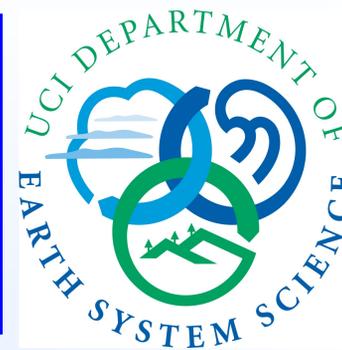




Simplifying and accelerating model evaluation by NASA satellite data

Charlie Zender, Pedro Vicente, and Wenshan Wang
Department of Earth System Science, University of California, Irvine



BACKGROUND

Geoscientific model results are increasingly evaluated by comparison to products derived from NASA satellite measurements. The satellite data are archived in HDF format, which is now a superset of the netCDF format employed by most models. Putting NASA-stewarded (HDF) data and model-generated (netCDF) data on a common grid, in the same format, for numerical comparison can be arduous. Many researchers desire a common toolkit for both HDF and netCDF data that would 1. simplify and accelerate the independent analysis of both data formats, 2. exploit the strengths of netCDF's simple API and underlying HDF data format with easy-to-use tools, 3. ease evaluations of model predictions by NASA-stewarded data.

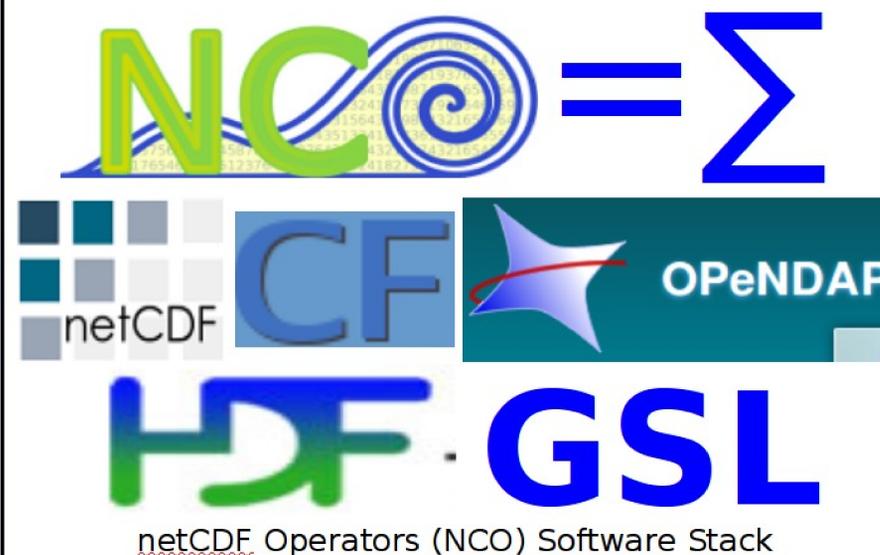
SIMPLIFY

Eliminate unnecessary complexity intercomparing multiple models and measurements. Traditional analysis requires loops over ensemble members/scenarios to evaluate and attribute observed changes. Some analyses that evaluate CMIP5 models against NASA observations:

MATERIALS

ACCESS Project

Our ACCESS project develops software to manipulate and analyze HDF and netCDF datasets. We added HDF5 group support to the netCDF Operators (NCO¹²). We are also creating NCO wrappers for HDF4 files. The NCO homepage is <http://nco.sf.net>.



ACCELERATE

Vectorize Analysis Using HDF/netCDF4 Groups

```
# Old Analysis: Loop over ensemble members
for run in '1 2 3 4 5 6 7 8'; do
  ncdiff CCSM_${run}.nc MODIS.nc CCSM_${run}_MODIS.nc
done

# New Analysis: Vectorize ensemble members by groups
ncecat CCSM_*.nc CCSM_all.nc
ncdiff CCSM_all.nc MODIS.nc CCSM_minus_MODIS.nc

# Old Analysis: Loop over ensemble members & models
for model in 'CESM GISS ECHAM ...'; do
  for run in '1 2 3 4 5 6 7 8'; do
    ncdiff ${model}_${run}.nc MODIS.nc \
      ${model}_${run}_minus_MODIS.nc
  done
done

# New Analysis: Vectorize ensembles into groups
ncecat CCSM_all.nc GISS_all.nc ECHAM_all.nc ... CMIP5.nc
ncdiff CMIP5_all.nc MODIS.A2012.nc CMIP5_minus_MODIS.nc
```

File-level Analysis of Native HDF Formats

```
# NCO works natively on many HDF files:
Ncwa -a nTimes MSLERLSTL3.he5 MSLERLSTL3.nc # HDF5 EOS2
ncra MOD10CM.A2007*.hdf MOD10CM_avg.nc # HDF EOS
ncwa -a nTimes SBUV2_2006.h5 SBUV2_avg.nc # HDF5
```

MILESTONES

- M2: Robust installs MacOS, RHEL, Windows 07/12
- M4: Single-level hierarchies in ncks operator 10/12
- M5: Nested hierarchies in ncks 01/13
- M7: MODIS/CERES snow/albedo scripts 01/13
- M8: Nested hierarchies in ncra operator 04/13
- M10: Verify CMIP5 hierarchical analyses 07/13
- M11: Nested hierarchies remaining operators 07/13
- M12: Wrappers for MODIS/CERES HDF data 10/13
- M15: Present simplification at AGU ESSI 12/13
- M16: Wrappers for other (MISR?) HDF data 01/14
- M17: Consolidated NCO/HDF release 04/14

STATUS

Accomplishments: Generic operators for hierarchical datasets: subset, hyperslab, move, rename, aggregate, flatten, dismember, inherit, regexp, broadcast, pack/unpack, arithmetic. Fully documented with graphics².

```
# Extract variables from groups g1, g2
ncks -g g1,g2 -v v1,v2 in.h5 out.h5

# Hyperslab tropical latitudes of g1
ncks -g g1 -d lat,-30.,30. in.h5 out.h5

# Move data from /g1/g1g1 to /g2
ncks -G g2:-2 -g g1g1 in.h5 out.h5

# Aggregate flat data into groups
ncecat -gag 1985.h5 1986.h5 1985_1986.h5

# Weighted average all groups
ncwa -a lat,lon -w area in.h5 out.h5

# Time-mean all groups
ncra 1985.h5 1986.h5 1985_1986_avg.h5

# Unpack HDF convention, repack netCDF
ncpdq -hdf_upk -P xst_new in.h5 out.nc

# Dismember then check CF-compliance
ncdismember in.nc $TMPDIR cf
```

REFERENCES

- Zender, C. S. (2008), Analysis of Self-describing Gridded Geoscience Data with netCDF Operators (NCO), Environ. Modell. Softw., 23(10), 1338-1342, doi:10.1016/j.envsoft.2008.03.004.
- Zender, C. S. (2013), netCDF Operator (NCO) User Guide, Version 4.3.6, <http://nco.sf.net/nco.html>.
Research supported by NASA ACCESS NNX12AF48A

