

SOARS 2019 WORKSHOPS

COMPUTATIONAL THINKING AND DATA SCIENCE

5/25/2019

Keith E. Maull, PhD

Workshop Motivation



“ The publication of research papers is slowly changing to adapt to the digital age. We envision that in the near future (5–10 years), scientists will use **radically new tools to author papers and disseminate information** about the process and products of their research. These tools will **document** and **publish the computational workflow** as well as all the associated **digital objects** (data, software, etc.) that **form the basis of a paper**.

Gil, Y., et al. (2016), Toward the Geoscience Paper of the Future: Best practices for documenting and sharing research from data to software to provenance, Earth and Space Science, 3, 388–415, doi:[10.1002/2015EA000136](https://doi.org/10.1002/2015EA000136).

2019 Computational Thinking and Data Science Workshops

SECTION A(M)

- **Introduction** to computational thinking
- Use **algorithms** solve problems (GCD, search, etc.)
- Develop intermediate solutions with **pseudocode** and converting that pseudocode to **running Python code**
- Apply basic knowledge of tools (Jupyter) and resources

SECTION P(M)

- **Deeper dive** into Python and the data science stack
- Implement **working solutions** to common tasks (data manipulation, graphing, etc.)
- Develop **advanced strategies** and working knowledge of **platforms, tools and workflows** (Jupyter, Python libraries, etc.)

Tools we will use ...

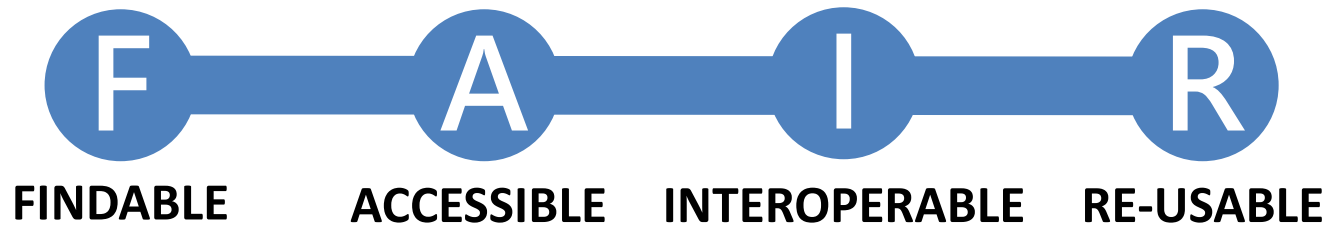
SECTIONS A(M)+P(M)



GitHub



We want our work to be ...



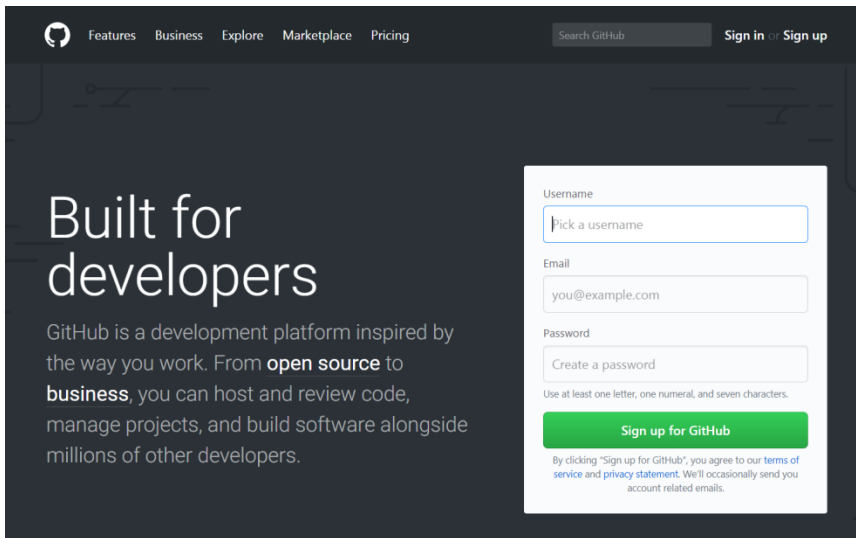
Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018.

<https://doi.org/bdd4>

3 things you need to do ...

1

GET A FREE GITHUB ACCOUNT

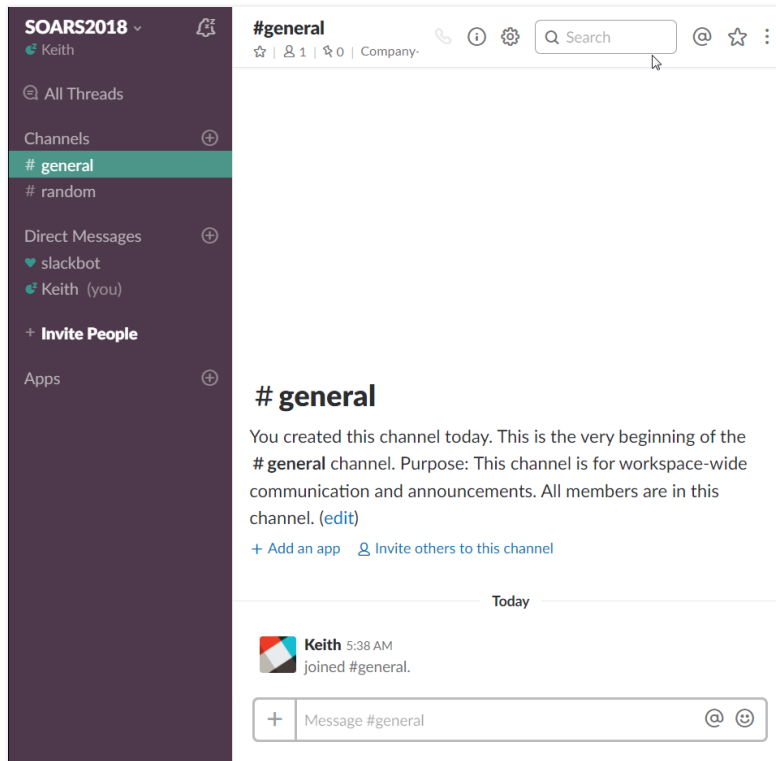
A screenshot of the GitHub website's sign-up page. The page has a dark background with the GitHub logo and navigation links (Features, Business, Explore, Marketplace, Pricing) at the top. The main heading is "Built for developers". Below it, a paragraph describes GitHub as a development platform inspired by the way you work, from open source to business. On the right side, there is a white sign-up form with fields for Username (placeholder: "Pick a username"), Email (placeholder: "you@example.com"), and Password (placeholder: "Create a password"). Below the password field, there is a note: "Use at least one letter, one numeral, and seven characters." A green button labeled "Sign up for GitHub" is at the bottom of the form. Below the button, there is a small disclaimer: "By clicking 'Sign up for GitHub', you agree to our terms of service and privacy statement. We'll occasionally send you account related emails."

- Go to Set up a Github account at <https://github.com> (make sure to use your .edu email address!)
- Send me (kmaull@ucar.edu) your Github ID
- Bookmark the repository for this year's workshop (<https://git.io/fj00z>)
- Install GithubDesktop (<https://desktop.github.com>) for Win/Mac
- Browse around Github and watch a tutorial!

3 things you need to do ...

2

JOIN OUR SLACK CHANNEL



- Go to <https://soarshq19.slack.com> and set up a free account on our channel
- Send me (kmaull@ucar.edu) if you have trouble getting on to the channel
- Watch a video / tutorial about slack (there is one when you login)

3 things you need to do ...

3

FAMILIARIZE / RE-ACQUAINT YOURSELF WITH PYTHON



- Skim a few of the online resources in the syllabus (on Github)
- <https://python.org!>
- You do **not** need to install Python on your computer, but if you do, install Anaconda from <https://anaconda.com>

Things to watch for ...

1

INFORMATION ABOUT OUR JUPYTER ENVIRONMENT



- I should be sending an email out soon
- You can look at **prior workshops** to see what Jupyter is all about
- You will **not** need to install anything, **but** if you **want to play on your own machine**, do **contact me**

Things you might also do ...

- Read the paper:
 - *Gil, Y., et al. (2016), Toward the Geoscience Paper of the Future: Best practices for documenting and sharing research from data to software to provenance, Earth and Space Science, 3, 388–415, doi:[10.1002/2015EA000136](https://doi.org/10.1002/2015EA000136).*
- Explore the awesome Github repository `Python for the Geosciences`:
 - <https://git.io/vhLhr>
- Check out Jupyter Notebooks examples
 - <http://nb.bianp.net/sort/views/>
- Check out some of the prior workshops:
 - <https://git.io/v6XyK> (2016)
 - <https://git.io/vHqof> (2017)
 - (2018)
- Check out the website for this year:
 - <https://ncar.github.io/soars19cdws/>

Some inspiration ...

Here is a paper ...
that actually started as a
pre-print

The screenshot shows the journal's website with the article title "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG". The article is marked as a preprint under review. The page includes a sidebar with navigation links, a search bar, and various metrics. The article's abstract and author information are visible.

Atmospheric Chemistry and Physics
An interactive open-access journal of the European Geosciences Union

EGU

EGU.eu | EGU Journals | EGU Highlight Articles | Contact | Imprint |

Submit a manuscript | Manuscript tracking

Discussion papers

Abstract | Assets | Discussion | Metrics

<https://doi.org/10.5194/acp-2018-118>
© Author(s) 2018. This work is distributed under the Creative Commons Attribution 4.0 License.

Research article

02 May 2018

Review status

This discussion paper is a preprint. It is a manuscript under review for the Journal Atmospheric Chemistry and Physics (ACP).

Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG

Benjamin S. Grandey¹, Daniel Rothenberg², Alexander Avramov^{2,3}, Qinjiao Jin², Hsiang-He Lee¹, Xiaohong Liu⁴, Zheng Lu⁴, Samuel Albeni^{5,6}, and Chien Wang^{2,1}

¹Center for Environmental Sensing and Modeling, Singapore - MIT Alliance for Research and Technology, Singapore
²Center for Global Change Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
³Department of Environmental Sciences, Emory University, Atlanta, Georgia, USA
⁴Department of Atmospheric Science, University of Wyoming, Laramie, Wyoming, USA
⁵Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, New York, USA
⁶Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Gif-sur-Yvette, France

Received: 01 Feb 2018 – Accepted for review: 01 May 2018 – Discussion started: 02 May 2018

Abstract. We quantify the effective radiative forcing (ERF) of anthropogenic aerosols modelled by the aerosol-climate model CAM5.3-MARC-ARG. CAM5.3-MARC-ARG is a new configuration of the Community Atmosphere Model version 5.3 (CAM5.3) in which the default aerosol module has been replaced by the two-Moment, Multi-Modal, Mixing-state-resolving Aerosol model for Research of Climate (MARC). CAM5.3-MARC-ARG uses the default ARG aerosol

Journal metrics

Clarivate Analytics	IF 5.318
Clarivate Analytics	IF 5-year 5.896
Scopus	CiteScore 5.72
CWTS	SNIP 1.330
SJR	SJR 3.207
CWTS	IPP 4.756
h5-index	89

Search articles

Search

Author

Download

PDF (6128 KB)

XML

Short summary

Anthropogenic emissions of aerosol particles likely cool the climate system. We investigate...

Citation

BibTeX

EndNote

Share

Twitter, Facebook, Google+, LinkedIn

Some inspiration ...

This paper includes
analysis ... done with
some **data** and **code**



Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG

Benjamin S. Grandey et al.

Supplement

<https://doi.org/10.5194/acp-2018-118-supplement>

Data sets

Data for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"

B. S. Grandey

<https://doi.org/10.6084/m9.figshare.5687812>

Model code and software

Configuration and analysis for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"

B. S. Grandey

<https://doi.org/10.5281/zenodo.1239247>

MARC - Model for Research of Aerosols and Climate

A. Avramov, D. Rothenberg, Q. Jin, S. Garimella, B. Grandey, and C. Wang

<https://doi.org/10.5281/zenodo.1117370>

Some inspiration ...

The data is findable here:

<https://doi.org/10.6084/m9.figshare.5687812>

The screenshot shows a Figshare dataset page. At the top, there is a list of files: `so4_elev_p17c_marc_2000.nc`, `so4_surf_p17c_marc_1850.nc`, and `so4_surf_p17c_marc_2000.nc`. Below the file list, there are buttons for **Cite**, **Download all (16.41 GB)**, **Share**, **Embed**, and **+ Collect (you need to log in first)**. The page title is **Data for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"**. It indicates **Version 2** was published on 31.01.2018, 20:08 by Benjamin S. Grandey. The page has 203 views, 38 downloads, and 1 citation. The introduction states that the data accompany a manuscript titled "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG" by B. S. Grandey, D. Rothenberg, A. Avramov, Q. Jin, H.-H. Lee, X. Liu, Z. Lu, S. Albani, and C. Wang. The files contain input data and output data associated with the CESM-CAM5 climate model simulations described in the manuscript. A list of categories includes Atmospheric Sciences, Atmospheric Aerosols, Atmospheric Radiation, Climate Science, and Climate Change Processes. The keywords are aerosol and radiative forcing.

so4_elev_p17c_marc_2000.nc
so4_surf_p17c_marc_1850.nc
so4_surf_p17c_marc_2000.nc

Download all (16.41 GB) Share Embed + Collect (you need to log in first) 1 / 191

Data for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"

Version 2 31.01.2018, 20:08 by Benjamin S. Grandey

203 views 38 downloads 1 citations

Introduction:

These data accompany the manuscript titled "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG" by B. S. Grandey, D. Rothenberg, A. Avramov, Q. Jin, H.-H. Lee, X. Liu, Z. Lu, S. Albani, and C. Wang. The files contain input data and output data associated with the CESM-CAM5 climate model simulations described in the manuscript.

Please also see <https://github.com/grandey/p17c-marc-comparison/> for details of the experimental design, model configuration, data management, and

CATEGORIES

- Atmospheric Sciences
- Atmospheric Aerosols
- Atmospheric Radiation
- Climate Science
- Climate Change Processes

KEYWORD(S)

aerosol radiative forcing

Some inspiration ...

The code is also
findable here:

<https://doi.org/10.5281/zenodo.1239247>

A screenshot of a GitHub repository page for 'grandey / p17c-marc-comparison'. The page shows the repository's configuration and analysis for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG". It includes statistics such as 138 commits, 1 branch, 3 releases, and 1 contributor. A list of files and their commit messages is displayed, including 'analysis_draft2017b', 'analysis_prelim', 'config_simulations', 'input_data_p17c', 'manage_data', 'user_nl_cam', '.gitignore', 'LICENSE.md', 'README.org', 'experimental_design.org', and 'status.org'. The README section is partially visible at the bottom, showing the title and DOI: 10.5281/zenodo.1239247.

grandey / p17c-marc-comparison

Configuration and analysis for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"

138 commits 1 branch 3 releases 1 contributor MIT

Tag: v0.5.1 New pull request Create new file Upload files Find file Clone or download

File	Commit Message	Time Ago
analysis_draft2017b	Minor changes in preparation for v0.5.1 release	24 days ago
analysis_prelim	Removing notes_analysis_prelim_p17c.docx (moving offline)	6 months ago
config_simulations	Moving experimental design section from README to separate file	7 months ago
input_data_p17c	Re-retrieving input data files	10 months ago
manage_data	Minor changes in preparation for v0.5.0 publication via Zenodo	4 months ago
user_nl_cam	Timing simulations: adding user_nl_cam for p17c_t_marc_r1	7 months ago
.gitignore	Renaming analysis_for_draft2017b/ to analysis_draft2017b/	6 months ago
LICENSE.md	Simplifying LICENSE.md to MIT License	6 months ago
README.org	Minor changes in preparation for v0.5.1 release	24 days ago
experimental_design.org	figs: sulphate -> sulfate; exp design: MARC link; README: updating au...	4 months ago
status.org	Minor changes in preparation for v0.5.0 publication via Zenodo	4 months ago

README.org

Configuration and analysis for "Effective radiative forcing in the aerosol-climate model CAM5.3-MARC-ARG"

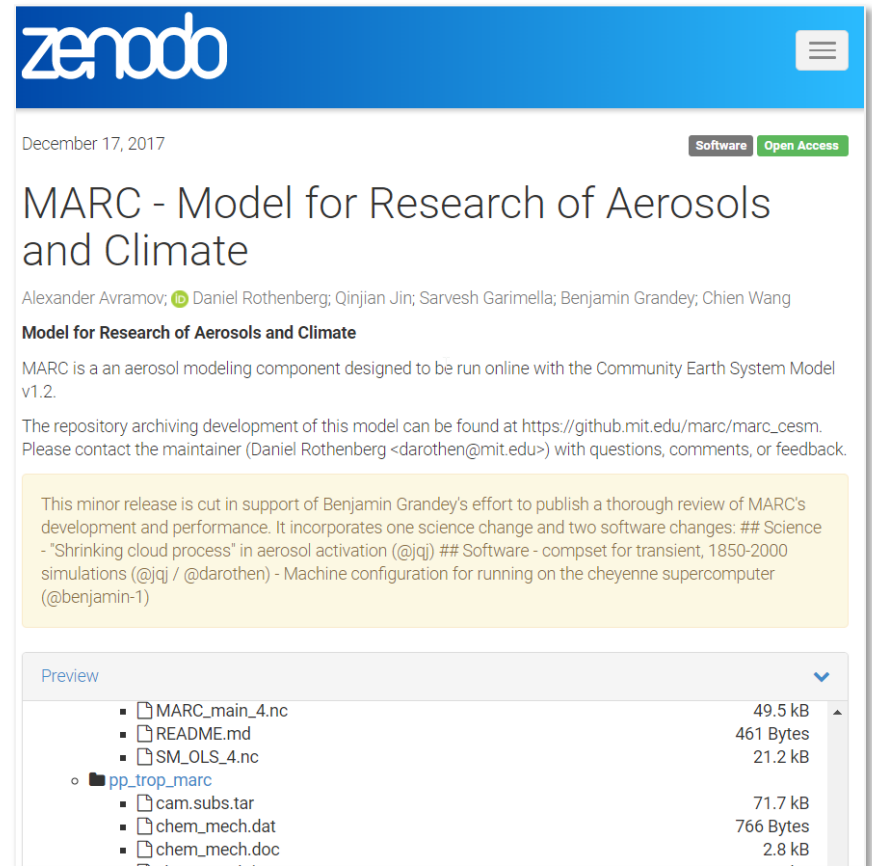
DOI: 10.5281/zenodo.1239247

Introduction

Some inspiration ...

The model code upon
which this work is
build is here ...

<https://doi.org/10.5281/zenodo.1117370>



The screenshot shows the Zenodo repository page for the MARC model. The page has a blue header with the Zenodo logo. Below the header, the date "December 17, 2017" is displayed on the left, and "Software" and "Open Access" badges are on the right. The title "MARC - Model for Research of Aerosols and Climate" is prominently displayed. Below the title, the authors "Alexander Avramov, Daniel Rothenberg, Qinqian Jin, Sarvesh Garimella, Benjamin Grandey, Chien Wang" are listed. The subtitle "Model for Research of Aerosols and Climate" is also present. The description states that MARC is an aerosol modeling component designed to be run online with the Community Earth System Model v1.2. It provides a link to the GitHub repository for archiving development and contact information for the maintainer, Daniel Rothenberg. A yellow box contains a note about a minor release supporting a review of MARC's development and performance, mentioning science and software changes. At the bottom, a "Preview" section shows a file tree with items like MARC_main_4.nc, README.md, SM_OLS_4.nc, and a subdirectory pp_trop_marc containing cam.subs.tar, chem_mech.dat, and chem_mech.doc, along with their respective sizes.

zenodo

December 17, 2017 Software Open Access

MARC - Model for Research of Aerosols and Climate

Alexander Avramov, Daniel Rothenberg, Qinqian Jin, Sarvesh Garimella, Benjamin Grandey, Chien Wang

Model for Research of Aerosols and Climate

MARC is an aerosol modeling component designed to be run online with the Community Earth System Model v1.2.

The repository archiving development of this model can be found at https://github.mit.edu/marc/marc_cesm. Please contact the maintainer (Daniel Rothenberg <darothern@mit.edu>) with questions, comments, or feedback.

This minor release is cut in support of Benjamin Grandey's effort to publish a thorough review of MARC's development and performance. It incorporates one science change and two software changes: ## Science - "Shrinking cloud process" in aerosol activation (@jqj) ## Software - compset for transient, 1850-2000 simulations (@jqj / @darothern) - Machine configuration for running on the cheyenne supercomputer (@benjamin-1)

Preview

- MARC_main_4.nc 49.5 kB
- README.md 461 Bytes
- SM_OLS_4.nc 21.2 kB
- pp_trop_marc
 - cam.subs.tar 71.7 kB
 - chem_mech.dat 766 Bytes
 - chem_mech.doc 2.8 kB

Some inspiration ...

And if that wasn't enough ...

Acknowledgements

This research is supported by the National Research Foundation of Singapore under its Campus for Research Excellence and Technological Enterprise programme. The Center for Environmental Sensing and Modeling is an interdisciplinary research group of the Singapore-MIT Alliance for Research and Technology. This research is also supported by the U.S. National Science Foundation (AGS-1339264) and the U.S. Department of Energy, Office of Science (DE-FG02-94ER61937). The CESM project is supported by the National Science Foundation and the Office of Science (BER) of the U.S. Department of Energy. We acknowledge high-performance computing support from Cheyenne ([doi:10.5065/D6RX99HX](https://doi.org/10.5065/D6RX99HX)) provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We thank Natalie Mahowald for contributing dust model code, optical tables, a soil erodibility map, and advice, all of which have aided the development of CAM5.3-MARC-ARG.

Wow!

A person wearing a red and yellow jumpsuit is captured mid-leap from a rocky cliff edge covered in snow. The person is falling horizontally against a clear blue sky. In the background, a range of jagged, snow-capped mountains stretches across the horizon under a soft, golden light, suggesting dawn or dusk. The overall scene conveys a sense of adventure and freedom.

You got this!