

Methods, Challenges, and Opportunities in Measuring the Impact of CI Projects



Richard Gerber

NERSC HPC Department Head
NERSC Senior Science Advisor
ECP Director of Hardware & Integration

September 17, 2019

Impacts of HPC/CI

This is a broad space, HPC has impacts everywhere

Challenge is measuring and communicating to stakeholders in an understandable, meaningful, impactful way



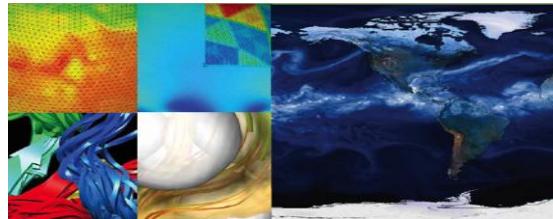
NERSC: Mission HPC for DOE Office of Science Research



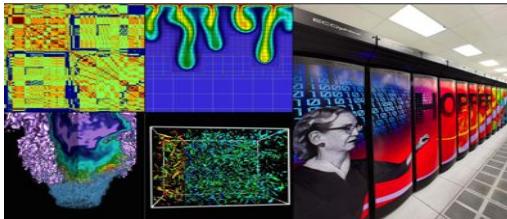
U.S. DEPARTMENT OF
ENERGY

Office of
Science

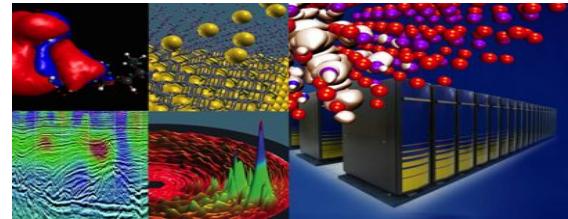
Largest funder of physical science
research in the U.S.



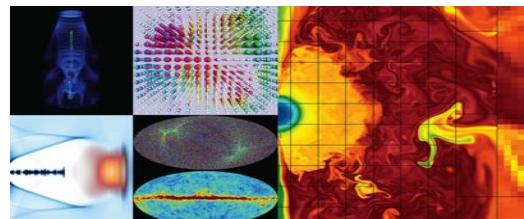
Bio Energy, Environment



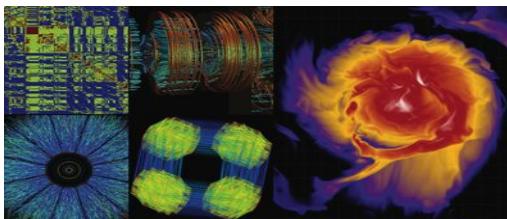
Computing



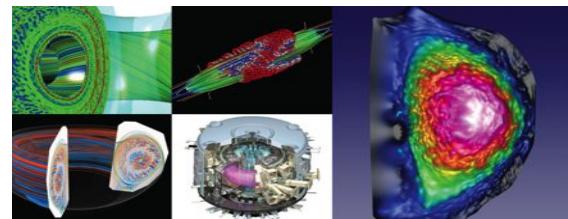
Materials, Chemistry, Geophysics



Particle Physics, Astrophysics



Nuclear Physics



Fusion Energy, Plasma Physics

9K users, 1K projects, 50 states, 40 countries, universities & national labs



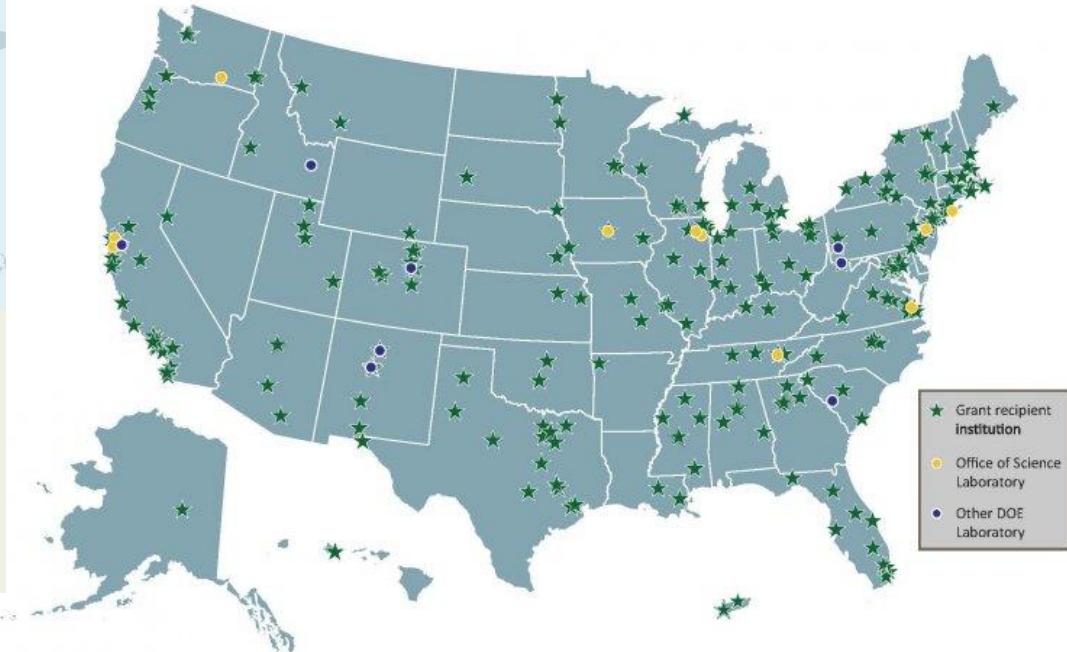
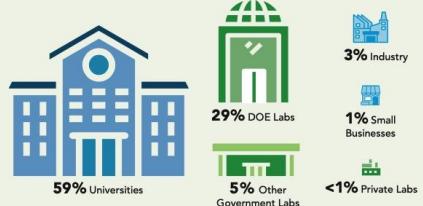
NERSC Has a Broad National and International Impact



Office of Science funded research projects make up the bulk of NERSC users.



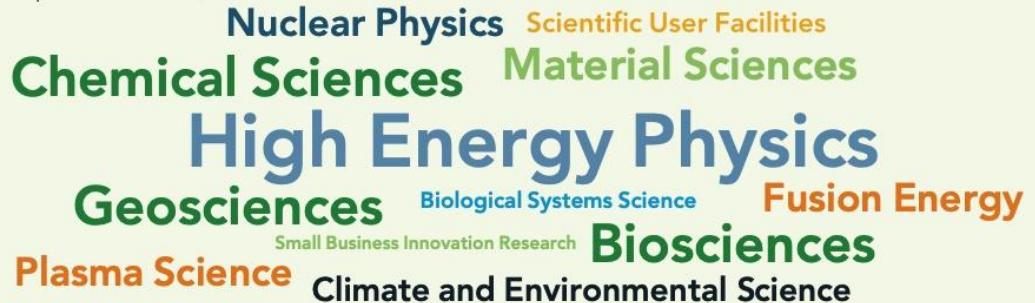
~9,000 ANNUAL USERS FROM ~800 Institutions + National Labs



NERSC by the Numbers

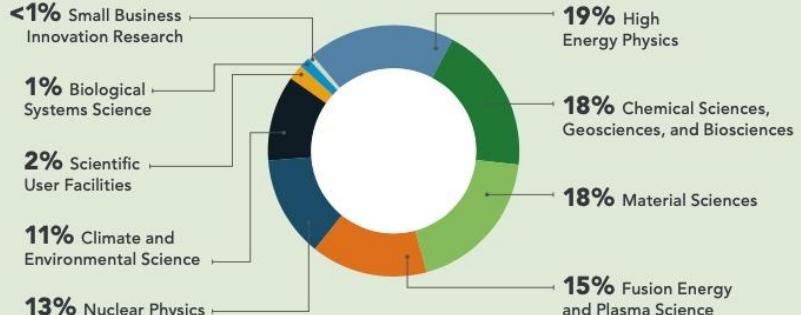
Top Science Disciplines

(By computational hours used)



~1,000 Projects

Breakdown of Compute Used by DOE Program



>2,000
Scientific Journal Articles
per Year

Making effective use of exascale systems has required a significant investment in software – The Exascale Computing Project



Maintain international leadership in HPC



Promote the health of the US HPC industry



Deliver a **sustainable software ecosystem** used and maintained for years to come



Ensure that exascale systems can be used to deliver **mission-critical applications**



EXASCALE COMPUTING PROJECT

7-year, \$1.8B

US Department of Energy project funded 1000+ people at national labs, universities, US industries

This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.



Application Development

- Develop and enhance the predictive capability of applications, **25 applications, 6 Co-Design Centers**

Software Technology

- Deliver expanded and vertically integrated software stack, **70 unique products**

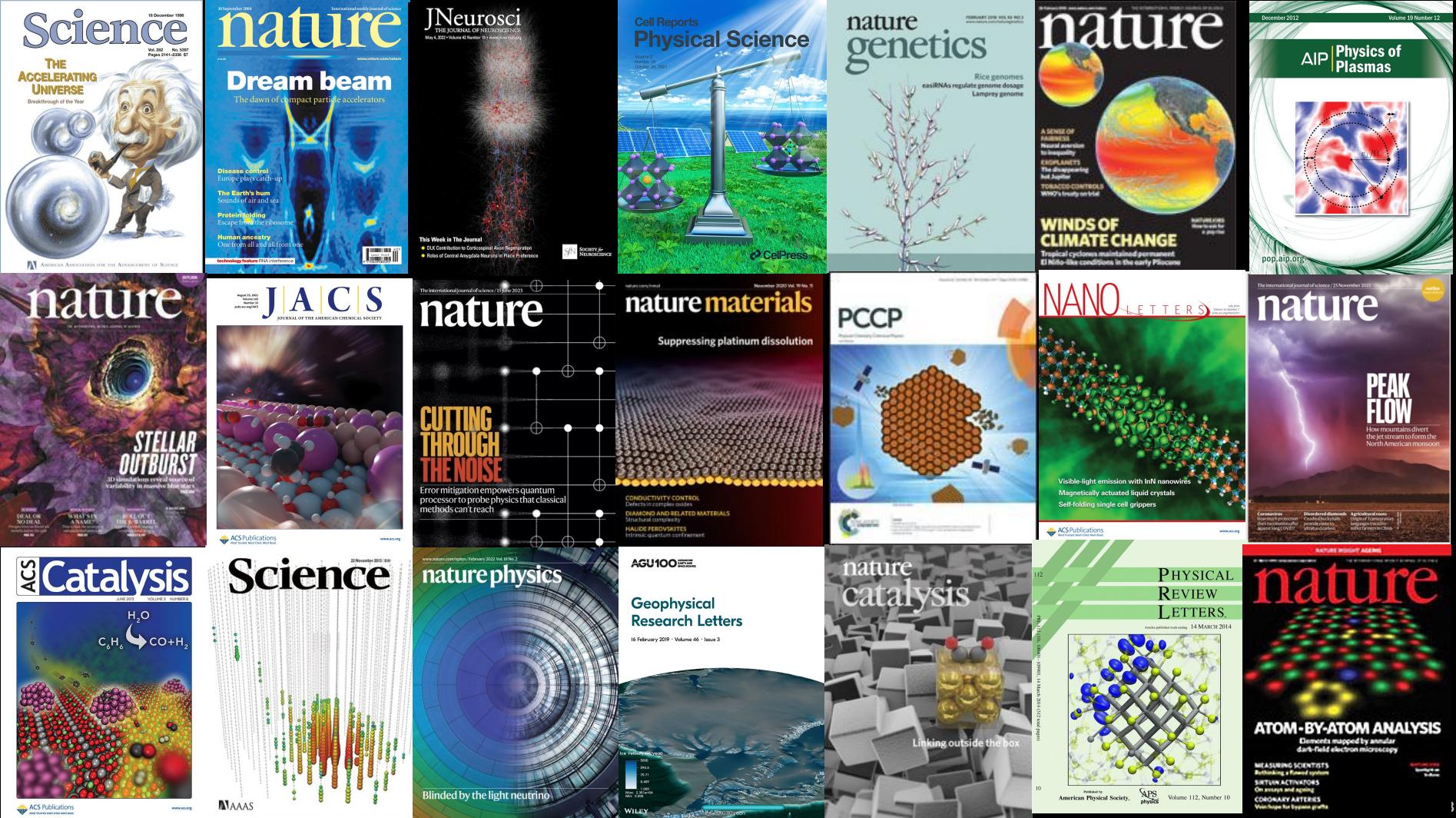
Hardware and Integration

- Application integration and software deployment to facilities, exascale node and system design, **6 US HPC vendors**



Office of Science

Example: Publications



Measuring Scientific Productivity



- Tracking publications is a standard way of measuring scientific impact and productivity; they provide
 - Tangible results enabled by NERSC's systems, services, and expertise.
 - A measure of impact and are reported to DOE
 - The basis for science highlights communicated to stakeholders
 - Motivation for staff
 - Additional understanding of how our systems are used
 - Key input to the project renewal review process
- NERSC is acknowledged in 2,000+ refereed scientific publications each year
 - Many high impact journals
 - Nature (12), Nature family of journals (124), Science (20), PRL (53), PNAS (12)



Tracking Publications

We want a system that automatically identifies and validates publications that were based on work that used NERSC, and associates them with the appropriate users and projects **in a reliable and accurate manner**

- Users can enter publications directly into IRIS and link them to people and projects
- For other pubs, we have a system that goes part of the way towards this goal (development continues) using three data sources
 - Online citation databases (Scopus & Web of Science) (primary source)
 - Project allocation renewal submissions from users (secondary, but significant)
 - Manual input (rare)
- All entries are validated via Crossref
 - If no entry found via Crossref, the publication does not get into our database
 - Each publication is guaranteed to exist and have correct dates

Collecting Publications

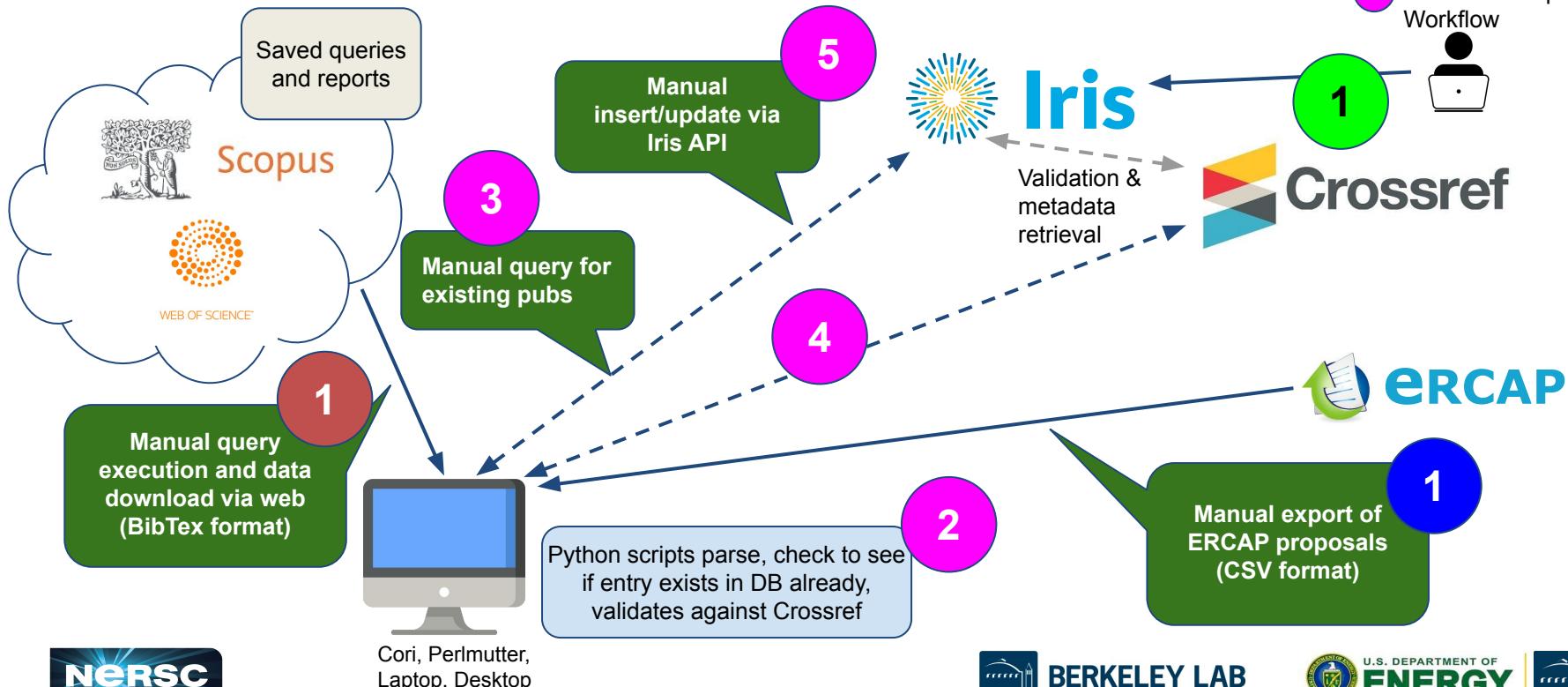
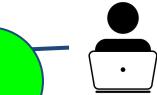
Our users are very good at including the string “National Energy Research Scientific Computing Center” in their acknowledgements!



- Online citation DB
- ERCAP Proposals



- Common Script Workflow



My Account

My Projects

Reports

Publications

Server Logins

Vetting Records (S)

Storage Type Quotas

Active Users

Reserves

Usage Tracking Report

Training

Admin Tools

Utilities

Center

Resources

Settings

Logout

Publications

Select Columns

Filter Report

Select Rollups

Report Results

Run Report

+ Link Publication(s)

	Publication Title	Publication Authors	Publication Journal	Publication DOI	Publication Date	Publication Users	Publication Accounts
<input type="checkbox"/>	Observations Indicate That Clouds Amplify Mechanisms of Southern Ocean Heat Uptake	Ariel L. Morrison, Hansi A. Singh, Philip J. Rasch	Journal of Geophysical Research: Atmospheres	10.1029/2021JD03...	2022-02-04	✉ Morrison, Ariel ✉ Rasch, Philip	✉ m1199
<input type="checkbox"/>	Subtropical Marine Low Stratiform Cloud Deck Spatial Errors in the E3SMv1 Atmosphere Model	Brunke, Michael A.; Ma, Pou2010Lun; Reeves Eyre, J. E. Jack; Rasch, Philip J.; Sorooshian, Armin; Zeng, Xubin	Geophysical Research Letters	10.1029/2019GL08...	2019-11-03	✉ Rasch, Philip	
<input type="checkbox"/>	Climate response of the South Asian monsoon system to anthropogenic aerosols	Ganguly, Dilip; Rasch, Philip J.; Wang, Hailong; Yoon, Jin-Ho	Journal of Geophysical Research: Atmospheres	10.1029/2012JD01...	2012-06-08	✉ Rasch, Philip	
<input type="checkbox"/>	Fast and slow responses of the South Asian monsoon system to anthropogenic aerosols	Ganguly, Dilip; Rasch, Philip J.; Wang, Hailong; Yoon, Jin-ho	Geophysical Research Letters	10.1029/2012GL05...	2012-08-21	✉ Rasch, Philip	
<input type="checkbox"/>	A Partial Coupling Method to Isolate the Roles of the Atmosphere and Ocean in Coupled Climate Simulations	Garuba, Oluwayemi A.; Rasch, Philip J.	Journal of Advances in Modeling Earth Systems	10.1029/2019MS0...	2020-09-01	✉ Rasch, Philip	✉ m1199
<input type="checkbox"/>	Disentangling the Coupled Atmosphere-Ocean-Ice Interactions Driving Arctic Sea Ice Response to CO ₂ Increases	Garuba, Oluwayemi A.; Singh, Hansi A.; Hunke, Elizabeth; Rasch, Philip J.	Journal of Advances in Modeling Earth Systems	10.1029/2019MS0...	2020-11-06	✉ Rasch, Philip	✉ m1199
<input type="checkbox"/>	The Separate Physics and Dynamics Experiment (SPADE) framework for determining resolution awareness: A case study of microphysics	Gustafson, William I.; Ma, Po-Lun; Xiao, Heng; Singh, Balwinder; Rasch, Philip J.; Fast, Jerome D.	Journal of Geophysical Research: Atmospheres	10.1002/jgrd.50711	2013-08-12	✉ Rasch, Philip	
<input type="checkbox"/>	Aerosols in the E3SM Version 1: New Developments and Their Impacts on Radiative Forcing	Hailong Wang, Richard C. Easter, Rudong Zhang, Po-Lun Ma, Balwinder Singh, Kai Zhang, Dilip Ganguly, Philip J. Rasch, Susannah M. Burrows, Steven J. Ghan, Sijia Lou, Yun Qian, Yang Yang, Yan Feng, Mark Flanner, L. Ruby Leung, Xiaohong Liu, Manish Shrivastava, Jian Sun, Qi Tang, Shaocheng Xie, Jin-Ho Yoon	Journal of Advances in Modeling Earth Systems	10.1029/2019MS0...	2019-12-06	✉ Rasch, Philip	✉ e3sm
<input type="checkbox"/>	CondiDiag1.0: a flexible online diagnostic tool for conditional sampling and budget analysis in the E3SM atmosphere model (EAM)	Hui Wan, Kai Zhang, Philip J. Rasch, Vincent E. Larson, Xubin Zeng, Shixuan Zhang, Ross Dixon	Geoscientific Model Development	10.5194/gmd-15-3205-2022	2022-04-19	✉ Rasch, Philip	
<input type="checkbox"/>	Improved Simulation of the QBO in E3SMv1	Jadwiga H. Richter, Chih-Chieh Chen, Qi Tang, Shaocheng Xie, Philip J. Rasch	Journal of Advances in Modeling Earth Systems	10.1029/2019MS0...	2019-10-01	✉ Rasch, Philip	
<input type="checkbox"/>	Impact of numerical choices on water conservation in the E3SM Atmosphere Model Version 1 (EAM v1)	Kai Zhang, Philip J. Rasch, Mark A. Taylor, Hui Wan, Lai-Yung Ruby Leung, Po-Lun Ma, Jean-Christophe Golaz, Jon Wolfe, Wuyin Lin, Balwinder Singh, Susannah Burrows, Jin-Ho Yoon, Hailong Wang, Yun		10.5194/gmd-2017-293	2017-12-14	✉ Rasch, Philip	

Reports

Publications

Run Report

With the data in IRIS, we can run reports and directly link to journal articles.

Select Columns 2

Filter Report 1

Select Rollups

Report Results 33

+ Link Publication(s)

	Publication Journal	count of Publication DOI
<input type="checkbox"/>	Nature Communications	277
<input type="checkbox"/>	Nature	80
<input type="checkbox"/>	Nature Materials	41
<input type="checkbox"/>	Nature Physics	28
<input type="checkbox"/>	Nature Chemistry	18
<input type="checkbox"/>	Nature Climate Change	18
<input type="checkbox"/>	Nature Catalysis	17
<input type="checkbox"/>	Nature Energy	14
<input type="checkbox"/>	Nature Biotechnology	11
<input type="checkbox"/>	Nature Microbiology	10
<input type="checkbox"/>	Nature Nanotechnology	9
<input type="checkbox"/>	Nature Geoscience	7



BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF
ENERGY

Office of
Science

The system is still far from complete and accurate!

- Incomplete coverage in literature searches databases
- Errors returned by literature searches
- Users don't acknowledge NERSC
- No good automated way to associate projects with pubs
- Derived products and artifacts are not included (e.g. library of computed climate scenarios used by other researchers)
- Difficult to measure impact of different publications