

Introduction

- Model evaluation has always been part of model development- what we're proposing is continuing expanding and improving that process (don't want to trash what's already been done)
- Individual modelling groups have always evaluated their models but with the advent of CMIP multimodel evaluation to inform development and projections has become increasingly useful.
- Multimodel evaluation has only been done episodically for the phases of CMIP and IPCC assessment. The idea here is to move to a standard on going multimodel evaluation process
- Why do want to do this? The expansion of the CMIP ensemble and greater associated model dependence means as well as the improvement of the observing system (reanalyses, satellite records etc)there is ever more need to evaluate and understand the models (e.g. model documentation Tom Philips papers, ESdocs),
- Why do want to do this? Systematic biases-e.g. double ITCZ, Southern Ocean bias (refs?) & accounting for more complex models that increase model uncertainty, allows everyone to have a fast quick look across variables
- Our suggestions leverage the efforts of different model groups by standardising evaluation metrics across MMEs –this is a community effort that won't be solved in isolation, with more metrics and complex diagnostics there's a greater need for a such a systematic approach

Cont.

- Making use of more and more observations and now even evaluating model projections against present day Obs (hindcasts)
- The concept model democracy? The multimodel mean has historically outperformed any individual model (Gleckler papers) does this. The recognition of model interdependence represents a shift away from biased multimodel mean estimates
- There are no good and bad models- all models have value
- In AR5 performance based weighting schemes were tested, there are numerous issues recognised with this (e.g. ensemble size Flato et al. etc, model interdependence Sanderson/Abramowitz papers) progressing le
- Improving on past CMIP evaluation (model weighting, emergent constraints etc)
- Using variability to estimate sensitivity (Hasleman papers)- emergent constraints could be viewed as an extension of this...intro to new physical ECs (Hall, Sherwood papers) and biogeochemical ECs(Cox, Wenzel, Kwiatkowski)