International Cooperation For Developing and Sharing Environmental Data, and Their Use

GHASSEM R ASRAR
JOINT GLOBAL CHANGE RESEARCH INSTITUTE
PACIFIC NORTHWEST NATIONAL LABORATORY@
UNIVERSITY OF MARYLAND

14 March 2018

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

Outline

- Introduction
  - Need for science-based information
  - Earth-human systems science and Earth observations
- Example 1: Global Observing System of Systems
  - Policy and practice for governing partnerships
  - Policy and practice for data/information stewardship and sharing
- Example 2: Weather and Climate Monitoring and Prediction
  - Policy and practice for collecting, managing and sharing source data
  - Practice and policy for forecasting, evaluating and sharing resulting information
- Example 3: Modeling and Predicting Earth System Federation
  - Collecting, processing and sharing of data for models
  - Simulating, evaluating and sharing of the resulting data/information
- Summary and Conclusions

Need for Science-Based Information

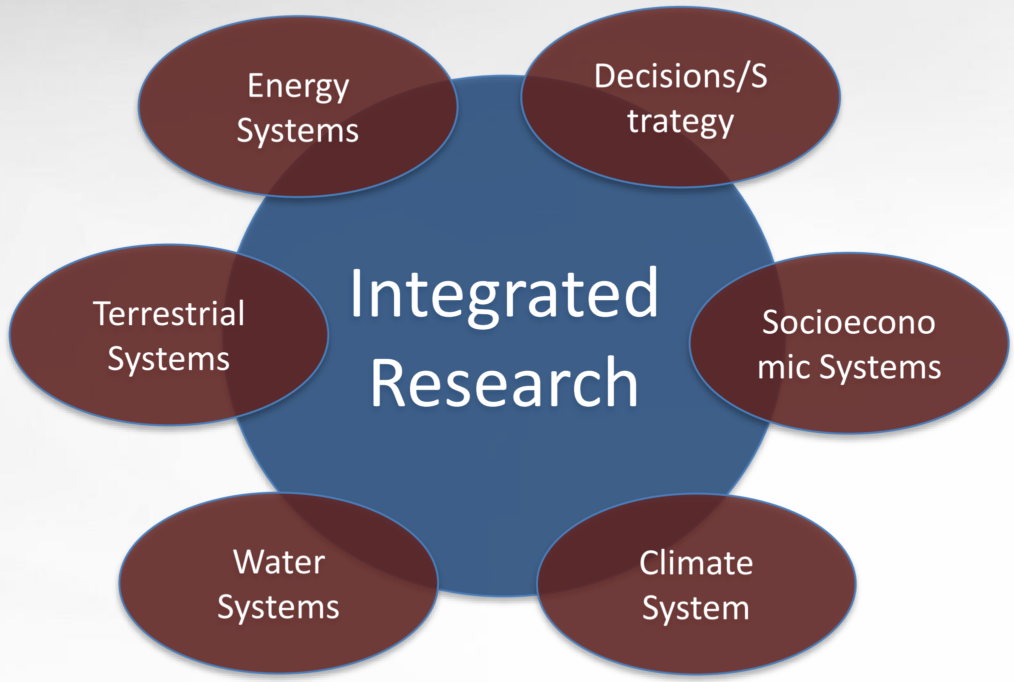
- Increasing cross-scales(i.e. time and space), cross-sectors (e.g. food, water, energy) complexity an interdependencies of many societal-environmental challenges;
  - Land Use Land Cover: cropping, forest cover, wildfire, energy production, conservation, feedbacks, etc.
  - Water availability, quality and use: demand, surface and sub-surface supply, recharge rates, conflict potential, etc.
  - Energy access and use: seasonal changes in demand, supply mix, renewable integration, relative prices, reliability, etc.
  - Urbanization dynamics: size, density, location on marginal lands, health issues, quality of life, rapid growth, etc.
  - …
- All are affected by local to global conditions including demographics, markets, security threats, and other environmental stressors…
- Managing risks and opportunities will benefit from integrating data and information from multiple sources with ability to analyze inter-relationships, feedbacks, tradeoffs, etc.

Fully Integrated Inquiry and “What If” Scenarios Analysis

- How will growth and changes in the global demand for energy, water, and agricultural goods affect each other and impact social stability?
- What are the factors that drive these changes?
  - Energy supply sources and demand sectors
  - Socio-economics (economics, demographics, and migration)
  - Technology (generation, conversion, transmission)
  - Policies (efficiency, conservation, new technologies)
  - Climate and weather extremes
- How do these drivers interact (positive or negative feedbacks)? What are the potential impacts/ consequences changes?
- **Why, When, Who, What if scenarios….?**

Integrated Earth-human System Research & Analysis

- **What changes are occurring now** in environmental, social, institutional, and technological systems?
- **What risks and opportunities could result in the future** from the interactions of environmental changes with evolving social, institutional, and technological systems?
- **What are the options for responding** to, and increasing resiliency in the face of, changing environmental, social, and technological conditions?
- **What are the options for reducing adverse impacts** on the environment?

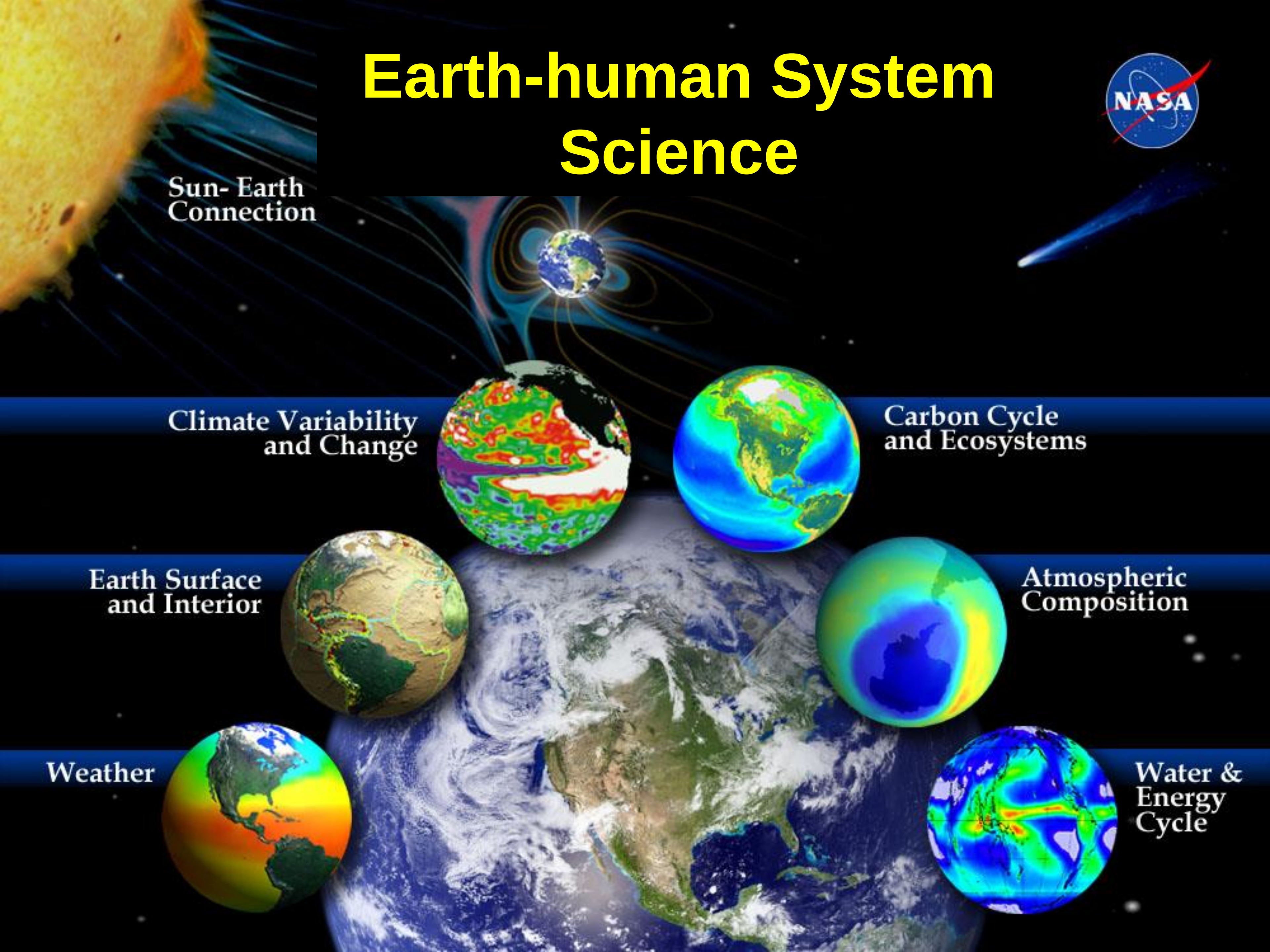


Maintaining disciplinary focus, but conducting interdisciplinary and integrated research to answer these complex questions.

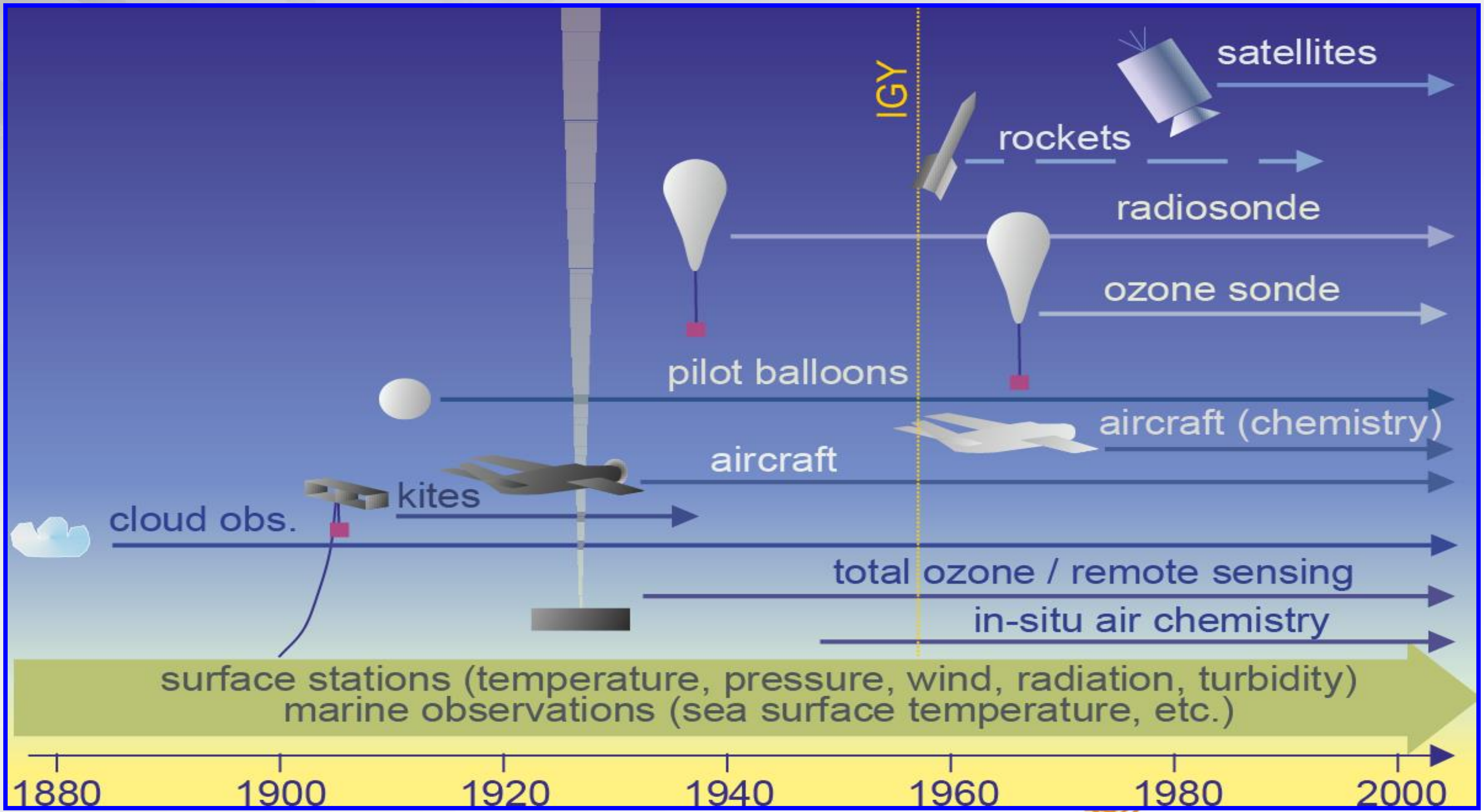
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

Earth-human System Science



The Changing Earth Observing Systems & Technologies



Courtesy, S. Brönnimann

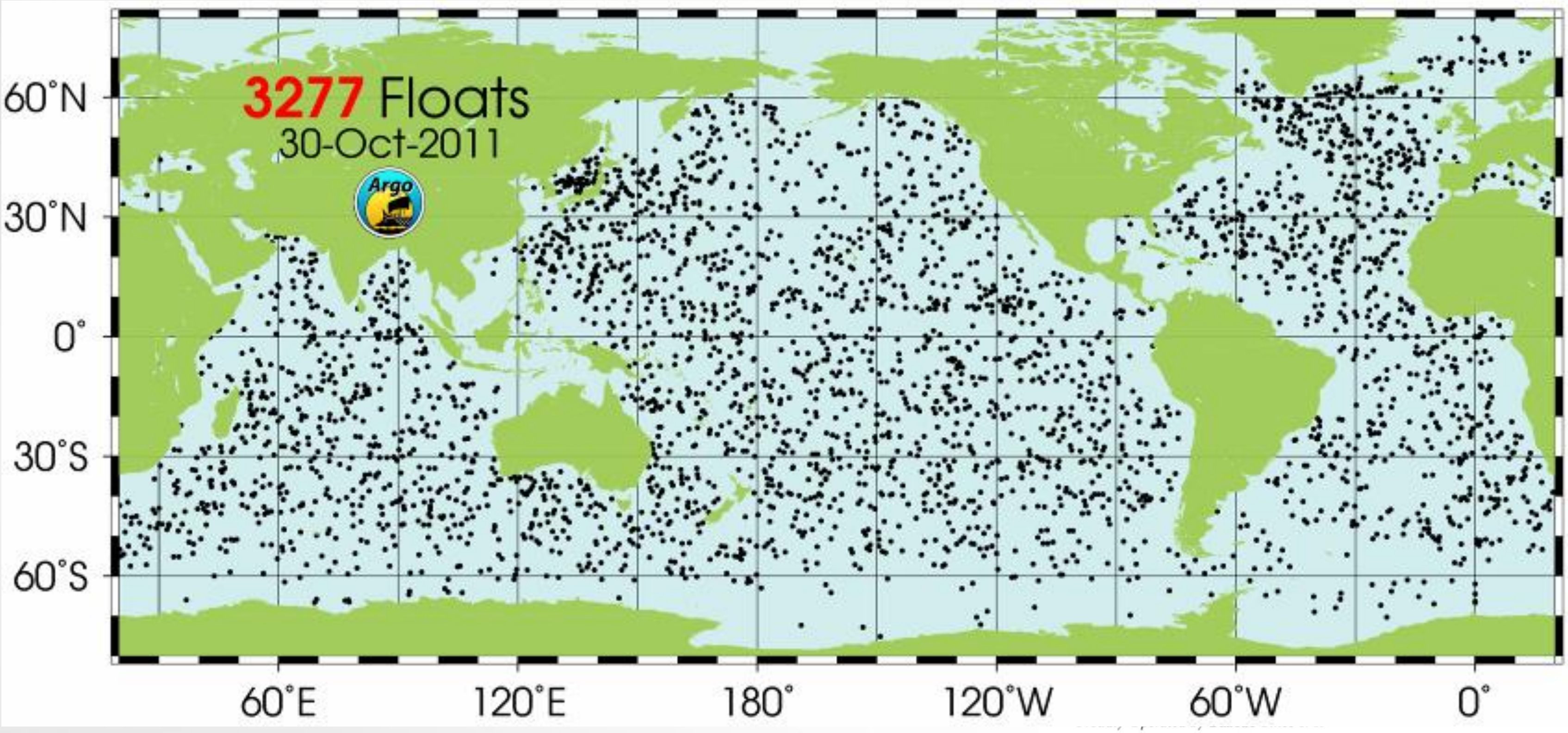
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

International Earth Observing Systems

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

International Ocean Observing Systems



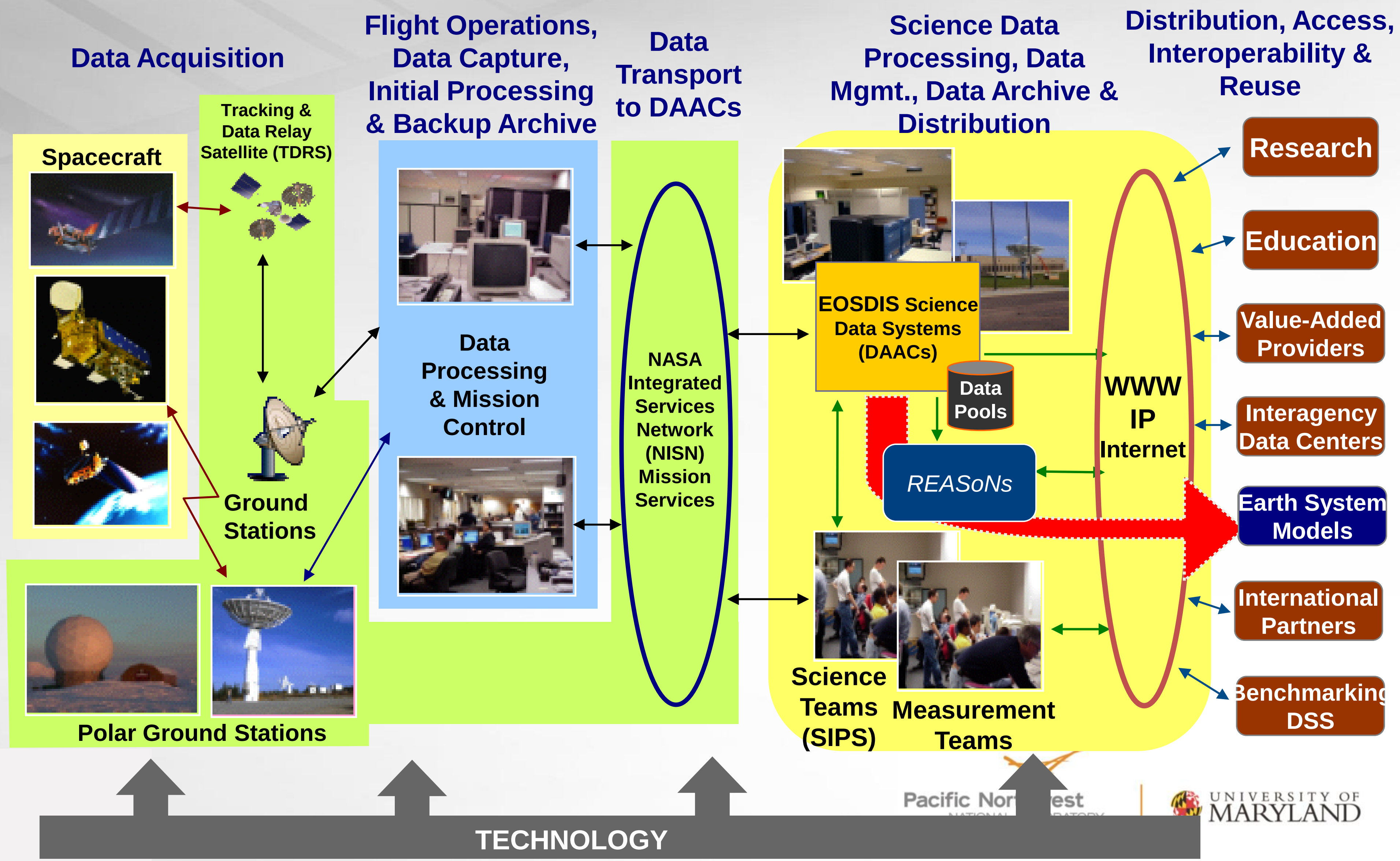
Scientific & Technical Challenges

- **Multiple Scales-** time, space and processes
- **Complex Systems and Feedbacks**- energy, water, food, climate
- **Multiple Data Sources**- consistency in definitions of the same phenomenon by different sources
- **Data Validation, and Quality Assurance**
- **Uncertainty Analysis and Characterization**
- **Data Curation, Stewardship and Dissemination/Sharing**
- **Computation, Visualization and Analysis Capabilities**—particularly for integrated Earth-human systems models/data analyses
- **Lack of Data**- specially for socio-economics aspects

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

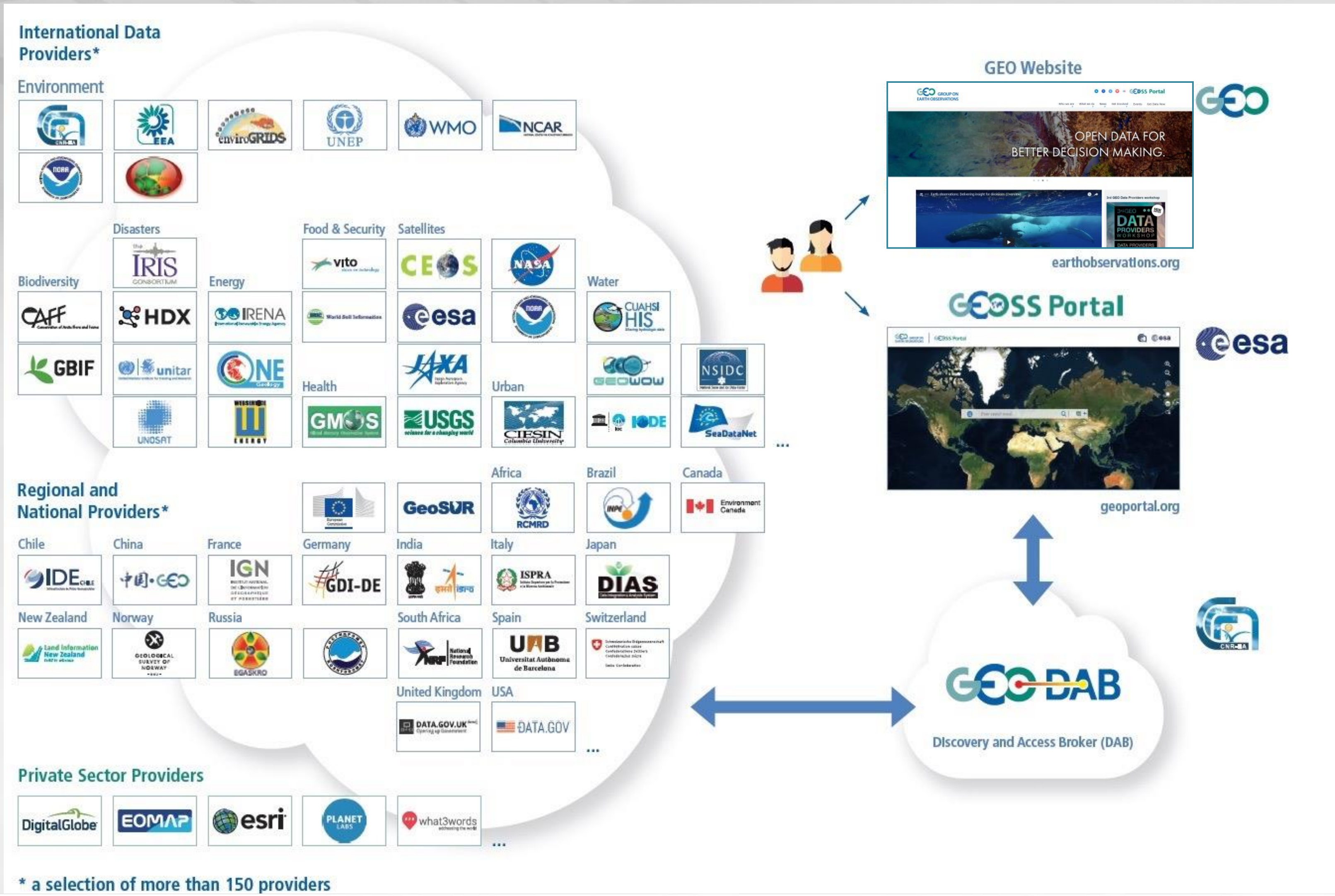
UNIVERSITY OF MARYLAND

A Comprehensive Data Acquisition, Analysis & Dissemination System



Pacific Northwest National Laboratory | University of Maryland

The Global Earth Observing System of Systems (GEOSS) Platform



GEOSS Data Sharing Principles

- Full and Open Exchange of Data -- Open by Default;
- Data and Products at Minimum Time Delay and at Minimum Cost; and/or
- Free of Charge or Cost of Reproduction.

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

GEO in Numbers

7 continents

8 societal benefit areas

12 years

60 work programme activities

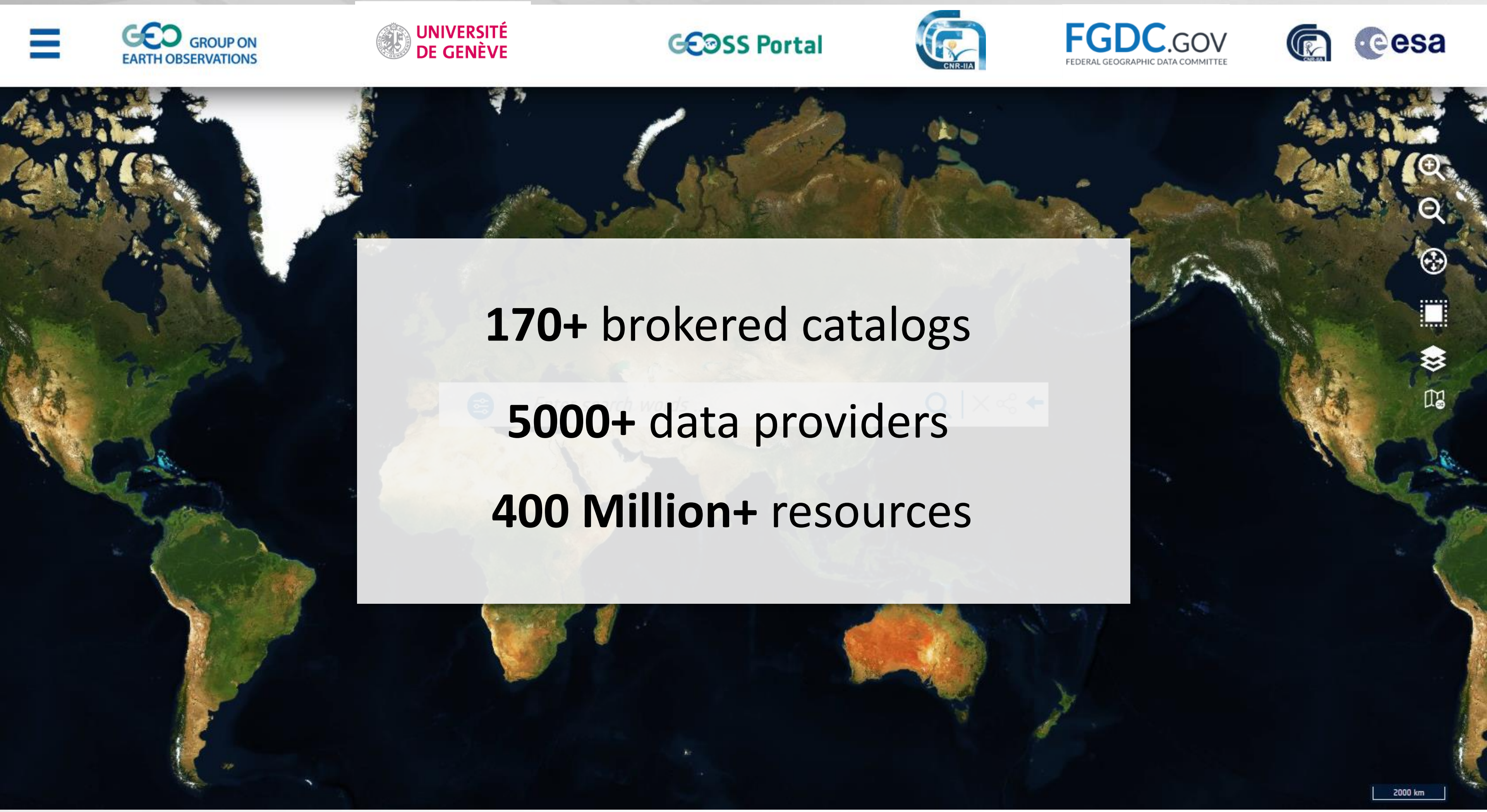
105 members

115 participating organizations

165 data providers

400,000,000 Earth observations

The GEOSS Platform



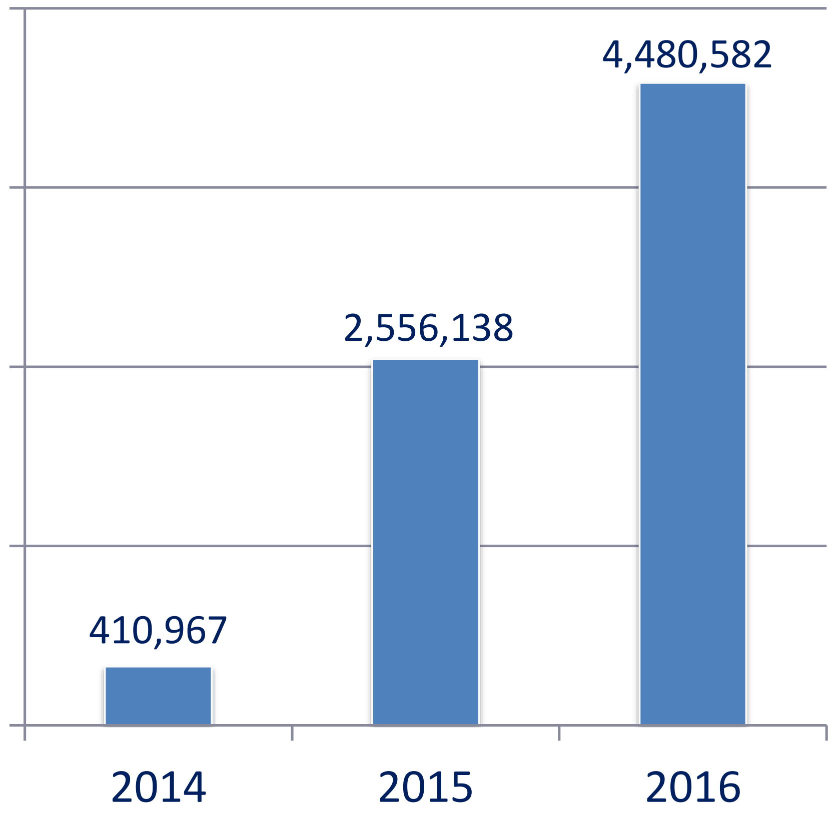
170+ brokered catalogs

5000+ data providers

400 Million+ resources

GEOSS Platform Use

(# of Queries -- People and Machines)

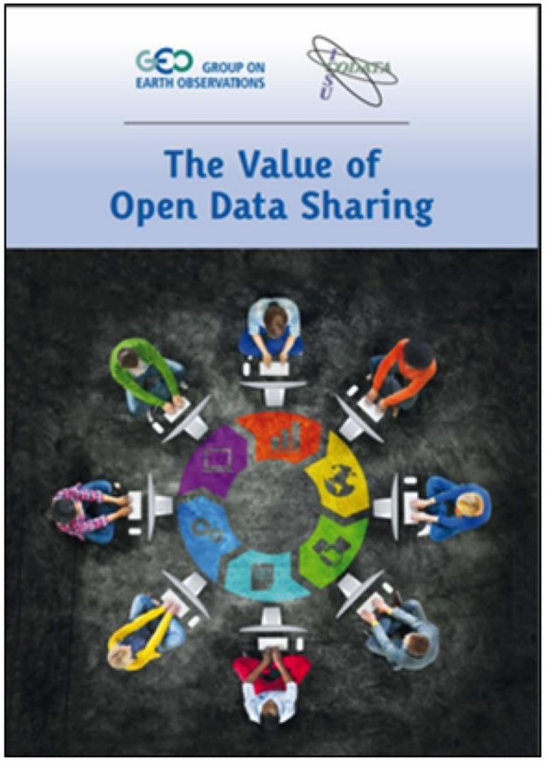


Data Sharing

Open Data for the Benefit of Humankind

How do countries benefit from open data?
There are many diverse opportunities and benefits from providing open data for unrestricted use worldwide.

- ECONOMY
- SOCIETY
- RESEARCH & INNOVATION
- EDUCATION
- GOVERNANCE



Top Institutional Use (2016)

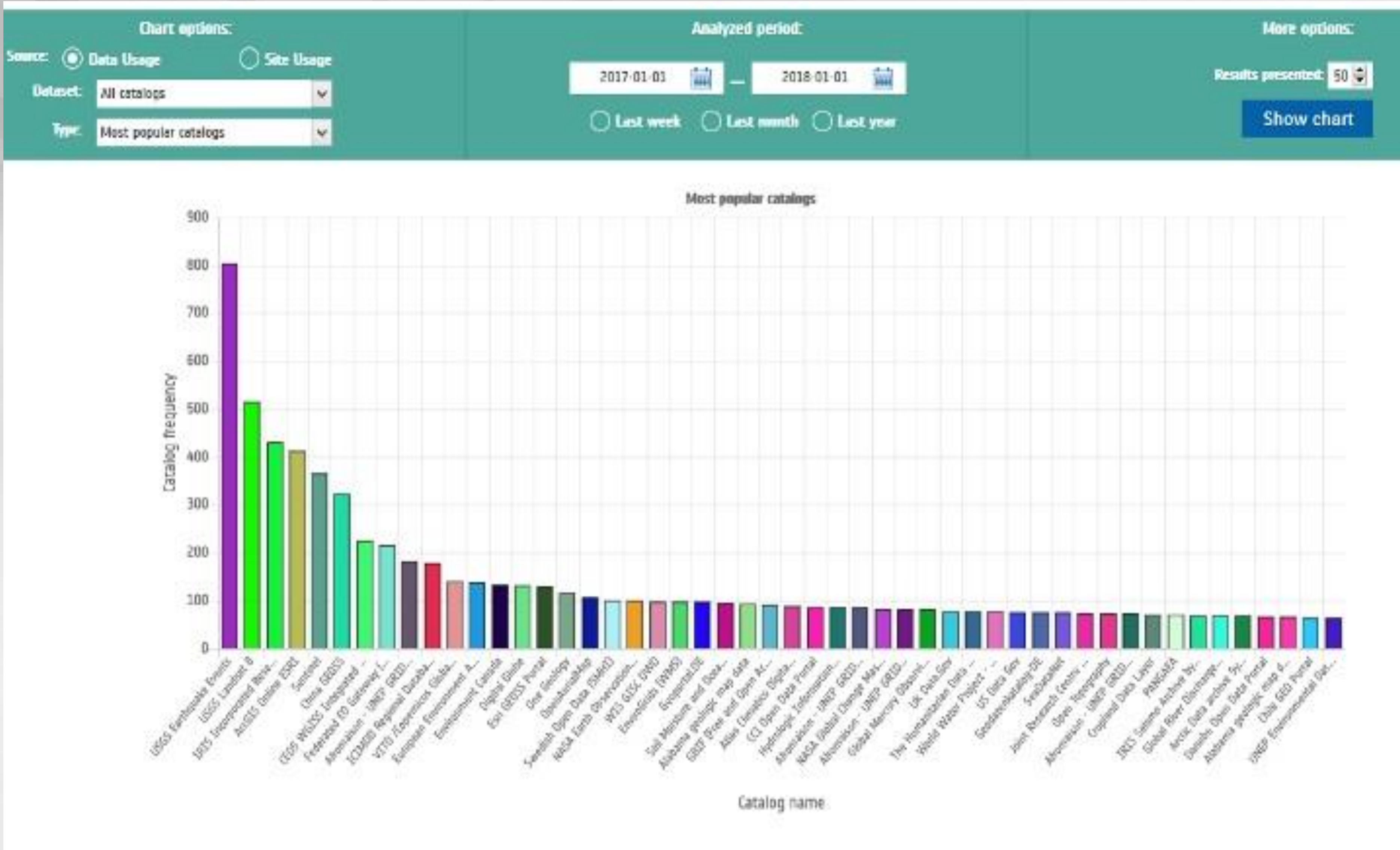
Entities Submitting Queries	Number of Queries	Order of Magnitude
WMO Information System	Very High	> 1 M
Cloud Infrastructures	High	> 100 K < 1 M
ESRI ArcGIS Online	Medium	> 10 K < 100 K
US GEO	Medium	> 10 K < 100 K
EC H2020 Projects	Medium	> 10 K < 100 K
China GEO	Medium	> 10 K < 100 K
CEOS Water Portal	Low	> 1 K < 10 K
UNEP GRID	Low	> 1 K < 10 K

Int. Orgs **GEO Members** **Private Sector** **Other**

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

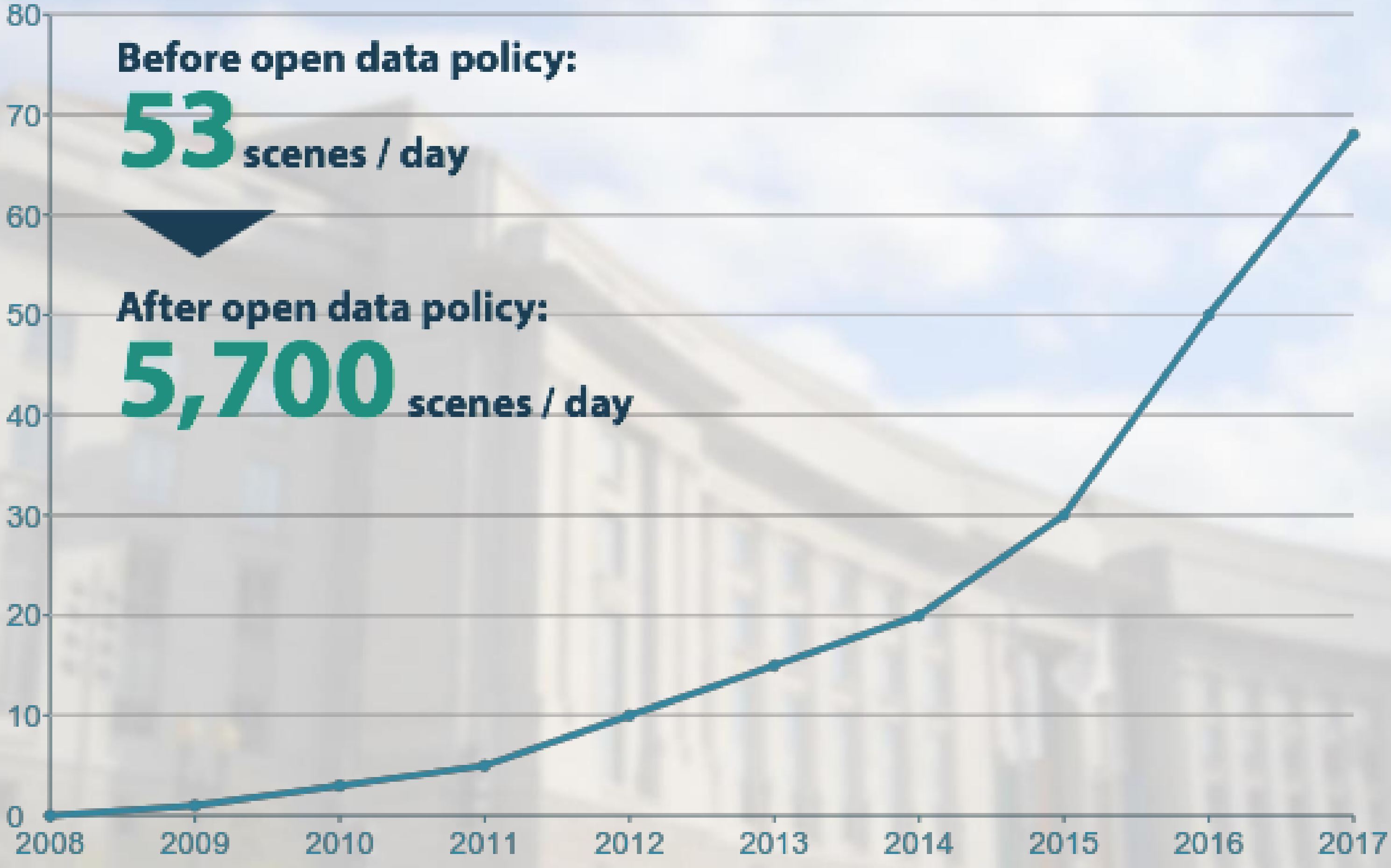
UNIVERSITY OF MARYLAND

Most Popular Data Catalogs (2017)

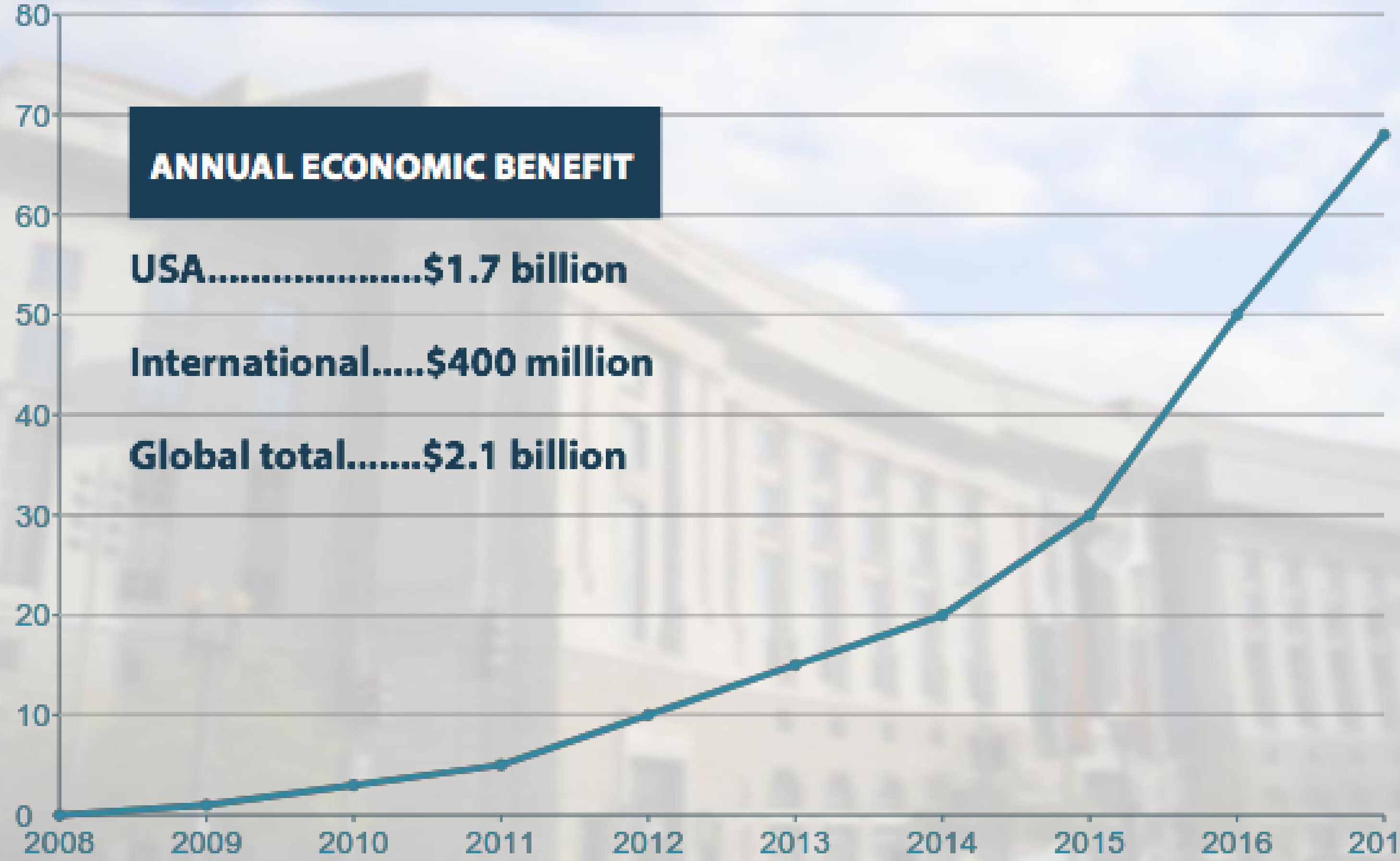


Proudly Operated by **Battelle** Since 1965

MILLIONS OF LANDSAT SCENE DOWNLOADS



MILLIONS OF LANDSAT SCENE DOWNLOADS



"The economic value of just one year of Landsat data far exceeds the multi-year total cost of building, launching, and managing Landsat satellites and sensors." **Landsat Advisory Group of the National Geospatial Advisory Committee**

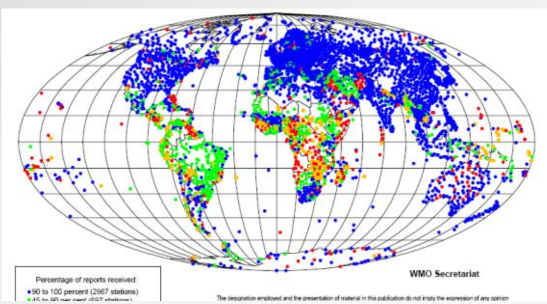
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

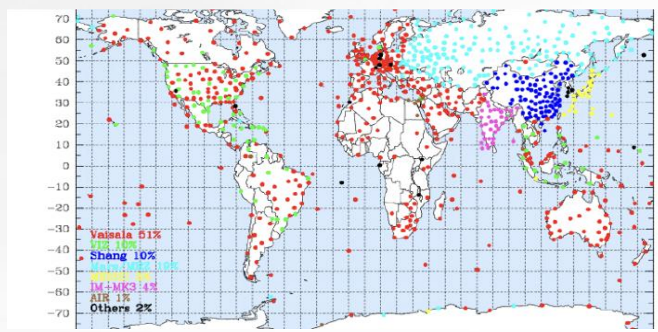
World Meteorological Organization: Members, Services and Data Policy

- UN Specialized Agency on weather, climate & water
- 191 Members, HQ in Geneva
- 2nd oldest UN Agency, 1873-
- Coordinates work of 200 000 national meteorological & hydrological experts, academia
- Co-Founder and host agency of IPCC (1st World Climate Conference)
- Co-Founder of UNFCCC (2nd World Climate Conference)
- Global Observing Networks >10000 stations, and operational weather satellites

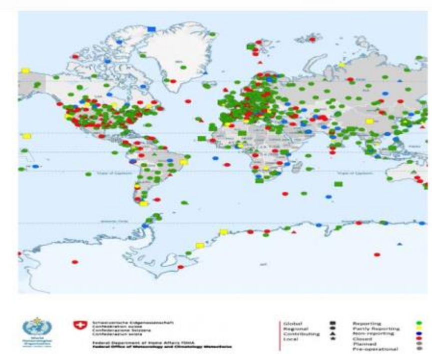
Surface observations



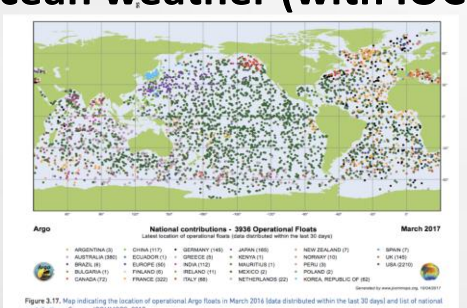
Balloon soundings

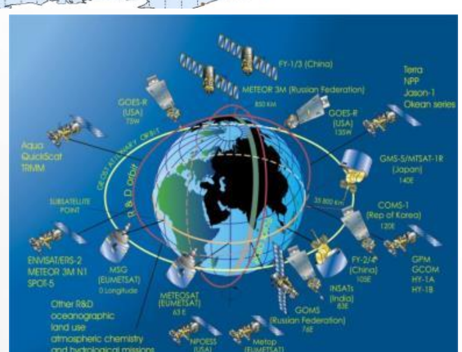


Air quality/greenhouse gases



Ocean weather (with IOC UNESCO)





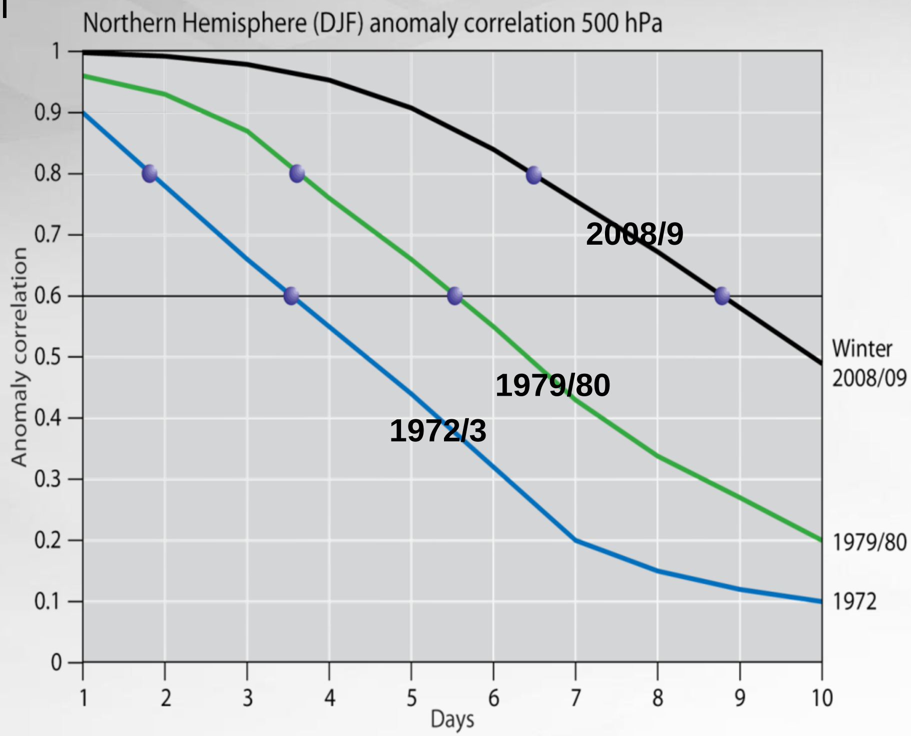
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

Enhancing Weather Forecast Skill

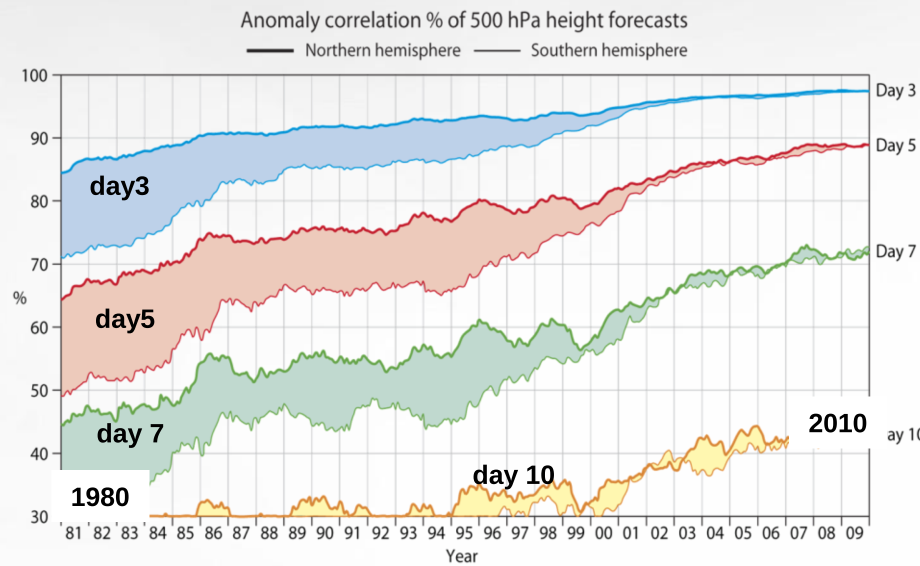
NH winter 500Z anomaly correlation

skill



Day of forecast

NH & SH 500Z ACC through the years



Year

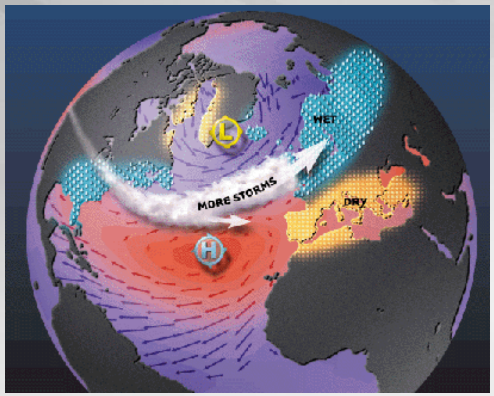
Courtesy of ECMWF

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

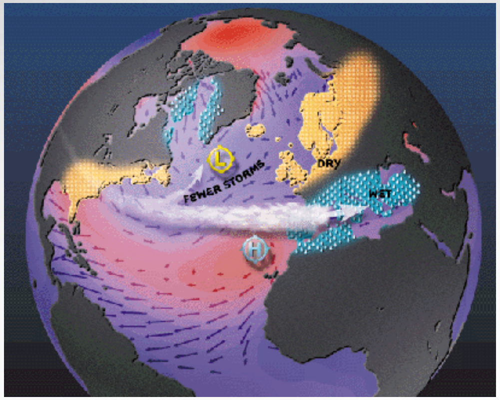
UNIVERSITY OF MARYLAND

Month – Seasons: The North Atlantic Oscillation

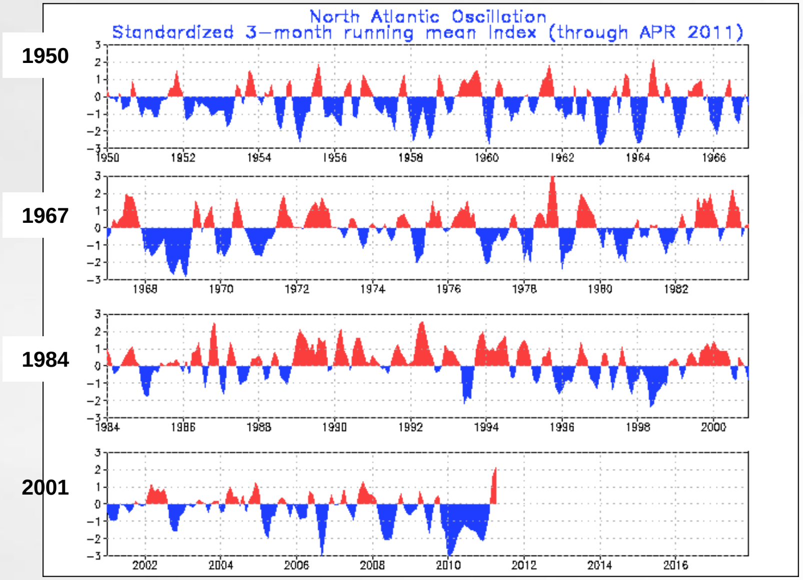
Positive NAO phase



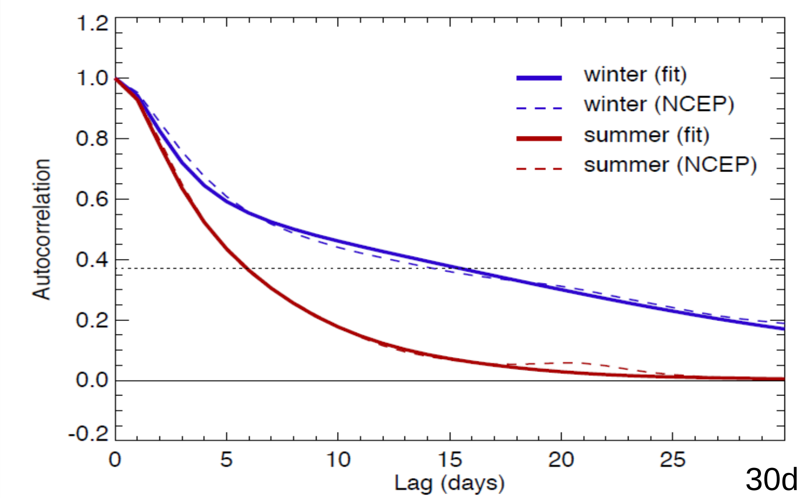
Negative NAO phase



3-month running mean of NAO index 1950-date



Autocorrelation



Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

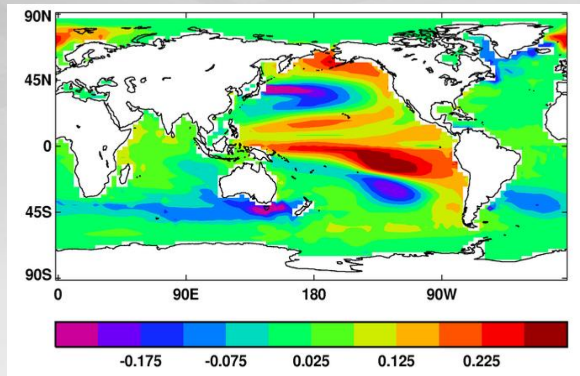
The University of Reading

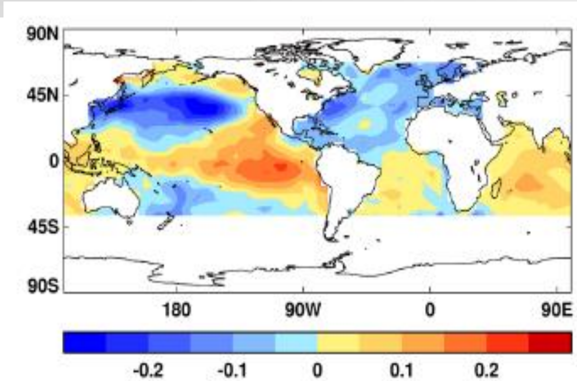
Natural Earth System Decadal Variability

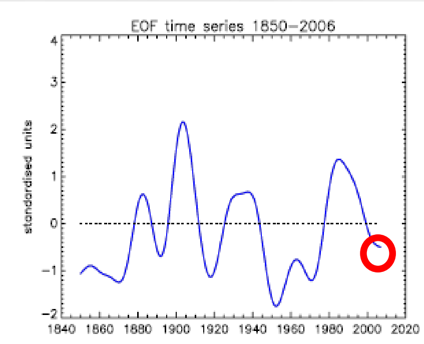
Model

Observation

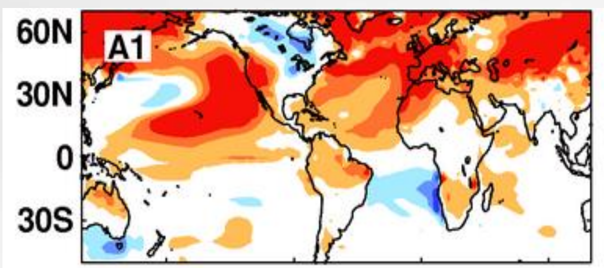
Pacific Decadal Oscillation

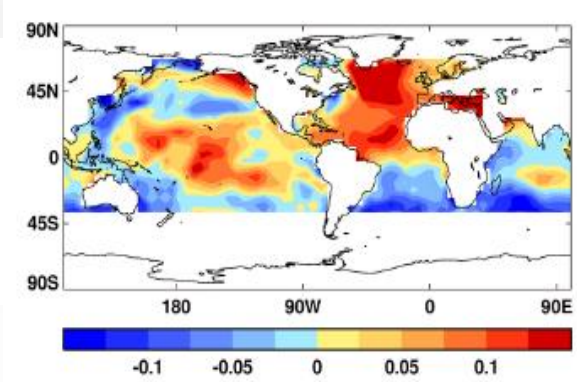


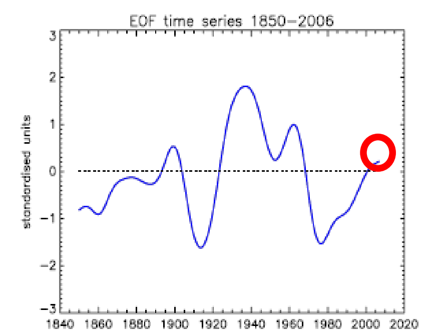




Atlantic Multidecadal Oscillation



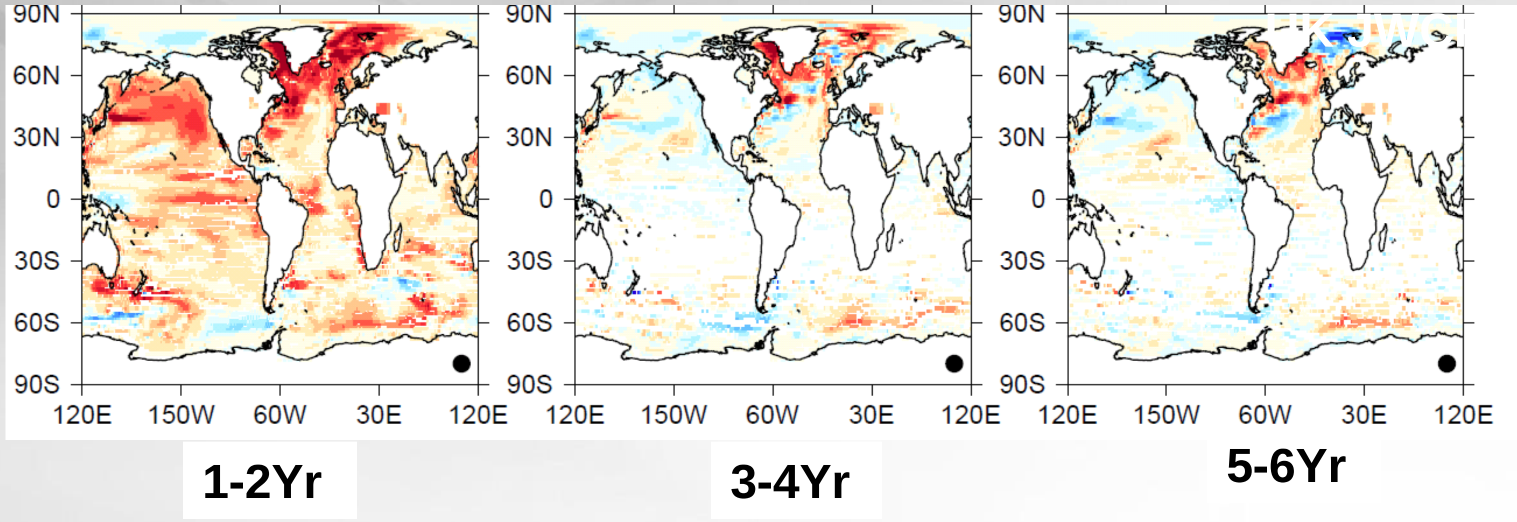




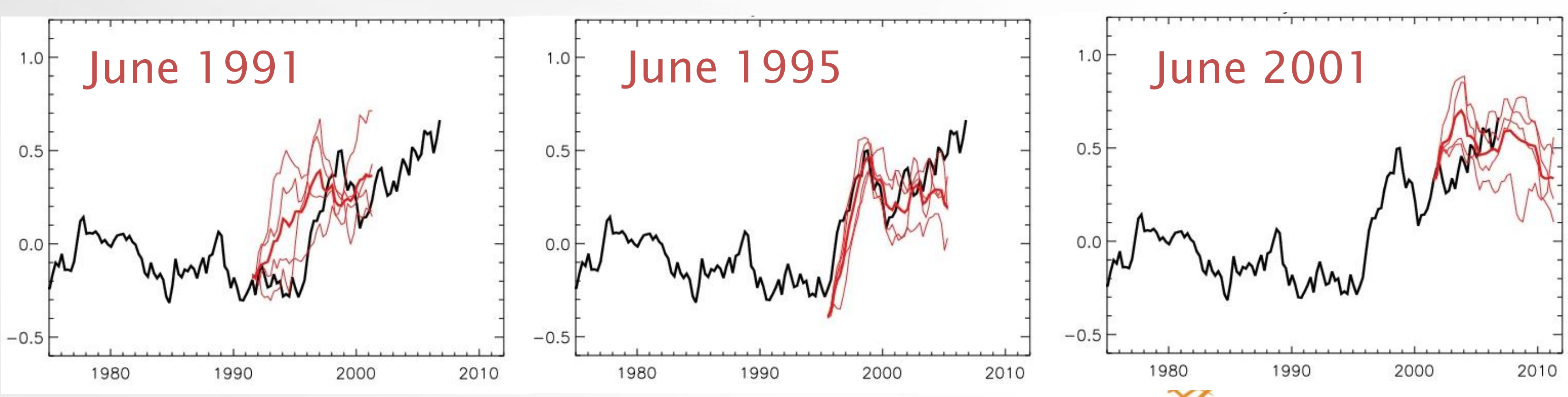
Knight et al 2005, Parker et al 2007

Earth System Prediction Skill on 1-10 Year Time-scale

Heat in top 100m ocean: Improvement in Skill from initialisation



Hindcast predictions of 500m heat content in Atlantic sub-polar gyre



Courtesy of UK MetOffice

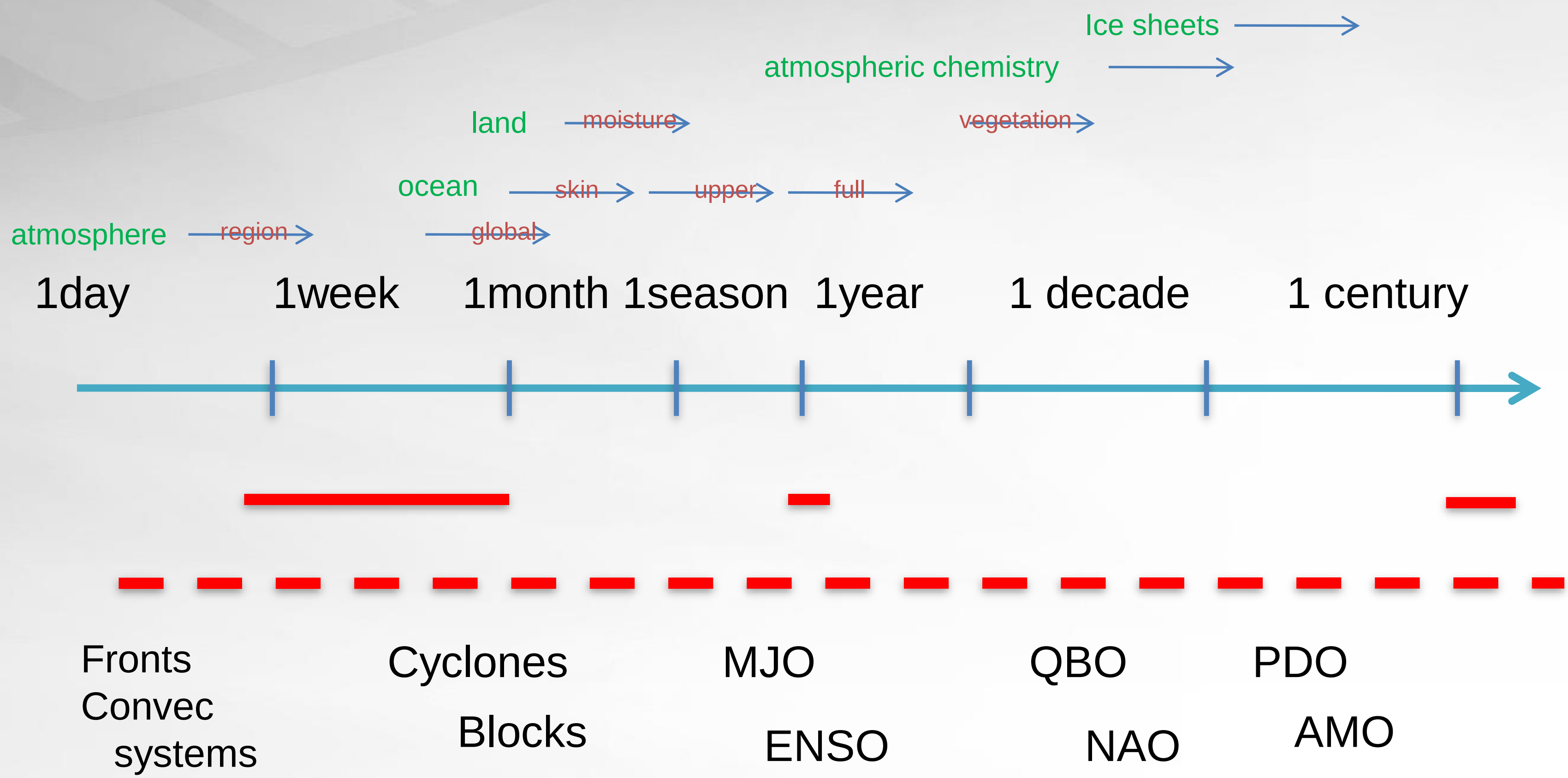
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

Data Acquisition, Access and Sharing Policy

WMO's principle of free and unrestricted exchange of hydrological (Resolution No. 20) and meteorological and related (Resolution No. 40) data and products;

1. Members shall provide on a free and unrestricted basis those hydromet data and products which are necessary for the provision of services in support of the protection of life and property and for the well-being of all people;
2. Members should also provide additional hydromet data and products, where available, which are required to sustain programs and projects of WMO, other United Nations agencies, ICSU and other organizations of equivalent status, related to water resources research at the global, regional and national levels and, furthermore, to assist other Members in the provision of hydromet services in their countries;
3. Members should provide to the research and education communities, for their non-commercial activities, free and unrestricted access to all hydrological data and products exchanged under the auspices of WMO;
4. Respecting (2) and (3) above, Members may place conditions on the re-export 3, for commercial purposes, of these hydro- logical data and products, outside the receiving country or group of countries forming a single economic group;
5. Members should make known to all Members, through the WMO Secretariat, those data and products which have such conditions as in (4) above;
6. Members should make their best efforts to ensure that the conditions placed by the originator on the additional hydro- logical data and products are made known to initial and subsequent recipients;

Seamless Prediction of Earth System Across Time and Space



The University of Reading

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

Earth System Grid Federation (ESGF): A Framework for Collaboration on Earth System Modeling

- **ESGF is an international collaboration of centers working together to manage and provide access to Earth/climate system data, models and observations**
- **Started a decade ago, it is now the world premier data focused technology infrastructure in support of Earth system science**
- **Spanning a wide range of institutions in Europe, North America, Asia, Australia, with funding from a variety of agencies**
- **Open Access-ESGF archive is available to all users, except commercial applications. Data is made available readily and can be corrected locally, if needed.**

is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA

NCI

canarie

NOAA · NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION · U.S. DEPARTMENT OF COMMERCE

NASA

NSF

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

UNIVERSITY OF MARYLAND

Earth System Grid Federation (ESGF): A Framework for International Collaboration in Earth System Modeling

The ESGF value emerges from being a federation of sites, users have access to data, metadata and services that are provided by multiple sites distributed around the world. Interoperability across sites is based on;

- **Sites operate and provide services that are using standard protocols and APIs (Secured Socket Layer-SSL, OpenID, Wireless Intrusion Protected System- WIPS, ...)**
- **Services are configured to know about their peers (e.g. search, IdPs, CoGs,..)**
- **Sites establish trusting relationships through SSL certificates**

The ESGF is currently a federation of:

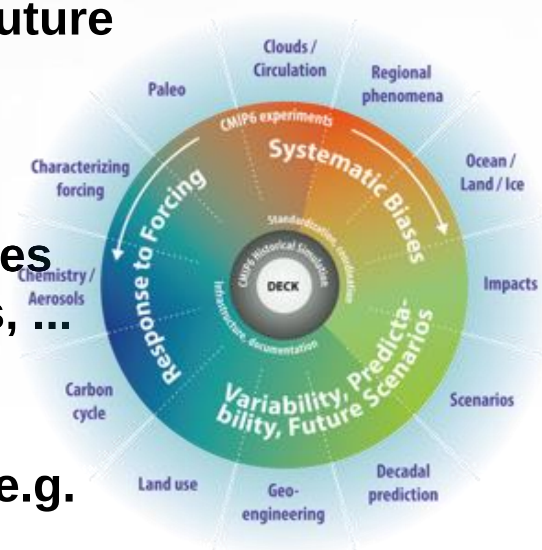
- **31 data nodes, 11 index nodes-IdPs**
- **23 data projects, 198,298 datasets, 3,309,528 files (non- replica, latest version)**
- **147 CoG projects, 15,571 users across all CoG sites**

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

Earth System Grid Federation (ESGF): Scientific & Technical Challenges/Opportunities

ESFG Supports the Coupled Model Intercomparison Project (CMIP-6) for Intergovernmental Panel on Climate Change (IPCC) Assessments:

- 10X the CMIP5 data volumes (25-40 PB of data)
- Increased users, number of ESGF nodes, and experiments
- Integrate a larger number of observational datasets
- Critical for validation and scoring of climate models
- More Obs4MIPs data collections, data providers
- Scale performance of services to match the expected future data volumes
- Searching, downloading, sub-setting, processing
- Use new architectures and more modularization
- Interoperate with other data infrastructures and agencies such as NASA DAACs, NOAA, USGS,DOE Copernicus, ...
- Provide reliable, efficient data exploration, analysis, visualization and learning
- Take advantage of commodity hardware and services, e.g. Cloud-based services.



Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

Earth System Grid Federation (ESGF): Data Development Priorities & Services

- Easy to deploy, upgrade and use by developers and other users
- Enable/expand server-side APIs and client toolkits to interact with the data
- Enable server-side data processing, sub-setting, re-gridding,..
- Proper citation/identification of data via DOIs for overall system security
- Improve overall documentation

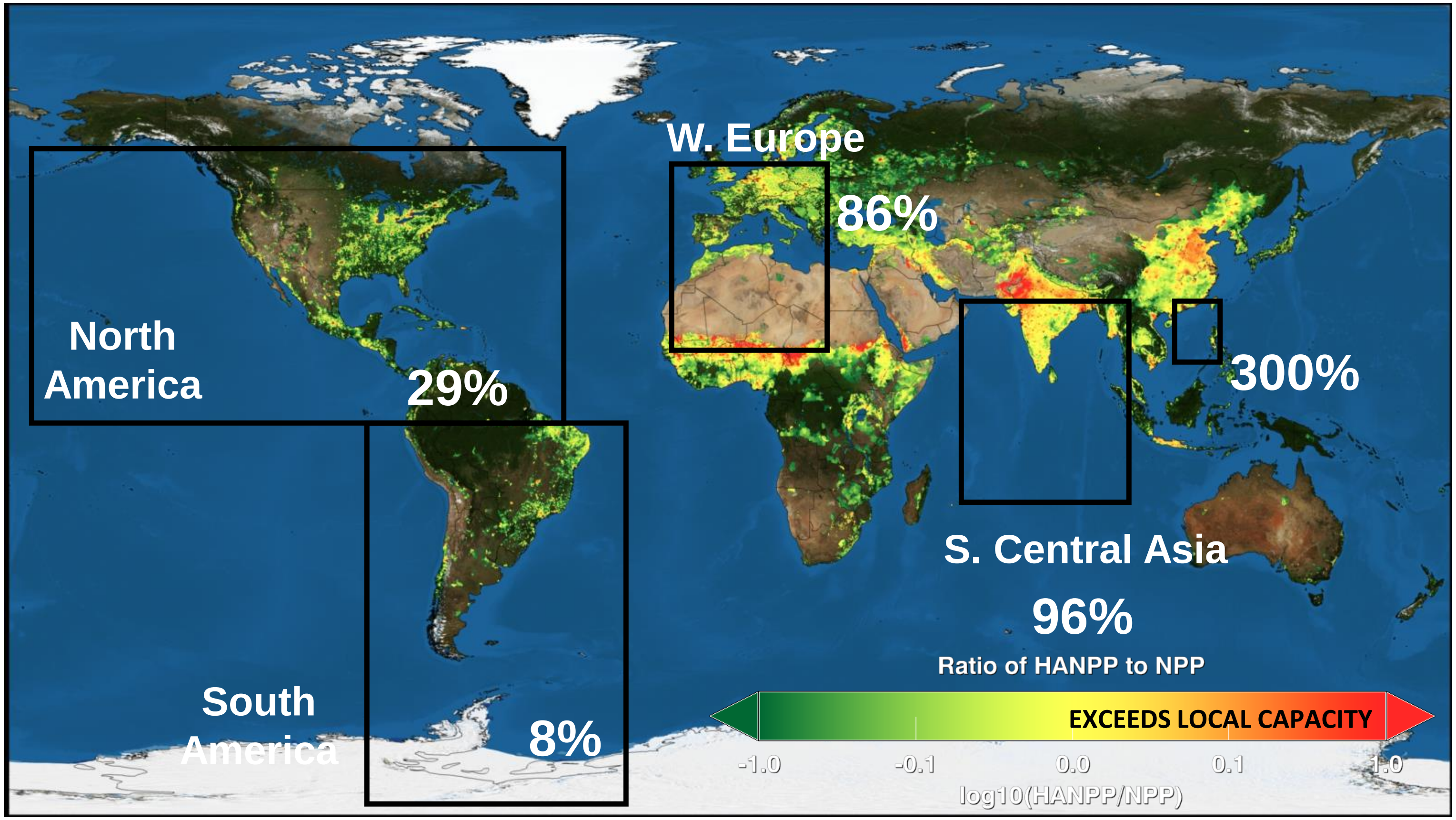
- Persistent Identifiers (PIDs)
  - Provides persistent identifiers for all files and datasets
- Early Citation
  - Provides citable identification for model and experiment granularities before DOIs are assigned
- Digital Object Identifiers (DOIs)
  - Provides fully citable DOI at dataset level, after mandatory information has been populated
- Errata
  - Hosted service to log issues with data at the dataset level, part of ES-DOC eco-system Connected to CoG via the PID landing page
- Metadata
  - Hosted service provides detailed documentation for experiments, models, also part of ES-DOC eco-system

Role of Research in Data Development

- Providing advice on **best datasets** for various purposes, and their merits and limitations.
- Advocating for **improved observations and analysis** suitable for human-Earth systems studies.
- **Developing new datasets**: Develop new products and datasets, analytical and diagnostic techniques, high level derived products, for use in understanding and analyzing Earth-human systems.
- Identifying high priority research needs such as **datasets for use in evaluating Earth-human systems models**, and specifically those used for environmental assessments and future outlooks.
- Promoting sound **data stewardship**, including data archiving, management, and access.
- Helping make datasets **accessible** and usable, and promote data quality and uncertainty characterization.

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
UNIVERSITY OF MARYLAND

Geography of Food-Energy: Human Appropriation of Net Primary Production as Percent of Terrestrial Ecosystems Supply



M. L. Imhoff et al., Nature 429, 870, 2004
M. L. Imhoff et. al., JGR, VOL. 111, 2006

Summary & Conclusions

- The complex socioeconomic and environmental opportunities/challenges transcend individual disciplines and nations, and addressing them benefit greatly from multi-national collaborations.
- Using advances in science, technology and engineering was essential in building the global Earth observing and information systems, that now results in highly diverse and large, but open datasets available for use worldwide.
- The three examples benefited greatly from a wide range of disciplines and sponsors over multiple decades, and they were all built on principles of open data sharing/access.
- They all developed and promoted use of data standards, formats, documentation, quality assurance, and calibration and evaluation based on the use of national/international standards in data curation, stewardship and dissemination/sharing.
- Partnerships and collaborations were used to:
  - Share expertise, resources and experience
  - Develop capacity and infrastructure where needed
  - Sustain, improve and expand the capacity and infrastructure over multiple decades
- The major challenges divided into manageable tasks which helped with timely progress
- Slowest progress made in nations/regions that had greatest needs for data/information
- Slower than expected progress made on adopting and implementing uniformly data stewardship/sharing principles.

Thank You!

Special thanks to:

Mr. Rob Masters and Stefano Belfoire, for providing WMO related information

Dr. Barbara Ryan, for providing GEO/GEOSS related information

Drs. Michel Rixen of WCRP and K. Taylor LLNL/DOE and V. Balaji of GFDL/NOAA for sharing the ESGF/CMIP related information.