Educational and Scientific Applications of Climate Model Diagnostic Analyzer

Seungwon Lee, Chengxing Zhai, Terence Kubar, Benyang Tang, Lei Pan, and Juilin Li
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA, USA
Seungwon.Lee@jpl.nasa.gov

Jia Zhang and Qihao Bao
Department of Computer Science
Carnegie Mellon University
Moffett Field, CA, USA
Jia.Zhang@sv.cmu.edu

Abstract— **Climate Model Diagnostic Analyzer (CMDA) is a web-based information system designed for the climate modeling and model analysis community to analyze climate data from models and observations. CMDA provides tools to diagnostically analyze climate data for model validation and improvement, and to systematically manage analysis provenance for sharing results with other investigators. CMDA utilizes cloud computing resources, multi-threading computing, machine-learning algorithms, web service technologies, and provenance-supporting technologies to address technical challenges that the Earth science modeling and model analysis community faces in evaluating and diagnosing climate models. As CMDA technology and infrastructure have matured, we have developed the educational and scientific applications of CMDA. Educationally, CMDA supported the summer school of the JPL Center for Climate Sciences in 2014, 2015, and 2016. In the summer school, the students work on group research projects where CMDA provide datasets, analysis tools, and provenance support utility tools. Each student is assigned to a virtual machine with CMDA installed in Amazon Web Services. Scientifically, we have developed several science use cases of CMDA covering various topics, datasets, and analysis types. Each of the science use cases is described in terms of a scientific goal, datasets used, the analysis tools used, scientific results discovered, an analysis result such as output plots and data files, and a link to the corresponding analysis service call with all the input arguments filled.**

Keywords—climate model evaluation; cloud computing; provenance; climate data; diagnostic analysis

I. INTRODUCTION

Understanding climate change at various temporal and spatial scales and related human influences becomes ever more important as the public's concerns about climate changes have steadily grown over the past years. However, achieving this goal has remained a tremendous challenge to the Earth Science community, manifested by the persistence of the large spreads in the model projections from sub-grid scale processes to global-scale variability [1]. These large uncertainties in the model prediction are due to the fact that the climate changes involve many complex feedback processes that are either poorly represented or not yet included in the models [1]. A better understanding of the processes and in return a better representation of the processes in the models are needed to improve the model fidelity. The process-level data analysis and model evaluation can potentially open a new direction in climate modeling that can significantly improve the model representation of the current climate and future climate prediction capabilities.

Both the National Research Council (NRC) Decadal Survey and the latest Intergovernmental Panel on Climate Change (IPCC) Assessment Report stressed the need for the comprehensive and innovative evaluation of climate models with the synergistic use of global observations in order to maximize the investments made in Earth observational systems and also to capitalize on them for improving our weather and climate simulation and prediction capabilities [1,2,3]. The abundance of satellite observations for fundamental climate parameters and the availability of coordinated model outputs from the Coupled Model Intercomparison Project Phase 5 and 6 (CMIP5 and CMIP6) for the same parameters offer a great opportunity to understand and diagnose model biases in climate models [4,5]. In addition, the Obs4MIPs efforts have created several key global observational datasets that are readily usable for model evaluations [6].

NASA, NOAA, and DOE have established a collection of data centers to store and distribute the rapidly growing satellite-based and ground-based sensor data and model-generated data. The Earth Science community also has developed a number of analysis tools to process such datasets. The examples of the community tools include NetCDF Operators (NCO) [7,8], NCAR Command Language (NCL) [9], Climate Data Operators (CDO) [10], Iris Python library [11], CF-Python [12], Program for Climate Model Diagnosis and Intercomparison (PCMDI) Metrics Package [13], Community Intercomparison Suite (CIS) [14], UV-CDAT [15], and ESMValTool [16]. All of these tools are designed to be downloaded and installed in a user's machine. And, they do not provide a mechanism to keep track of analysis history and to reproduce and share analysis provenance and result with others. With the exponential growth of datasets and analysis tools, scientists are struggling to keep track of their datasets (original or derived), tools, and analysis history and results, let along sharing them with others. Infrastructure tools to support sharing the analysis tools and knowledge drived from the data analysis are needed.

In response to the community's need, we have developed an online analysis sytem called Climate Model Diagnostic

978-1-5386-1996-4/17 $31.00 © 2017 IEEE
DOI 10.1109/BigDataCongress.2017.57

Analyzer (CMDA) [17,18]. CMDA is a collection of online web services for multi-variable climate model diagnostic analysis through the comprehensive use of multple observaitonal data, reanalysis data, and model outputs. CMDA provides an online collaboraborative environment, where scientists can design their analysis workflow, organize their work efficiently, and share their work with others. A CMDA server hosts all its analysis tools as web services that can be assessible through a web browser on a user machine and therefore does not require a local installation of CMDA in a user's machine. The server service approach ensures quality-controlled performance with required computational resources and data storage.

In this paper, we describe the main technologies and components of CMDA in Section II. We describe the educational use of CMDA in Section III, and the scientific use of CMDA in Section IV. We summarize the perceived impact and future development and infusion direction of CMDA in climate model evaluation in Section V.

II. TECHNOLOGIES USED FOR CMDA

A. *Web Service Technology*

Many of research codes are written in a non-general and non-scalable way, making it difficult to share with others. In addition, the programming languages and libraries used by the code often require a local software installation, environment configuration, software license (e.g. Matlab and IDL), making it difficult for others to adopt the tool. We have developed a methodology to transform an existing science application code in various programming languages into a web service. A web service on a server approach is chosen because it not only lowers the learning curve and remove the adoption barrier of the tool but also enables instantaneous use compared to offline standalone applications, avoiding the hassle of local software installation and environmen incompatibility. The web service technology also has a simple and flexible environment with a rich set of open source packages.

The CMDA methodology of transforming an existing science application code into a web service has the following steps. (1) We wrap an existing science application code with a Python caller. The Python caller treats the application as a process, defines where to put the outputs of the child process, spawns off the child process, captures the stdout and stderr of the child process. At the end , the science application looks like a python application. If an application is written in license-required programming languages like Matlab and IDL, we either translate it into an equivalent free-license version (Octave for Matlab) or rewrite it into a Python code so we can avoid the license requirement. (2) We use Flask, an open source light-weight web development framework for Python applications, to create an entry point code for a web service. The entry code parses input arguments from a client (a web service), call a Python application, passes input arguments ot the Python application, retrieve return values from the Python application, and pass them to a client. It follows a REST-ful (Representational State Transfer) style, where scoping information (what data to operate) is placed in a URI (Uniform Resource Identifier) while method information (what to do with the data) is conveyed in an HTTP (Hypertext Transfer Protocol) method [19]. (3) We separate application traffic from static HTTP traffic. We use Gunicorn to provide WSGI service application traffic for web service scoping and method information, while we use Tornado to provide web service static HTTP traffic for web service results. (4) We design and implement a web browser interface for a web service and implement it using JavaScript.

B. *Online Collaborative Environment*

We have built an online collaborative environment in CMDA, where a community can build, share, search, and recommend web services for climate data analytics and organize their execution history. The key functionalities are CMDA service and dataset search and recommendation and CMDA service execution history management. We have applied mature semantic web techniques and machine learning techniques to build an intelligent search facility [20]. Furthermore, we used the most recent web techniques (including HTML5, JavaScript, Apache Lucene, Play framework) and modern software engineering methodologies (including Extreme Programming, Agile technique, and Scrum) to develop a scalable, extensible, and interoperable online environment [21,22]

A system with strong query and recommendation facility requires an underlying semantics model. We have developed both static semantics and behavioral semantics: static semantics describes the functionalities and goals of a CMDA service, and behavioral semantics describes the required circumstances when a CMDA service can behave, including input and output parameters, pre- and post-constraints, and historical usage patterns. Based on the service semantics model, we have developed a method that can automatically extract aforementioned semantic metadata from CMDA services. In order to support reproducibility, we have developed a provenance model to record and track scientists' activities and behaviors. With the execution history stored in a database, we developed a system to search executions with many search conditions and to reproduce the results with the found execution history.

C. *Analysis Capabilities*

We have developed a set of analysis tools that cover a broad range of analysis types to support model evaluations. Analysis capabilities currently supported by CMDA are (1) the calculation of annual and seasonal means of physical variables, (2) the calculation of time evolution of the means in any specified geographical region, (3) the calculation of an anomaly time series of a physical variable by subtracting its seasonal climatological mean, (4) the calculation of correlation between two variables with a time lag if needed, (5) the calculation of difference between two datasets, (6) the conditional sampling of one physical variable with respect to another variable, (7) the calculation of the conditional probability density function of one physical variable with respect to another, (8) the analysis of Emipirical Orthogonal Function (EOF) to find dorminant modes of spatial variability in time series of a variable, (9) the random-forest based feature importance ranking of a variable with respect to dependences on other variables, and (10) the regridding of datasets with specified horizontal and vertical resolutions. Note that some of

This work was supported by NASA ROSES AIST and CMAC program.

the analysis capabilities represent more than one analysis tool because some tools are designed for 2D (2-dimensional, latitude and longitude) variables only or 3D (3-dimensional, latitude, longitude, and pressure) variables only.

D. Datasets

On its data server, CMDA hosts three types of dataset: model outputs, observation data, and reanalysis data. The model outputs are from CMIP5 models, covering a broad range of atmosphere, ocean, and land variables from CMIP5 experiments such as historical runs, AMIP (Atmospheric Model Intercomparison Project) runs, and RCP (Representative Concentration Pathways) 4.5 experiment runs [4]. The observation data are from Obs4MIPs and a few ocean datasets from NOAA and Argo and serve as reference data for model evaluations. The reanalysis data are ERA-Interim outputs for several environmental variables in order to supplement observation datasets [23]. ERA-Interim is a global atmospheric reanalysis from a data assimilation system developed by European Centre for Medium-Range Weather Forecasts (ECMWF).

In addition to the data hosted on our data server, CMDA allows a user to upload their datasets to our server and use our analysis tools with the uploaded datasets. This feature is useful and critical for making the synergistic use of both our analysis system and the user analysis system. A user can start their analysis from our server, download the analysis result into their machine as a derived dataset, use the dataset with their own analysis tool on their machine, upload their analysis result dataset to our server, use our server for another analysis, and then download the result to their machine. This interactive use of our server with the user's machine makes our analysis appplication more broad and powerful because not all user specific analysis approaches that cannot be supported by analysis tools that we have in CMDA. In order to ensure the seamless exchange of data between the two systems, we make our output data files to be compliant with Climate and Forecast (CF) convention [24]. The CF-convention is a community standard required for CMIP output files. Our tools can read and process user uploaded data files as long as they are CF-convention compliant.

III. Educational Applications of CMDA

CMDA has been used as an educational tool for the summer school organized by JPL's Center for Climate Science in 2014, 2015, and 2016 [25]. The theme of the summer school is using satellite observations to advance climate models, which is well aligned with the main goal and capability of CMDA. The requirements of the educational tool are defined with the interaction with the school organizers, and CMDA is customized to meet the requirements accordingly. Since CMDA needs to be used simultaneously by over 30 users (students and instructors) during the school, we have imported CMDA to cloud computing resources provided by Amazon Web Services (AWS). The cloud-enabled CMDA provides each student with an independent computing resource and working environment while the user interaction with the CMDA system remains the same through its web-browser interface.

Fig. 1. JPL Center for Climate Sciences Summer School in 2016. CMDA provided datasets and tools for students to use for their group research project during the school.

In 2016 summer school, CMDA have provided students with 625 climate datasets and 18 analysis tools. The datasets covered are multi-year monthly gridded data from observations (44 datasets), reanalysis runs (13 datasets), and model runs (568 datasets). The analysis tools process one variable or multi variables for time average, spatial average, correlation, time variability, spatial variability, and conditional sampling. A one-hour session for the CMDA introduction was given, and immediately after the introduction the students were able to start using CMDA with a virtual machine assigned to them in AWS. The students used CMDA for two practice sessions, which lasted about 5 hours total, and were able to present their results of their group research project. Figure 1 shows the pictures of the 2016 summer school during the group research sessions and the student presentation session. During the research sessions, each student group met in a separate room, coordinated their work, performed analyses with CMDA, and discussed their approaches and results. During the student presentation session on the last day of the school, they presented their results as a group presentation.

The summer school in 2016 had six group research topics: (1) where is global warming?; (2) tropical variability and analysis of the El Nino-Southern Oscillation (ENSO) forcing in observations and models; (3) spatial and temporal variability of clouds and precipitation; (4) vegetation phenology and climate controls; (5) land water storage variability as a function of human and natural controls; and (6) sensitivity of equilibrium climate on physical parameterizations.

The description of each topic contains the introduction of the topic, the set of datasets to use, the geographical location suggested, questions to address, and approaches to take for the topic. Here is one example of the topic description.

- Topic: Tropical Variablity and analysis of the El Nino-Southern Oscillation (ENSO) forcing in observations and models
- Introduction: ENSO is a periodical variation in winds, SSTs, and cloud characteristics over the near equatorial

central/eastern Pacific Ocean, with remote influences nearly globally (teleconnections). The irregular periodicity of ENSO (2-7 years) continues to provide challenges; during some years a full-fledged El Nino never develops even in the midst of anomalously warm waters. In this topic, we use statistical techniques, including Empirical Orthogonal Function (EOF) Analysis, lead/lag assessment, and composite analysis to examine the variability of atmospheric and ocean variables, whether any insights can be gleaned about forcing versus response variables, or if ENSO represents a true ocean-atmosphere coupled system for which predictors and responses are difficult to disentangle.

- Datasets: ECMWF Sea-Surface Temperatures (SSTs), near surface winds, cloud fraction, ARGO Ocean Temperatures, NOAA NODC Ocean Heat Content, MODIS cloud top temperature, cloud optical depth (τ), and cirrus reflectance, CERES TOA (Top of Atmosphere) LW (Long-Wave) cloud forcing, and a subset of these variables from both historical coupled climate models and atmosphere-only climate models (AMIP models).
- Geographic Foci: The tropical band, mostly between 25°S and 25°N, and primarily along the Pacific
- Questions and Approaches:
  1. Starting with ECMWF-Interim and observational datasets, perform EOF analysis, beginning with a domain of [100°:290°, 25°S:25°N], and also examine the sensitivity of the domain. For datasets that precede 1998, identify moderate/strong El Nino and La Nina events from Principal Component (PC) Time Series, and compare with published results (e.g. NOAA). Describe the spatial structure of EOF1 (and possibly EOF2), and identify any geographical offsets of variables in terms of the Walker Circulation.
  2. Does the u-wind (eastward wind) ever lead the SST signal, and if so, what does this suggest about the lead/lag and forcing/response of these two variables? How are anomalously strong westerly winds (e.g. westerly wind bursts) geographically related to SST?
  3. Using CERES TOA LW Cloud Forcing and MODIS cloud top temperatures, examine the response of cloud heights as a function of ENSO. Explain in terms of the Stefan-Boltzmann Law, including the shift in the location and height of convection during El Nino events. Is there any signal in cirrus cloud reflectance?
  4. Next, evaluate the ability to simulate ENSO from several of the available historical coupled GCMs. Utilize techniques described above, and characterize fidelity in terms of spatial structure, timing, etc. Are AMIP models more skillful? Explain why or why not.
  5. Use ECMWF-Interim to divide the past ~30 years into decadal chunks. Describe location changes with time in terms of "Central" vs. "Eastern" ENSO events. Can precipitation off the U.S. West Coast be related to the type of El Nino event?

The summer school students presented their research result on the last day of the school. Figure 2 shows key results that the students presented for Topic 2. The key conclusions the students presented are

- Observations show a robust weakening of the Walker Circulation with increased precipitation over the eastern equatorial Pacific during El Niño.
- Coupled models results resemble observations, but AMIP simulations struggle.
- Eastern Pacific El Niño generates more intense equatorial precipitation, a more robust Pacific North American (PNA) pattern, and exhibits a stronger influence on U.S. rainfall than Central Pacific El Niño.

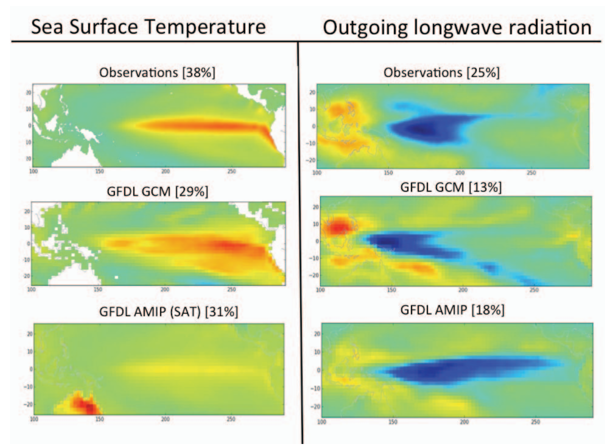

Fig. 2. Comparison of the first mode of the empirical orthogonal function analysis result with observations and models (GFDL GCM and GFDL AMIP) for two variables: sea surface temperature and outgoing longwave radiation flux. GFDL GCM, a coupled model, agrees with observations better than its uncoupled model, GFDL AMIP.

IV. Scientific Applications of CMDA

We have developed several science use cases of CMDA covering various topics, datasets, and analysis types. Each use case developed is described and listed in terms of a scientific goal, datasets used, the analysis tools used, scientific results discovered from the use case, an analysis result such as output plots and data files, and a link to the corresponding analysis service call with all the input arguments filled. Table 1 list the science use cases developed for CMDA users to learn about the analysis capabilities and available datasets of CMDA and the kind of results they can obtain from CMDA.

As an example, we describe the use case of "evauation of NCAR CMA5 model with MODIS total cloud fraction." In the use case, we used an CMDA analysis service called "the difference plot of two variables" and used NCAR CAM model output and MODIS instrument observation data for total cloud fraction. Figure 3 shows the result found using the service and datasets. The result shows that the CAM5 model overall does a fairly decent job at simulating total cloud cover, though simulates too few clouds especially near and offshore of the eastern ocean basins (Atlantic, Pacific, and Australian) where low clouds are dominant, and also further away in the trade cumulus/shallow cumulus regimes. Some of these differences were documented in [26]. These biases are commonly found in other CMIP5 models. This science use case demonstrates how the CMDA analysis services can be used to evaluate climate models using observational datasets as a reference dataset.

TABLE I. SCIENCE USE CASES FOR CMDA

Science Topic	Analysis Services Used	Datasets Used
Sea Surface Temperature over a hot spot region	Scatter and histogram plots of two variables; 2D variable map	AMSR-E SST; ECMWF-interim SST
Northern hemisphere West Pacific and East Pacific vertical velocity profiles	3D variable average vertical profile	ECMWF-interim vertical velocity; CAM5 vertical velocity
Climate-scale features of total cloud fraction and pressure vertical velocity	2D variable map; 3D variable 2D slice	MODIS total cloud fraction; ECMWF pressure vertical velocity
Relative humidity at 700 hPa	3D variable 2D slice	ECMWF-interim relative humidity
Comparison of two precipitation observational datasets	Difference plot of two variables	TRMM precipitation and GPCP precipitation
Evauation of NCAR CMA5 model with MODIS total cloud fraction	Difference plot of two variables	NCAR CAM total cloud fraction and MODIS total cloud fraction
Correlation maps of TRMM precipitation and Grace water land storage	Time-lagged correlation map	TRMM precipitation and Grace equivalent water height over land
Identifying the SST variability associated with ENSO (El Nino and La Nina)	Empirical orthogonal function	ECMWF-interim sea surface temperature

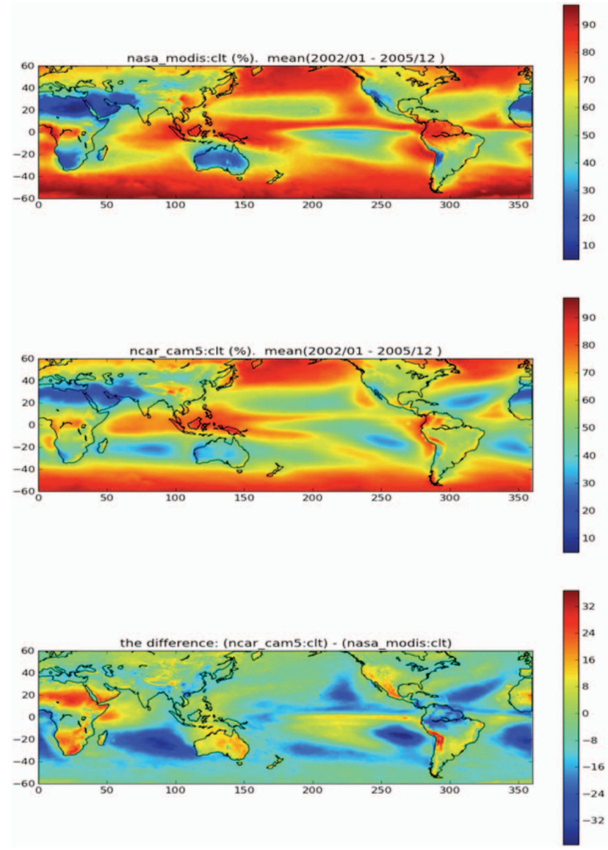

Fig. 3. Evaluation of NCAR CAM5 model with MODIS total cloud fraction. The top panel shows the total cloud fraction map generated with NASA MODIS instrument observation dataset. The middle panel shows the total cloud fraction map generated with NCAR CAM5 model output. The bottom panel shows the difference between the NCAR CAM5 and MODIS.

V. CONCLUSIONS

Rapidly growing datasets and analytics services in Earth Science challenge individual Earth scientists in organizing their work and concurrently challenge the whole community in sharing the datasets and tools and derived knowledge. With the community recognizing the need of infrastructure systems to address those challenges, some systems are under development but with a marginal impact so far in terms of tool adoption by the community and tool functionality. CMDA is designed to address the community need in a lightweight and easy-to-use and easy-to-maintain manner, with a focused domain of climate data analysis. CMDA provides a space where Earth scientists can organize their work efficiently and at the same time, share their work with others. With the projected exponential growth of the datasets and analytics tools, the goal of CMDA is to significantly ease the burden of individual scientists, increase their productivity, and as the whole community, to increase the scientific return of the NASA and NOAA's Earth science investments.

In order to expand the user base of CMDA, we have developed the educational and scientific applications of CMDA. With its successful educational application in the JPL Center for Climate Sciences Summer School in 2014, 2015, and 2016, we have gained real user experiences and feedback, which helped us identify the areas to improve CMDA usabilty and capability. We also have developed science use cases by working with local scientists, who are engaged in climate data analysis, have used CMDA for their research work, and have given us direct feedback on the CMDA analysis tools and datasets. Currently, we are developing more science use cases that involve multiple analysis steps and elaborate scientific approaches and results. By developing more applications, we expect that we will make CMDA more useful and more impactful for the community's need.

ACKNOWLEDGMENT

The authors acknowledge support from their institutions and funding sources. A part of the work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. A part of the work was carried at Carnegie Mellon University – Silicon Valley. We thank the NASA ROSES CMAC and AIST programs for funding this work.

REFERENCES

[1] Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5), 2014: Climate Change 2013: The Physical Science Basis.

[2] NRC, 2007: Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, National Research Council, National Academies Press, 2007.

[3] NRC, 2012: Earth Science and Applications from Space: A Midterm Assessment of NASA's Implementation of the Decadal Survey; Committee on the Assessment of NASA's Earth Science Program; Space Studies Board; Division on Engineering and Physical Sciences; National Research Council

[4] Taylor, K.E., R.J. Stouffer, and G.A. Meehl, 2012: An overview of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc., 93, 485-498.

[5] Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E., 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958.

[6] Teixeira, J., D. Waliser, R. Ferraro, P. Gleckler, T. Lee, and G. Potter, 2014: Satellite Observations for CMIP5: The Genesis of Obs4MIPs, Bull. Amer. Meteor. Soc., 95, 1329-1334

[7] NCO, 2017: NetCDF Operators (Version 4.6.4) [Software], available at: http://nco.sourceforge.net/.

[8] Zender, C. S., 2008: Analysis of self-describing gridded geoscience data with netCDF Operators (NCO), Environ. Model. Softw., 23, 1338–1342.

[9] NCL, 2016: The NCAR Command Language (Version 6.3.0) [Software], Boulder, Colorado, UCAR/NCAR/CISL/TDD, available at: http://dx.doi.org/10.5065/D6WD3XH5 (last access: May 2016), 2016.

[10] CDO, 2016: The Climate Data Operators (Version 1.7.2) [Software], Max-Planck-Institut for Meteorologie, Germany, available at https://code.zmaw.de/projects/cdo.

[11] Iris, 2017: Iris Python library for Meteorology and Climatology (Version 1.12.0) [Software], available at http://scitools.org.uk/iris/index.html.

[12] CF-Python, 2016: The Python CF Package (Version 1. 3.2) [Software], available at https://cfpython.bitbucket.io/.

[13] Gleckler, P., Doutriaux, C., Durack, P., Taylor, K., Zhang, Y., Williams, D., Mason, E., and Servonnat, J., 2016: A More Powerful Reality Test for Climate Models – Eos, Eos, 97, doi:10.1029/2016EO051663.

[14] Watson-Parris, D. et al., 2016: Community Intercomparison Suite (CIS) v1.4.0: a tool for intercomparing models and observations, Geosci. Model Dev., 9, 3093-3110.

[15] Williams, D. N., et al. 2013: The Ultra-scale Visualization Climate Data Analysis Tools (UV-CDAT): Data analysis and visualization for geoscience data, *Computer, 46*(9), 68–76, doi:10.1109/MC.2013.119.

[16] Eyring, V. et al., 2016b: ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP, Geosci. Model Dev., 9, 1747–1802.

[17] Lee, S. et al., Climate model diagnostic analyzer, IEEE International Conference on Big Data, Santa Clara, CA, October 29-November 1, 2015.

[18] Climate Model Diagnostic Analyzer (CMDA) running instance can be found in http://cmda-test.jpl.nasa.gov:9033.

[19] Richardson, L. and Sam Ruby, RESTful Web Services, O'Reilly, ISBN 978-0-596-52926-0, May 2007.

[20] G. Bort, "The High Velocity Web Framework for Java and Scala", 2015, accessed, Available from: https://www.playframework.com/.

[21] Apache, "Apache Lucene", accessed on: Jul. 16, 2015, Available from: https://lucene.apache.org.

[22] K. Schwaber and J. Sutherland, "The Scrum Guide", 2011, accessed: http://www.scrum.org/storage/scrumguides/Scrum_Guide.pdf.

[23] Hewson, T., 2017: Use and verification of ECMWF products in member and co-operating states (2016), ECMWF Technical Memoranda, 797.

[24] CF Convention: http://cfconventions.org/

[25] JPL Center for Climate Scineces Summer School Site: http://climatesciences.jpl.nasa.gov/events/summer-school.

[26] Kubar, T. L., G. L. Stephens, M. Lebsock, V. E. Larson, and P. A. Bogenschutz, 2015: Regional Assessments of Low Clouds against Large-Scale Stability in CAM5 and CAM-CLUBB Using MODIS and ERA-Interim Reanalysis Data. *Journal of Climate*, **28,** 1685-1706.