### The NASA Reanalysis Ensemble Service

**Advanced Capabilities for Integrated Reanalysis Access and Intercomparison**

Jian Li1, Glenn S. Tamkin1, Savannah L. Strong1, Roger L. Gill1, John H. Thompson2, Daniel Q. Duffy3, John L. Schnase4

1 COMCEPTUAL AND INFORMATION SCIENCES AND TECHNOLOGY OFFICE
2 NASA CENTER FOR CLIMATE SIMULATION
3 NASA GODDARD SPACE FLIGHT CENTER

**Abstract**

NASA’s efforts to advance climate analytics-as-a-service are making new capabilities available to the research community: (1) A full-featured Reanalysis Ensemble Service (RES) comprising monthly means data from multiple reanalysis datasets, accessible through an enhanced set of extraction, analytic, arithmetic, and intercomparison operations. The operations are made accessible through NASA’s climate data analytics web services and our client-side Climate Data Services Python library. (2) A cloud-based, high-performance Virtual Real-Time Analytics Testbed supporting select climate variables. This near-real-time capability enables advanced technologies like Spark and Hadoop-based MapReduce analytics over native NetCDF files. (3) A Web Processing Service (WPS)-compliant web interface to our climate data analytics service that enables interoperability with next-generation systems such as the Earth System Grid Federation (ESGF).

### Features

1. Extended capabilities that enable single and multiple reanalysis area average, vertical average, re-gridding, standard deviation, and ensemble averages
   - Convenient, one-stop shopping for commonly used data products from multiple reanalyses, including basic sub-setting and arithmetic operations (e.g., avg, sum, max, min, var, count, anomaly)
   - Full support for the MERRA-2 reanalysis dataset in addition to, ECMWF ERA-Interim, NCEP CFSR, JMA JRA-55, and NOAA/ESRL 20CR.
   - The ability to compute and visualize multiple reanalysis for ease of inter-comparisons

2. Automated tools to retrieve and prepare data collections for analytic processing

3. Docker-ready RES application to deploy across platforms

4. Supporting analytic services for NASA Global Modeling and Assimilation Office (GMAO) Forward Processing datasets

5. Basic uncertainty quantification services that combine heterogeneous ensemble products with comparative observational products (e.g., reanalysis, observational, visualization)

6. Jupyter notebook-based distribution mechanism designed for client use cases that combines CDSLib documentation with interactive scenarios and personalized project management

7. New API that supports full temporal, spatial, and grid-based resolution services with sample queries

### Cross-Platform Docker Container

- Deployable to multiple hosts.
- Includes:
  - Application tier (Map/Reduce)
  - Web services (PHP)
  - Client (CDSLib/Python)
  - Configuration files
  - Subset of MERRA-2 monthly means

### Automated Analytic Data Provisioning (Loader)

- 1) Order Data
- 2) Download Data
- 3) Sequence Data
- 4) Validate Data
- 5) Load Data
- 6) Transmit Data

### RES Jupyter Notebook

For Additional Information:
- jian.li@nasa.gov
- glenn.s.tamkin@nasa.gov
- john.l.schnase@nasa.gov
- daniel.q.duffy@nasa.gov

---

**Abbreviations**

- ESGF: Earth System Grid Federation
- CDSLib: Climate Data Services Library
- WPS: Web Processing Service
- RES: Reanalysis Ensemble Service
- MERRA-2: Modern-Era Retrospective Analysis for Research and Applications, Version 2
- GMAO: NASA Center for Climate Simulation
- NCEP: National Center for Environmental Prediction
- ECMWF: European Centre for Medium-Range Weather Forecasts
- JMA: Japan Meteorological Agency
- NOAA: National Oceanic and Atmospheric Administration
- MODIS: Moderate Resolution Imaging Spectroradiometer
- LANDSAT: Landsat
- GPCP: Global Precipitation Climatology Project
- GMAO: NASA Modeling and Assimilation Office
- MM5: Mesoscale Model 5
- 20CR: Twentieth-Century Reanalysis
- NCEP: National Centers for Environmental Prediction
- CFSR: Climate Forecast System Reanalysis
- ERA: European Reanalysis
- NARR: North American Regional Reanalysis