Breakout Group 3: CMIP Evaluation Tools
Its Not a Beauty Contest….

Key Scientific Advances Since AR5

- We have maturing publically available open source CMIP evaluation tools that are currently being tested by reproducing CMIP5 evaluation results
- A subset of them will evaluate models as the output is published to ESGF for the DECK and CMIP6 historical simulations to provide early feedback to the modeling groups and wider community
- More community acceptance of the value of these open source evaluation tools and performance metrics and the value of standardized performance metrics.
- Quality controlled observational datasets now aligned with CMIP data structures (e.g. obs4mips, ana4mips) that have been developed through international collaboration for use in CMIP evaluation. This archive is growing.
- Traceability and provenance of models, evaluation codes, and datasets is increasingly appreciated and adopted.

Shortcomings: Gaps and Opportunities

- CMIP-endorsed MIPs provide analysis suites to be integrated into CMIP evaluation tools
- A namelist should be incorporated into the CMIP evaluation suite associated with each MIP.
- Examples being set now regarding traceability and provenance of models should be expanded to other codes
- There are gaps in observational data and ways in using existing observational data and domain gaps in model evaluation that require enhanced collaboration and funding opportunities across modeling and observational communities to address
- Scientific studies on key evaluation metrics and diagnostics should be reproducible and the algorithms submitted as part of the publication process, adding uptake of these diagnostics into the evaluation suite
- Systematic comparison of results from existing tools
- Community-based process needed to establish a tiered list of critical metrics across Earth system components
- More widespread dissemination of process-oriented diagnostics from the peer reviewed literature into the public domain
- International efforts have started to define process-oriented metrics and frameworks that are beginning to be disseminated in CMIP evaluation (e.g. WMO MJO Task Force, NOAA MAPP Model Diagnostics TF)

Long-term Perspective

- Increase transparency to accelerate scientific understanding of climate
- We need to stay ahead of an accelerating global change with global flat funding and these tools will help
- Improve the accessibility and uptake of the CMIP model evaluation tools within model development teams
- Lower the barriers to access and broaden the community using the the CMIP evaluation tools and model (e.g. observational and impact communities, policy makers, public)
- Latch onto technological advances that make model diagnosis more efficient
- Systematically grow the standard for model evaluation to further accelerate advance of earth system understanding