Reducing Reanalyses Data Wrangling with CREATE

Laura Carriere, Gerald L. Potter, Mark A. McInerney, Judy Hertz, Yingshuo Shen, Thomas Maxwell, Eric Winter, Julien Peters, Savannah Strong

Climate Model Data Services (CDS)
NASA GODDARD SPACE FLIGHT CENTER

ABSTRACT

Scientists engaged in reanalyses—essentially re-forecasts of past weather using the latest forecast models—are interested in reproducing the success of the Coupled Model Intercomparison Project (CMIP5). They are studying reanalysis differences and uncertainties to improve reanalysis techniques. Reanalysis data also allows interdisciplinary scientists to compare their datasets (e.g., biodiversity, water planning, wind power) with 30 or more years of gridded climate data. These research efforts require large sets of monthly and hourly data, formatted identically to facilitate comparisons. NASA’s Climate Model Data Services (CDS) is collaborating with the world’s five major reanalysis projects to collect this data and present it through Distribution, Visualization, Analytics, and Knowledge services, resulting in the Collaborative REAnalysis Technical Environment (CREATE).

CREATE-V

Based on code from UCAR’s Climate Inspector, CDS is building a visualization tool for interdisciplinary and reanalysis scientists to support exploration of variables by reanalysis, date, and level. The tool is used TDS and OpenLayers and provides the ability to display images of commonly used monthly and 6-hourly variables from multiple reanalyses simultaneously.

Data

CDS has converted the data to the standard ESGF format of one variable per file using code (obs4MIPs.py) developed in house to streamline analytic workflows. Griddata innovations and observations, data used in each reanalysis model step, will also be processed and published to support reanalysis model improvements.

For Additional Information

http://cds.nasa.gov
laura.carriere@nasa.gov

UV-CDAT Remote Visualization

CDS has built a remote visualization system that provides authorized users access to the CREATE-IP data using UV-CDAT. The data is available via OpenDAP and ESGF (when it comes back online). Users can visualize the data using the local graphics card and send only the rendered image back to the user’s system. This provides fast visualization across the continental U.S. and allows for the quality assessment of data residing at the Goddard Space Flight Center by a scientist in California.

Comparing two time slices of Specific Humidity over northern California from two reanalyses to assess data quality. The bottom plot shows the difference between the two (MERRA minus ERA-Interim).

The ability to remotely visualize the full CREATE-IP dataset will assist scientists in identifying differences in reanalysis algorithms and processing. For example, above we can see the December 2014 atmospheric river fading out more slowly in MERRA than it does in the ERA-Interim reanalysis.

Summary

CREATE will provide access to the full range of reanalysis datasets—including all relevant variables, innovations, and observations—through multiple distribution and visualization services. Access to the data will be provided through the CREATE-IP project space on ESGF/COG. Collaborations with existing reanalysis organizations include reanalysis.org and the European Copernicus Climate Change project. It is anticipated that this collection of data, services, and science collaborations will support future work in reanalysis intercomparison and interdisciplinary science.

For Additional Information

http://cds.nasa.gov
laura.carriere@nasa.gov

Data

CDS has converted the data to the standard ESGF format of one variable per file using code (obs4MIPs.py) developed in house to streamline analytic workflows. Griddata innovations and observations, data used in each reanalysis model step, will also be processed and published to support reanalysis model improvements.

For Additional Information

http://cds.nasa.gov
laura.carriere@nasa.gov

UV-CDAT Remote Visualization

CDS has built a remote visualization system that provides authorized users access to the CREATE-IP data using UV-CDAT. The data is available via OpenDAP and ESGF (when it comes back online). Users can visualize the data using the local graphics card and send only the rendered image back to the user’s system. This provides fast visualization across the continental U.S. and allows for the quality assessment of data residing at the Goddard Space Flight Center by a scientist in California.

Comparing two time slices of Specific Humidity over northern California from two reanalyses to assess data quality. The bottom plot shows the difference between the two (MERRA minus ERA-Interim).

The ability to remotely visualize the full CREATE-IP dataset will assist scientists in identifying differences in reanalysis algorithms and processing. For example, above we can see the December 2014 atmospheric river fading out more slowly in MERRA than it does in the ERA-Interim reanalysis.

Summary

CREATE will provide access to the full range of reanalysis datasets—including all relevant variables, innovations, and observations—through multiple distribution and visualization services. Access to the data will be provided through the CREATE-IP project space on ESGF/COG. Collaborations with existing reanalysis organizations include reanalysis.org and the European Copernicus Climate Change project. It is anticipated that this collection of data, services, and science collaborations will support future work in reanalysis intercomparison and interdisciplinary science.

For Additional Information

http://cds.nasa.gov
laura.carriere@nasa.gov

Data

CDS has converted the data to the standard ESGF format of one variable per file using code (obs4MIPs.py) developed in house to streamline analytic workflows. Griddata innovations and observations, data used in each reanalysis model step, will also be processed and published to support reanalysis model improvements.

For Additional Information

http://cds.nasa.gov
laura.carriere@nasa.gov