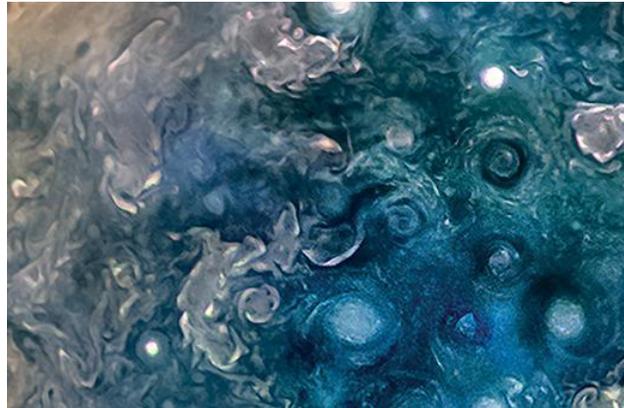**SCIENCE**

National Aeronautics and
Space Administration

Science Mission Directorate

NASA HQ SMD Strategic Initiative for Scientific Data and Computing

HPAC Fall Meeting 2018

Jeffrey Hayes

Program Executive, Mission Operations

Astrophysics & Heliophysics

Division, SMD

December 18, 2018

Strategic Plan for Scientific Data and Computing

It is our intent to develop a 5-year strategy that has four overall goals:

- Improve discovery and access for all SMD data to immediately benefit science data users and improve the overall user experience.

- Leverage current technology for the discovery, access, and effectiveness of NASA's data, as well as enable new technology and analysis techniques for scientific discovery.

- Identify large-scale and cross-disciplinary/division science users and use cases to inform future science data system capabilities.

- Champion robust theory programs that are firmly based on NASA's observations.

Motivation

SMD aims to answer some of science's most difficult questions: Is there life outside of Earth? How does the universe work? What changes are happening on our planet?

To answer these questions effectively, data is imperative. Each division within SMD produces, examines, and catalogs significant amounts of data to reinforce scientific objectives and provide scientific findings and data to millions of people. Within 5 years, all SMD divisions are predicted to generate **more than 100PB of data annually.**

The expected increase in the Directorate's archival needs presents opportunities for cutting-edge scientific discover as well as noteworthy challenges in data analysis, management, and access in the future.

SMD seeks to take a more strategic view of its science data systems, including high-end computing, to promote more efficient and effective data management across SMD divisions, as well as enable cross-disciplinary discovery and analysis of science data.

Development of a SMD strategic data management strategy that aligns advances in information technology with the unique needs of science data systems and computing will inform technology investments and provide a roadmap for how SMD can partner with other organizations, both within NASA and external to the agency, to enable greater scientific discovery.

Case Study: Earth Science Data System Evolution (EOSDIS)

The current architecture will not be cost effective as the annual ingest rate increases from 4 to 50PB/year.

EOSDIS is developing open source cloud native software for reuse across the agency, throughout the government, and for any other user.
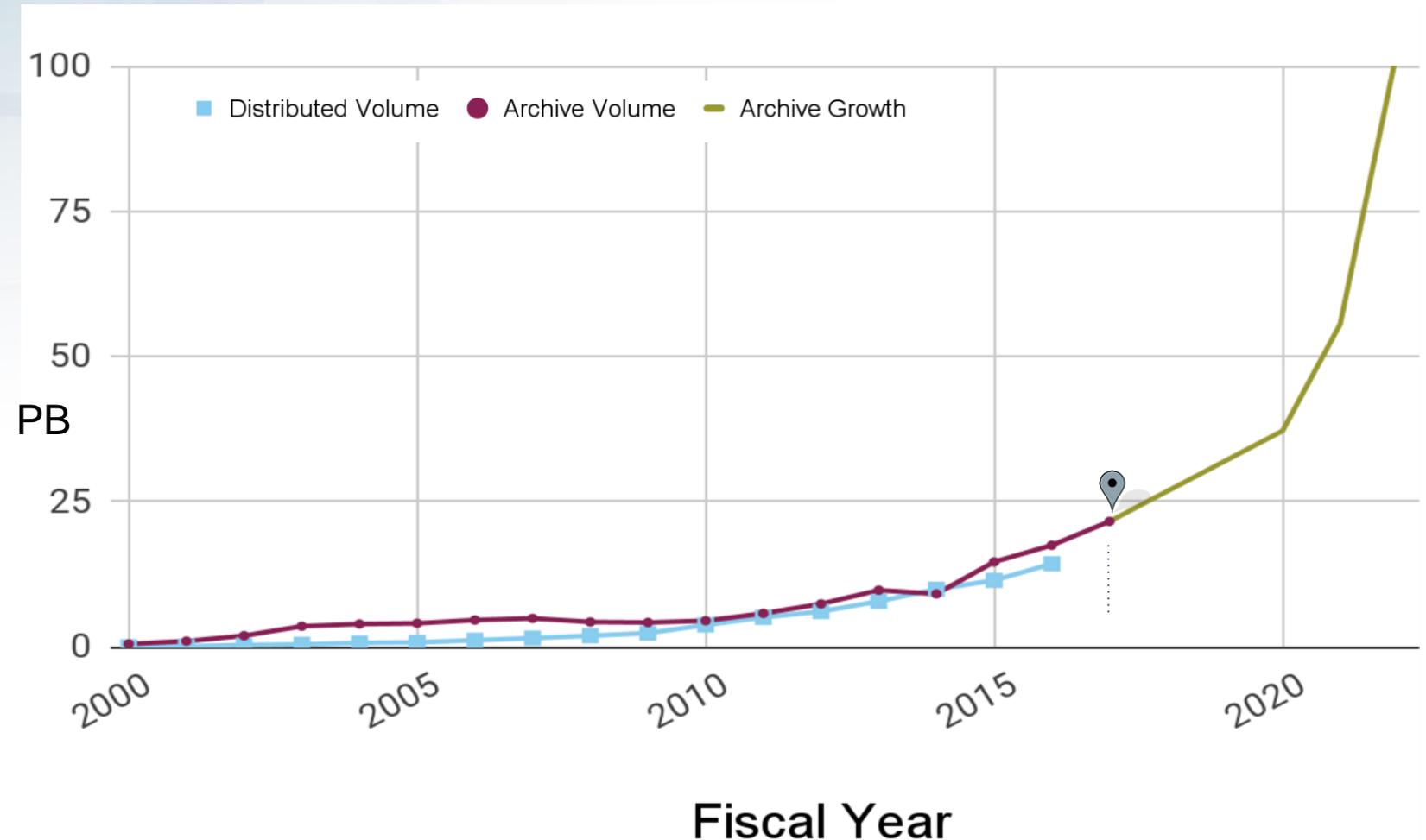
Cloud offers benefits like the ability to analyze data at scale, analyze multiple data sets together easily and avoid lengthy expensive moves of large data sets allowing scientists to work on data "in place"

SMD Strategic Initiative for Scientific Data and Computing

- Historically, SMD's management of data and computing resources has been conducted based on the specific needs of each mission or division, with limited consideration for enabling inter-disciplinary research.

- In late 2014 an Ad-hoc committee, the Big Data Task Force (BDTF), was chartered through the NASA Advisory Council (NAC) to study and identify best practices from what NASA, the science community, and other agencies are doing in big data. The final report of the BDTF's findings were presented to the NAC in late 2017.

- During the SMD Senior Leadership retreat in May 2017, strategic data management and science computing across SMD was identified as priorities for assessment and action over the following year.

- To continue to be on the forefront of groundbreaking scientific discovery, SMD has committed to develop a new SMD-wide Strategic Plan for Scientific Data and Computing that will be implemented across the five divisions: Earth Science, Planetary Science, Heliophysics, Astrophysics, and the Joint Agency Satellite Division.

- A small team comprised of representatives from each of the SMD divisions was charted to develop the strategy to enable greater scientific discovery by leveraging advances in information technology for SMD science computing and data archives.

SMD's Approach to Meet Goals

The core principles behind NASA SMD's efforts are :

- Continued free and open access to scientific data

- Improved ease of use and discoverability

- Enhanced science applications and new use cases

- Incorporates best practices and "state of the art" through partnerships

To ensure we successfully meet our goals, we've engaged experts specializing in NASA-related research, management of big data across the government, and information technology to understand their current states along with the opportunities and challenges they present. We also solicited feedback from outside experts on best practices for discovery and access for all SMD data and whether the current structure inhibits science opportunities.

SMD's Efforts to Meet Goals

Beginning in early 2018, the SDMWG started laying the groundwork for SMD's strategic plan. The working group began to collect input from key stakeholders, including but not limited to: members of the scientific community, academic institutions, other government agencies, the private sector, professional societies, the general public, advocacy groups, and international collaborators.

To facilitate input, the following efforts were used as forums of information sharing:

- Archives Processing and Data Exploitation Meeting, August 2018

- Workshop on Maximizing the Scientific Return of NASA Data, October 2018

- Request for Information (RFI): Strategic Plan for Scientific Data and Computing, September 2018 – November 2018

Through this outreach, the SMDWG coalesced around five key themes from stakeholders: **open-source**, **cloud computing**, **data management**, **metadata standardization**, and **machine learning/artificial intelligence**.

Next Steps

- Evaluation of RFI responses to brief SMD Senior Management of analysis and results

- Briefing OMB, OSTP, and Congress on the promise offered by adopting current techniques

- Broader community inputs via the National Academy of Sciences, Science Advisory Committees, and various topical meetings, e.g., AGU, AAS

- Soliciting via the ROSES NRA pilot programs to implement aspects of the strategic data management
  - NASA sponsored open source code development and cross-discipline data analysis

Resources

- [Science @ NASA](#)
- [NASA Open Data Portal](#)
- [Earth Data](#)
- [Heliophysics Data](#)
- [Planetary Data](#)
- [Astro Data](#)
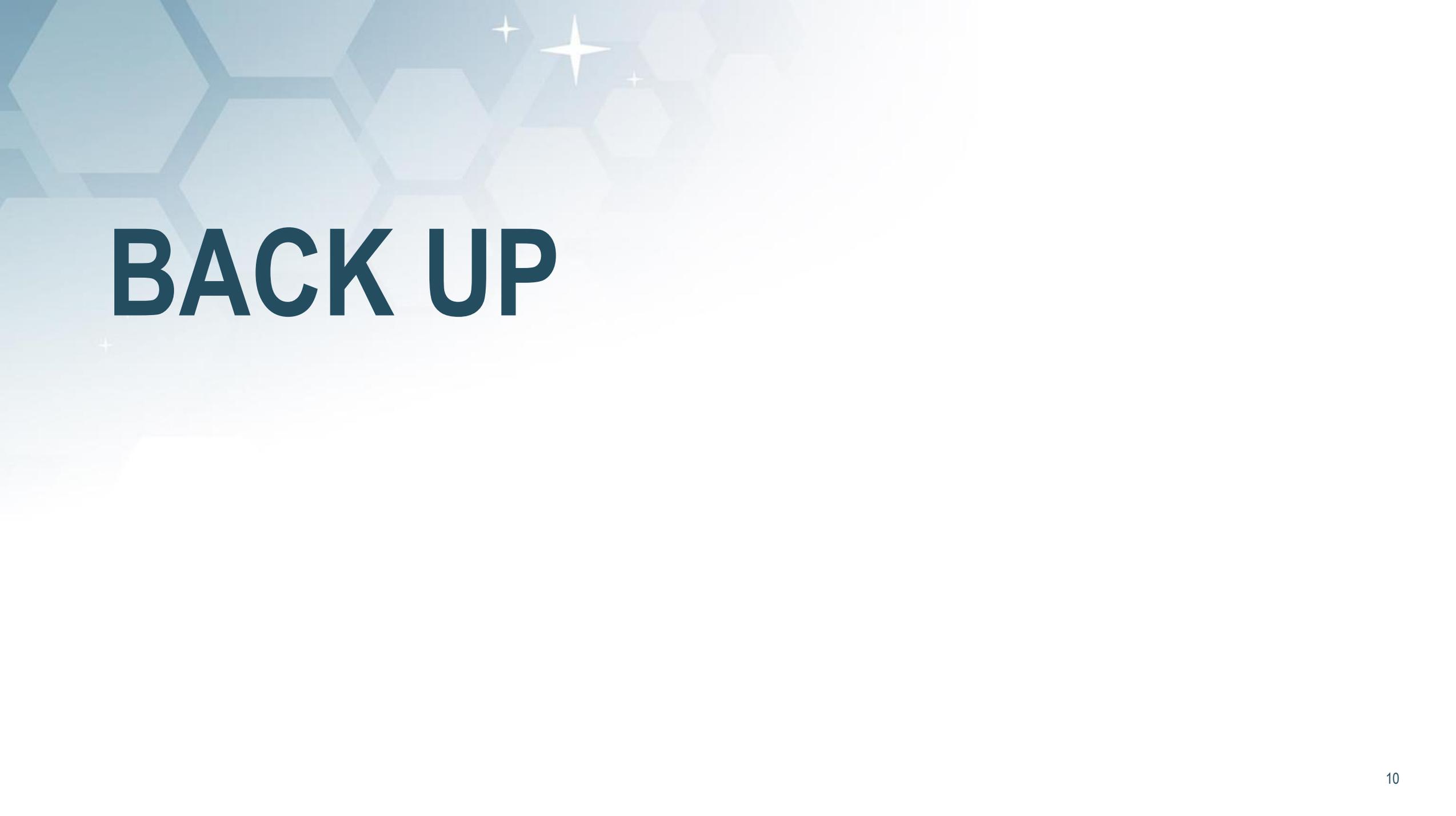
BACK UP

Working Group Members

Name	Affiliation
Ellen Gertsen, Co-Chair	NASA HQ - SMD Front Office
Kevin Murphy, Co-Chair	NASA HQ - ESD
Rebecca McCauley Rench	NASA HQ – PSD
Jeffrey Hayes	NASA HQ - HPD
Pat Knezek	NASA HQ - APD (NSF Detailee)
Bill Knopf	NASA HQ - PSD
Janet Kozyra	NASA HQ – HPD (IPA)
Tsengdar Lee	NASA HQ - ESD
Viet Nguyen	NASA HQ - JASD
Alex Young	NASA GSFC - Science Activation
Jared Leisner	NSA HQ - HPD
Horace Mitchell	NASA GSFC - Science Activation

Hubble Space Telescope Publication Statistics

HST Publication Rates

http://archive.stsci.edu/hst/bibliography/pubstat.html
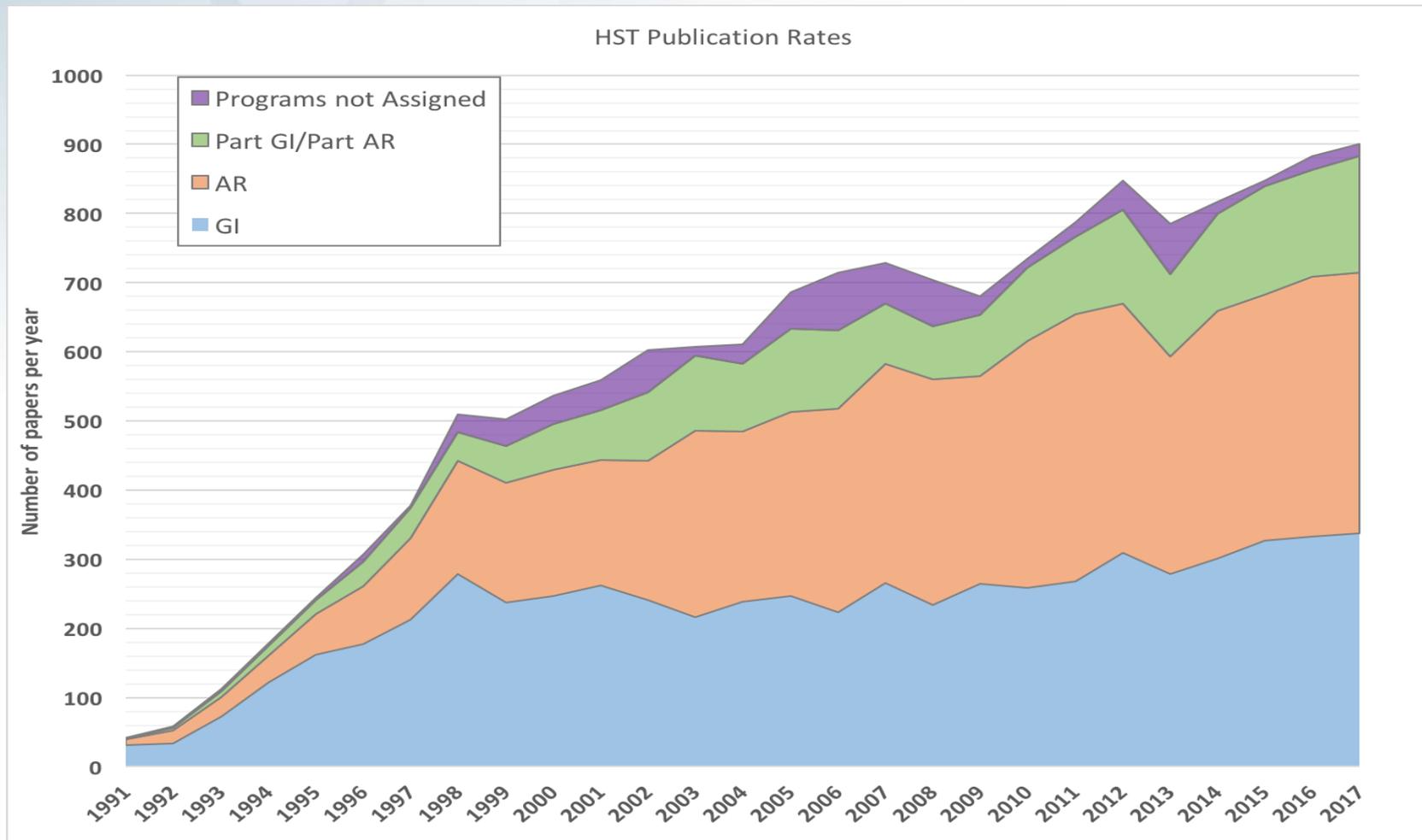
Moderate Resolution Imaging Spectroradiometer (MODIS) Publication Statistics

Publications

The MODIS Publications section is a complete reference of peer reviewed MODIS-related publications that have been written by the scientific community.

Currently the MODIS publications database houses 12820 unique publications, with 12757 related abstracts

To view a listing of citations, please use the links below. If you choose to view a listing by either the year, a specific authoring team member, or discipline simply hover over that link. If you prefer to search the publication database by keywords, please use hover over the search links.

View Publications By Year			
2018 (1222)	2017 (1780)	2016 (1721)	2015 (1188)
2014 (1374)	2013 (617)	2012 (820)	2011 (499)
2010 (734)	2009 (505)	2008 (624)	2007 (466)
2006 (354)	2005 (314)	2004 (156)	2003 (132)
2002 (86)	2001 (42)	2000 (46)	1999 (30)
1998 (37)	1997 (25)	1996 (16)	1995 (6)
1994 (8)	1993 (2)	1992 (6)	1991 (6)
1989 (2)	1987 (2)		

https://modis.gsfc.nasa.gov//sci_team/pubs/