Climate Data Analysis Tools (CDAT)
Overview

• Promote community wide sharing of:
  – Data access, handling, and manipulation
  – Diagnostic development
  – Analysis and visualization

• Originally developed to promote achieving and diagnosing of simulation data.

• Open-source based on Python
Data Handling and Management

- Core module: Climate Data Management System 2 (CDMS 2)
  - Multiple I/O formats: NetCDF, HDF, PP, GrADS-GRIB, ASCII, Binary
  - Lies on top of strong numerical package but made metadata smart: MV2 – layer on top of Numpy which preserves metadata
  - NetCDF Climate and Forecast (CF) convention – metadata designed to promote the processing and sharing of simulation and observation data
Climate Model Output Rewriter 2 (CMOR 2)

- The "Climate Model Output Rewriter" (CMOR) can be used to produce CF-compliant NetCDF files that fulfill the requirements of many of the climate community's standard model experiments and observations.
- Used as the metadata schema for the Earth System Grid Federation (ESGF)

<table>
<thead>
<tr>
<th>CMOR Table Amon: Monthly Mean Atmospheric Fields and Some Surface Fields</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amon</strong></td>
</tr>
<tr>
<td>All Saved on the Atmospheric Grid</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Priority</th>
<th>long name</th>
<th>units</th>
<th>comment</th>
<th>questions</th>
<th>output variable name</th>
<th>standard name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Near-Surface Air Temperature</td>
<td>K</td>
<td>near-surface (usually, 2 meter) air temperature.</td>
<td>tas</td>
<td>tas</td>
<td>air_temperature</td>
</tr>
<tr>
<td>1</td>
<td>Surface Temperature</td>
<td>K</td>
<td>“skin” temperature (i.e., SST for open ocean)</td>
<td>ts</td>
<td>ts</td>
<td>surface_temperature</td>
</tr>
<tr>
<td>1</td>
<td>Daily Minimum Near-Surface Air Temperature</td>
<td>K</td>
<td>monthly mean of the daily-minimum near-surface (usually, 2 meter) air temperature.</td>
<td>tasmin</td>
<td>tasmin</td>
<td>air_temperature</td>
</tr>
<tr>
<td>1</td>
<td>Daily Maximum Near-Surface Air Temperature</td>
<td>K</td>
<td>monthly mean of the daily-maximum near-surface (usually, 2 meter) air temperature.</td>
<td>tasmax</td>
<td>tasmax</td>
<td>air_temperature</td>
</tr>
<tr>
<td>1</td>
<td>Sea Level Pressure</td>
<td>Pa</td>
<td>not, in general, the same as surface pressure</td>
<td>ps</td>
<td>ps</td>
<td>air_pressure, at sea_level</td>
</tr>
<tr>
<td>1</td>
<td>Surface Air Pressure</td>
<td>Pa</td>
<td>not, in general, the same as mean sea-level pressure</td>
<td>ps</td>
<td>ps</td>
<td>surface_air_pressure</td>
</tr>
</tbody>
</table>
Community Contributed Analysis Packages

- Numpy/MA/MV
- Genutil (developed at LLNL)
  - Commonly used functions to compute correlation, covariance, auto-
    correlation, auto-covariance, lagged correlation, lagged covariance,
    mean absolute difference, root mean square, standard deviation,
    variance, geometric mean, median, percentiles, linear regression,
    etc.
- Cdutil (developed at LLNL)
  - Climate data specific utilities such as spatial, area weighing,
    climatology diagnostics, departures, etc.
- Community contributed packages
  - Pyclimate
  - SciPy
- Over 100 software packages contributed from the climate community
UV-CDAT Supports Multiple Regridding Tools

- UV-CDAT expands the choice of CDAT regridding tools by adding:
  - LibCF for nodal interpolation of curvilinear data in N-dimensions
  - ESMF for cell (conservative) or nodal (linear or quadratic) interpolation of curvilinear data in 2D or 3D
    - Leverages parallel ESMF library: domain decomposition performed on the fly by ESMF
    - Also supports regridding from and to unstructured grids
    - Handles well tripolar grid, displaced pole, and other types of grid warping

Tripolar of ocean model (detail)

Conservative regridding with ESMF

Valid fractional areas/volumes act as masking
Additional Grid Awareness

- Multiple grid software built-in: ESMF/ESMP, Gridspec, Ugrid, SCRIP, etc.
The Visualization and Control System (VCS)
1D and 2D Plots
• Constantly harvesting new diagnostics
• Some samples:
  – (a) WK
  – (b) Thermo
  – (c) Taylor-diagrams
  – (d) Performance plots