Diagnostic and benchmarking of CMIP models

Veronika Eyring (DLR)

WGCM 17 session

October 1-3 2013

Victoria, Canada

DLR

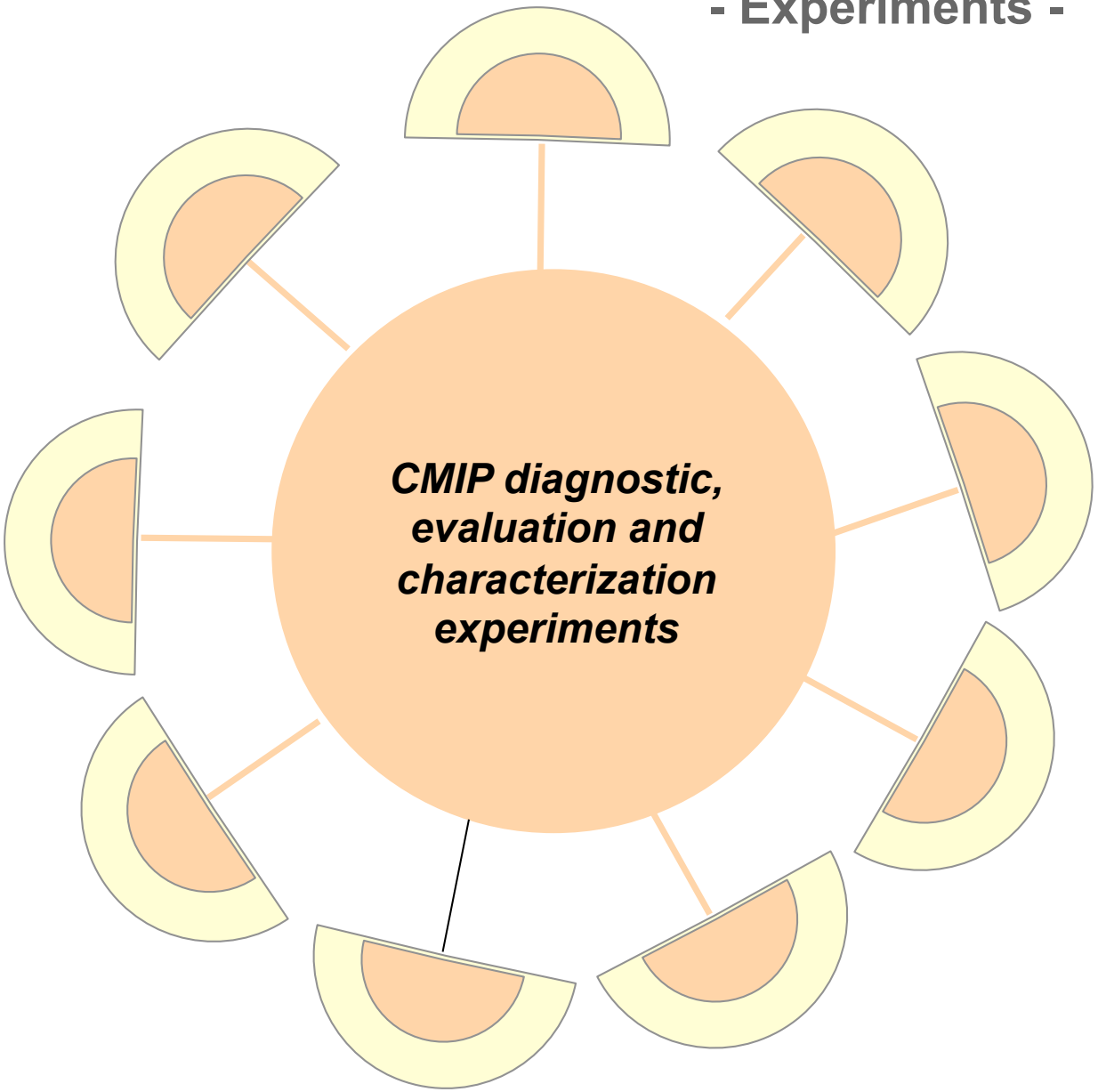
CMIP5 Survey Scientific Gaps : Model evaluation, performance metrics

In many responses the need for a **more routine evaluation of the CMIP model simulations** was pointed out**.**

- **Need to ensure model evaluation occurs before the runs for IPCC are undertaken.**
- **Detailed and systematic model evaluation** during the development process could be facilitated by CMIP.
- **Obs4MIPs** seen as very positive to improve regular model-obs evaluation, **should be grown.**
- The continued push for **standard performance metrics**, **readily published and viewable on a central website** (also for providing guidance for impact analyses).
- **Process-oriented evaluation** to understand model biases and error compensation.
- **Centrally coordinated model assessment**
- **Code repository** (e.g. at WGNE/WGCM Climate Metrics Panel Website)
- **Standardized diagnostic and performance metrics package that runs on the ESGF**

Coordination with CMIP Panel
(experiments, variable list, etc.)

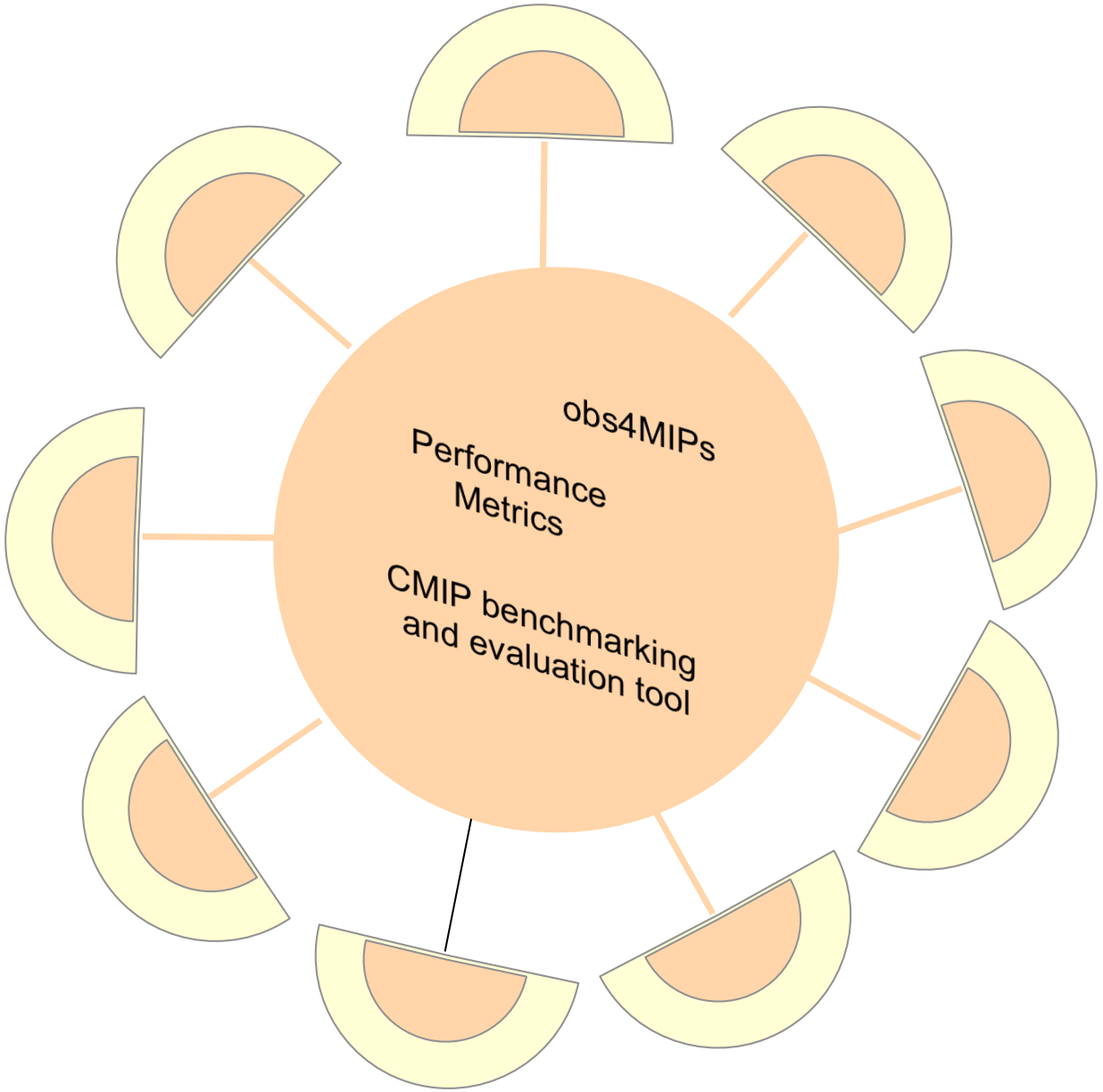
CMIP Structure
- Experiments -



There seems to be broad agreement so far that CMIP would benefit from a more distributed structure

CMIP Structure

Routine Evaluation and Benchmarking Integrated Part of CMIP6



CMIP Panel to work in close collaboration with obs4MIPs, MIPs and the WGNE/WGCM climate metrics

MIPs to provide:

- Recommendations for model diagnostics to evaluate models
- Recommendations for observations for model evaluation
- Recommendations for performance metrics
- If possible: code to be included into the CMIP benchmarking and evaluation tool that should run routinely on CMIP6 models as soon as the runs are submitted

CMIP benchmarking and evaluation tool
- from “Aspen proposal” -

Overall, there seems to be agreement in the community that having a benchmarking tool would be highly desirable.

- The CMIP Panel will work in close collaboration with obs4MIPs, the MIPs and the WGNE/WGCM climate metrics panel to develop a **CMIP benchmarking and evaluation tool** that could be run **directly on the ESGF** as soon as any model simulations are submitted.
- The MIPs are encouraged to provide
  - recommendations for model diagnostics, performance metrics, and observations for model evaluation,
  - and if possible code that could be included in the tool.
- One objective of this is to **aid the model development process** by providing feedback concerning **systematic model errors** and the relative merits of individual models.

Meehl et al., EOS, subm., 2013

DLR

CMIP Benchmarking and Evaluation Tool

… A long way still to go.

… Requires a community effort to make it happen.

DLR

Motivation

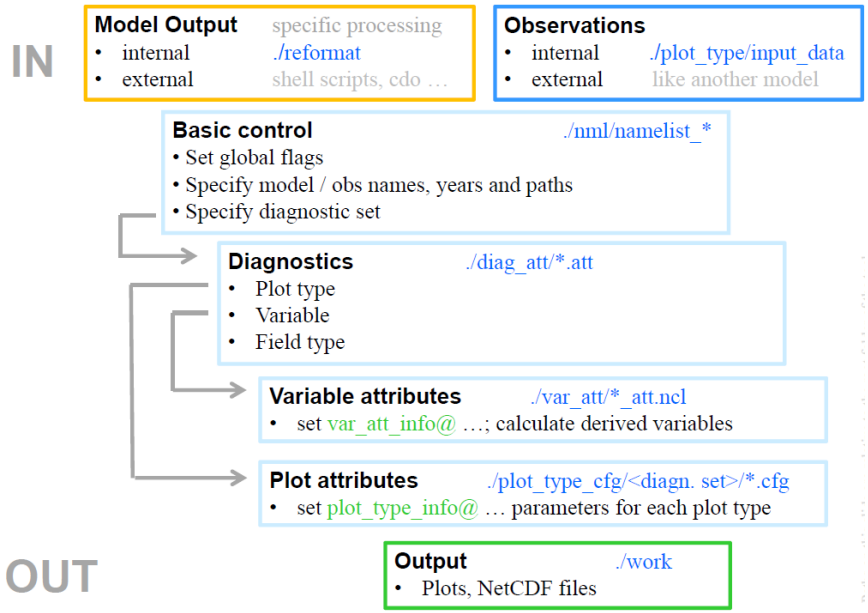
- **Facilitate the evaluation of complex Earth System Models, e.g.**
  - Allows quick looks at standard diagnostic plots & output diagnostic variables.
  - Allows to easily compare new simulations (e.g. sensitivity runs or runs with a new model versions) to existing runs and to observations.
- **Raise the standard for model evaluation**
  - Include additional diagnostics of ongoing evaluation activities so that we don't have to start from scratch each time
  - Implement more and more observations, account for uncertainty
  - Ensures progress
  - Allows to assess quickly where we stand with a new set of model simulations by developing standard namelists that reproduce specific paper, reports etc.
- **Facilitates participation in and analysis of Model Intercomparison Projects (MIPs)**
  - Allows to easily compare models participating in CMIP and CMIP Satellite MIPs.
- **Expandable and extensible**
  - Use synergies with ongoing projects to expand the tool
  - Useful for model groups & those analyzing models
  - Useful for model development

DLR

Development of an Earth System Model Evaluation Tool

Within EMBRACE: DLR, SMHI & EMBRACE partners in collaboration with NCAR, PCMDI, GFDL

- **Open Source:** **Python** Script that calls **NCL (NCAR Command Language)** and other languages (e.g. R, fortran)
- **Input:** CF compliant NetCDF model output (CMIP standards)
- **Observations:** Can be easily added
- **Extensible:** easy to (a) read models (b) process output [diagnostic] with observations and (c) use a standard plot type (e.g. lat-lon map)
- Easy to install



Current developments include

- Essential Climate Variables, e.g.
  - Sea-Ice
  - Temperatures
  - Water Vapor
  - Radiation
  - CO2
  - Ozone
- Tropical variability (incl. Monsoon, ENSO, MJO)
- Southern Ocean
- Continental dry biases and soil-hydrology-climate interactions (e.g. Standardized Precipitation Index)
- Atmospheric CO2 and NO2 budget
- More Observations (e.g., obs4MIPs, ESA CCI)
- Statistical measures of agreement

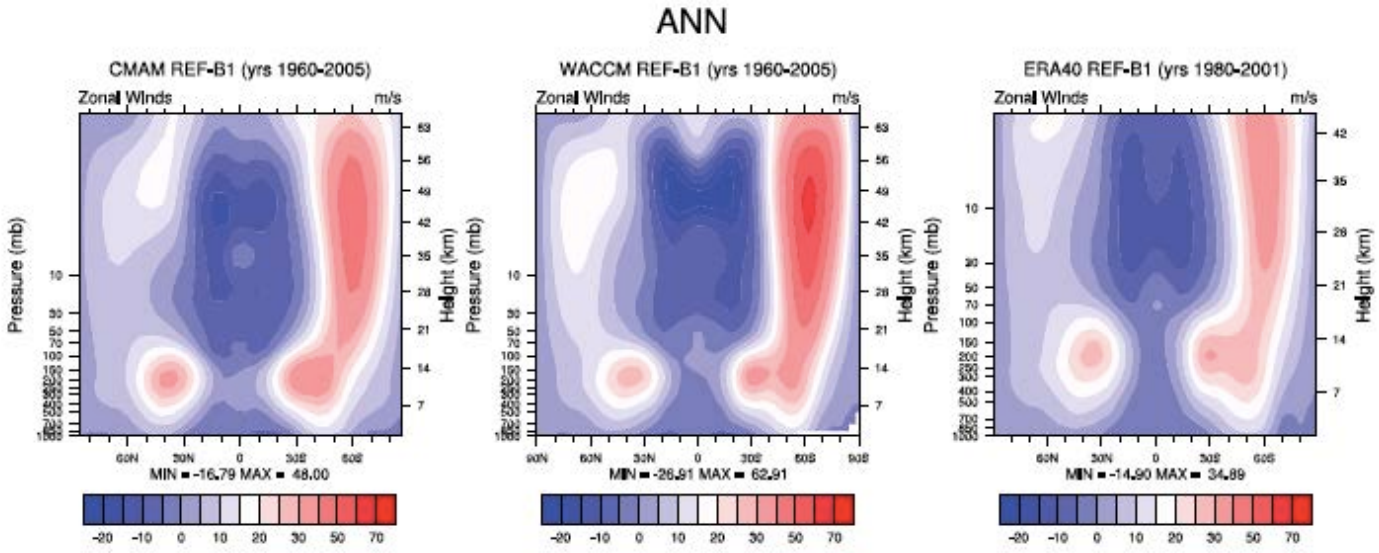
Goal: Standard namelists to reproduce certain reports or papers (e.g., Ch9 AR5, Massonnet et al., 2012; Anav et al., 2012; Cox et al., 2013; Eyring et al., 2013)

EMBRACE Consortium
SMHI
KNMI
ECWMF
ULUND
MPG
DLR
FMI
ETH
UEA
UNEXE
METUK
UCL
CNRS
NERC
WUR
USPLIT
IMPERIAL
VUAMS

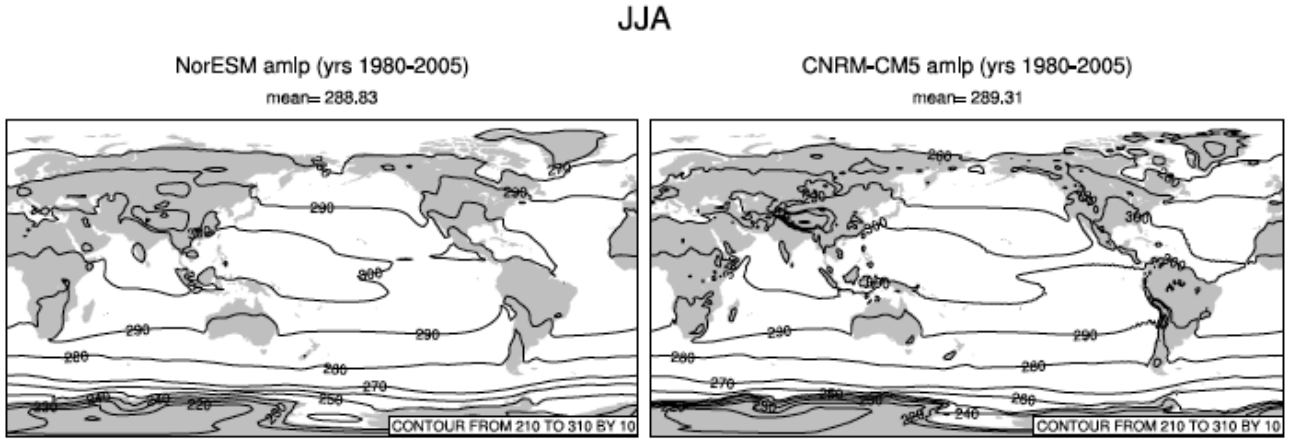
Examples of diagnostics implemented in the ESMValTool

Selected Essential Climate Variables

CCMVal-2 Models and Reanalysis: Zonal mean zonal wind cimatology



CMIP5 models: Contour plot of June–August surface air temperature (K)

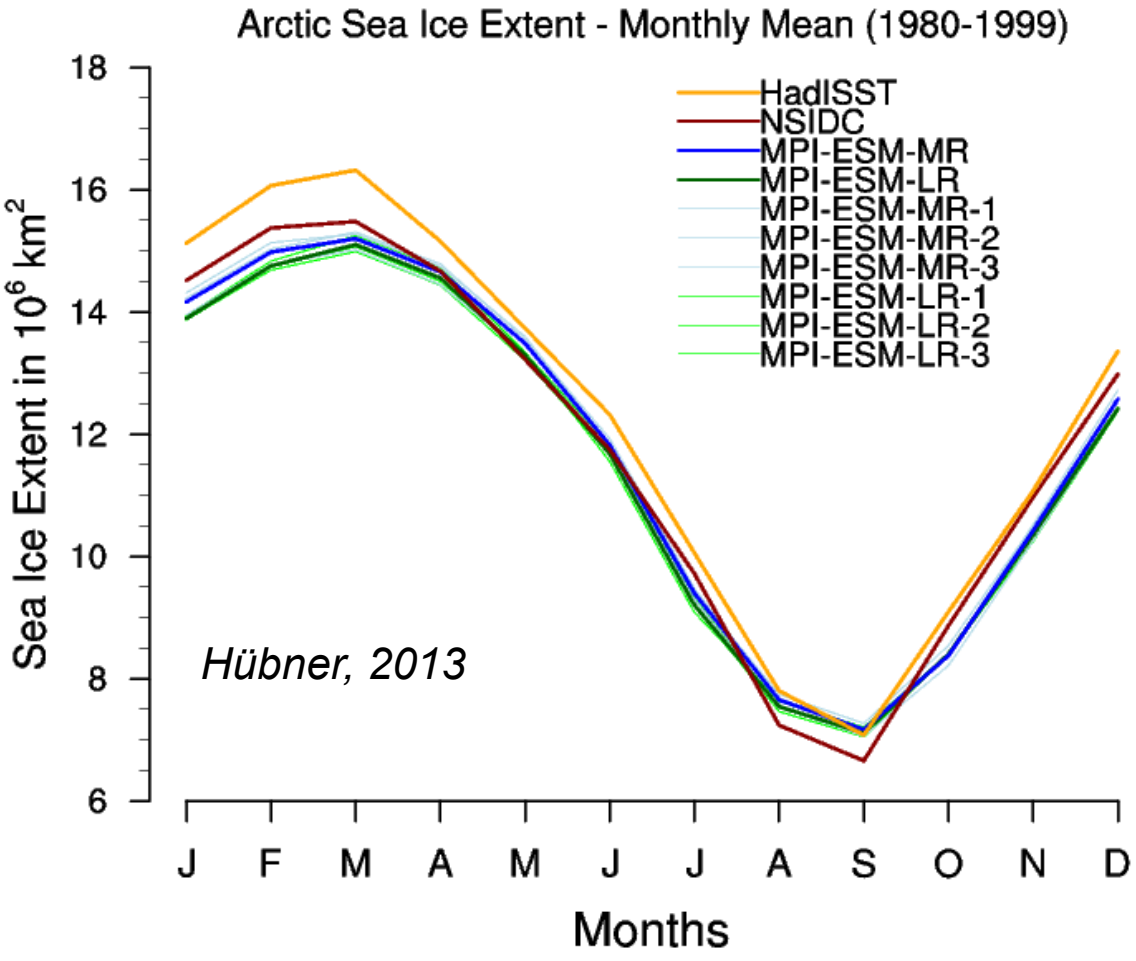


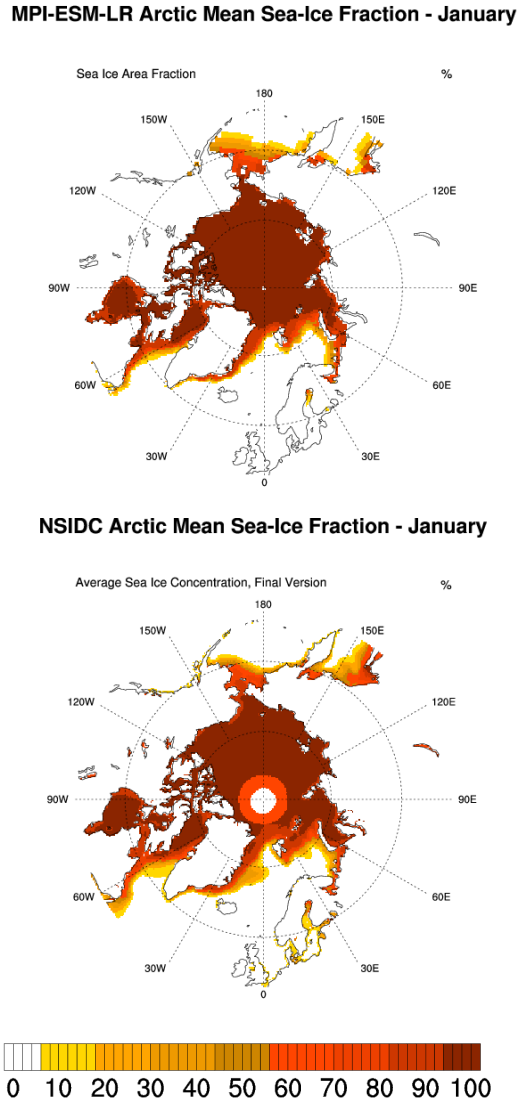
CMIP5 Models

DLR

Examples of diagnostics implemented in the ESMValTool

Sea-Ice

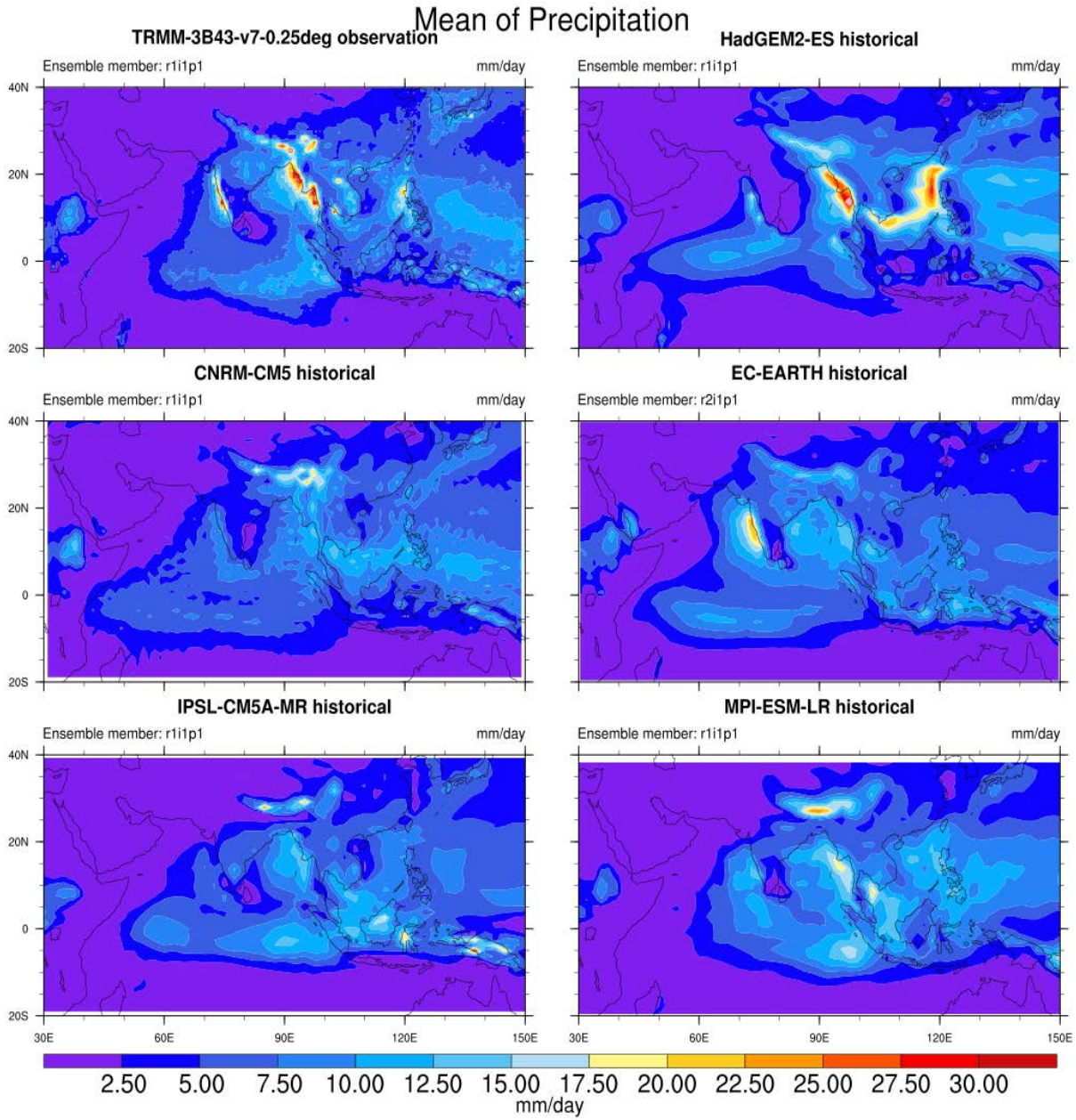




Further work: add the main sea ice variables (concentration, thickness and drift) and compare them to observations focusing on both mean state and variability, compute performance metrics (Massonnet et al., 2012)

Examples of diagnostics implemented in the ESMValTool

Tropical variability (here Monsoon, next ENSO in collaboration with E Guilyardi)



Implementation of East Asian Monsoon diagnostics into the **Earth System Model Evaluation Tool (ESMValTool)** that has been provided by the UK MetOffice.

Daily precipitation amount during the monsoon season (June-Sep) for CMIP5 models and satellite observations from TRMM.

- None of the models captures both precipitation maxima along the Indian and the Indochina west coasts.
- The increase in precipitation induced by orographic lifting across the Himalayan mountain range is reproduced by the models.

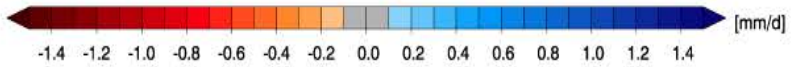
Task 4.2.4: Continental dry biases, soil hydrology-climate interactions

Evaluation of **evapotranspiration (ET)** using LandFlux-EVAL Synthesis dataset (GEWEX/ILEAPS; http://www.iac.ethz.ch/url/research/LandFlux-EVAL)

annual **CMIP5** **CMIP3**

ET

P

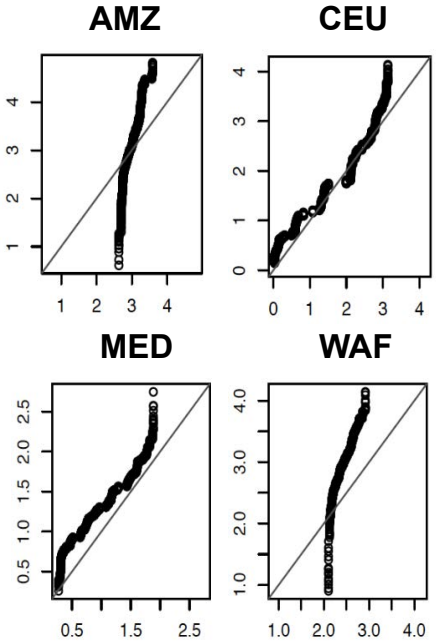


Too high ET on annual scale (too wet)
Linked to respective precipitation biases
Biases larger in CMIP5 than CMIP3

JJA **CMIP5**

Opposite in JJA:
Dry bias in many areas

Ranked monthly ET values (qq plots)

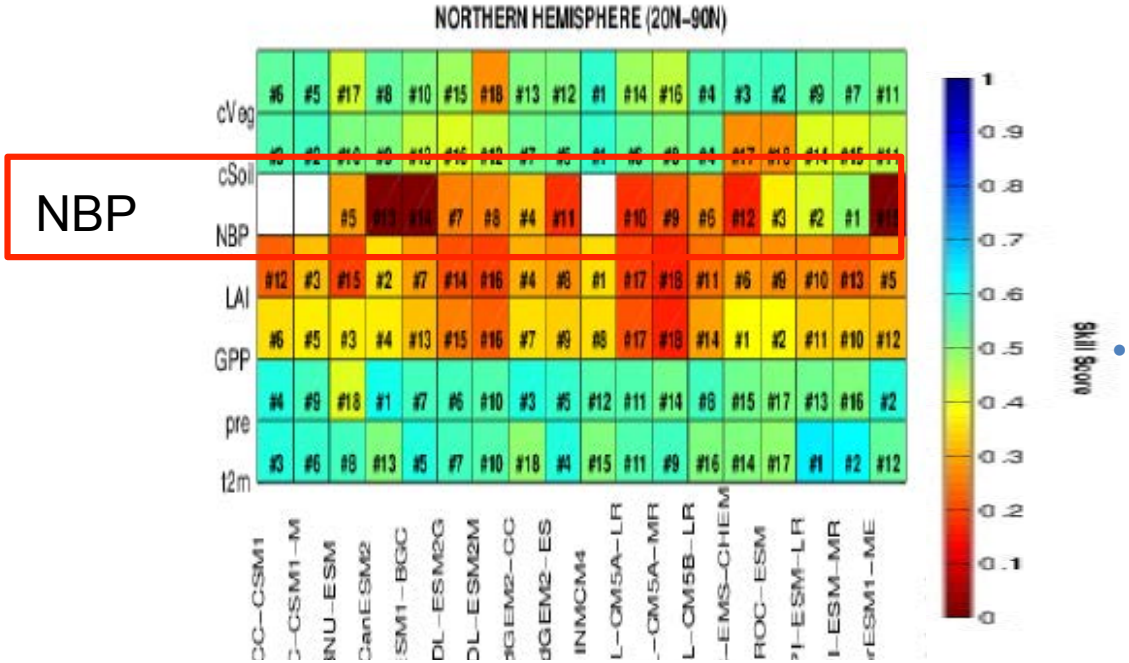


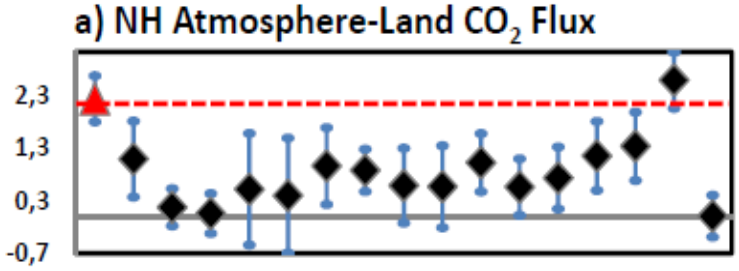
Systematic biases in pdfs: high ET↑ ; tropics low ET ↓

(Mueller and Seneviratne, in prep.)

Examples of diagnostics implemented in the ESMValTool

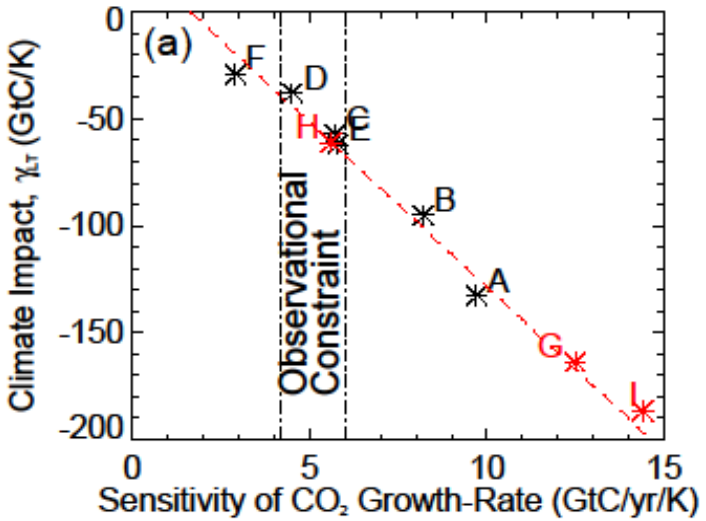
Performance Metrics, CO2 and Emergent Constraints (work in progress)

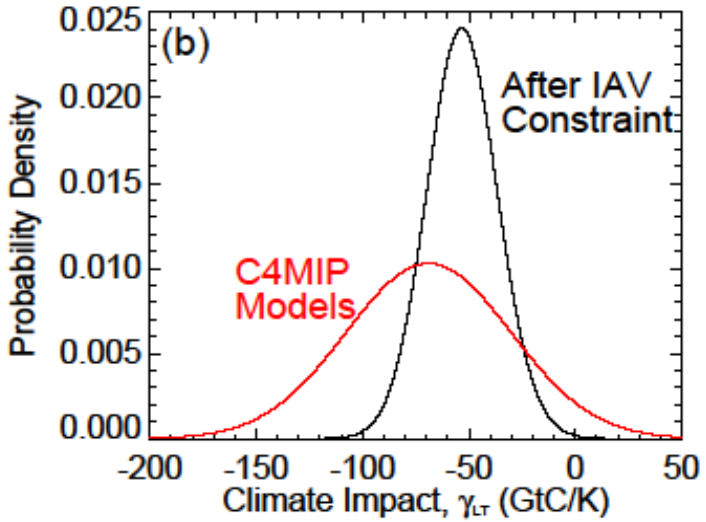




- In the NH over land almost all models systematically underestimate the sink, reflecting absence of sinks due to nitrogen deposition or poor representation of forest regrowth.

Anav et al., 2013

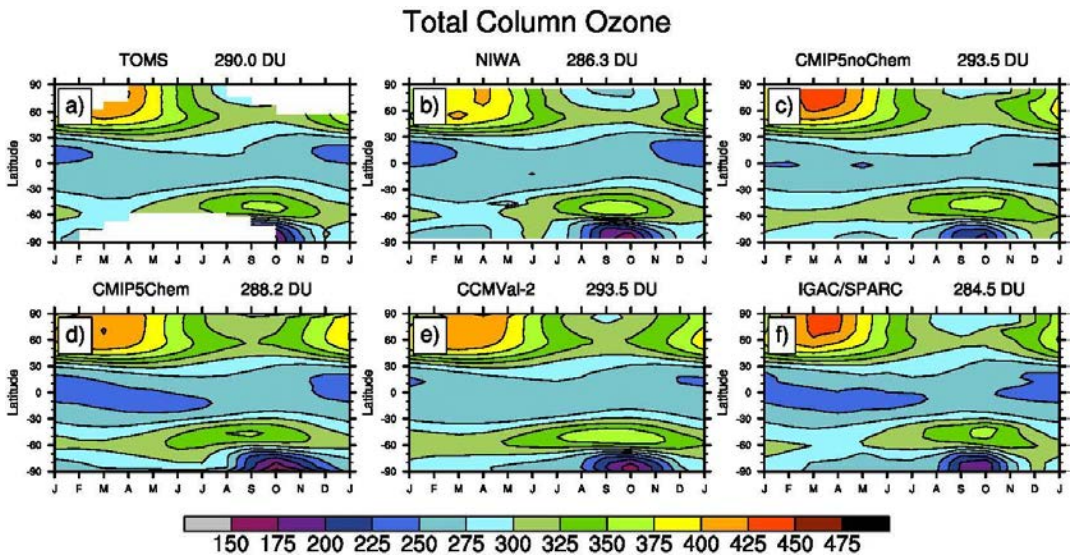




Cox et al., 2013

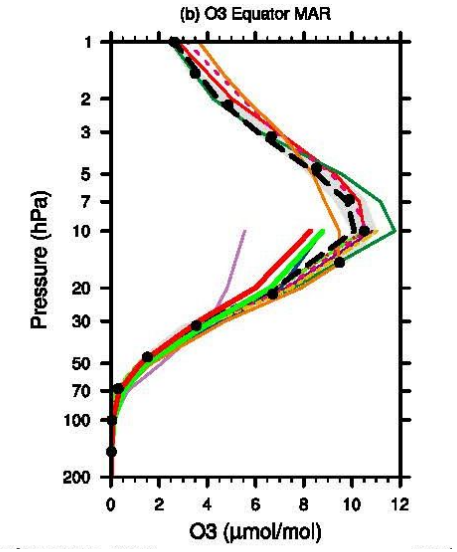
Examples of diagnostics implemented in the ESMValTool

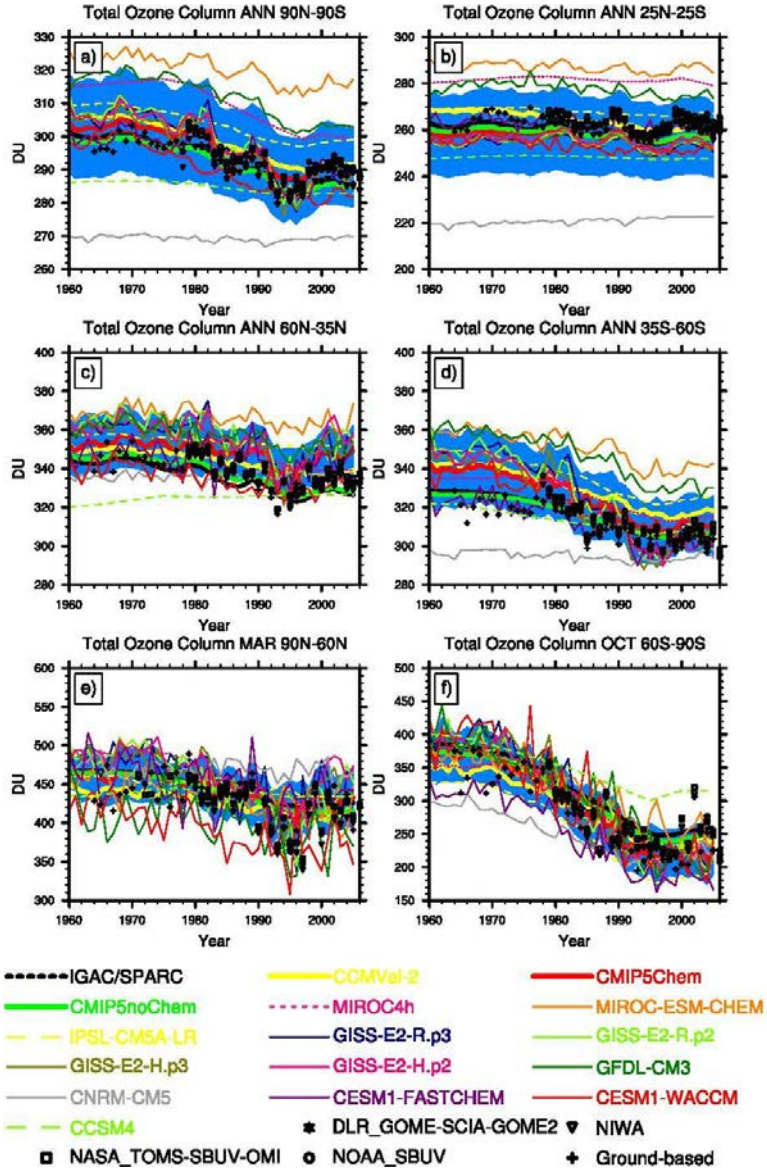
Ozone and associated climate impacts



Eyring et al., JGR, 2013

Vertically Resolved Datasets





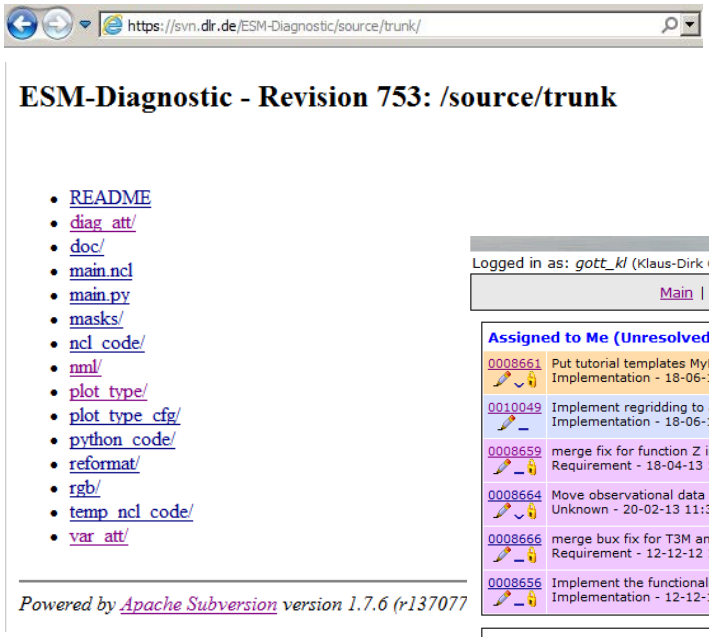
DLR

International cooperation for community development

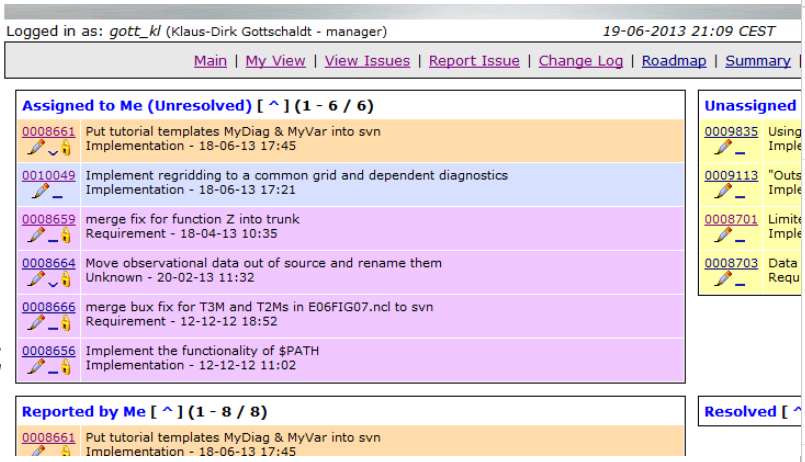
Options to contribute beyond EMBRACE

Join the core development team with full access to:

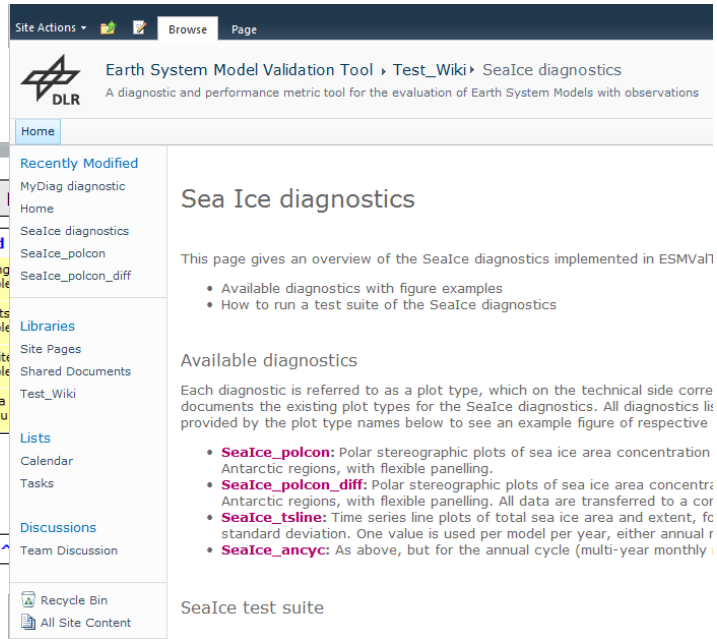
Subversion repository



Mantis bug tracker



Teamsite & Wiki



1. Implement your changes in a snapshot of the ESMValTool (tarball or checkout from repository)

2. Give us your diagnostics „as is“

Your contribution is very welcome!

Please contact us.

Summary and Outlook

- **An Evaluation Tool** is available that facilitates the complex evaluation of ESMs and their simulations submitted to international Model Intercomparison Projects (e.g., CMIP, C4MIP, CCMI)
- **The tool is developed under a subversion controlled repository**
  - Allows multiple developers from different institutions to join the development team
  - Broad community-wide and international participation in the development is envisaged
  - Collaboration with the metrics panel and PCMDI
- **Current extensions**
  - Atmospheric dynamics, biogeochemical processes, cryosphere and ocean
  - Need for a wide range of observational data to be implemented into the tool
  - Observations should ideally be provided with uncertainty and a technical documentation (e.g. similar to those from obs4MIPs) and in CF compliant netCDF as the models
  - We will work on improving the comparability between models and observations
  - Improve statistical comparison and visualization
- **Regular releases to the wider international community**
  - Further develop and share the diagnostic tool and routinely run it on CMIP output and according observations (obs4MIPs) on the Earth System Grid Federation (ESGF).
  - Release of ESMVal tool to the public at the end of the EMBRACE project (Oct 2015) => will be contribute to metrics panel code repository

DLR