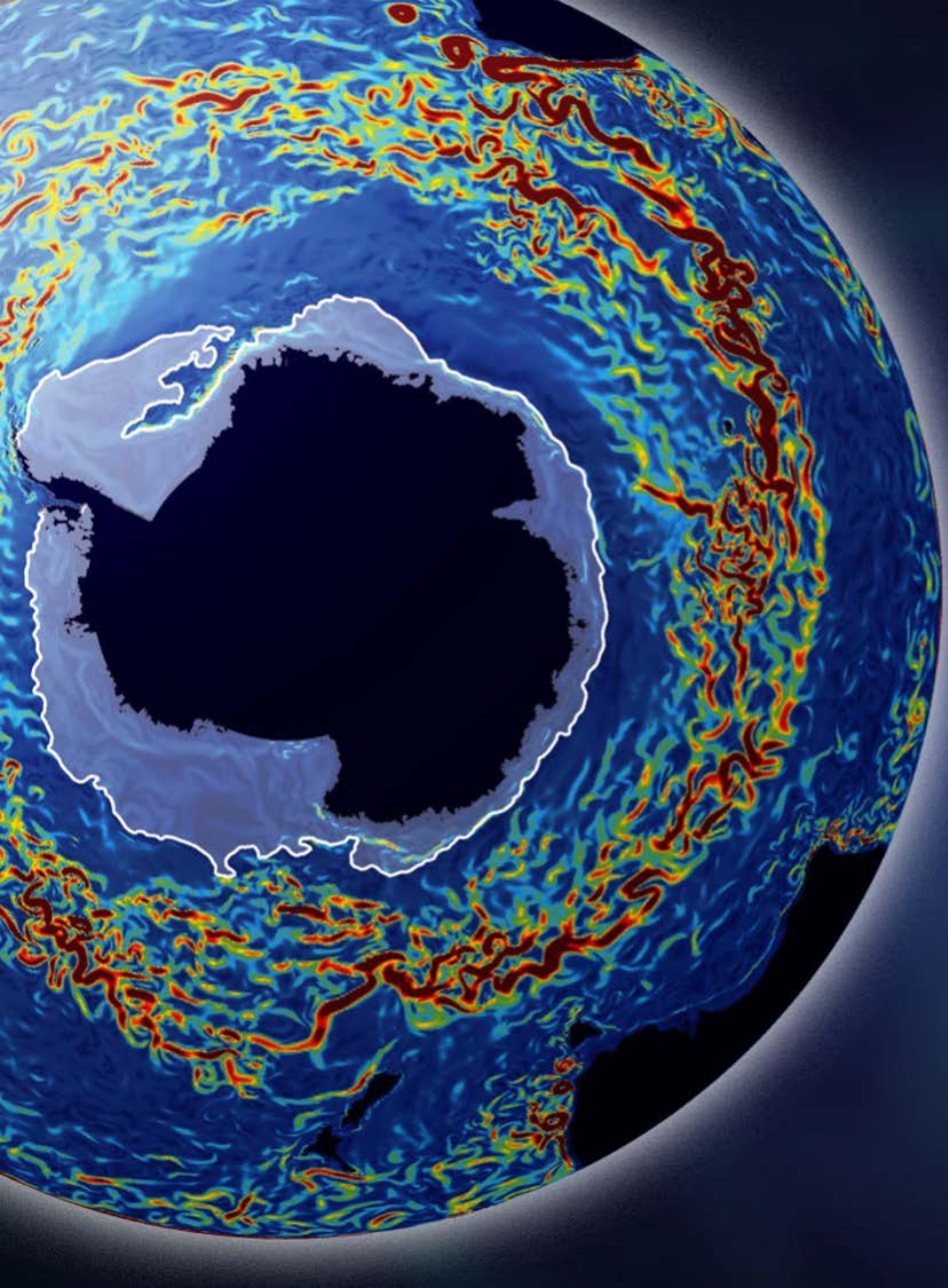
An Example ….



Process Oriented Diagnostics to Understand Antarctic Circumpolar Current Transport in Climate Models: Beyond the Total (*in prep*)

Becki Beadling, PhD

Stephen Griffies, Graeme MacGilchrist, John Krasting,
Jan-Erik Tesdal, and Marion Alberty.



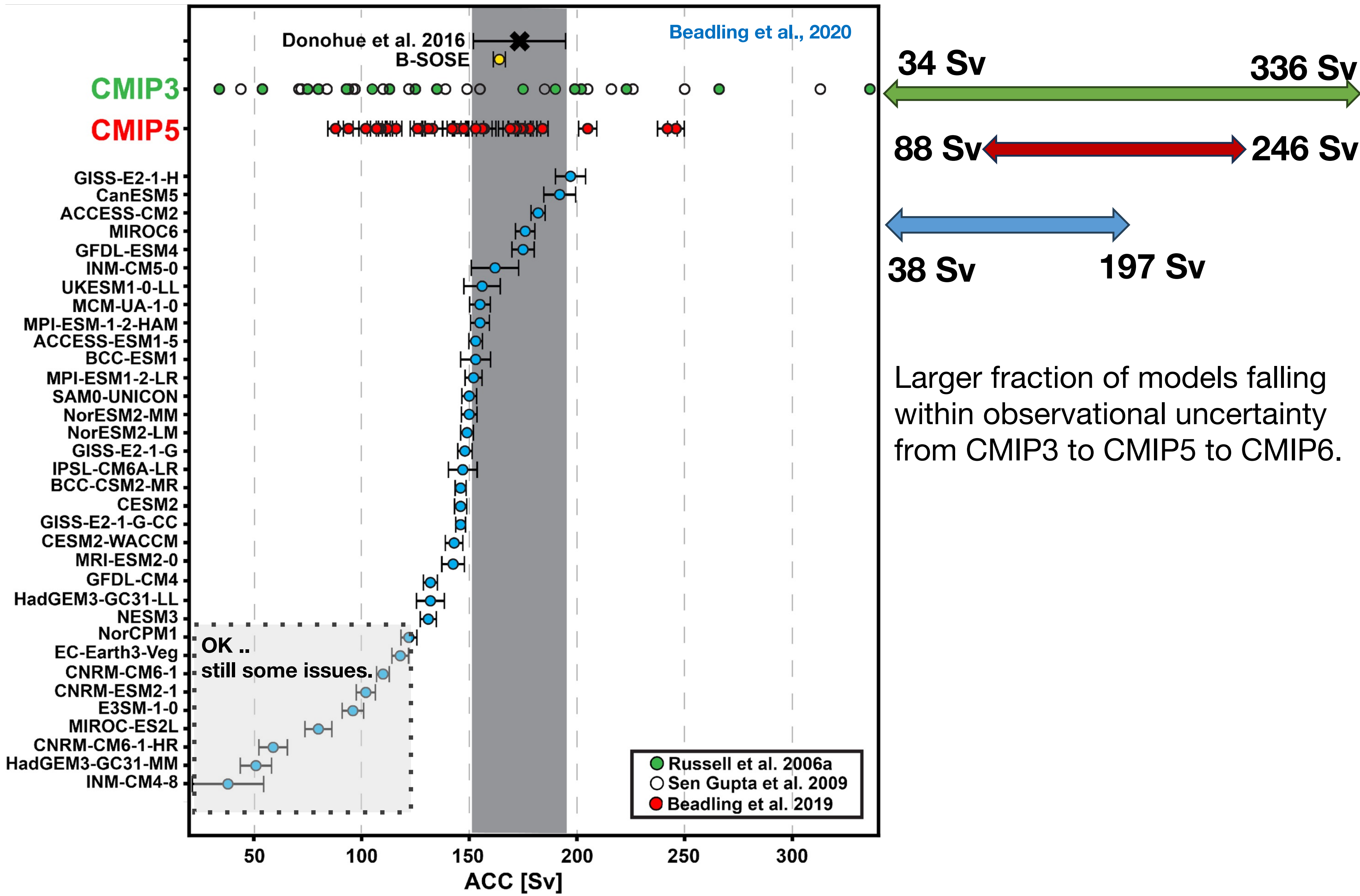
Why do we care about the ACC?

Primary conduit for inter-basin exchange

Vertical & horizontal structure intimately tied to the transport of heat, freshwater, nutrients, and carbon between the subpolar Southern Ocean and mid-latitudes and from the abyss to the surface

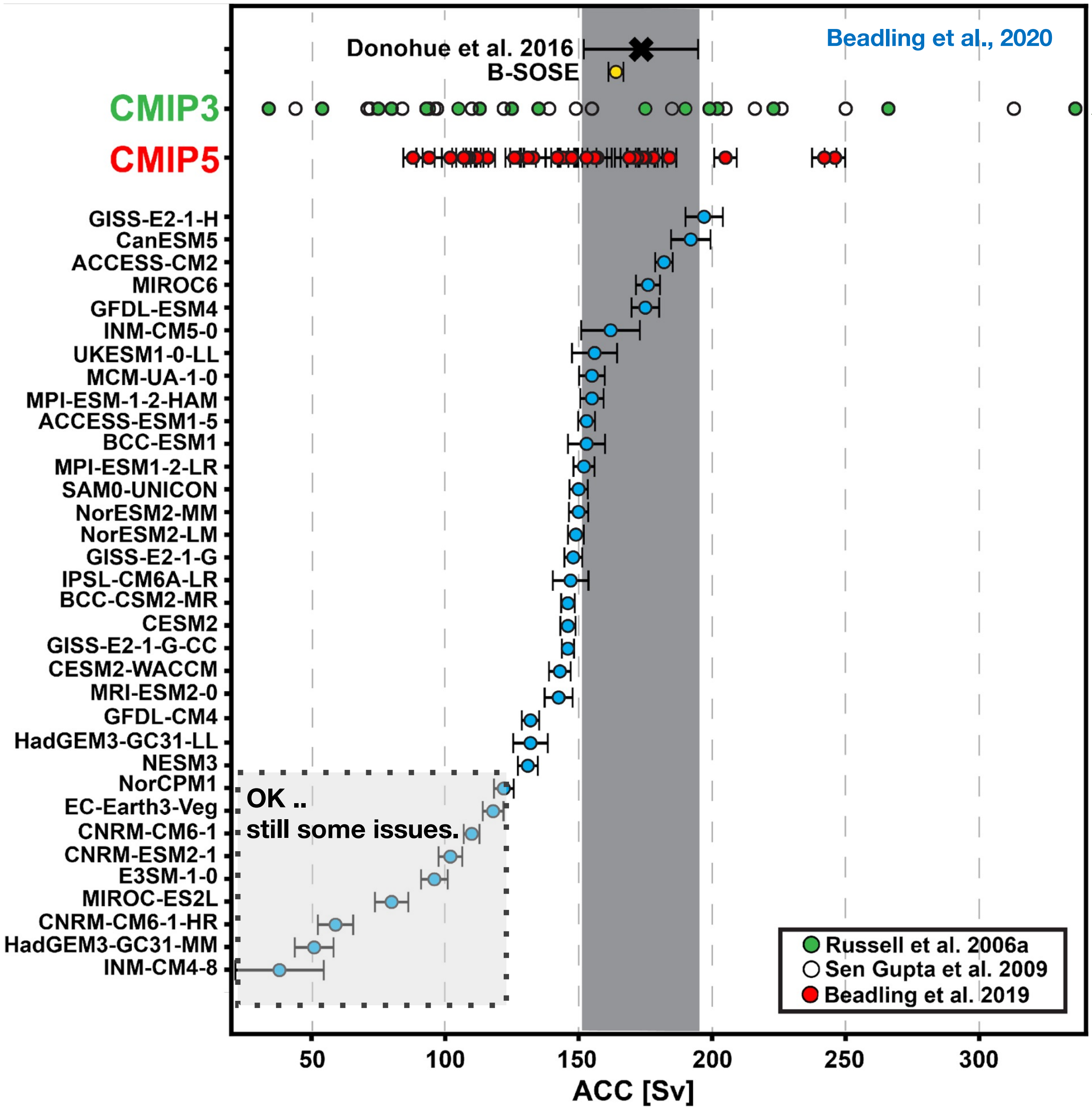
An emergent feature of the complex dynamics of the Southern Ocean … it is a good *first-pass* metric to look at to assess model performance.

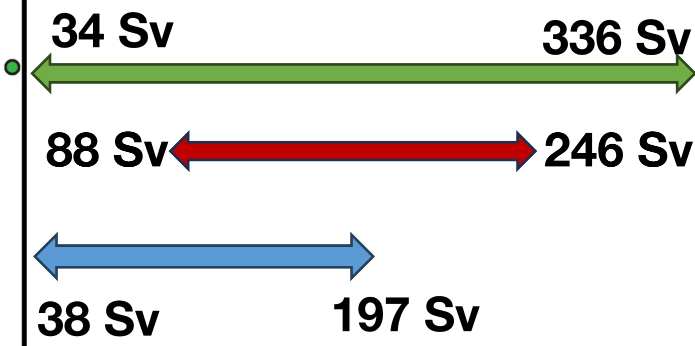
Multiple generations of climate models have **struggled to accurately capture** the *total transport* of the ACC through the Drake Passage ….



Larger fraction of models falling within observational uncertainty from CMIP3 to CMIP5 to CMIP6.

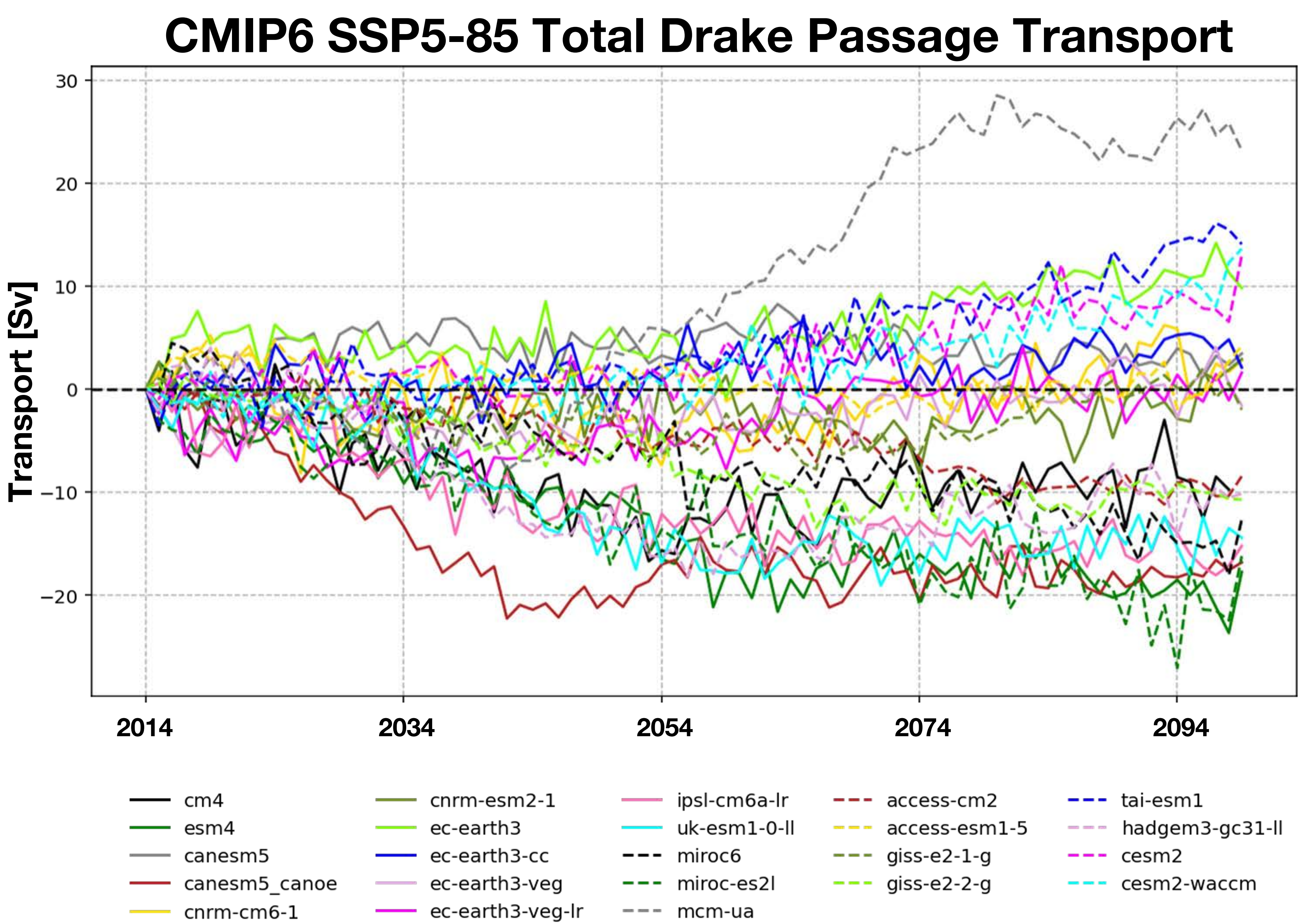
Multiple generations of climate models have **struggled to accurately capture** the *total transport* of the ACC through the Drake Passage ….



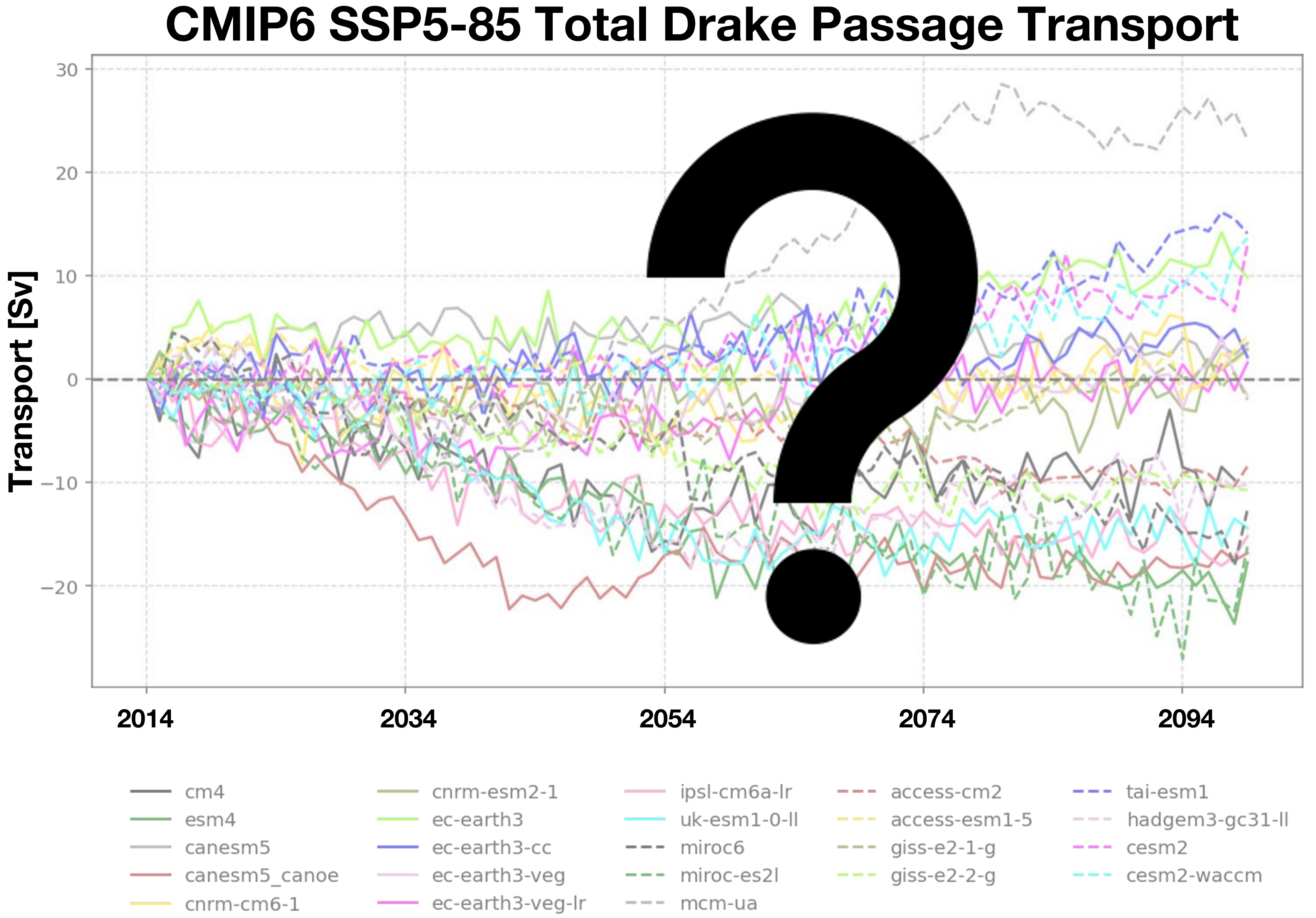


Larger fraction of models falling within observational uncertainty from CMIP3 to CMIP5 to CMIP6.

But … are they actually getting this total transport "right" for the right reasons?



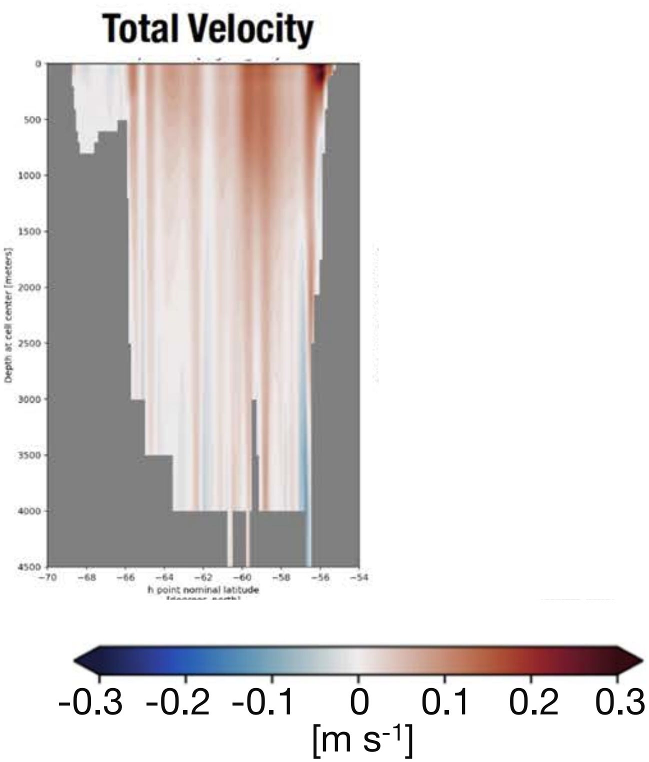
CMIP6 SSP5-85 Total Drake Passage Transport
Transport [Sv]
30
20
10
0
−10
−20
2014
2034
2054
2074
2094
cm4
esm4
canesm5
canesm5_canoe
cnrm-cm6-1
cnrm-esm2-1
ec-earth3
ec-earth3-cc
ec-earth3-veg
ec-earth3-veg-lr
ipsl-cm6a-lr
uk-esm1-0-ll
miroc6
miroc-es2l
mcm-ua
access-cm2
access-esm1-5
giss-e2-1-g
giss-e2-2-g
tai-esm1
hadgem3-gc31-ll
cesm2
cesm2-waccm



CMIP6 SSP5-85 Total Drake Passage Transport
Transport [Sv]
30
20
10
0
−10
−20
2014
2034
2054
2074
2094
?
cm4
esm4
canesm5
canesm5_canoe
cnrm-cm6-1
cnrm-esm2-1
ec-earth3
ec-earth3-cc
ec-earth3-veg
ec-earth3-veg-lr
ipsl-cm6a-lr
uk-esm1-0-ll
miroc6
miroc-es2l
mcm-ua
access-cm2
access-esm1-5
giss-e2-1-g
giss-e2-2-g
tai-esm1
hadgem3-gc31-ll
cesm2
cesm2-waccm

What insight on model performance and spread across CMIP6 models can we gain from **moving beyond the total** and decomposing the total ACC transport through the Drake Passage into various flow components?

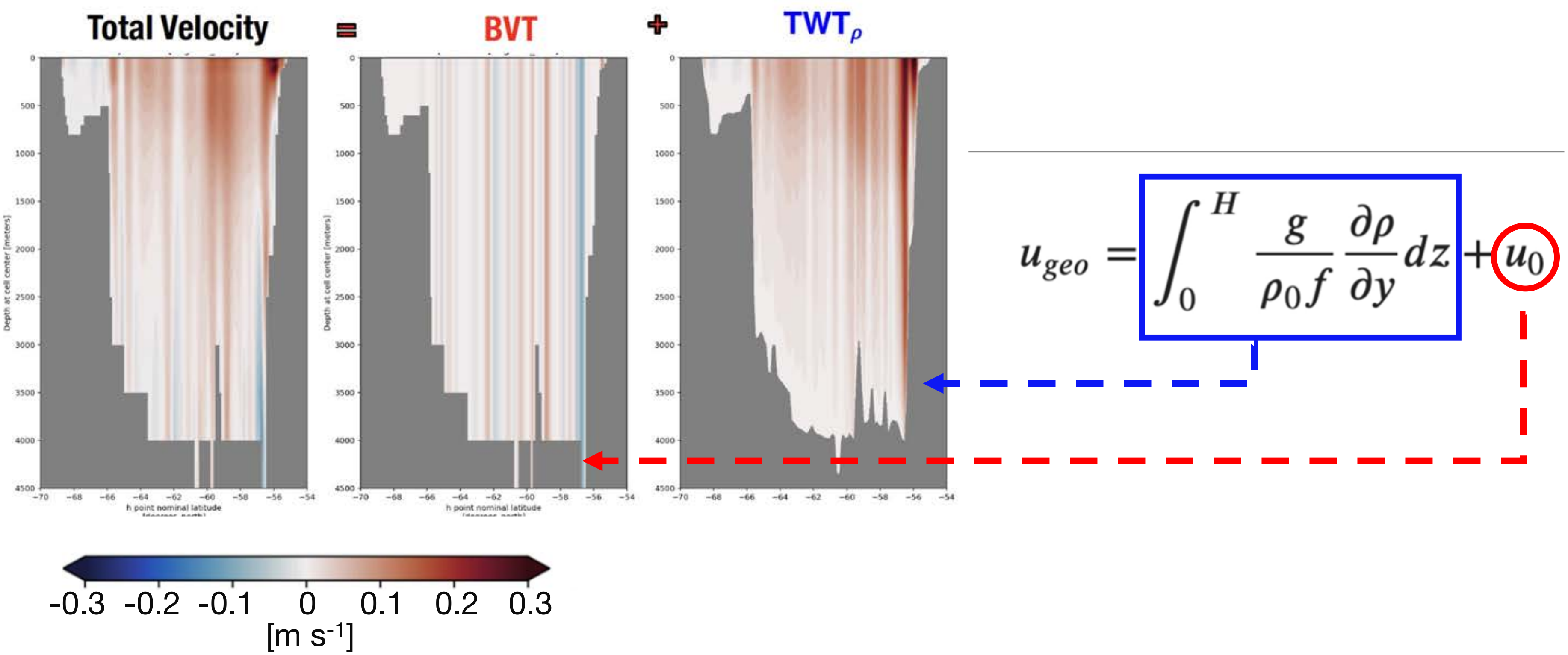
GFDL - CM4 piControl



Total velocity field

What insight on model performance and spread across CMIP6 models can we gain from **moving beyond the total** and decomposing the total ACC transport through the Drake Passage into various flow components?

GFDL - CM4 piControl



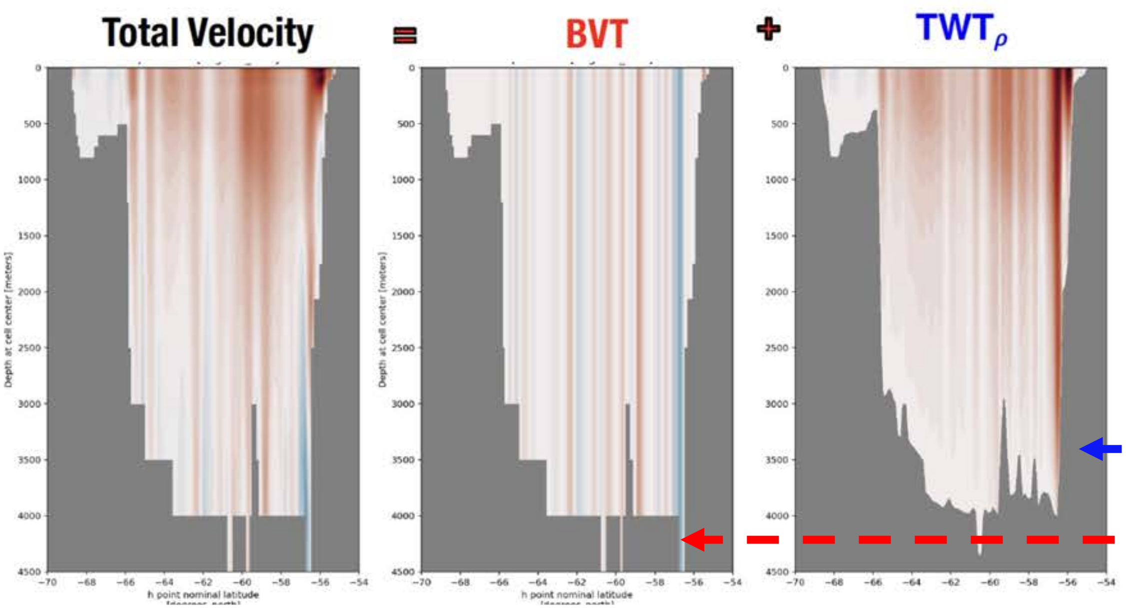
Total velocity field

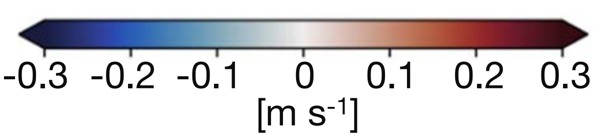
Bottom velocity transport

Thermal wind transport

What insight on model performance and spread across CMIP6 models can we gain from **moving beyond the total** and decomposing the total ACC transport through the Drake Passage into various flow components?

GFDL - CM4 piControl





$$u_{geo} = \int_0^H \frac{g}{\rho_0 f}\frac{\partial \rho}{\partial y}dz + u_0$$

$$\frac{\partial \rho}{\partial y} = \rho_0 * \left(-\alpha * \frac{\partial \theta}{\partial y} + \beta * \frac{\partial S}{\partial y}\right)$$

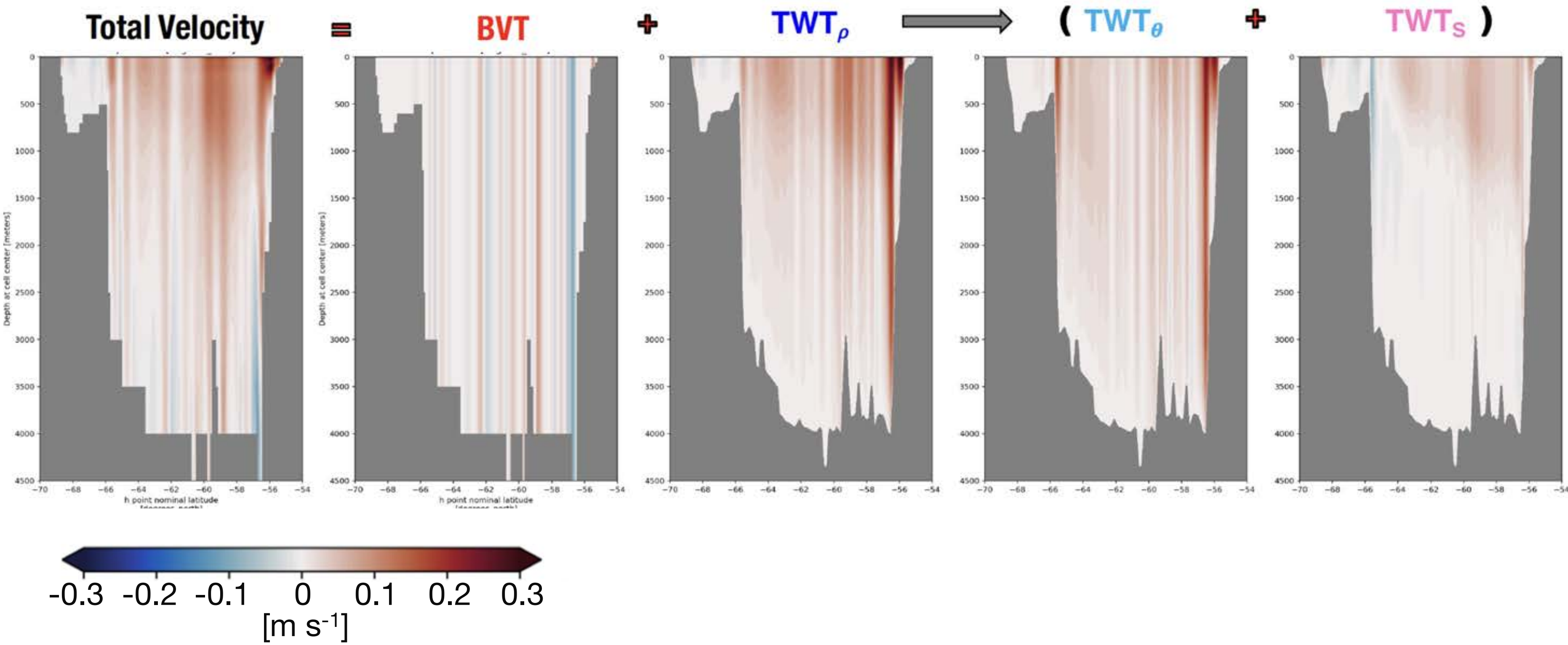
Total velocity field

Bottom velocity transport

Thermal wind transport

What insight on model performance and spread across CMIP6 models can we gain from **moving beyond the total** and decomposing the total ACC transport through the Drake Passage into various flow components?

GFDL - CM4 piControl



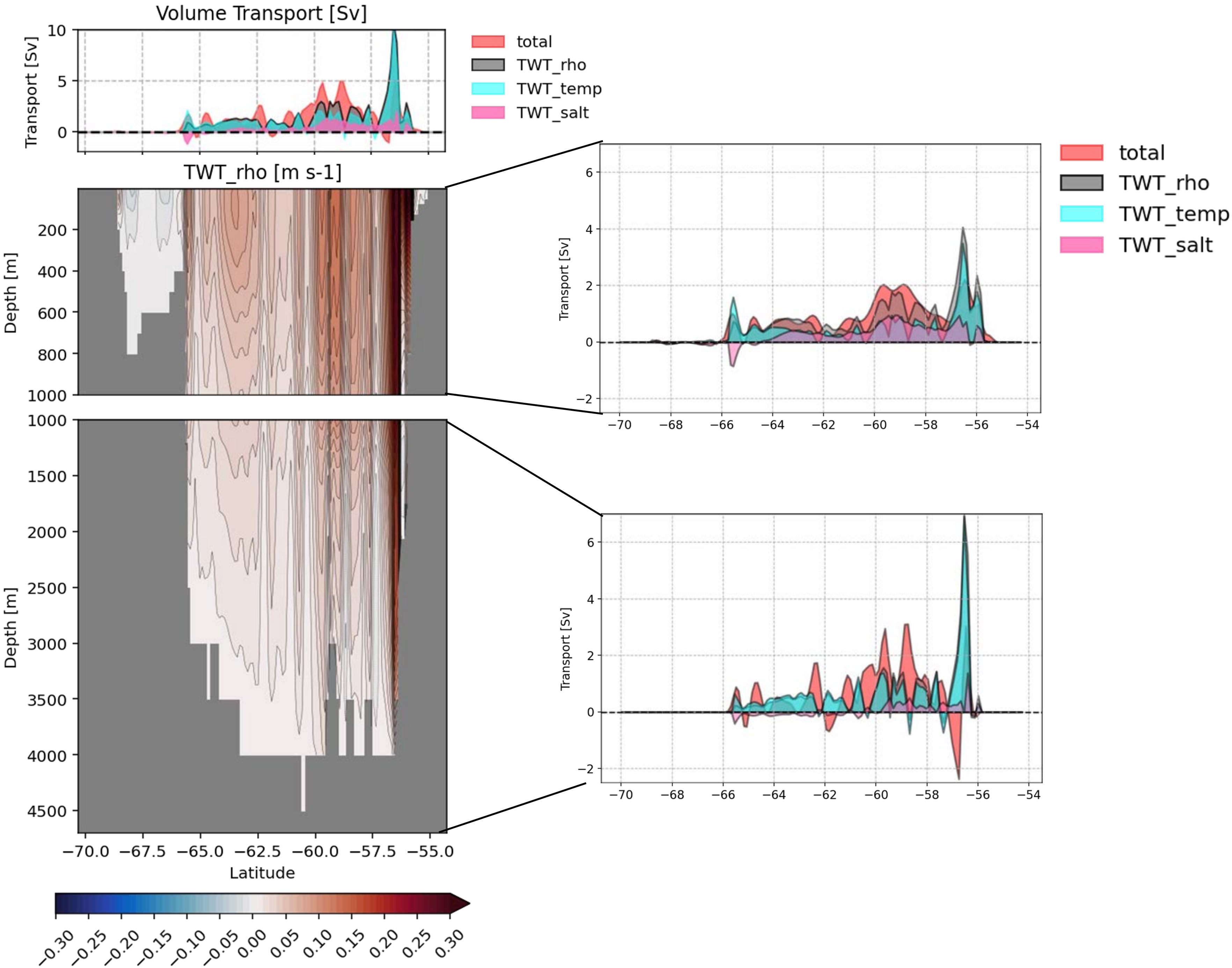
Total velocity field

Bottom velocity transport

Thermal wind transport

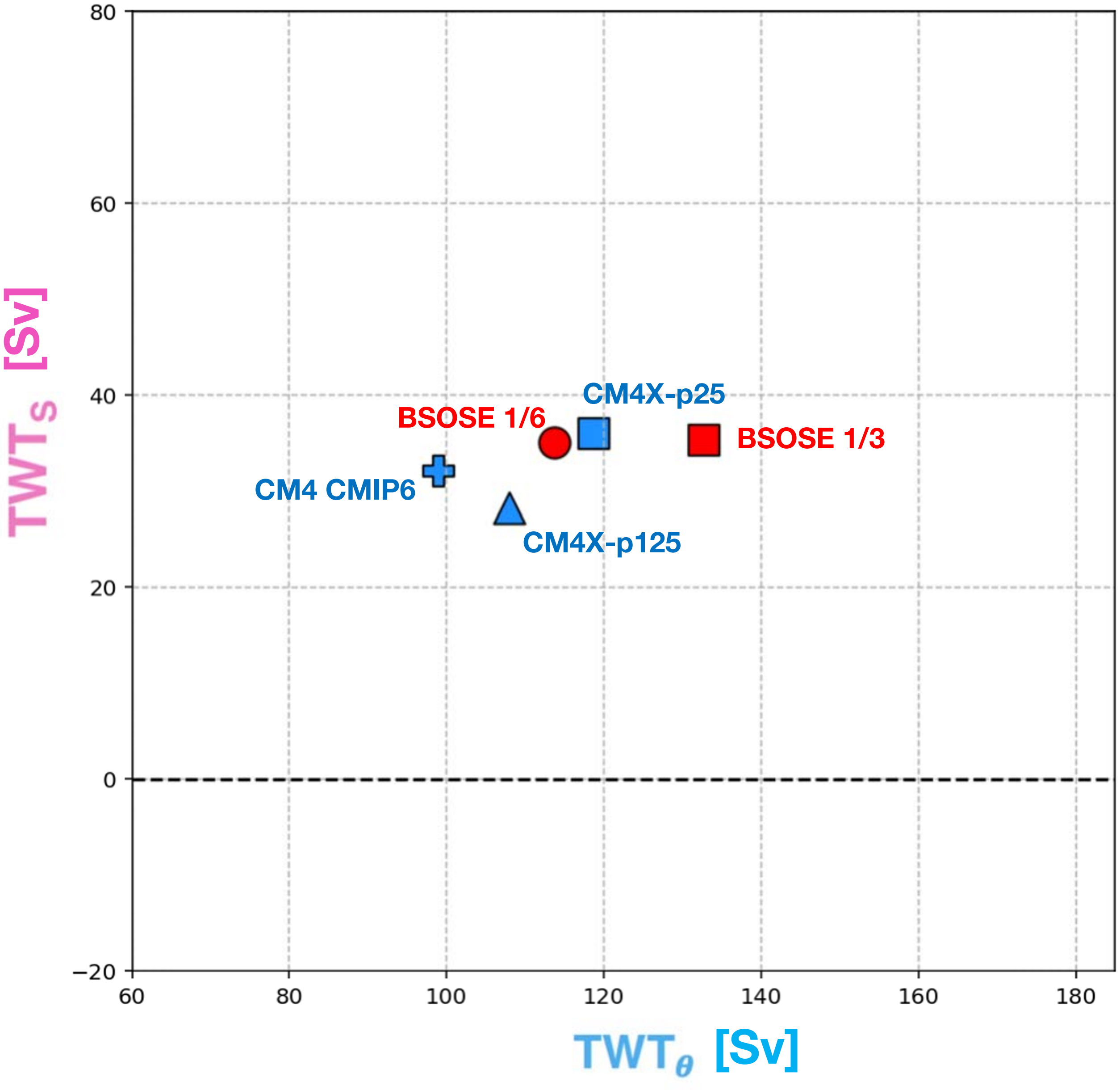
Thermal wind transport temperature contribution

Thermal wind transport salinity contribution



Volume Transport [Sv]
Transport [Sv]
total
TWT_rho
TWT_temp
TWT_salt
TWT_rho [m s-1]
Depth [m]
Latitude
Transport [Sv]

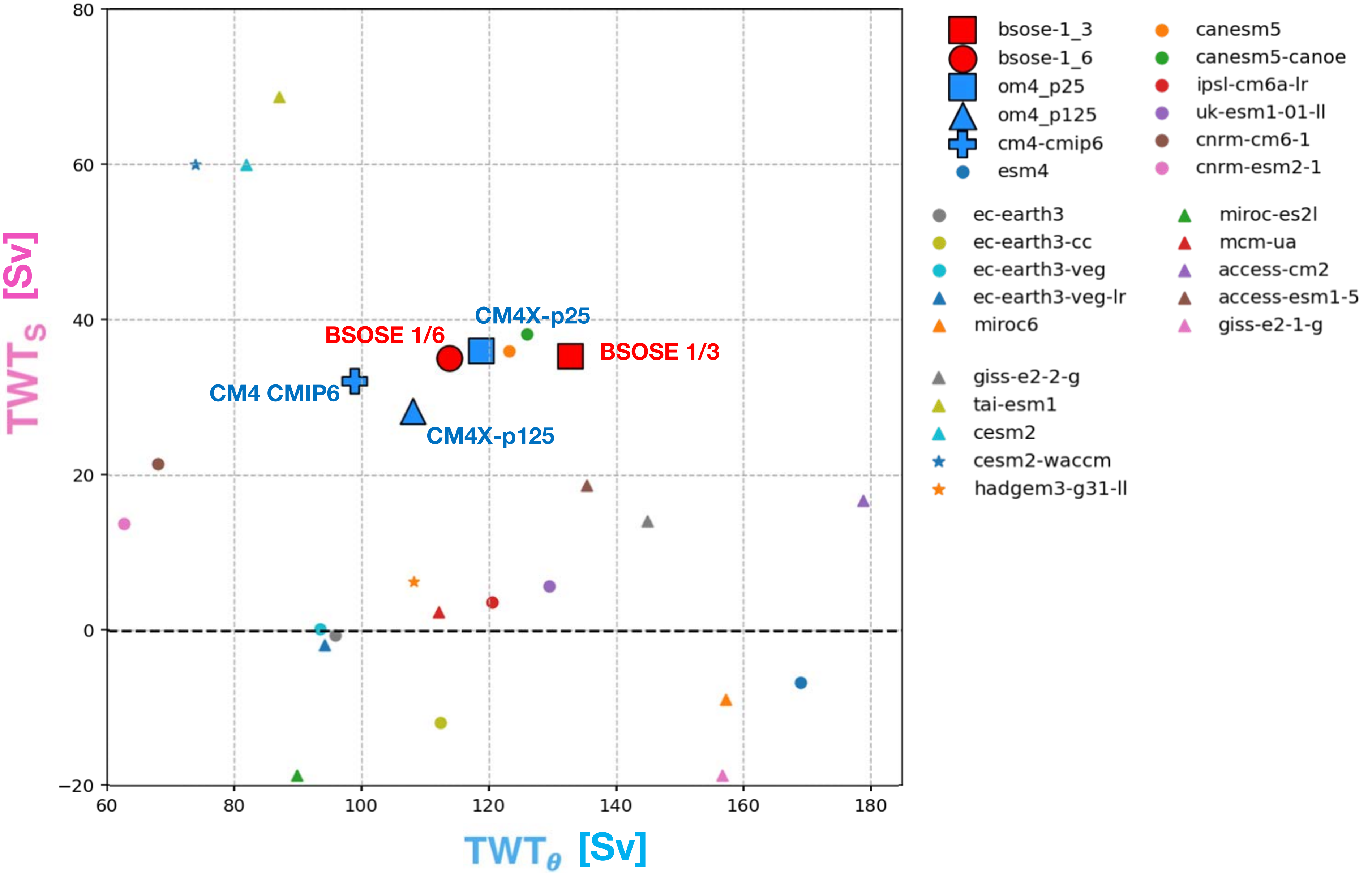
ACC breakdown in CM4X-p25/p125 development



Both configurations diverge from CMIP6 in their temperature contribution:

CM4X-p25 20 Sv > CM4 CMIP6

Very large spread across CMIP6 models in the contribution from T and S to thermal wind transport



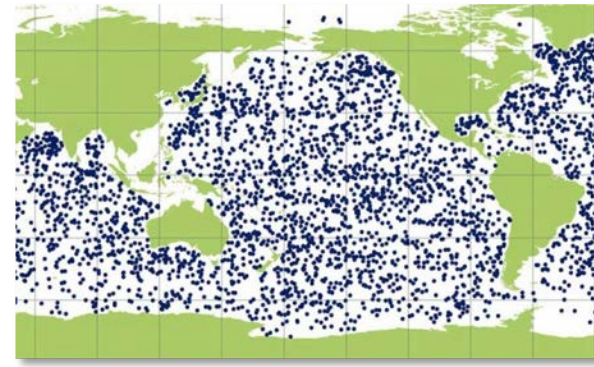
key points (ACC)

- Decomposing the total ACC transport can provide further insight into understanding model performance and spread.
- CMIP6 models show a very large spread in the transport contributions from temperature and salinity.
- The CM4X configurations show clear differences in ACC transport strength and variability relative to CMIP6 – mostly linked to a stronger transport associated with meridional temperature gradients.
- Understanding this spread in mean-state representation and how the individual components are projected to evolve may allow us to constrain our understanding of future ACC transport.

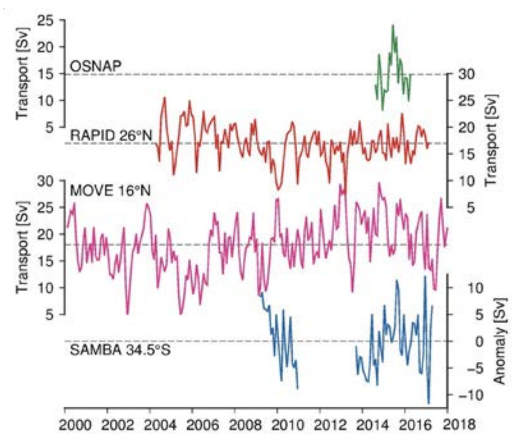
Overview:
Challenges & advances in diagnostic capabilities for ocean processes

Challenges in diagnostic capabilities for ocean processes

A relatively (on climate timescales) short and imperfect observational record.



Argo (top 2000 m) since ~ 20 years
Deep Argo (to 4000 m) since < 10 years
Biogeochemical Argo (+ SOCCOM) < 10 years
World Ocean Atlas ~ 30 years (sparse prior to Argo!)



AMOC =< 20 years

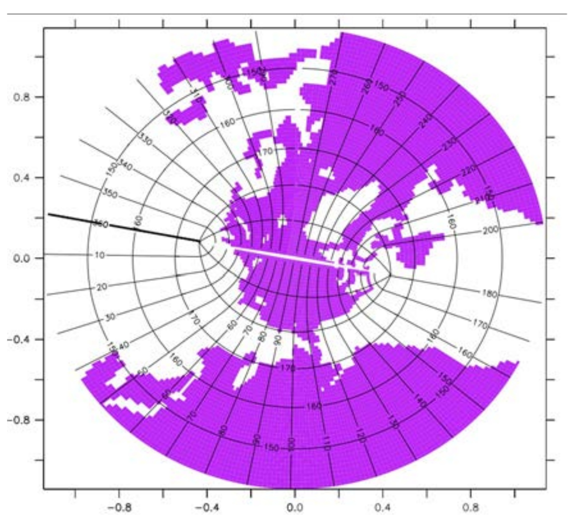
Frajka-Williams et al., 2019

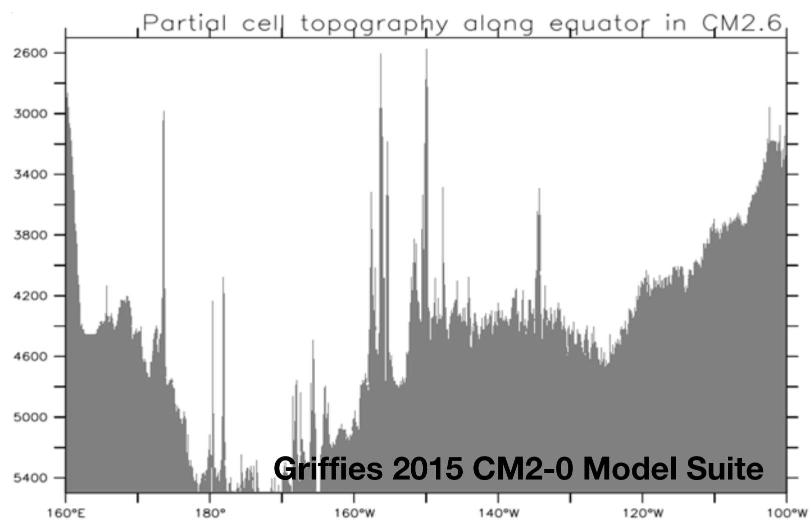
Polar processes temporally & spatially sparse due to observational challenges.

Challenges in diagnostic capabilities for ocean processes

Challenges from a usability standpoint:

- complex horizontal grids
- varying vertical coordinates (layered, z-star, sigma, rho, hybrid)
- partial grid cells
- high computational cost with size of model output
- Many calculations must be done on native grids
- Crucial need to consider model drift in evaluation (requires piControl)
- Need for communication and coordination to ensure appropriate diagnostics are saved and available.





Griffies 2015 CM2-0 Model Suite

Software advances to meet usability challenges
(analysis & education / training)

xarray

dask

GCM

momlevel

python™

jupyter

Intake-ESM

Climatematch Academy

PANGEO
A community platform for Big Data geoscience

Iris

Moving Forward

Advanced and continued coordination between those developing diagnostics & modeling centers.

- Ensure availability of variables at appropriate time frequencies for processes of interest.
- Many diagnostic capabilities exist "in-house" at modeling centers and are actively used for development … need to avoid re-inventing the wheel and make these open and interoperable with other models / in flexible open-source languages.

Moving Forward

Advanced and continued coordination between those developing diagnostics & modeling centers.

- Ensure availability of variables at appropriate time frequencies for processes of interest.
- Many diagnostic capabilities exist “in-house” at modeling centers and are actively used for development … need to avoid re-inventing the wheel and make these open and interoperable with other models / in flexible open-source languages.

Continued development of standards and best-practices for diagnostic development.

Moving Forward

Advanced and continued coordination between those developing diagnostics & modeling centers.

- Ensure availability of variables at appropriate time frequencies for processes of interest.
- Many diagnostic capabilities exist “in-house” at modeling centers and are actively used for development … need to avoid re-inventing the wheel and make these open and interoperable with other models / in flexible open-source languages.

Continued development of standards and best-practices for diagnostic development.

Continued development on software side to handle increasing size of observational datasets and high-resolution ocean model output.

Thank you

Special acknowledgement to fellow CMIP7 Model Benchmarking Task Team members whom have been in active discussions regarding available diagnostic packages and capabilities, and to Dr. John Krasting at NOAA's Geophysical Fluid Dynamics Laboratory for discussions on the efforts of NOAA's Model Diagnostic Task Force and diagnostic development best practices.

Climate Model Benchmarking members

Birgit Hassler	**2022–**	**Co-lead**	**DLR**	**Germany**
Forrest Hoffman	**2022–**	**Co-lead**	**ORNL**	**USA**
Rebecca Beadling	2022–	Member	Temple University	USA
Ed Blockley	2022–	Member	UK Met Office	UK
Jiwoo Lee	2022–	Member	PCMDI/LLNL	USA
Valerio Lembo	2022–	Member	ISAC	Italy
Jared Lewis	2022–	Member	Climate Resource Pty Ltd	Australia
Jianhua Lu	2022–	Member	SYSU & SML	China
Luke Madaus	2022–	Member	Jupiter Intelligence, Inc.	USA
Elizaveta Malinina	2022–	Member	Environment Canada	Canada
Brian Medeiros	2022–	Member	NCAR	USA
Wilfried Pokam Mba	2022–	Member	University of Yaoundé I	Cameroon
Enrico Scoccimarro	2022–	Member	CMCC Foundation	Italy
Ranjini Swaminathan	2022–	Member	University of Reading	UK

rebecca.beadling@temple.edu