LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-742817

Enabling Reanalysis Research Using the Collaborative REAnalysis Technical Environment (CREATE)

G. L. Potter, L. Carriere, J. A. Hertz, M. Bosilovich, D. Q. Duffy, T. J. Lee, D. N. Williams

December 6, 2017

Bulletin of the American Meteorological Society

Disclaimer

Enabling Reanalysis Research Using the Collaborative REAnalysis Technical Environment (CREATE)

Gerald L. Potter[1]

Laura Carriere[1]

Judy Hertz[1]

Michael Bosilovich[2]

Daniel Duffy[1]

Tsengdar Lee[3]

Dean N. Williams[4]

[1]NASA Goddard Spaceflight Center/NASA Center for Climate Simulation, Greenbelt, Maryland

[2]NASA Goddard Spaceflight Center/Global Modeling and Assimilation Office, Greenbelt, Maryland

[3]NASA Headquarters, Washington D.C.

[4]Lawrence Livermore National Laboratory, Livermore, California

Corresponding Author: Gerald L. Potter, email: gerald.potter@nasa.gov

Abstract

This paper describes repackaging and consistent distribution of the world's major atmospheric and oceanic reanalyses. It also presents examples of the usefulness of examining multiple reanalyses. This service will make it much easier for anybody using reanalysis to access multiple data sets using an approach similar to that of the Coupled Model Intercomparison Project Phase 5 (CMIP5) data. Experienced users as well as students will find the standardized formatted data convenient to use.

Introduction

Modern atmospheric and oceanic reanalyses are valuable assets for atmospheric research and climate monitoring (Kalnay et al. 1996). Now that most reanalysis records are more than 36 years long, the data has become more useful for climate modeling research (CCSP, 2008). See Table 1 for a list of acronyms. For investigators who need to use multiple reanalyses, a common challenge is that the data is distributed at various sites and often in different formats. The NASA Collaborative Reanalysis Technical Environment (CREATE) provides access to the data in one location in a standard format (one variable per file, and standardized meta-data in the CMIP5 style). The collection includes monthly and 6-hourly data from the seven major atmospheric reanalyses: CFSR (Saha et al. 2010), ERA-Interim (Dee et al. 2011), MERRA (Rienecker et al. 2011), MERRA-2 (Gelaro et al. 2017), JRA-25 (Onogi et al., 2007), JRA-55 (Kobayashi et al., 2014), and 20CRv2c (Compo et al. 2011). An ancillary portion of CREATE includes eight ocean reanalyses: NCEP/CFSR, CMCC/C-GLORSv5 (Storto, et al. 2016), ECMWF/ORA-S4 (Balmaseda et al. 2013), ECMWF/ORAP5.0 (Zuo, et al. 2015), University of Hamburg/GECCO2 (Köhl 2015), GFDL/ECDA (Zhang et al. 2007), NOAA/GODAS (Saha et al. 2010), and MOVE/MRI.COM-G2i (Toyoda et al. 2016). The ocean state variables were similarly reformatted but then also regridded to a common horizontal and vertical grid. This approach facilitated calculating an ensemble average and spread that is also published alongside the native gridded data. A third reanalysis product is a global hourly 0.5-degree land surface air temperature data set constructed by Wang & Zeng, (2013). All three data sets are distributed

through the Earth System Grid Federation (ESGF) in a format consistent with the CMIP style described by Cinquini et al. (2014).

Background

The CMIP3/CMIP5 effort was extremely successful in providing the climate research community with accessible data and provides a template for how comparing climate models benefits climate prediction (Taylor et al. 2012). CREATE provides the community with an easily accessible multiple-reanalysis data set patterned after the success of the CMIP5. Once the reanalyses have been standardized and placed in a common location, it is relatively easy to compare them or select the one most suited for a study. One of the important benefits of reanalysis intercomparison is to reveal uncertainty for fields where the various reanalyses disagree and confidence in those fields where they agree.

We have chosen to process a common subset of monthly and 6-hour reanalysis output similar to the variables chosen for the standard output of climate modeling in the CMIP5 experiments. Tables 2, 3 and 4 show the list of variables selected. The 3D fields are published on pressure levels that include the CMIP5 standard levels as well as on the native horizontal grid. This standardized processing requires that the variable names be changed to the CMIP5 (Climate and Forecast [CF] compliant) names and that the dimension bounds be set. The processing also provides a standardized set of metadata to allow the ESGF software queries to perform searches across all the reanalyses.

Processing – Atmospheric Reanalyses

Figure 1 is a schematic summary of the processing necessary to standardize each reanalysis to a common format. Except for NASA's MERRA and MERRA-2, each reanalysis required a different data retrieval method. In addition, ERA-interim, JRA-25, JRA-55, and CFSR all required conversion from GRIB to NetCDF. None of these processing steps in themselves were particularly difficult, but each required a unique set of procedures. The 6-hour data processing was similar to the monthly data, but it was necessary to process the data into monthly chunks to allow the file size to be manageable for downloading. Once the data was organized, it was subjected to the Climate Model Output Rewriter (CMOR), where units, axis labeling, and standardized metadata were modified to conform with the CMIP5 standards (Williams 2013). This processing is now largely automated, allowing newly processed reanalyses to be added to CREATE shortly after the data is made available by the reanalysis centers.

Bosilovich et al. (2009) demonstrated that a number of errors in individual analyses are reduced when combined as an ensemble average. We have recently included the ensemble mean of the individual reanalyses and standard deviation for each month of key fields from the atmospheric reanalyses. A preliminary view of the Multiple Reanalysis Ensemble (MRE) suggests a similar result for a few selected variables in most of the reanalyses used to make the ensemble mean. To perform the ensemble mean, it was necessary to regrid each reanalysis to a standard horizontal and vertical grid. This data is also included in the CREATE ESGF site.

Ocean Reanalyses

The ocean reanalysis intercomparison work is a collaboration with the Climate and Ocean: Variability, Predictability and Change (CLIVAR) Global Synthesis and Observations Panel (GSOP). The objective was to develop ensemble averages and spread products for monthly, three-dimensional potential temperature, salinity, and zonal and meridional velocities from eight ocean reanalyses.

The monthly fields from each ocean reanalyses were horizontally regridded and vertically interpolated to a common 1x1-degree latitude-longitude grid and common World Ocean Atlas 2009 depths. These data, along with the native grids, are also published on the CREATE ESGF site and adhere to the same common CF standards as the atmospheric reanalyses. In addition to the regridded and native gridded data, we have provided the ensemble mean and standard deviation for each variable as additional products.

Land Surface Air Temperature

The reanalysis-based hourly land surface air temperature data set (Wang and Zeng 2013) was also reformatted to adhere to the CMIP5 standards and then published in the CREATE ESGF repository. The primary purpose of creating this data set was to produce a bias-corrected hourly land 0.5x0.5-degree surface air temperature product using both reanalysis and the Climate Research Unit high-resolution surface air temperature data set (New et al. 2002).

Scientific Use Cases

One of the primary benefits for gathering the reanalyses onto one site, and formatting them similarly, is the ease with which a user can both access the data and analyze it with a single workflow. Since the data is organized by variable, it is a simple task to retrieve the same variable from several reanalyses and extract all or part of the time series. The data can be regridded to a common grid if desired, and for the three-dimensional data, all the atmospheric reanalyses include the common CMIP5 pressure levels.

One of the motivations for comparing reanalyses is to determine what variables are the most reliable estimates of the real world. For example, Figure 2 shows the zonally averaged precipitation anomaly for five of the CREATE atmospheric reanalyses for the period 1980–2012 compared with GPCP. All the reanalyses suggest an increase in tropical precipitation sometime after 2000, whereas GPCP does not show this as clearly. The figure shows a large variation in decadal variability and more reliable short term-variation as indicated by the relatively consistent signal of the 1997–98 El Niño (CCSP, 2008).

Another example that demonstrates possible tests of reliability is the ensemble average and spread of the eight ocean reanalyses' salinity at 5 meters for January 1998 shown in Figure 3. The spread among the ensemble in the central Pacific near the 1998 El Niño maximum indicates some disagreement among the ensemble members.

Example of 6-Hour History Data

141 The 6-hour files are very large, with a single month requiring 4 gigabytes and 37 years requiring
142 1.7 terabytes. In order to study a single event, it is possible to select a time and region subset
143 using the THREDDS OPeNDAP server either from the NASA Center for Climate Simulation
144 (NCCS) THREDDS data server or from ESGF (additional information can be found here:
145 https://cds.nccs.nasa.gov/data/faq/). None of the reanalysis data centers provide this type of data
146 service. For example, it is simple to download a subset of the precipitation for a series of
147 atmospheric rivers that deluged California in January 2017. Figure 4 shows the average of four
148 grid points surrounding the latitude and longitude of San Francisco, California. In general, all the
149 reanalyses shown captured the precipitation during these events but there are some differences in
150 structure during the peak times.
151
152

Data Access

Downloading reanalysis data

A convenient way to download the CREATE data is to access the ESGF/CoG site. It is first necessary to register with ESGF at https://www.earthsystemcog.org/projects/cog/doc/esgf/faq/login . Once the credentials have been obtained, then log in and go to https://esgf.nccs.nasa.gov/projects/create-ip/.

For those unfamiliar with the ESGF/CoG see: https://www.earthsystemcog.org/projects/cog/tutorials_web. For additional help, there is a "?" to the left of the "Search" button on https://esgf.nccs.nasa.gov/search/create-ip/ to aid in searching and downloading variables.

Once data is selected and placed in the cart, there are several options for downloading, including HTTP, WGET, and OPeNDAP. For large data sets, such as 6-hour atmospheric temperature, the OPeNDAP option is often the most convenient. For OPeNDAP-enabled analysis or visualization tools that are unable to work with ESGF's credentialed access, a second option is to use the NCCS/THREDDS server: https://dataserver.nccs.nasa.gov/thredds/catalog/bypass/CREATE-IP/reanalysis/catalog.html.

Visualization

Another option is to view the data before downloading. The CREATE-V website (https://cds-cv.nccs.nasa.gov/CREATE-V/) has implemented a quick look comparison capability that allows a user to display atmospheric data (monthly or 6-hour) or ocean data (monthly).

Numerous variables are available for display with options to select date, level, color map, and scale. Figures 5 and 6 show screenshots of the CREATE-V interface displaying multiple reanalyses and the ability to produce anomalies of the selected variable by picking a location on the map. The anomalies are calculated on-the-fly using a prototype of a climate data analytics tool in development at the NCCS. Developing this interface was considerably simplified by standardizing the data. Improvments are planned for this site including calculating differences and anomalies and simple contour interval selection. This tool complements the WRIT tool developed at NOAA (Smith et al. 2014) and the KNMI tools found at: https://climexp.knmi.nl/start.cgi?id=someone@somewhere.

Example Using Jupyter Notebooks to Perform Analytics

It is also possible to perform calculations and create plots without directly downloading the data. An example of such access is demonstrated in the sample Jupyter Notebook https://esgf.nccs.nasa.gov/projects/create-ip/use_cases. This example can be run on computers running Linux or MacOS X after installing Anaconda (https://www.continuum.io/) and following the instructions at https://github.com/UV-CDAT/uvcdat/wiki/Install-using-Anaconda. Data is accessed through the NASA/NCCS THREDDS server mentioned above. The notebook contains a very simple Python script that opens and extracts surface temperature from two different reanalyses, calculates the monthly global mean, and plots a graph for the reanalysis period. Those familiar with Python may want to modify the script or add other reanalyses.

Updates

For studies of recent events using reanalysis data, we have automated processing of the atmospheric data. As each center completes their processing of a new month, the NCCS runs automatic processing scripts to download and reformat the data and make it available through ESGF and THREDDS shortly thereafter. This allows for studying events such as the la El Niño and the unusually intense precipitation on the U.S. West Coast during the winter/spring of 2017. The selected data is published at monthly and 6-hour intervals (Table 1). The 20CRv2c, JRA-25 and MERRA are not automatically updated because all three reanalyses ended in or before 2016.

Current Science Applications for CREATE

Recently, the Stratosphere-troposphere Processes and their Role in Climate (SPARC) Stratospheric Reanalysis Intercomparison Team (S-RIP) performed a comprehensive assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses (Davis et al., 2017). The work makes use of the atmospheric CREATE data in determining the utility of stratospheric water vapor from reanalyses.

Future Plans

We hope to include the ERA20C and ERA5. We also hope to add additional value-added products such as the resultant wind velocity.

Another contribution to CREATE will be the inclusion of the reanalysis increments. The reanalysis centers typically do not release these fields, but they would contribute significantly to understanding of the reanalyses error estimates. The difficulty with using these test statistics is that the reanalysis centers do not use the same standards for generating the fields (Gregow 2014).

Summary

This paper summarizes the CREATE project. The primary goals of CREATE are to serve the CMIP and reanalysis science community with easy access to the major atmospheric and ocean reanalysis products to facilitate comparison and to aid in evaluation of the various reanalysis efforts operating today. One of the primary benefits of using CREATE for accessing reanalysis data is the time saved in the processing. For most reanalysis data sets, accessing many variables from the original site and reformatting the data for a consistent workflow requires a time commitment on the order of weeks. Using CREATE to access the data reduces the time required to prepare the data for analysis to a few hours. CREATE is closely coupled to reanalysis.org and

230 is coordinated with the WCRP Task Team for the Intercomparison of ReAnalyses (TIRA).

231

232

233 **References**

234 CCSP, 2008: Reanalysis of Historical Climate Data for Key Atmospheric Features: Implications
235 for Attribution of Causes of Observed Change. *A Report by the U.S. Climate Change*
236 *Science Program and the Subcommittee on Global Change Research* [Randall Dole, Martin
237 Hoerling, and Siegfried Schubert (eds.)]. National Oceanic and Atmospheric
238 Administration, National Climatic Data Center, Asheville, NC, 156 pp.

239 Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean
240 reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.*, **139**, 1132–1161, doi:10.1002/qj.2063.

241 Bosilovich, M. G., D. Mocko, J. O. Roads, and A. Ruane, 2009: A Multimodel Analysis for the
242 Coordinated Enhanced Observing Period (CEOP). *J. Hydrometeorol.*, **10**, 912–934,
243 doi:10.1175/2009JHM1090.1.

244 Cinquini, L., and Coauthors, 2014: The Earth System Grid Federation: An open infrastructure for
245 access to distributed geospatial data. *Futur. Gener. Comput. Syst.*, **36**, 400–417,
246 doi:10.1016/j.future.2013.07.002.
247 http://linkinghub.elsevier.com/retrieve/pii/S0167739X13001477.

248 Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Q. J. Roy.*
249 *Meteorol. Soc.*, **137**, 1–28, doi:10.1002/qj.776.

250 Davis, S. M., and Coauthors, 2017: Assessment of upper tropospheric and stratospheric water
251 vapor and ozone in reanalyses as part of S-RIP. *Atmos. Chem. Phys*., **17,** 12743-12778,

doi.org/10.5194/acp-17-12743-2017

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Roy. Meteorol. Soc.*, **137**, 553–597, doi:10.1002/qj.828. http://doi.wiley.com/10.1002/qj.828..

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.*, **2**, JCLI-D-16-0758.1, doi:10.1175/JCLI-D-16-0758.1. http://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0758.1.

Gregow, H., 2014: COordinating Earth observation data validation for RE-analysis for CLIMAte ServiceS, (CORE CLIMAX). Unpublished Report, Deliverable D5.53, Procedure for comparing reanalyses , and comparing reanalyses to assimilated observations and CDRs. 1–45.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc.*, **77**, 437–471, doi:http://dx.doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2..

Köhl, A., 2015: Evaluation of the GECCO2 ocean synthesis: Transports of volume, heat and freshwater in the atlantic. *Q. J. Roy. Meteorol. Soc.*, **141**, 166–181, doi:10.1002/qj.2347.

New, M., D. Lister, M. Hulme, and I. Makin, 2002: A high-resolution data set of surface climate over global land areas. *Clim. Res.*, **21**, 1–25.

Onogi, K., and Coauthors, 2007: The JRA-25 Reanalysis. *J. Meteorol. Soc. Japan*, **85**, 369–432, doi:10.2151/jmsj.85.369.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1. http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00015.1 (Accessed July 9, 2014).

S. Kobayashi, Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. T., 2014: The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *J. Meteorol. Soc. Japan*, **93**, doi:10.2151/jmsj.2015-001. http://jmsj.metsoc.or.jp/EOR/JMSJ2015-01.html.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1. http://journals.ametsoc.org/doi/abs/10.1175/2010BAMS3001.1 (Accessed July 14, 2014).

Smith, C. A., G. P. Compo, and D. K. Hooper, 2014: Web-based Reanalysis Intercomparison Tools (WRIT) for analysis and comparison of reanalyses and other datasets. *Bull. Am. Meteorol. Soc.*, **95**, 1671–1678, doi:10.1175/BAMS-D-13-00192.1.

Storto, A., S. Masina, and A. Navarra, 2016: Evaluation of the CMCC eddy-permitting global ocean physical reanalysis system (C-GLORS, 1982-2012) and its assimilation components. *Q. J. Roy. Meteorol. Soc.*, **142**, 738–758, doi:10.1002/qj.2673.

Taylor, K. E., R. J. Stouffer, and G. a. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bull. Am. Meteorol. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1. http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-11-00094.1.

Toyoda, T., Y. Fujii, T. Yasuda, N. Usui, K. Ogawa, T. Kuragano, H. Tsujino, and M. Kamachi, 2016: Data assimilation of sea ice concentration into a global ocean–sea ice model with

293 corrections for atmospheric forcing and ocean temperature fields. *J. Oceanogr.*, **72**, 235–
294 262, doi:10.1007/s10872-015-0326-0.

295 Wang, A., and X. Zeng, 2013: Development of Global Hourly 0.5° Land Surface Air
296 Temperature Datasets. *J. Clim.*, **26**, 7676–7691, doi:10.1175/JCLI-D-12-00682.1.
297 http://dx.doi.org/10.1175/JCLI-D-12-00682.1.

298 Williams, D., 2013: Ultrascale Visualization of Climate Data. *IEEE Comput. Soc. Perspect.*, **46**,
299 70–78, doi:10.1109/MC.2013.119.
300 http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6487516.

301 Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg, 2007: System Design and Evaluation of
302 Coupled Ensemble Data Assimilation for Global Oceanic Climate Studies. *Mon. Weather*
303 *Rev.*, **135**, 3541–3564, doi:10.1175/MWR3466.1.
304 http://journals.ametsoc.org/doi/abs/10.1175/MWR3466.1.

305 Zuo, H., M. A. Balmaseda, and K. Mogensen, 2015: The new eddy-permitting ORAP5 ocean
306 reanalysis: description, evaluation and uncertainties in climate signals. *Clim. Dyn.*,
307 doi:10.1007/s00382-015-2675-1. http://link.springer.com/10.1007/s00382-015-2675-1.

308

309

310 **Tables**

311

20CRv2c	NOAA/CERES 20th Century Reanalysis
C-GLORSv5	Latest ocean reanalysis produced at the CMCC
CDO	Climate Data Operators
CF	Climate and Forecast (conventions)
CLIVAR	Climate and Ocean - Variability, Predictability, and Change
CMCC	Euro-Mediterranean Center on Climate Change
CMIP5	Coupled Model Intercomparison Project Phase 5
CMOR	Climate Model Output Rewriter
CoG	Earth System Community of Governance to support distribution of data
CREATE	Collaborative REAnalysis Technical Environment
ECDA	Ensemble Coupled Data Assimilation
ECMWF	European Centre for Medium-Range Forecasts
ERA	ECMWF Reanalysis
ESGF	Earth System Grid Federation
ESRL	Earth System Research Laboratory
FTP	File Transfer Protocol
GECCO2	German contribution of the Estimating the Circulation and Climate of the Ocean project
GES DISC	Goddard Earth Sciences Data and Information Center
GFDL	Geophysical Fluid Dynamics Laboratory
GODAS	NCEP Global Ocean Data Assimilation System (GODAS)

GRIB	GRIdded Binary
GSOP	Global Synthesis and Observations Panel
HTTP	Hypertext Transfer Protocol
JMA	Japan Meteorological Agency
JRA-25	Japanese 25-year Reanalysis
JRA-55	Japanese 55-year Reanalysis
MARS	Meteorological Archival and Retrieval System
MERRA	Modern-Era Retrospective analysis for Research and Applications
MERRA-2	Modern-Era Retrospective analysis for Research and Applications Version 2
MOVE	Multivariate Ocean Variational Estimator
MRE	Multiple Reanalysis Ensemble
MRI	Meteorological Research Institute
NASA	National Aeronautics and Space Administration
NCCS	NASA Center for Climate Simulation
NCEP	National Center for Environmental Prediction
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
OPeNDAP	Open-source Project for a Network Data Access Protocol
ORA-S4	Operational Ocean Reanalysis Version S4
ORAP5.0	Ocean Reanalysis Pilot 5
S-RIP	SPARC Reanalysis Intercomparison Project
SPARC	Stratosphere-troposphere Processes And their Role in Climate
THREDDS	Thematic Real-time Environmental Distributed Data Services

TIRA	WCRP Task Team for Intercomparison of Reanalyses
WGET	GNU package for retrieving files
WRIT	Web-based Reanalysis Intercomparison Tools

312

313 **Table 1.** List of acronyms

314

315

Variable Long Name	CMOR	20CRv2c	CFSR	ERA-I	MERRA	MERRA-2	JRA-25	JRA-55
Air Pressure at Convective Cloud Base	ccb		✓					
Air Pressure at Convective Cloud Top	cct		✓					
Air Temperature *	ta	✓	✓	✓	✓	✓	✓	✓
Cloud Area Fraction	cl				✓	✓		
Condensed Water Path	clwvi			✓				
Convective Precipitation	prc	✓	✓	✓	✓	✓		
Daily Maximum Near-Surface Air Temperature	tasmax	✓	✓			✓	✓	
Daily Minimum Near-Surface Air Temperature	tasmin	✓	✓			✓	✓	
Eastward Near-Surface Wind *	uas	✓	✓	✓	✓	✓	✓	✓
Eastward Wind *	ua		✓	✓	✓	✓	✓	✓
Evaporation *	evspsbl		✓	✓	✓	✓		✓
Geopotential Height *	zg	✓	✓	✓	✓	✓	✓	✓
Ice Water Path	clivi	✓		✓				
Mole Fraction of O3 *	tro3	✓	✓	✓	✓	✓	✓	✓

Near-Surface Air Temperature *	tas	✓	✓	✓	✓	✓	✓	✓
Near-Surface Relative Humidity	hurs	✓	✓				✓	
Near-Surface Specific Humidity	huss	✓	✓			✓	✓	
Northward Near-Surface Wind *	vas	✓	✓	✓	✓	✓	✓	✓
Northward Wind *	va	✓	✓	✓	✓	✓	✓	✓
Omega	wap	✓	✓	✓	✓	✓		
Precipitation *	pr	✓	✓	✓	✓	✓		✓
Relative Humidity *	hur	✓	✓	✓	✓	✓		✓
Sea Level Pressure *	psl	✓	✓	✓	✓	✓	✓	✓
Snowfall Flux	prsn			✓	✓	✓		
Specific Humidity *	hus	✓	✓	✓	✓	✓	✓	✓
Surface Air Pressure *	ps	✓	✓	✓	✓	✓	✓	✓
Surface Downward Eastward Wind Stress	tauu	✓	✓	✓	✓	✓		
Surface Downward Northward Wind Stress	tauv	✓	✓	✓	✓	✓		

Surface Downwelling Clear-Sky Longwave Radiation	rldscs		✓		✓	✓	✓	✓
Surface Downwelling Clear-Sky Shortwave Radiation	rsdscs		✓		✓	✓	✓	✓
Surface Downwelling Longwave Radiation	rlds	✓	✓	✓	✓	✓	✓	✓
Surface Downwelling Shortwave Radiation	rsds	✓	✓	✓	✓	✓	✓	✓
Surface Snow and Ice Sublimation Flux	sbl		✓					
Surface Temperature *	ts	✓	✓	✓	✓	✓		
Surface Upward Latent Heat Flux	hfls	✓	✓	✓	✓	✓		
Surface Upward Sensible Heat Flux	hfss	✓	✓	✓	✓	✓		
Surface Upwelling Clear-Sky Shortwave Radiation	rsuscs		✓			✓	✓	✓
Surface Upwelling Longwave Radiation	rlus	✓	✓	✓	✓	✓	✓	✓
Surface Upwelling Shortwave Radiation	rsus	✓	✓	✓		✓	✓	✓
TOA Incident Shortwave Radiation	rsdt		✓	✓	✓	✓	✓	✓

TOA Outgoing Clear-Sky Longwave Radiation	rlutcs		✓	✓	✓	✓	✓	✓
TOA Outgoing Clear-Sky Shortwave Radiation	rsutcs		✓	✓	✓	✓	✓	✓
TOA Outgoing Longwave Radiation	rlut	✓	✓	✓	✓	✓	✓	✓
TOA Outgoing Shortwave Radiation	rsut	✓	✓	✓	✓	✓	✓	✓
Total Cloud Fraction *	clt	✓	✓	✓	✓	✓	✓	✓
Transpiration (Land)	tran		✓			✓		
Water Vapor Path *	prw	✓	✓	✓	✓	✓	✓	✓

316

317 Table 2. The standard Climate and Forecast (CF) long names of the atmospheric variables

318 published in CREATE on ESGF. The data is monthly average.

319 * indicates the data is also 6-hour.

320

321

Variable Long Name	CMOR Variable	Realm	20CRv2c	CFSR
Evaporation from Canopy	evspsblveg	land		✓
Moisture in Upper Portion of Soil Column	mrsos	land		✓
Surface Runoff	mrros	land		✓
Temperature of Soil	tsl	land	✓	✓
Total Runoff	mrro	land		✓
Total Soil Moisture Content	mrso	land		✓
Water Content of Soil Layer	mrlsl	land		✓
Water Evaporation from Soil	evspsblsoi	land		✓
Liquid water Content of Snow Layer	lwsnl	landIce	✓	
Snow Area Fraction	snc	landIce	✓	✓
Snow Depth	snd	landIce	✓	✓
Surface Snow Amount	snw	landIce		✓
Sea Ice Area Fraction	sic	seaIce	✓	✓
Sea Ice Thickness	sit	seaIce	✓	✓
Snow Depth	snd	seaIce		✓

322

323 Table 3. The standard Climate and Forecast (CF) long names of the land surface variables

324 published in CREATE on ESGF. The data is monthly average.

325

Variable Long Name	CMOR variable	CFSR	CLIVAR/GSOP	C-GLORSv5	ORAP5	ORAS4	MOVE-G2i	GECCO2	ECDAv31	GODAS
Downward Heat Flux at Sea Water Surface	hfds	✓								
Sea Surface Height Above Geoid	zos	✓								
Sea Surface Temperature	tos	✓								
Sea Water Potential Temperature	thetao	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sea Water Salinity	so	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sea Water X Velocity	uo	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sea Water Y Velocity	vo	✓	✓	✓	✓	✓	✓	✓	✓	✓

326

327 Table 4. The standard Climate and Forecast (CF) long names of the ocean variables published in

328 CREATE on ESGF. The data is monthly average.

329

330 **Figure Captions**

331

332 Figure 1. A schematic view of the processing necessary to obtain, convert, and reformat
333 reanalysis data to the CMIP standards.

334

335 Figure 2. Zonal Average Precipitation Anomaly from 1980–2016 (mm/day) for five reanalyses.
336 compared to GPCP.

337

338 Figure 3. The January 1998 ocean ensemble salinity at 5 meters depth (psu), a: mean. b: standard
339 deviation.

340

341 Figure 4. 6-hour data from MERRA-2, CFSR, ERA-Interim, and JRA-55 precipitation in
342 mm/day for the month of January 2017. The data represents the average of ±1º around 38º N
343 latitude and 122º W longitude (San Francisco, California).

344

345 Figure 5. Screenshot of the CREATE-V interactive tool interface to display multiple reanalyses
346 side-by-side. See https://cds-cv.nccs.nasa.gov/CREATE-V/.

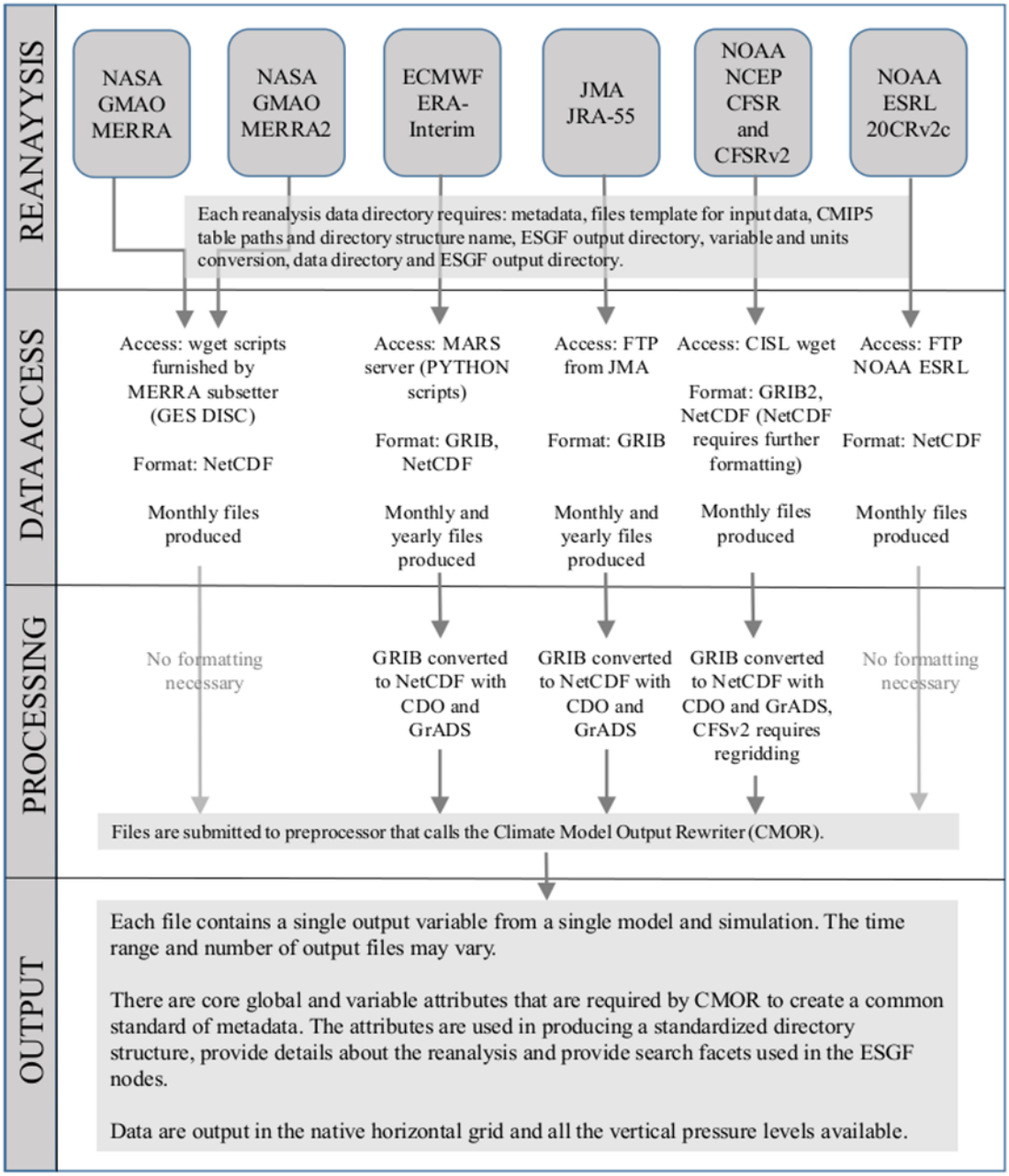
347

348 Figure 6. Using the CREATE-V visualization tool to zoom in on a region and select a point
349 producing the monthly anomaly and the average seasonal cycle of the selected location.

350

351 **Figures**

352



353

354 Figure 1. A schematic view of the processing necessary to obtain, convert, and reformat

355 reanalysis data to the CMIP standards.

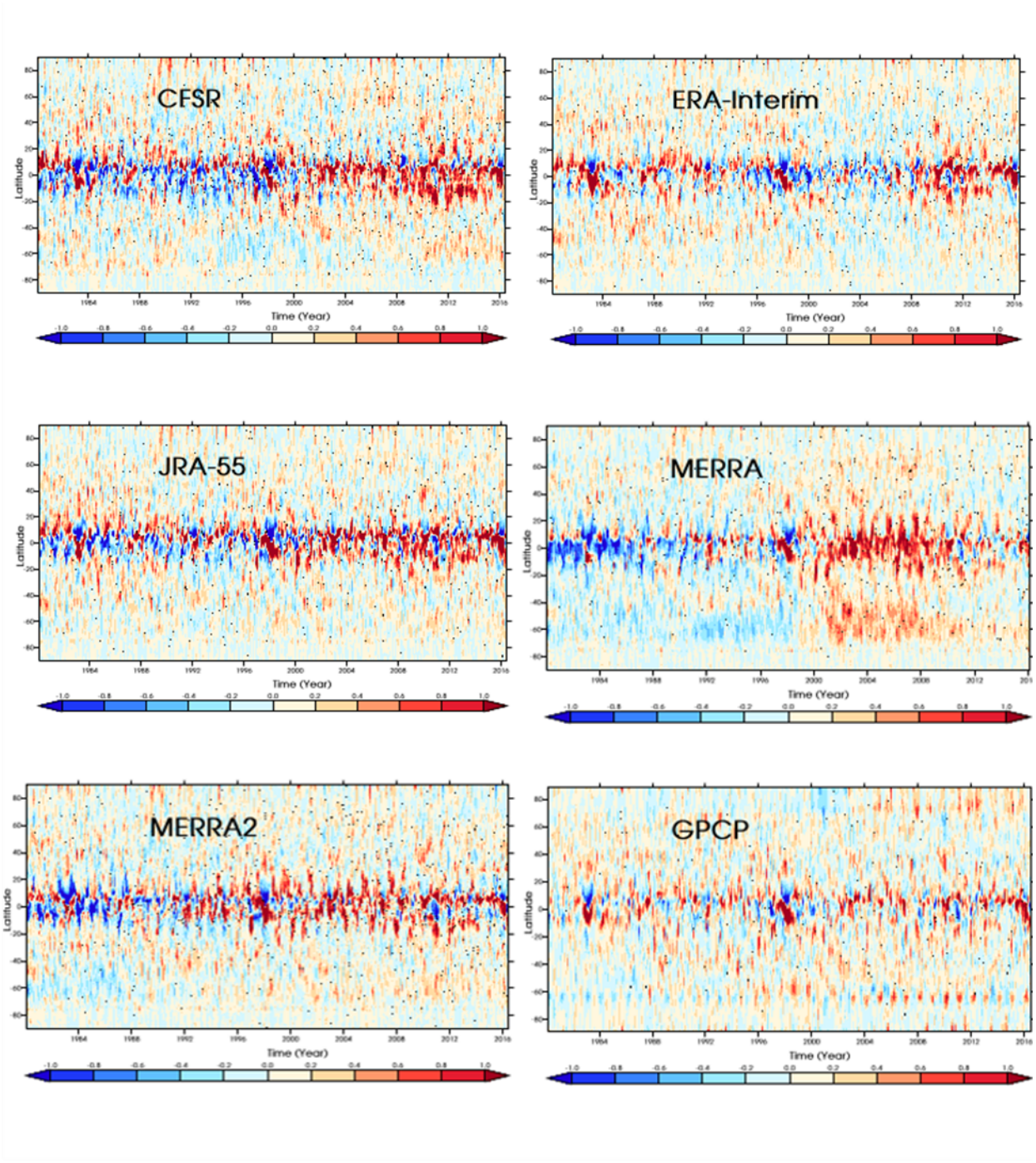


Figure 2. Zonal Average Precipitation Anomaly from 1980–2016 (mm/day) for five reanalyses. compared to GPCP

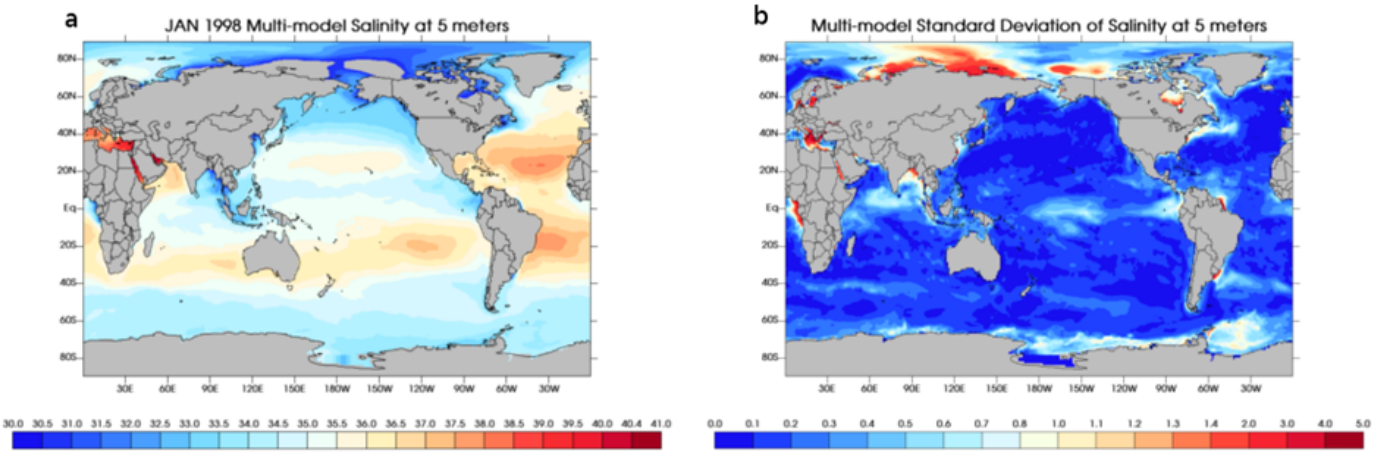
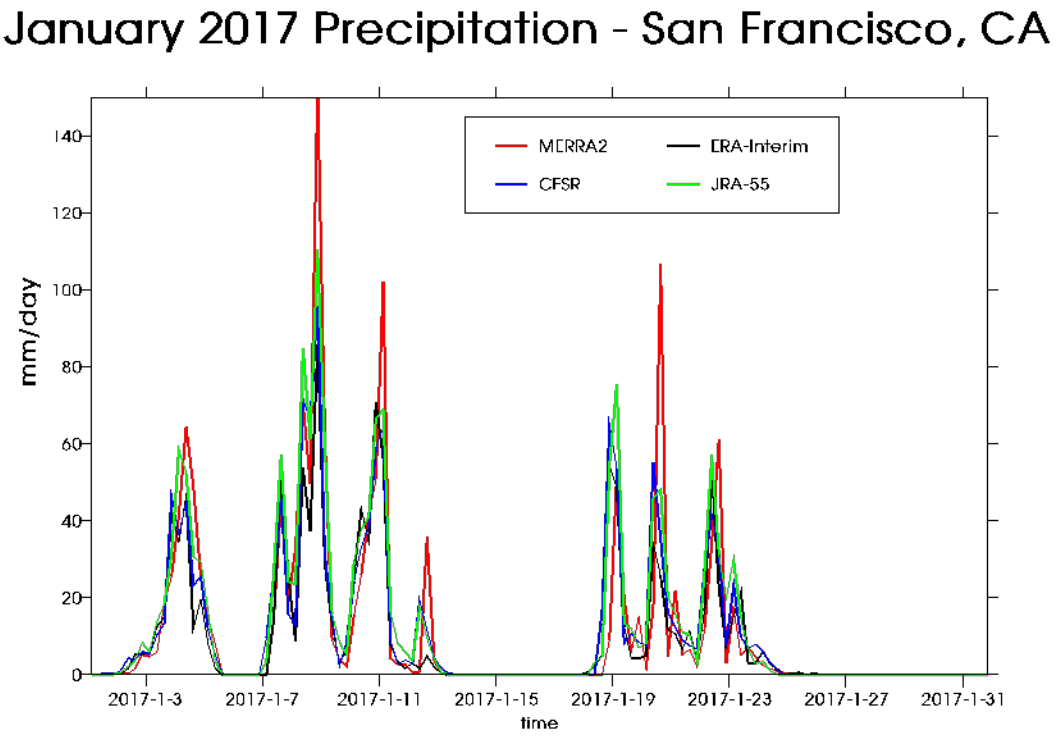


Figure 3. a: The January 1998 ocean ensemble mean salinity at 5 meters (psu). b: The ocean ensemble salinity spread (standard deviation).



366

367 Figure 4. 6-hour data from MERRA-2, CFSR, ERA-Interim, and JRA-55 precipitation in
368 mm/day for the month of January 2017. The data represents the average of ±1º around 38º N
369 latitude and 122º W longitude (San Francisco, California).

370

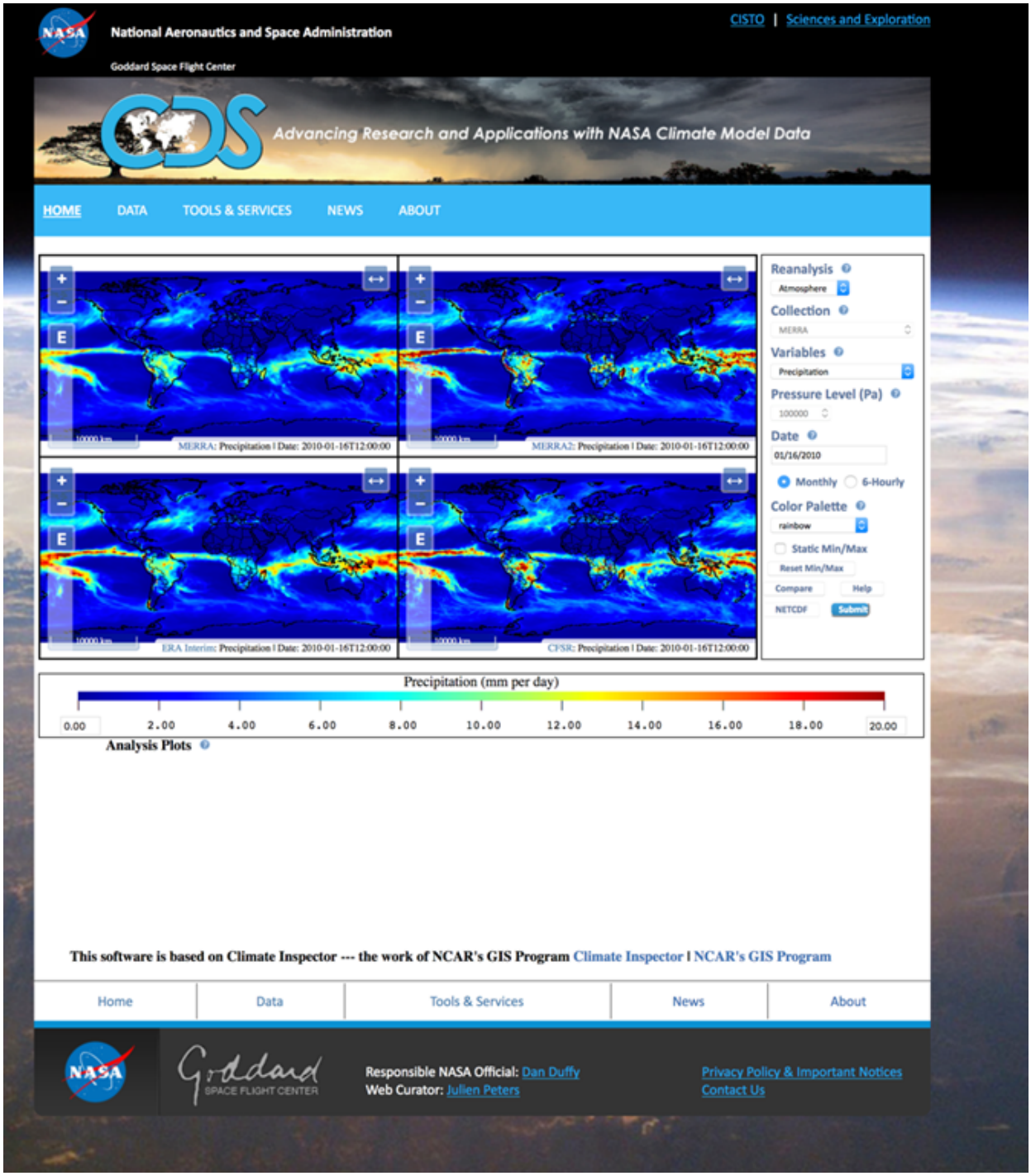


Figure 5. Screenshot of the CREATE-V interactive tool interface to display multiple reanalyses side-by-side. See https://cds-cv.nccs.nasa.gov/CREATE-V/.

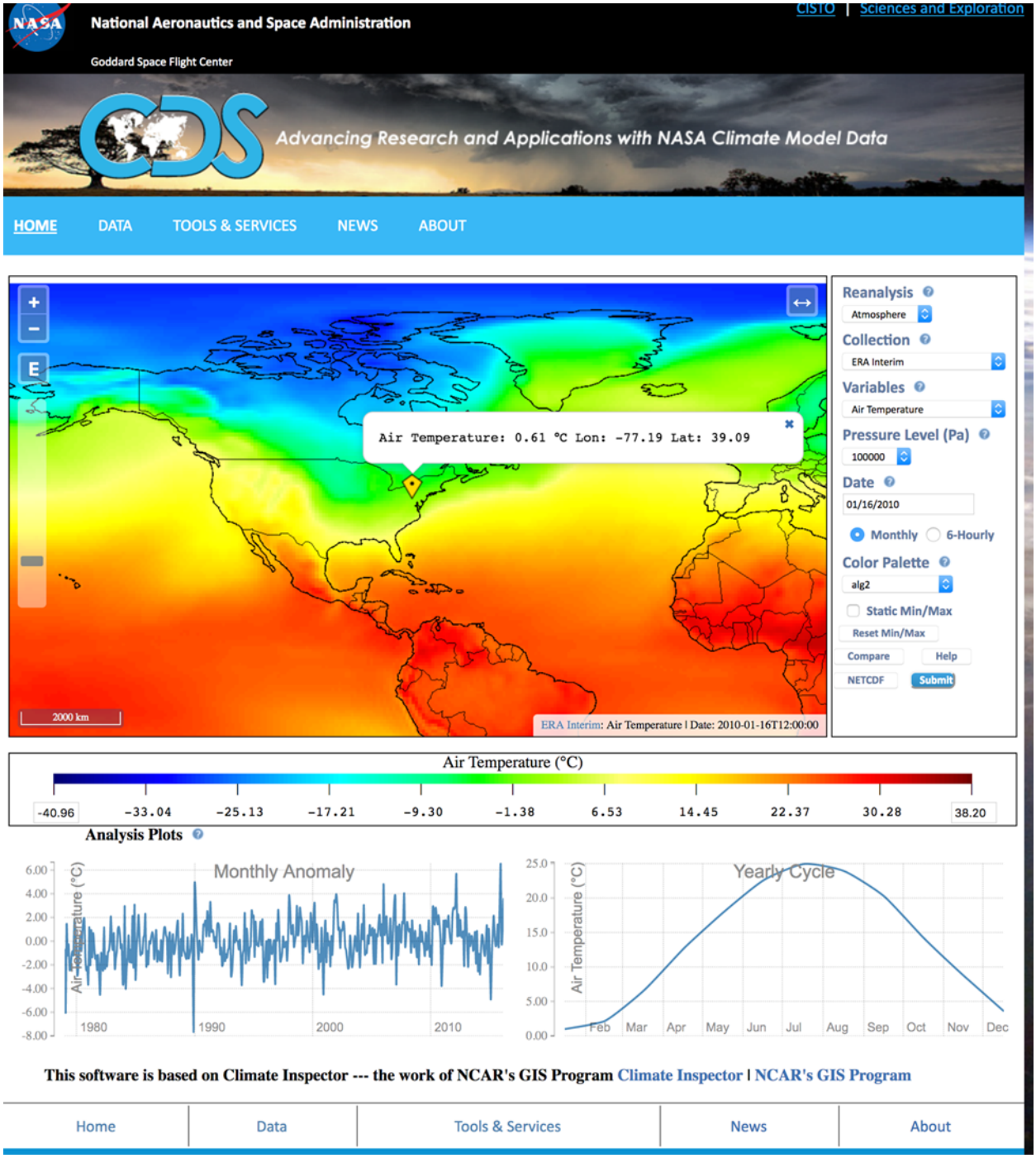


Figure 6. Using the CREATE-V visualization tool to zoom in on a region and select a point producing the monthly anomaly and the average seasonal cycle of the selected location.