An update on obs4MIPs

WDAC Observations for Model Evaluation Task Team

Peter Gleckler, co-chair, PCMDI and **Duane Waliser**, co-chair, JPL/NASA
Mike Bosilovich, GSFC/NASA
Helene Chepfer, IPSL
Carol Anne Clayson, WHOI
Veronika Erying, DLR
Robert Ferraro, JPL/NASA
Pierre-Phillipe Mathieu, ESA
Jerry Potter, GSFC
Roger Saunders, UKMO
Jörg Schulz, EUMETSAT
Karl Taylor, PCMDI
Jean-Noël Thépaut, ECMWF

Additional regular contributors: Otis Brown, Michel Rixen
Tsengdar Lee (NASA) and Renu Joseph (DOE)
Luca Cinquini (JPL) – CoG technical support
Denis Nadeu (PCMDI) – CMOR development
Paul Durack (PCMDI) – Data specifications
Sophie Cloché (IPSL) – CFMIP archive
Jim Biard (NCEI) and Matthias Tuma (WCRP) – beta testers
… and many others

WDAC7, Geneva, 26-27 March 2018

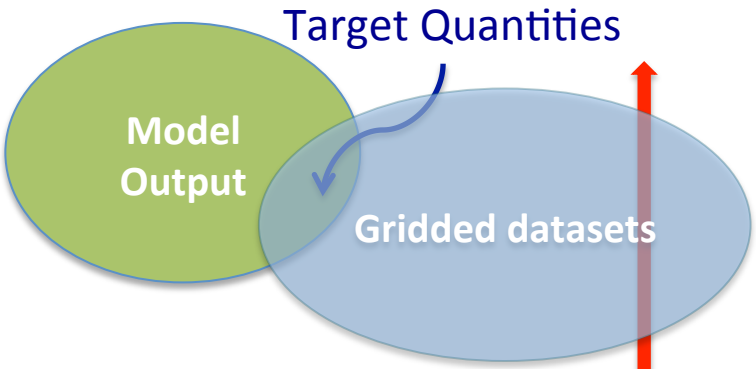
Presentation outline

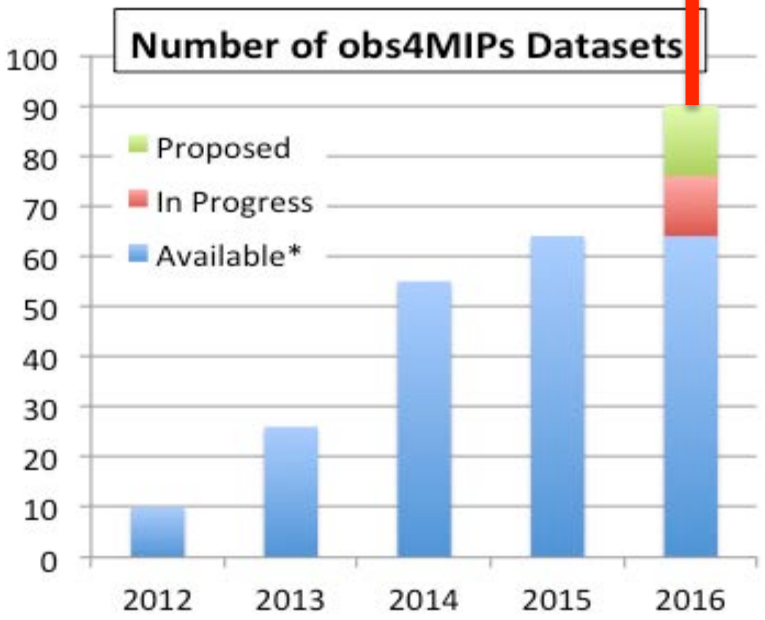
- Quick review, general status report
- Challenges, solutions in the works
- Strategic considerations for WCRP SP

obs4MIPs

https://www.earthsystemcog.org/projects/obs4mips/

- A project for identifying, documenting and disseminating observations for climate model evaluation in WCRP model intercomparisons, notably CMIP.
- Data (and tech notes) accessible with the distributed CMIP model output, adhering to same conventions
- Guided by the WCRP Data Advisory Council obs4MIPS Task Team



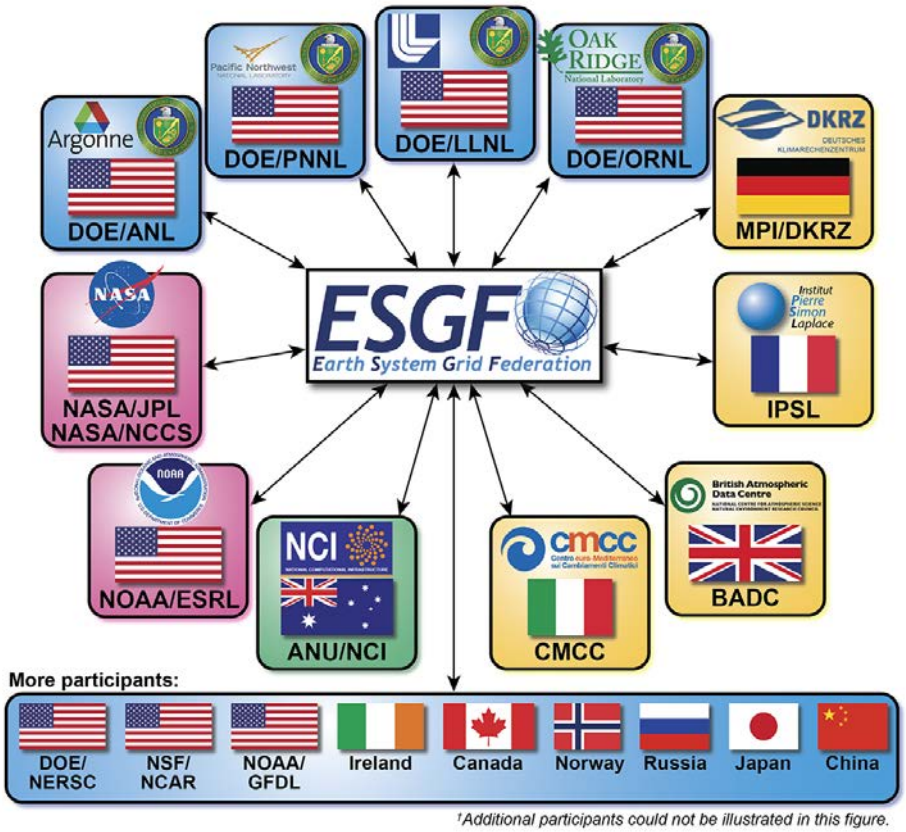


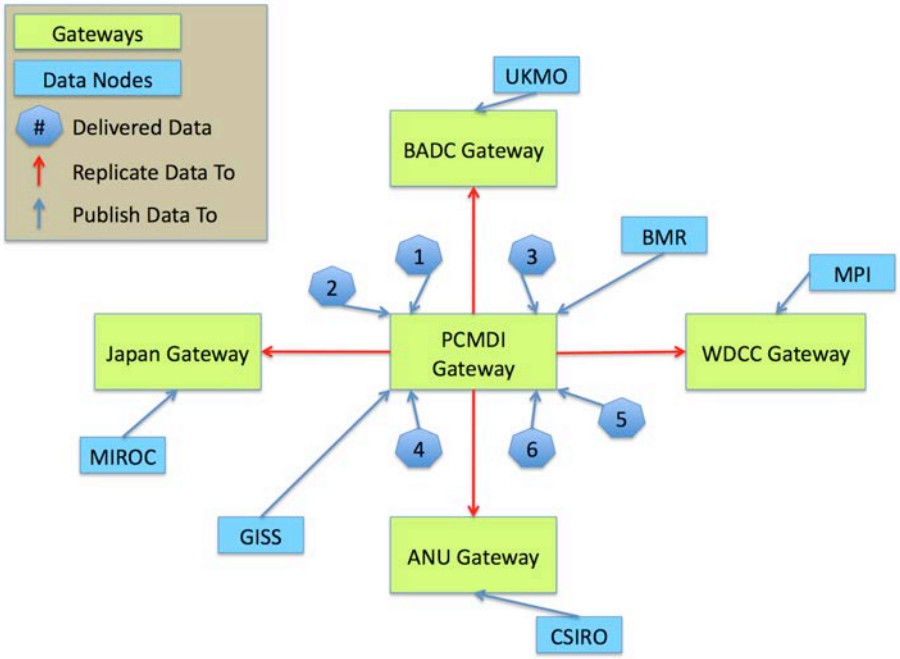
Complete (~125*)
In Progress* (~15)
Proposals from Data Call (~100)

…. and growing!

Data accessibility for WCRP projects

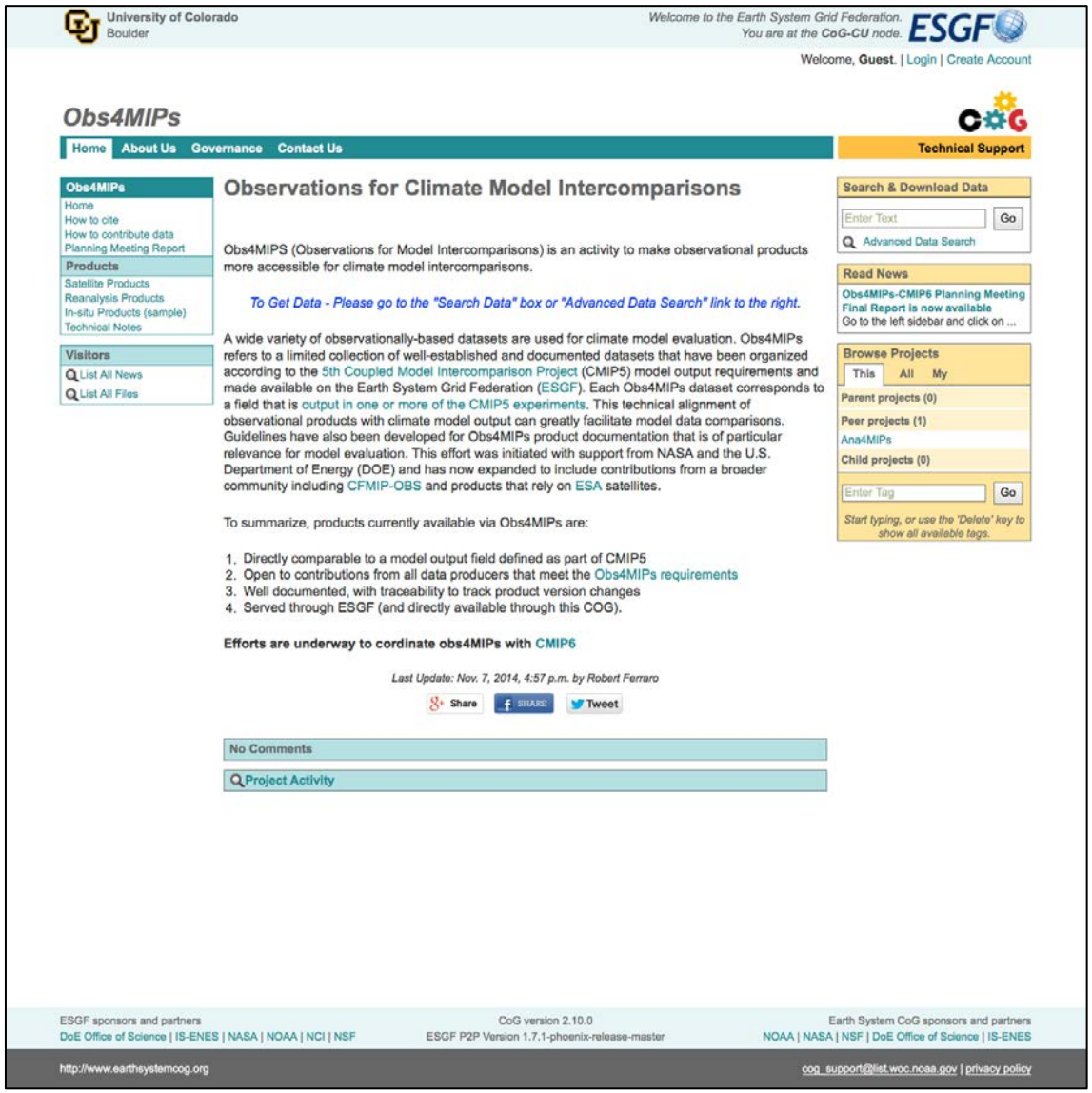
The Earth System Grid Federation (**ESGF**) is being used for **CMIP6** and other WCRP projects





[ESGF Dec 2016 conference report (Williams et al., 2017)]
[ESGF Dec 2017 Recap] (final report in preparation)

Data access and project connectedness



- Obs4MIPs data (and ana4MIPs) are available through the CoG
- CoG is integrated with ESGF
- CMIP6 will be hosted on the CoG, as are many other projects

Obs4MIPs leverages key protocols and infrastructure

- Experiment protocol and **CMIP standard output**
- **Climate Forecast (CF) Convention (as applied in CMIP)**
- **CMIP data governance** led by WGCM Infrastructure panel (WIP)
- Utilities to ensure data complies to CMIP structure: **CMOR3**
- Distribution/replication via **ESGF**, technically aligned with CMIP data
- **CoG**: Project information, data search/retrieval, integrating projects

Pros: Pulling together key efforts that serve WCRP projects

Cons: Vulnerable to weak links

obs4MIPs

obs4MIPs planning meeting for CMIP6

April 2014, NASA HQ but still relevant!

Selected consensus recommendations that applied to all of the meeting topic areas:

- **Expand the inventory**
- **Include more higher frequency** data (a "golden period"?)
- Reliable and defendable **error characterization**/estimation of observations
- Include datasets in support of **off-line simulators** (prime example: COSP—Cloud Feedback Model Intercomparison Project [CFMIP] Observation Simulator Package)
- **Collocated observations, including in-situ datasets** for processes level diagnostics
- **Precise definitions of data products** (what's actually being reported), including biases, and precise definitions of the model output variables are required

Ferraro et al. (2015) BAMS and full meeting report on CoG website

Progress since WDAC6

- 5 TT telecons, multiple presentations at scientific meetings
- Two manuscripts in-prep (#1 strategy, #2 data description governance)
- Further streamlining, e.g., **managing controlled vocabularies on GitHub**
- Datasets span across 7 ESGF nodes, complete listing generated daily

- **CMIP6/obs4MIPs data convention integration completed**

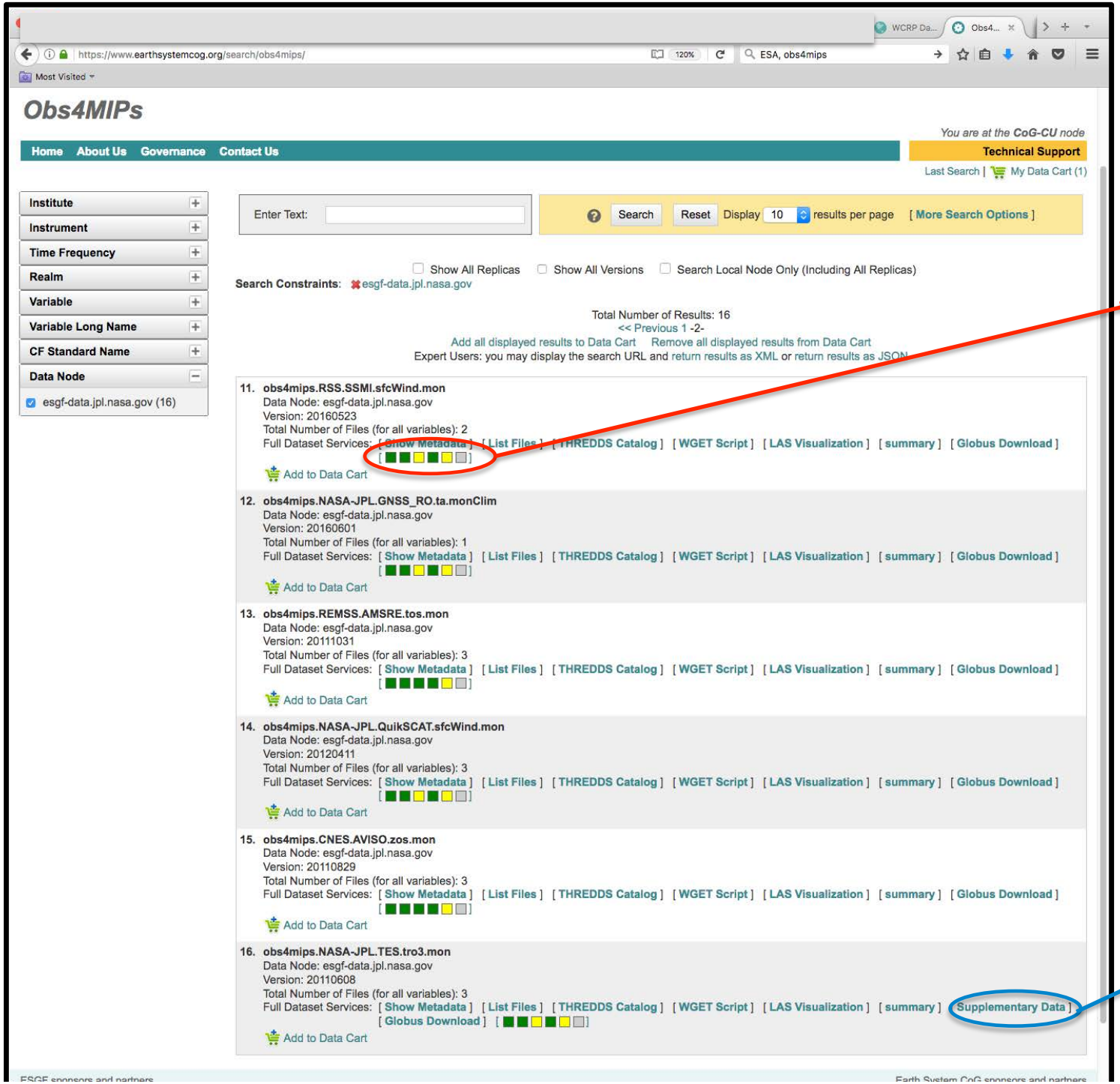
  (obs4MIPs Data Specifications ODS 2.1)

- Collaboration with ES-DOC to handle obs4MIPs tech notes, DOI's
- Two strategic efforts to broaden project scope prototyped:
  - obs4MIPs **data indicators** now established
  - **Mechanism for including Supplemental Information**

obs4MIPs Dataset Suitability & Maturity Indicators

Technical Requirements		***Dataset Suitability and Maturity***			***Comparison Complexity***
Meets obs4MIPs data technical requirements	Includes obs4MIPs technical note information	Closeness or robustness of measurement to observed reference quantity	Maturity with respect to climate model evaluation	Provision for robust uncertainty information	Complexity of Model Observation Comparison
Data suitably processed with CMOR and/or consistent with obs4MIPs standards	Complete technical note information provided	Measurement approach provides a very close relationship to observation quantity	Multiple peer-reviewed examples of application to climate model evaluation	Uncertainty information provided per retrieval/grid point	Comparison can be made directly with CMIP model output variable
Largely complete with minor metadata inconsistencies	Technical note information incomplete and/or could be improved	Measurement approach requires complex and/or non-linear retrieval methods and/or subjective inferences/definitions	One peer-reviewed example of application to climate or component model evaluation.	General uncertainty information given relative to the methodology and dataset as a whole - backed by actual field/in-situ validation exercises	Comparison requires some simple post processing of CMIP output variable(s) (e.g. vertical integral or ratio of two variables)
Non-compliant. Should be removed from database!	Technical note not provided	Measurement approach requires significant use/influence from complex or weakly constrained model and/or has significant ambiguity in definition(s)	No peer-reviewed xamples of application to model evlauation	No uncertainty information provided	Comparison requires complex processing of CMIP output (e.g. "simulator", budget calculation)

Indicators for all datasets will be determined by TT in consultation with dataset providers



obs4MIPs

TT will be responsible for confirming indicators (proposed by data providers) for all datasets

SI is “free form”, accessible from “best estimate” but not searchable independently

Coordinated CMIP/obs4MIPs global attributes, controlled vocabulary (CV), Registered Content (RC), and Data Reference Syntax (DRS)

Some predefined global attributes (there are many others)

	Sample
Variable ID #	**'sfcWind'**
Source_id *+	**'NOAA-NCEI-SeaWinds-1-2'**
Source_version_number * +	**'1.2'**
Institution_id +	**'NOAA-NCEI'**
Region* * #	**'global ocean'**
Frequency *	**'Amon'**
Nominal_resolution* #	**'1x1 degree'**
License	

* New obs4MIPs search facets

+ Registered Content (RC)

CV with pre-defined options maintained on github

2018 Timelines

Winter'18	Spring'18	Summer'18	Fall'18	Winter'19
Make existing obs4MIPs data consistent with ODS2.1				
	Prepare new datasets including from "1st Dataset Call"			
	Prototype enhanced CoG features			
		Propogate use of new CoG features across ESGF nodes (search facets, indicators and Supplemental Info)		
		Submit two papers (strategy and conventions)		
			Possible 2nd Dataset Call: Beyond Gridded?	

Summary and Perspective

- **obs4MIPs is well positioned to have substantial impact on CMIP6 and other WCRP MIPs**
- **WCRP SP: might obs4MIPs provide a basis for better connecting models and data?**
- **obs4MIPs TT has been able to address many recommendations (e.g., enable more data and information to be included)**
- **Key challenge continues: further "greasing the wheels" so that datasets can be *efficiently made to meet the obs4MIPs data specifications, and published on ESGF***
- **The goal for the coming year is to greatly expand the obs4MIPs archive, exploiting the new obs4MIPs data specifications (ODS2.1), supplemental information and data indicators**
- **Efforts to advance obs4MIPs remain focused on gridded datasets. Expanding the scope to include in-situ data will required new contributors prepared to do substantial work**

Summary and Perspective ii

- **Presuming obs4MIPs continues to successfully mature, there will likely be a need to revisit how it is supported and governed.**

- **While a great deal of infrastructure is in place, a broader obs4MIPs would require more people doing hands on work.**

- **The WGCM's WIP may serve as a good example of how to do this, and perhaps could be further leveraged.**

Links to related material – NEEDS TO BE UPDATED

- ESGF Dec 2016 conference report (pdf)
- CMIP6 data specifications (google docs)
- obs4MIPs data specifications
- obs4MIPs tables controlled vocabulary (github)
- Climate Model/obs Output Rewritter, CMOR (website)
- Draft user guide for preparing obs4MIPs (google docs)

EXTRAS

obs4MIPs: The 4 Commandments

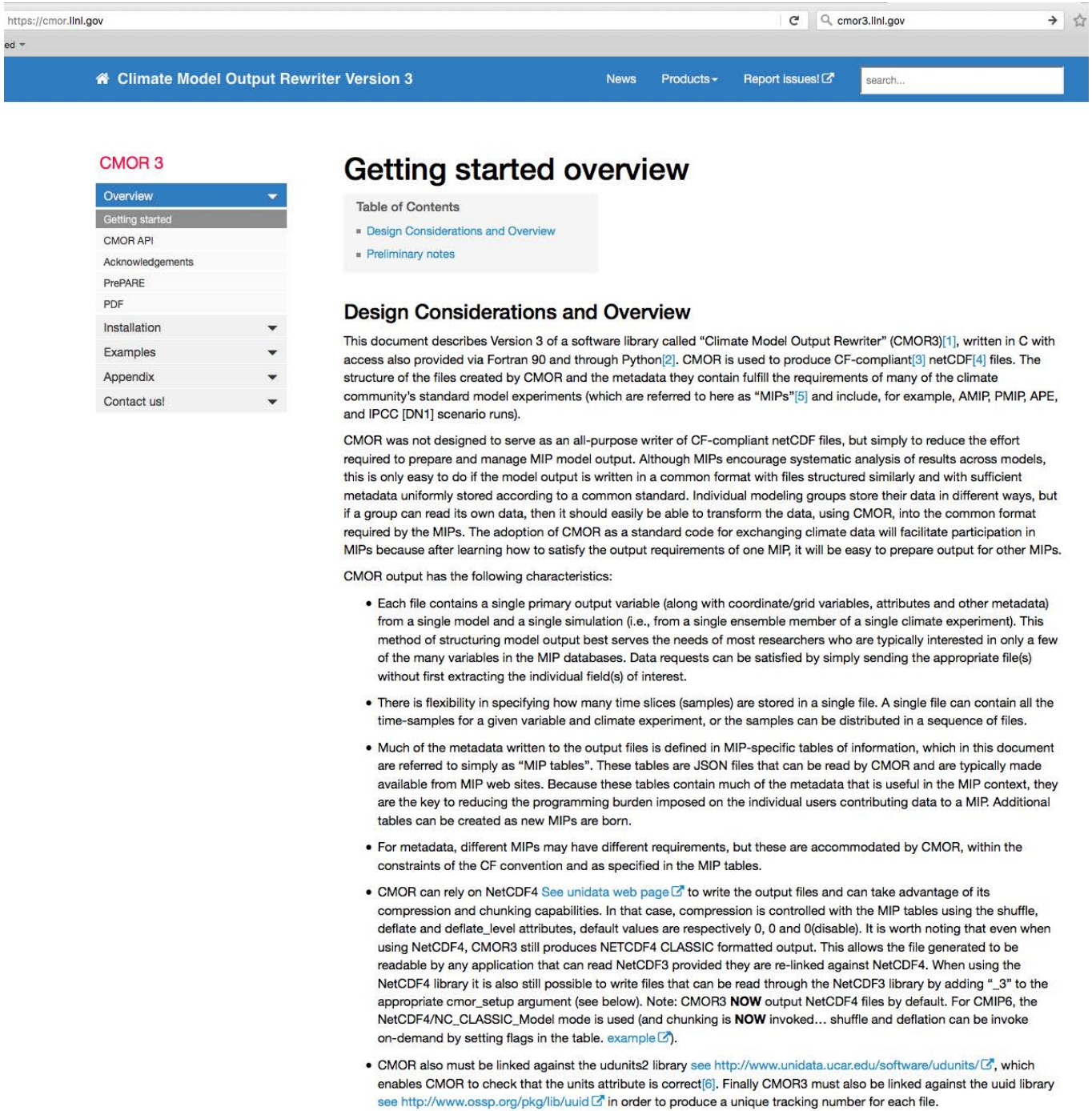
1. Use the CMIP* Standard Model Output as guideline for selecting observations

2. Observations to be structured in coordination with the CMIP output (e.g. NetCDF, CF Convention, **common vocabularies**)

3. Hosted side by side on the ESGF with CMIP model output

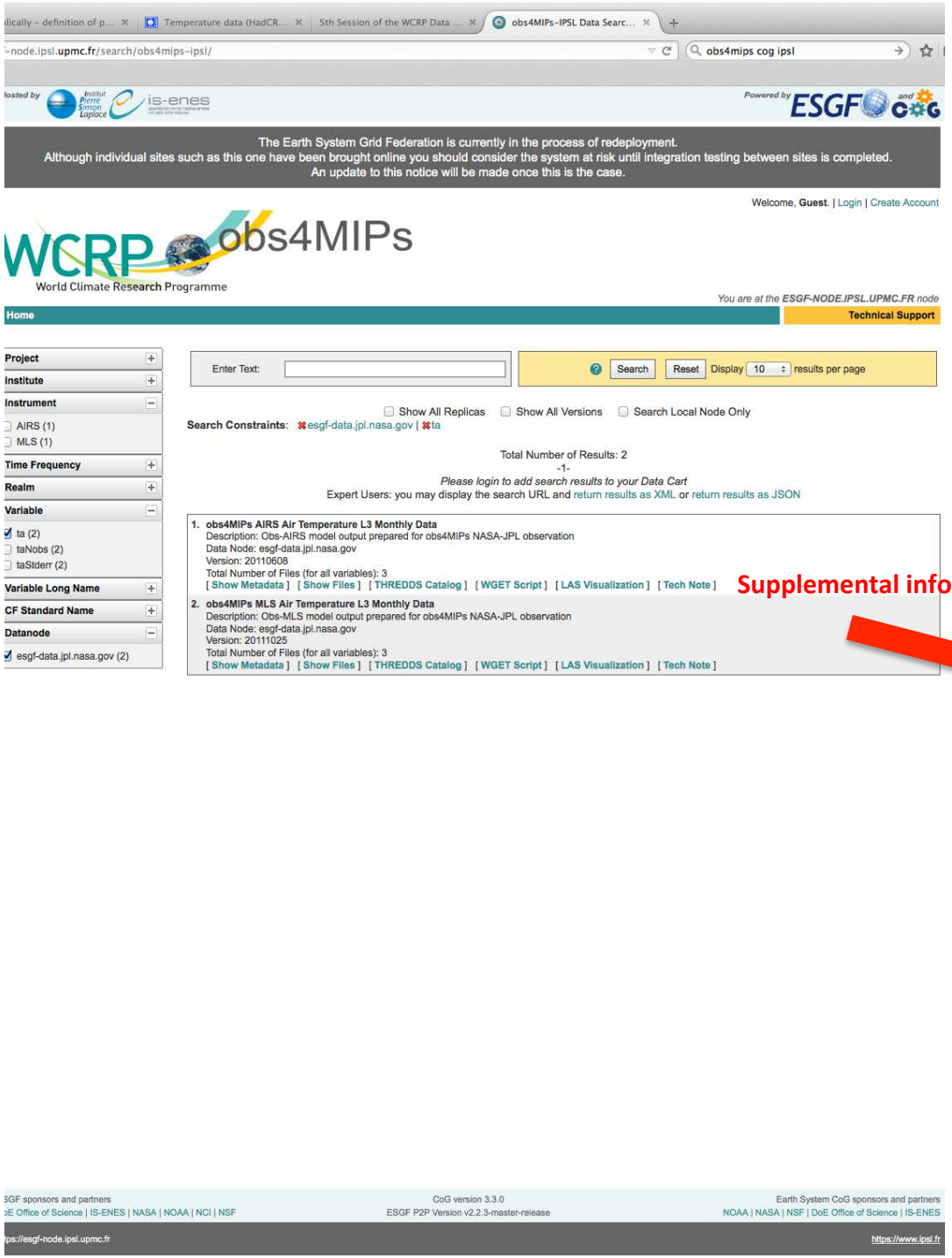
4. Include a Technical Note for each variable describing observation and use for model evaluation (at graduate student level)

*** obs4MIPs conventions have been updated to be consistent with CMIP6**



cmor.llnl.gov

- **CMOR3** now supports observational data!
- Strongly recommended for obs4MIPs – it is the easiest way to ensure compliance
- Support for CMIP and obs4MIPs users
- Open source, can be adopted for other purposes
- Being used by most CMIP participating modeling groups
- Available via Anaconda with python, C and fortran interfaces



Beyond the "best estimate" obs4MIPs dataset Supplemental Information

Strategy currently in preparation by WDAC task team

A possible path forward to deal with multiple unresolved issues within obs4MIPs, e.g.:

- Additional documentation
- Uncertainty data
- Data flags, masks
- Supporting validation data
- Detailed Maturity matrices

Organization of supplemental data will initially be more heterogeneous than associated "best estimate", and therefore not searchable from a higher level. It will be alongside best estimate as a zipped archive

Obs4MIPs: Status of Observation Holdings/ Submissions

Early 2018

Air Temperature
Specific Humidity
Air Temperature
Specific Humidity
Sea Surface Temperature
TOA Outgoing Longwave Radiation
TOA Outgoing Clear-Sky Longwave Radiation
TOA Outgoing Shortwave Radiation
TOA Outgoing Clear-Sky Shortwave Radiation
TOA Incident Shortwave Radiation
Surface Downwelling Longwave Radiation
Surface Downwelling Clear-Sky Longwave Radiation
Surface Upwelling Longwave Radiation
Surface Downwelling Shortwave Radiation
Surface Downwelling Clear-Sky Shortwave Radiation
Surface Upwelling Shortwave Radiation
Surface Upwelling Clear-Sky Shortwave Radiation
Total Cloud Fraction
Sea Surface Height Above Geoid
Precipitation - monthly
Precipitation – 3 hourly
Precipitation - daily
Precipitation - monthly
Near-Surface Wind Speed
Eastward Near-Surface Wind
Northward Near-Surface Wind
Leaf Area Index
Mole Fraciont of Ozone
Ambient Aerosol Opitical Thickness at 550 nm
Ambient Aerosol Opitical Thickness at 550 nm
Water Vapor Path
ISCCP Cloud albedo
ISCCP Cloud albedo
CALIPSO Scattering Ratio, srbox1
CALIPSO Scattering Ratio, srbox2
CloudSat Radar Reflectivity CFAD
CALIPSO Cloud Fraction
CALIPSO Clear Cloud Fraction
CALIPSO High Level Cloud Fraction
ISCCP Cloud Fraction
CALIPSO Low Level Cloud Fraction
CALIPSO Mid Level Cloud Fraction
CALIPSO 3D Clear Fraction
CALIPSO Total Cloud Fraction
CLOUDSAT Total Cloud Fraction
ISCCP Total Cloud Fraction
ISCCP Cloud Top Temperature
ISCCP Cloud Top Temperature
Missing data fraction
Overpasses
PARASOL Reflectance
Solar Zenith Angle
ISCCP Cloud Top Pressure
ISCCP Cloud Top Pressure
MISR CTH-OD Cloud Fraction
CALIPSO 3D Undefined fraction

Water Vapor Path
Fraction of Absorbed Photosynthetically Active Radiation
Snow area fraction
Ambient Aerosol Extinction Optical Thickness at 550 nm
Sea Ice fraction
Sea Ice fraction
Sea Surface Temp
Sea Surface Temp
Sea Surface Temp
TOA Outgoing Longwave Radiation
TOA Outgoing Longwave Radiation
TOA Outgoing Shortwave Radiation
TOA Outgoing Longwave Radiation
Mole Fraction of Ozone
albedo; Ratio of two variables
Ice Sheet Temperature?
Ambient Aerosol Opitical Thickness at 550 nm
Surface Aqueous Partial Pressure of CO2
dry_atmosphere_mole_fraction_of_carbon_dioxide
Near-Surface Wind Speed

Complete (~125*)

In Progress* (~15)

Submissions From Data Call (~90)

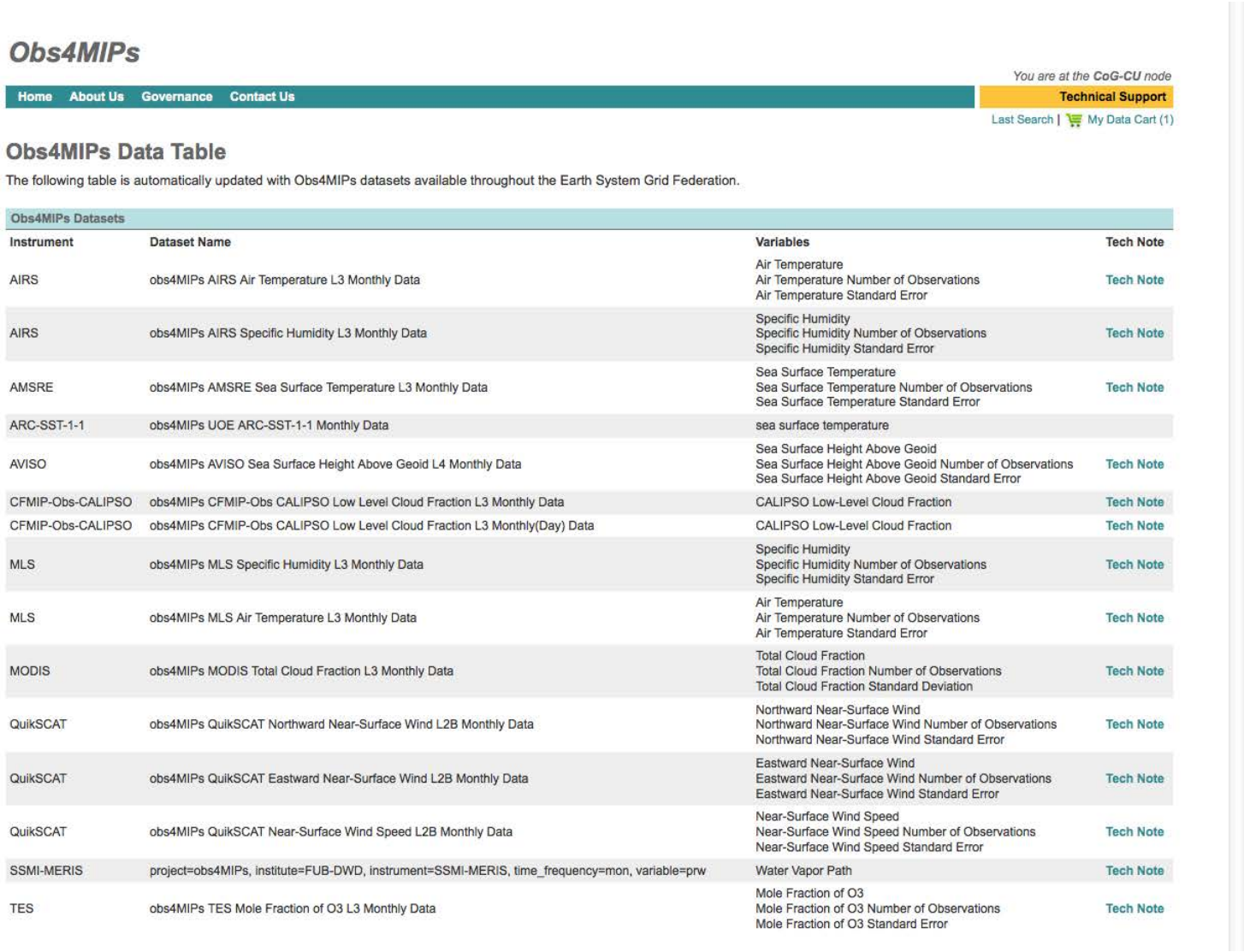
* A few technotes remain to be completed

Near-Surface Air Temperature
Air Temperature
Geopotential Height
Eastward Near-Surface Wind
Northward Near-Surface Wind
Near-Surface Wind Speed
Near-Surface Specific Humidity
Surface Downward Latent Heat Flux
Surface Downward Sensible Heat Flux
Near-Surface Air Temperature
Precipitation
Sea level pressure
Sea Surface Temp
Total Chlorophyll Mass Concentration
Infrared brightness temperatures (11 µm, 0.6 µm, 6.7 µm)
Leaf Area Index
normalized difference vegetation index
Fraction of Photosynthetically Active Radiation
Sea Surface Temperature
precipitation
air temperature
Burnt Area Fraction
Surface Snow Amount
Mid Tropospheric Humidity
Upper Tropospheric Humidity
Air Temperature
geopotential height
bending angle
refractivity
CLARA cloud_area_fraction; CFMIP 45
surface albedo
cloud_area_fraction
cloud top phase; area_fraction_of_liquid_cloud_water_particles_at_cloud_top
cloud top pressure; air_pressure_at_cloud_top
cloud optical thickness; atmosphere_optical_thickness_due_to_cloud
cloud ice water path; atmosphere_cloud_ice_content
surface_downwelling_shortwave_flux_in_air
surface downwelling clear sky shortwave flux
surface_downwelling_shortwave_flux_in_air
cloud ice water path; atmosphere_cloud_ice_content
cloud liquid water path
cci cloud area fraction (7x7 table); CFMIP 45 (tbd)
cloud area fraction
cloud top phase; area_fraction_of_liquid_cloud_water_particles_at_cloud_top
cloud top pressure; air_pressure_at_cloud_top
cloud optical thickness; atmosphere_optical_thickness_due_to_cloud
Sea Ice Area Fraction
Surface Temperature
Sea Ice Area Fraction
surface (2m) air temperature anomaly
Near-Surface Specific Humidity
Near-Surface Relative Humidity,
Near-Surface Air Temperature

Task Team Efforts/Organization

- Regular TT telecons (~monthly)
- All project management data accessible to TT via Google Docs
- Much of the TT effort has been helping to better define strategic path, but increasingly…
- Moving toward helping to shepherd datasets from proposal to publication on ESGF

Monitoring datasets across the distributed federation:
A new capability scans all ESGF nodes and generates a catalog of available datasets



Obs4MIPs

You are at the CoG-CU node

Home About Us Governance Contact Us | Technical Support

Last Search | My Data Cart (1)

Obs4MIPs Data Table

The following table is automatically updated with Obs4MIPs datasets available throughout the Earth System Grid Federation.

Obs4MIPs Datasets

Instrument	Dataset Name	Variables	Tech Note
AIRS	obs4MIPs AIRS Air Temperature L3 Monthly Data	Air Temperature Air Temperature Number of Observations Air Temperature Standard Error	Tech Note
AIRS	obs4MIPs AIRS Specific Humidity L3 Monthly Data	Specific Humidity Specific Humidity Number of Observations Specific Humidity Standard Error	Tech Note
AMSRE	obs4MIPs AMSRE Sea Surface Temperature L3 Monthly Data	Sea Surface Temperature Sea Surface Temperature Number of Observations Sea Surface Temperature Standard Error	Tech Note
ARC-SST-1-1	obs4MIPs UOE ARC-SST-1-1 Monthly Data	sea surface temperature	
AVISO	obs4MIPs AVISO Sea Surface Height Above Geoid L4 Monthly Data	Sea Surface Height Above Geoid Sea Surface Height Above Geoid Number of Observations Sea Surface Height Above Geoid Standard Error	Tech Note
CFMIP-Obs-CALIPSO	obs4MIPs CFMIP-Obs CALIPSO Low Level Cloud Fraction L3 Monthly Data	CALIPSO Low-Level Cloud Fraction	Tech Note
CFMIP-Obs-CALIPSO	obs4MIPs CFMIP-Obs CALIPSO Low Level Cloud Fraction L3 Monthly(Day) Data	CALIPSO Low-Level Cloud Fraction	Tech Note
MLS	obs4MIPs MLS Specific Humidity L3 Monthly Data	Specific Humidity Specific Humidity Number of Observations Specific Humidity Standard Error	Tech Note
MLS	obs4MIPs MLS Air Temperature L3 Monthly Data	Air Temperature Air Temperature Number of Observations Air Temperature Standard Error	Tech Note
MODIS	obs4MIPs MODIS Total Cloud Fraction L3 Monthly Data	Total Cloud Fraction Total Cloud Fraction Number of Observations Total Cloud Fraction Standard Deviation	Tech Note
QuikSCAT	obs4MIPs QuikSCAT Northward Near-Surface Wind L2B Monthly Data	Northward Near-Surface Wind Northward Near-Surface Wind Number of Observations Northward Near-Surface Wind Standard Error	Tech Note
QuikSCAT	obs4MIPs QuikSCAT Eastward Near-Surface Wind L2B Monthly Data	Eastward Near-Surface Wind Eastward Near-Surface Wind Number of Observations Eastward Near-Surface Wind Standard Error	Tech Note
QuikSCAT	obs4MIPs QuikSCAT Near-Surface Wind Speed L2B Monthly Data	Near-Surface Wind Speed Near-Surface Wind Speed Number of Observations Near-Surface Wind Speed Standard Error	Tech Note
SSMI-MERIS	project=obs4MIPs, institute=FUB-DWD, instrument=SSMI-MERIS, time_frequency=mon, variable=prw	Water Vapor Path	Tech Note
TES	obs4MIPs TES Mole Fraction of O3 L3 Monthly Data	Mole Fraction of O3 Mole Fraction of O3 Number of Observations Mole Fraction of O3 Standard Error	Tech Note

~125 DataSets