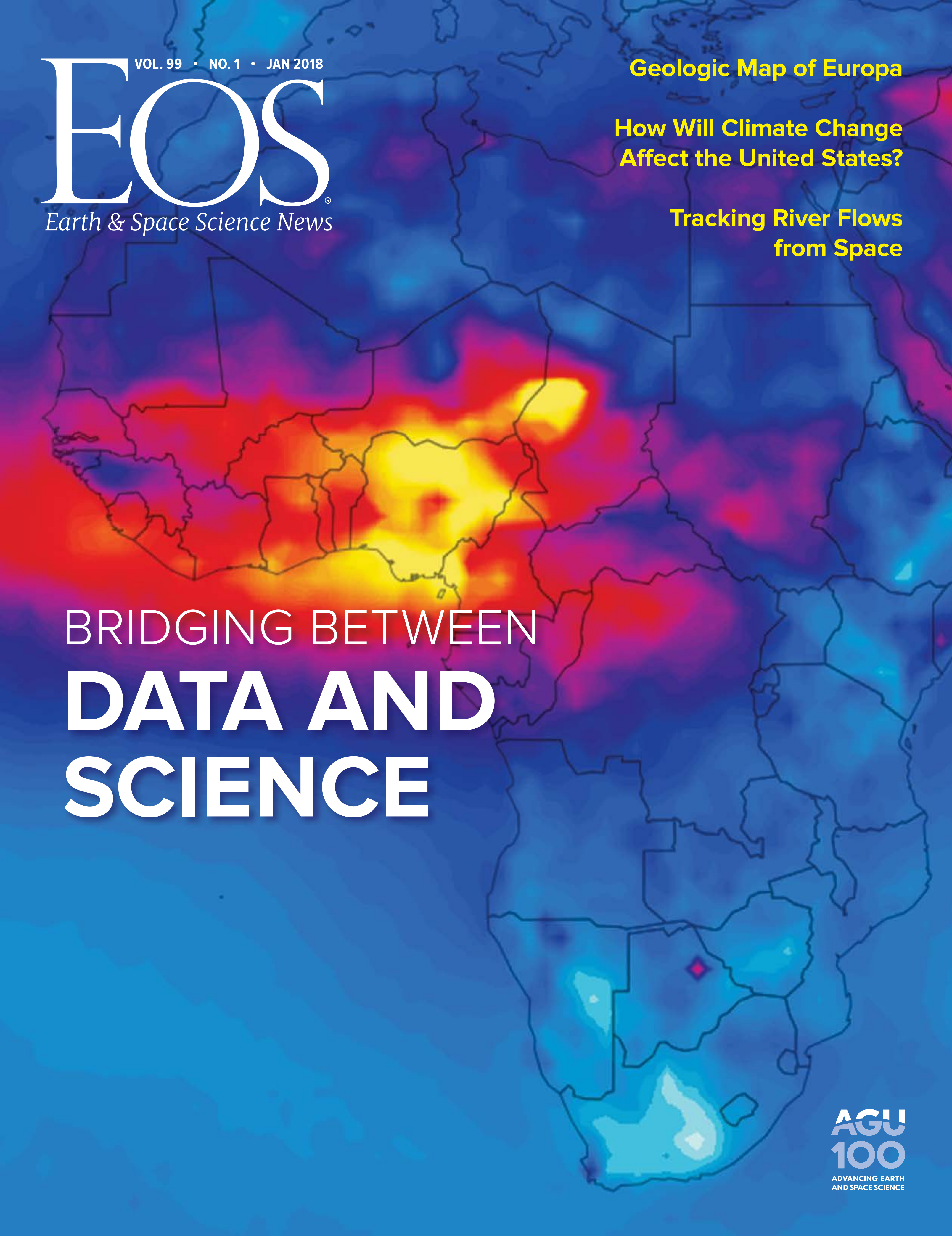


VOL. 99 • NO. 1 • JAN 2018

EOS®

Earth & Space Science News

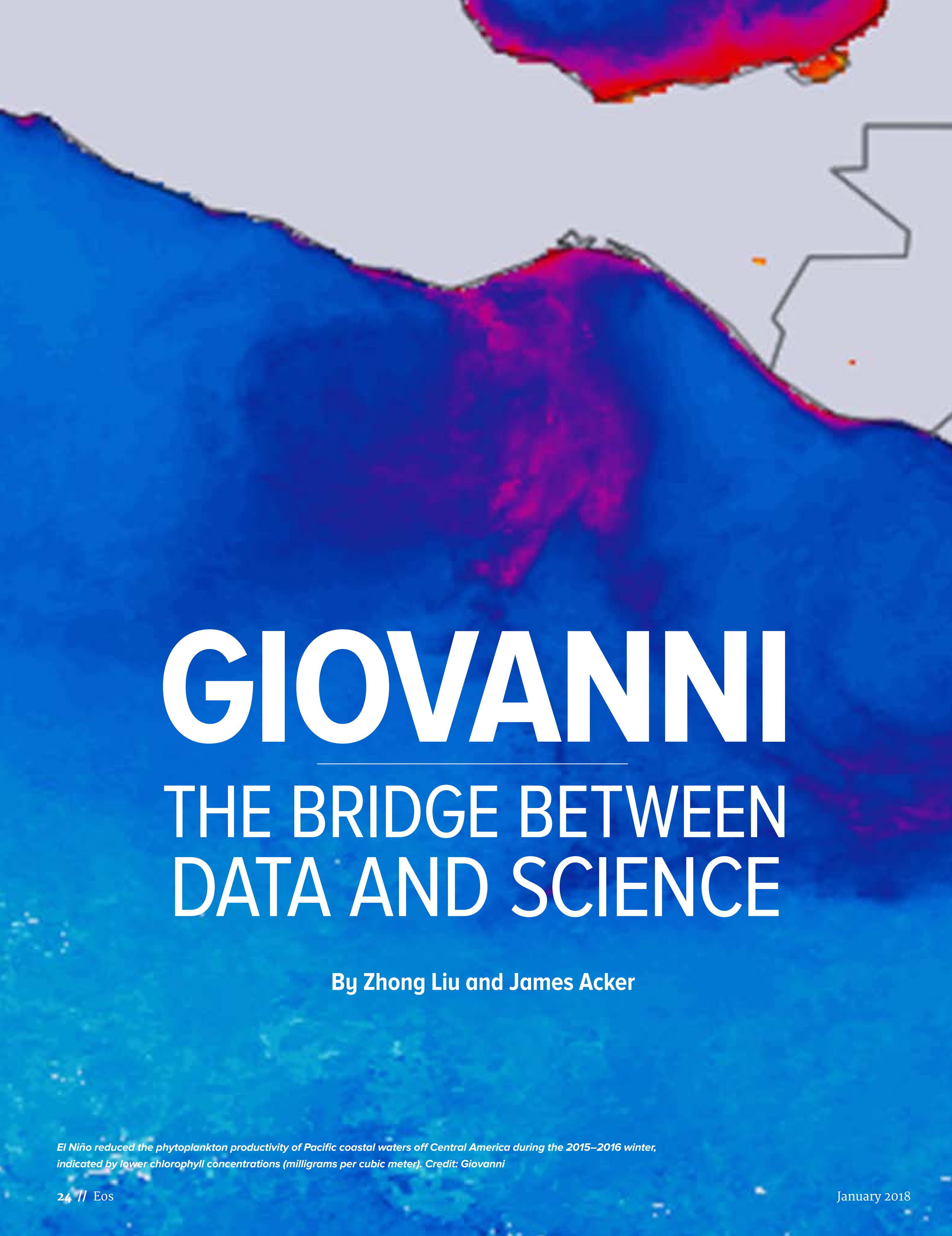
Geologic Map of Europa

How Will Climate Change Affect the United States?

Tracking River Flows from Space

BRIDGING BETWEEN DATA AND SCIENCE

AGU 100
ADVANCING EARTH AND SPACE SCIENCE

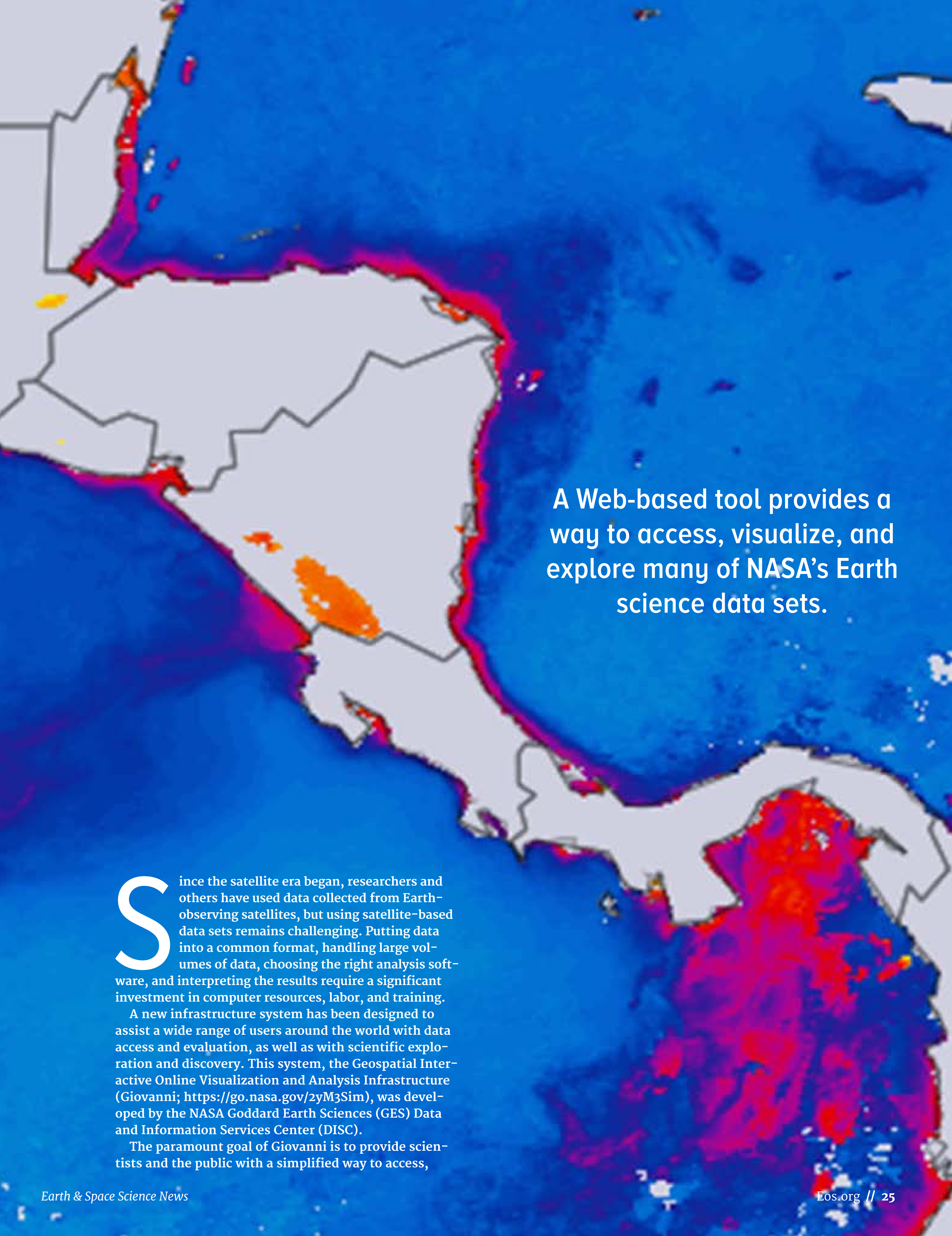


GIOVANNI

THE BRIDGE BETWEEN DATA AND SCIENCE

By Zhong Liu and James Acker

El Niño reduced the phytoplankton productivity of Pacific coastal waters off Central America during the 2015–2016 winter, indicated by lower chlorophyll concentrations (milligrams per cubic meter). Credit: Giovanni



A Web-based tool provides a way to access, visualize, and explore many of NASA's Earth science data sets.

Since the satellite era began, researchers and others have used data collected from Earth-observing satellites, but using satellite-based data sets remains challenging. Putting data into a common format, handling large volumes of data, choosing the right analysis software, and interpreting the results require a significant investment in computer resources, labor, and training.

A new infrastructure system has been designed to assist a wide range of users around the world with data access and evaluation, as well as with scientific exploration and discovery. This system, the Geospatial Interactive Online Visualization and Analysis Infrastructure (Giovanni; https://go.nasa.gov/2yM3Sim), was developed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC).

The paramount goal of Giovanni is to provide scientists and the public with a simplified way to access,

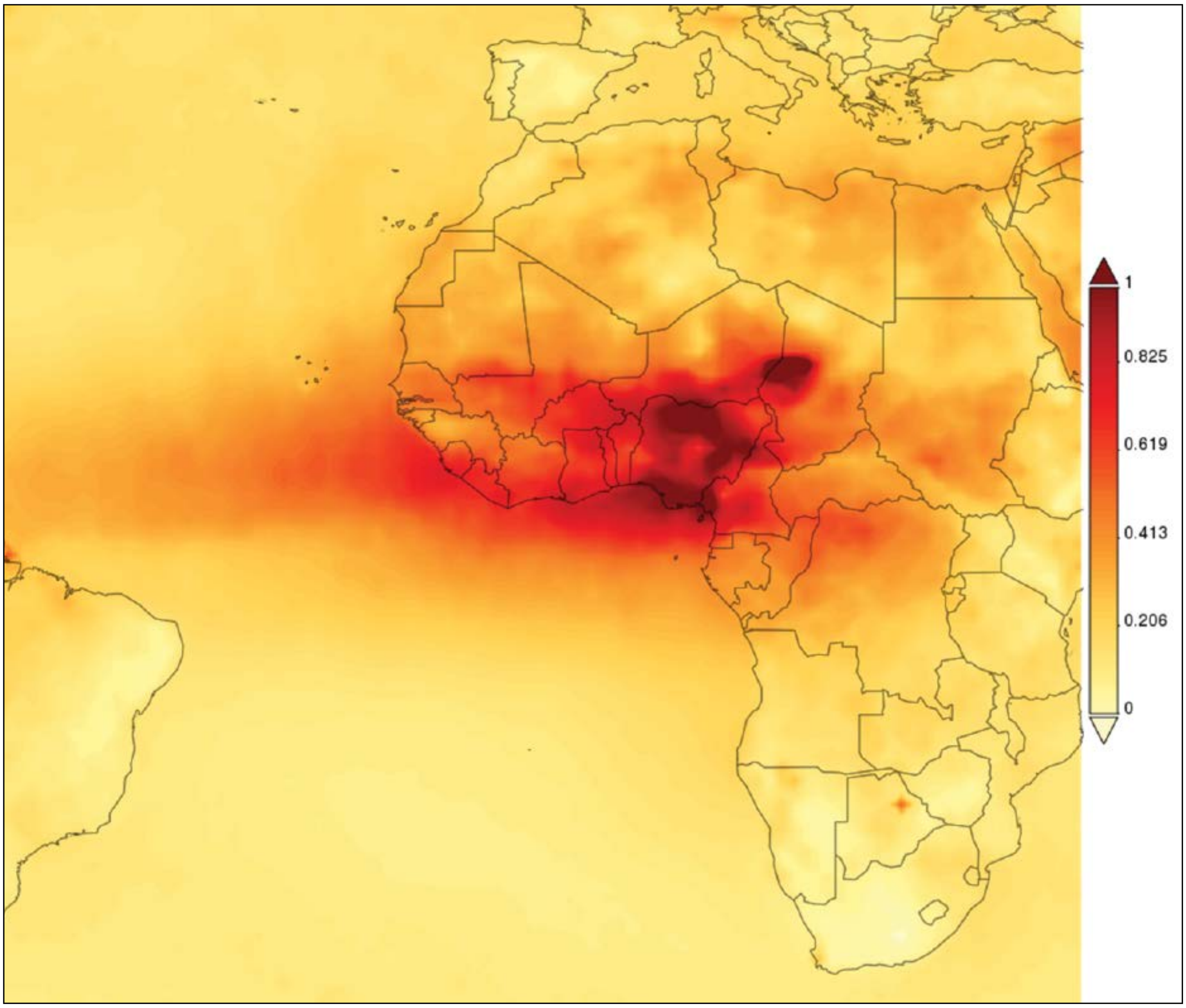


Fig. 1. This time-averaged satellite map of the March aerosol optical thickness off the coast of western Africa from 2003 to 2016 incorporates several of the new capabilities of NASA's Giovanni data visualization infrastructure. Credit: Giovanni

evaluate, and explore NASA satellite data sets. Here we describe the latest capabilities of Giovanni with examples (like the optical thickness map in Figure 1), and we discuss potential future plans for this innovative system.

Challenges of Using Satellite Data

Over Earth's vast oceans and remote continents, traditional large-scale, ground-based programs to observe the atmosphere, ocean, and land surface can be difficult and costly to deploy and maintain and are therefore impractical for providing adequate long-term observational data for research and applications. However, the need for large-scale observations is increasing as global observations become substantially more important for understanding global change processes like temperature and precipitation shifts.

Satellite instruments can overcome surface observation limitations by making repeated, synoptic observations of the Earth's land surface, ocean, and atmosphere. For example, NASA's Earth Observing System (EOS) is a global observation campaign consisting of a coordinated series of polar-orbiting satellites intended for long-term global observations, enabling improved understanding of Earth's geophysical systems.

However, many researchers find it challenging to access and use NASA data. Heterogeneous data formats, complex data structures, large-volume data storage, special programming requirements, diverse analytical software options, and other factors require a significant investment in time and resources, especially for novices.

By facilitating data access and evaluation, as well as promoting open access to create a level playing field for nonfunded scientists, NASA data can be more readily used for scientific discovery and societal benefits. Giovanni was developed to advance this goal. With Giovanni's assistance, researchers around the world have published more than 1,300 peer-reviewed papers in a wide range of Earth science disciplines and other areas.

A Brief History of Giovanni

Giovanni was initiated and developed for faster and easier access to and evaluation of data sets at GES DISC [*Liu*

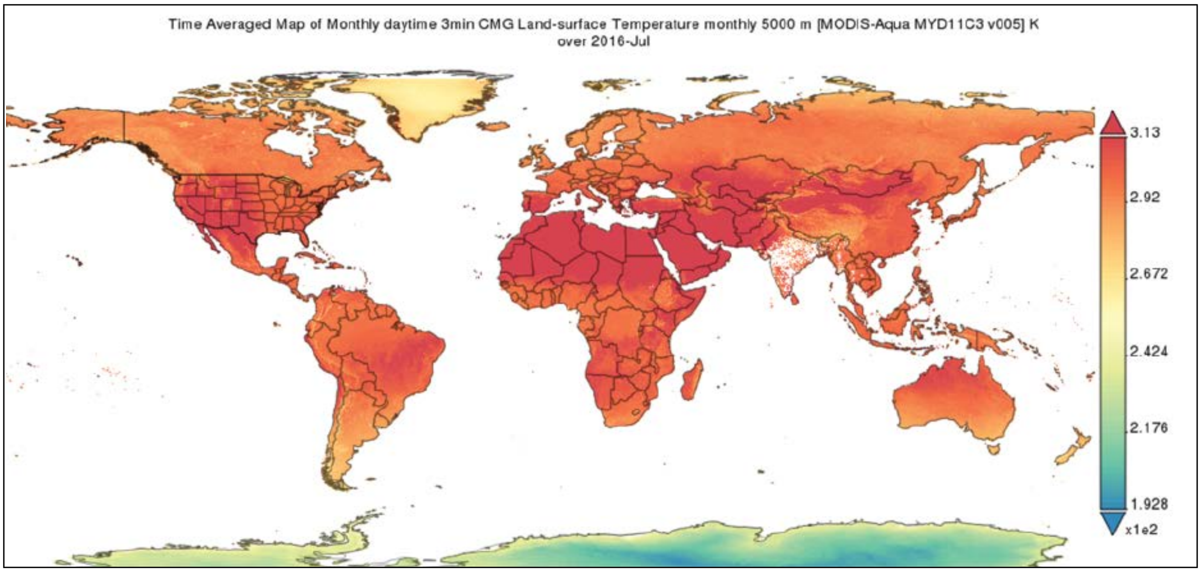


Fig. 2. July 2016, the hottest month on record for the globe. Shown are Moderate Resolution Imaging Spectroradiometer (MODIS) day surface temperatures (in kelvins). Credit: Giovanni

et al., 2007; *Acker and Leptoukh,* 2007; *Berrick et al.,* 2009]. The first implementation of Giovanni was an online visualization and analysis system for tropical rainfall data sets from NASA's Tropical Rainfall Measuring Mission (TRMM).

As the project gained popularity, scientists requested that more satellite data sets be included in Giovanni. To address this demand, we created multiple discipline- or mission-based data portals. The current Giovanni has evolved further, featuring a new unified Web interface to support interdisciplinary Earth system research, allowing synergistic use of data sets from different satellite missions.

A Wide Selection of Data Sets

Giovanni provides access to numerous satellite data sets, concentrated primarily in the areas of atmospheric composition, atmospheric dynamics, global precipitation, hydrology, and solar irradiance.

More than 1,600 variables are currently available in Giovanni. The Web interface has keyword and faceted search capabilities for locating variables of interest. For

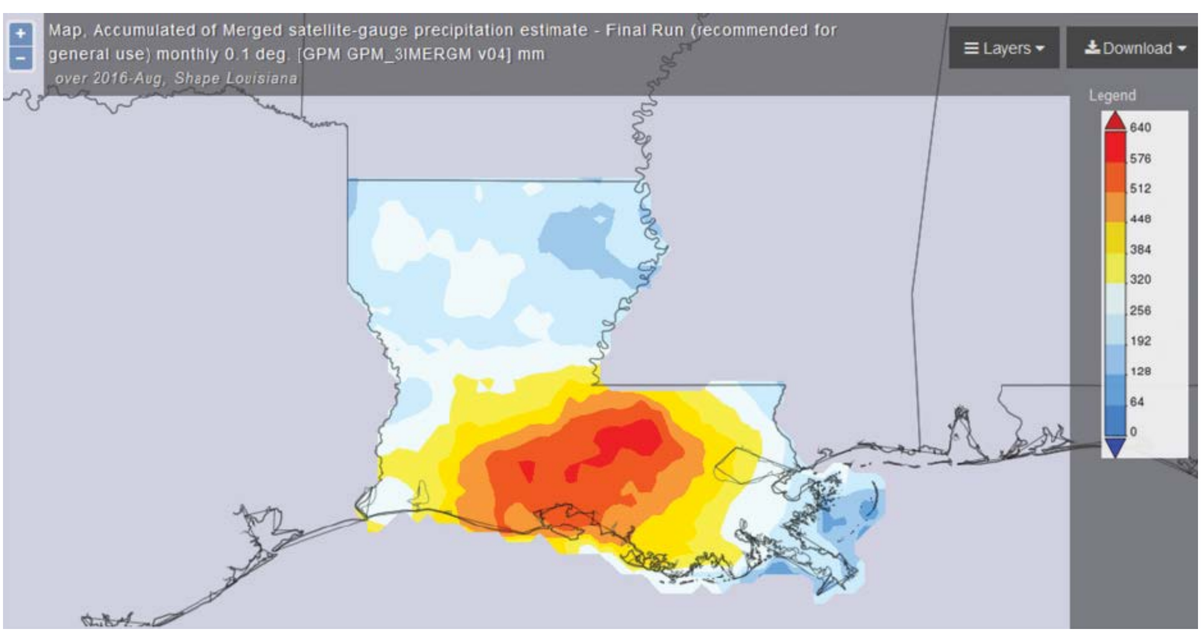


Fig. 3. Accumulated rainfall (millimeters) from GPM Integrated Multisatellite Retrievals (IMERG) Final Run (version 4), showing a record-breaking flood event in Louisiana in August 2016. Credit: Giovanni

example, a search for "precipitation" returns more than 100 related variables. A user performing a faceted search can filter for variables based on satellite missions (TRMM, Global Precipitation Measurement (GPM)), instruments, spatial or temporal resolution, or other categories.

The operating lifetimes of low-Earth-orbiting satellites are often quite limited (on the order of 5 years), far less than the 30 years recommended by the World Meteorological Organization for developing climatology data sets. Some users, however, may still wish to conduct preliminary studies with these satellite data sets to obtain information on spatial distribution and interseasonal variation. Giovanni provides the capability to derive climatologica l maps and time series based on user-defined time periods.

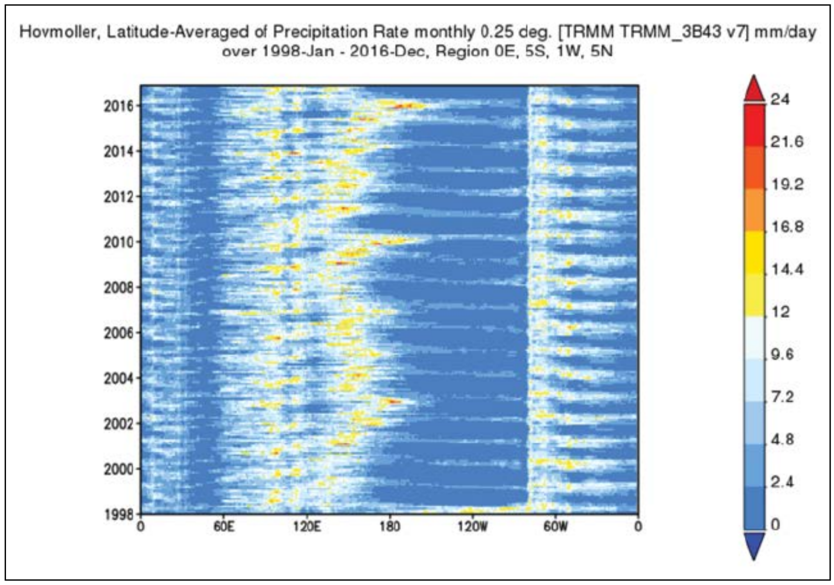


Fig. 4. Hovmöller diagram of TMPA monthly precipitation (millimeters per day) in the tropical region (5°S–5°N) showing El Niño–Southern Oscillation events between 1998 and 2016. Credit: Giovanni

Analytical Features

Giovanni includes many commonly used analytical and plotting capabilities for capturing spatial and temporal characteristics of data sets. Mapping options include time averaging (Figure 2), animation, precipitation accumulation (Figure 3), time-averaged overlay of two data sets, and user-defined climatology. For time series, options include area averaged, differences, seasonal, and Hovmöller diagrams (Figure 4).

Cross sections, applicable to 3-D data sets from NASA's Atmospheric Infrared Sounder (AIRS) instrument and Modern-Era Retrospective Analysis for Research and Applications (MERRA) data analysis program, include latitude-pressure, longitude-pressure, time-pressure (Figure 5), and vertical profile.

For data comparison, Giovanni has built-in processing code for data sets that require measurement unit conversion and regridding. Commonly used comparison functions include map and time series differences, as well as correlation maps and *X*-*Y* scatterplots (area averaged or time averaged). Zonal means and histogram distributions can also be plotted.

Visualization Features

Visualization features include interactive map area adjustment, animation, interactive scatterplots, data

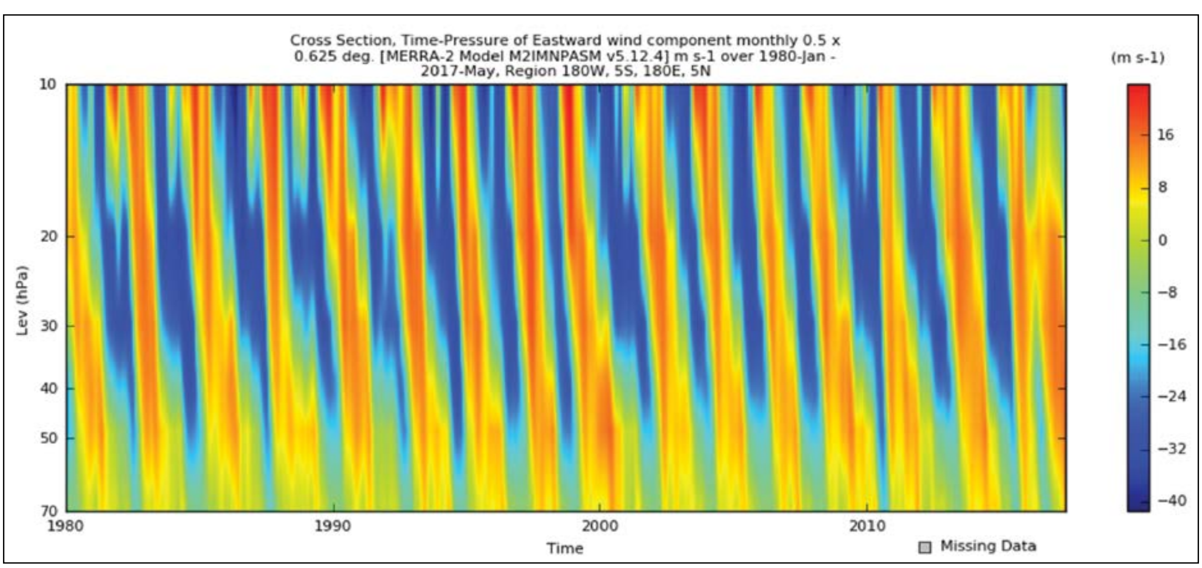


Fig. 5. Quasi-biennial oscillation (QBO) seen from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2), between 1980 and 2017. Credit: Giovanni

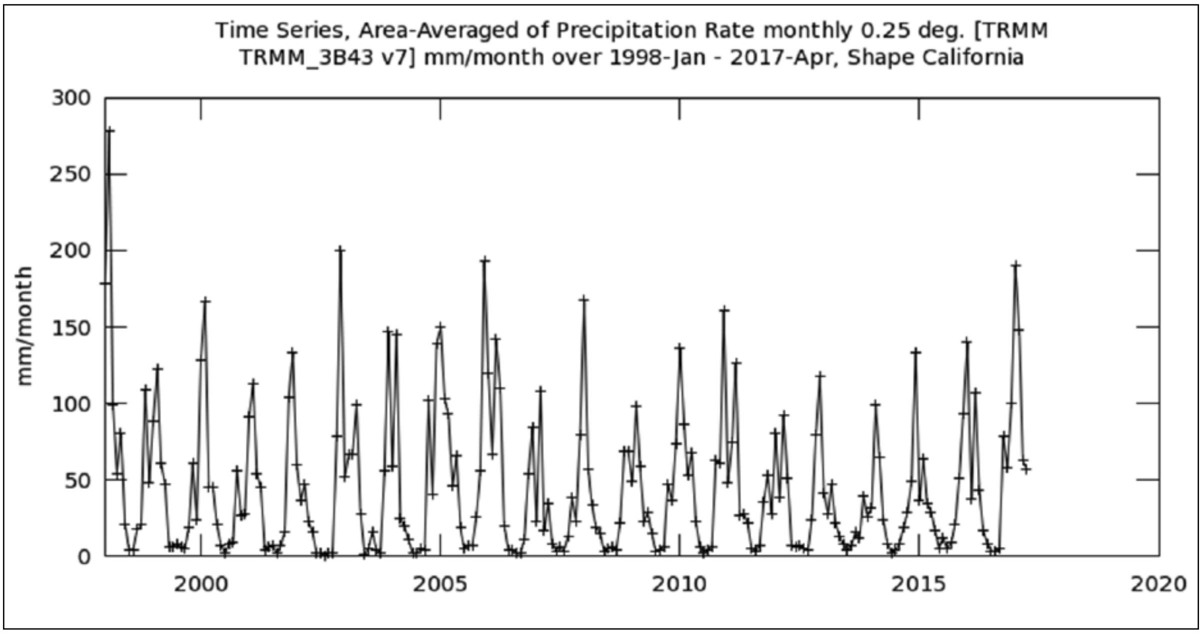


Fig. 6. Time series of area-averaged TMPA monthly precipitation (millimeters per month) for California, showing record-breaking droughts (2012–2015), followed by 2016–2017, the wettest winter ever recorded in northern California. Credit: Giovanni

range adjustment, choice of color palette, contouring, and scaling (linear or log). The on-the-fly area adjustment feature allows a user to examine a results map interactively and in detail without replotting data.

Giovanni also provides animations, which are helpful for tracking the evolution of an event or of seasonal changes. Interactive scatterplots allow identification and geolocation of a point of interest in a scatterplot. Adjustments of any of these plots provide customized options to users.

Formats Facilitate Many Applications

To support increasing socioeconomic and geographic information system (GIS) activities in Earth sciences, we have added shapefiles (a geospatial vector data format) for countries, states in the United States, and major watersheds around the world. Available functions for these shapefiles are time-averaged (Figure 2) and accumulated maps, area-averaged time series (Figure 6), and histograms. Land-sea masks have recently been added.

All data files involved in Giovanni processing are listed and available for download in the lineage page generated simultaneously with the visualization. Available output image formats are PNG, GeoTIFF, and Keyhole Markup Language (KMZ), and they can be used for different appli-

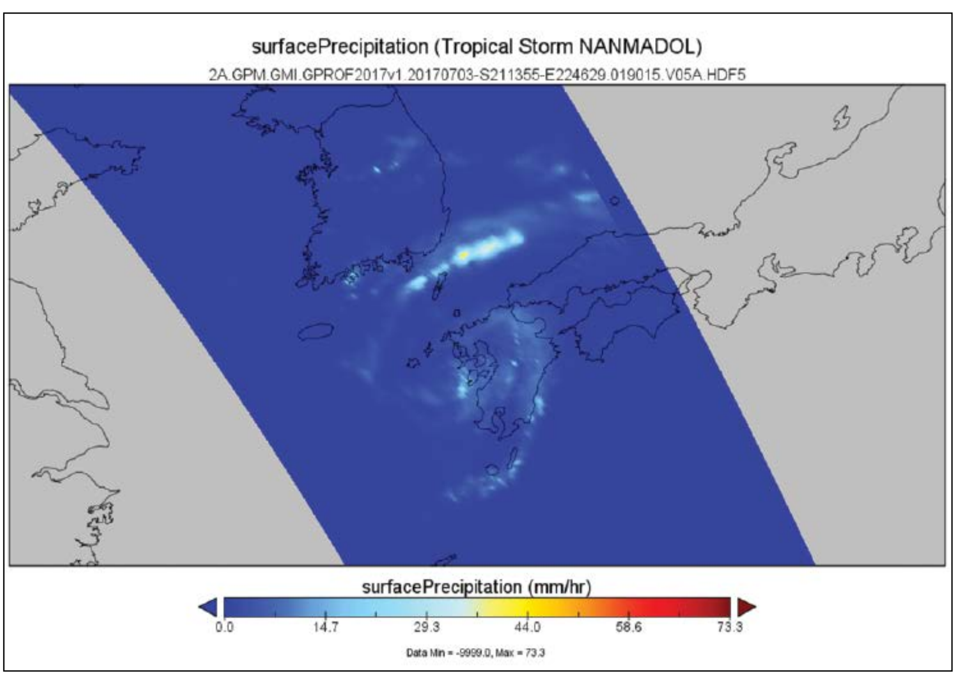


Fig. 7. A sample of satellite orbital data sets from GPM's Microwave Imager (GMI) showing surface precipitation of Tropical Storm Nanmadol on 3 July 2017. Credit: NASA Panoply

cations and software packages. For example, KMZ files are conveniently imported into Google Earth, where a rich collection of overlays is available.

All input and output data are available in the Network Common Data Form (NetCDF) formats, which can be handled by many off-the-shelf software packages. Furthermore, users can bookmark URLs generated by Giovanni processing for reference, documentation, or sharing with other colleagues.

Future Plans

With the latest features and applications, Giovanni simplifies accessing, evaluating, and exploring NASA satellite data sets. Despite these achievements, we still need to improve Giovanni to accommodate increasing demand for more analytical and plotting capabilities, more data sets, and advanced information technologies to make data exploration simple and productive.

Future plans include visualization and analysis of satellite orbital data sets (Figure 7), more data sets from other data centers, additional analytical methods and visualization, and analysis of multisatellite and multisensor measurements.

Data sets in Giovanni currently consist of variables mapped on uniform space-time grid scales, so nongridded or satellite orbital data sets remain largely untapped, even though they commonly provide higher spatial resolution. Adding orbital data sets to Giovanni could aid research requiring increased data resolution and coverage.

Data sets from other data centers and satellite missions will further enhance Giovanni for better understanding of Earth as an integrated system. Barriers still exist in the development of Giovanni for interdisciplinary studies and intercomparison among data sets. For example, terminologies in data sets can vary significantly between Earth science communities, requiring coordinated efforts to reach consensus and develop standards for uniform data products.

The NASA-wide User Registration System (URS) is also expected to enhance the Giovanni user experience. For example, with URS, users can set frequently used preferences in their profiles, record and retrieve their personal history of data set exploration, and establish their own data collections.

Data product developers can upload their test data and compare them with observations and other well-established data sets in Giovanni to identify issues in their products, a capability useful for improving data quality. Giovanni developers will also be able to better understand their users through profiles and other statistics collected from URS, so that they can develop more user-friendly services.

In summary, a wide variety of new features is available now in Giovanni, but it remains a work in progress. Creating a community tool with such a large scope is challenging, and fully realizing this tool requires active participation from the user community. We encourage users to provide their opinions as Giovanni continues to evolve.

Acknowledgments

We recognize the team effort of all past and current members at GES DISC for their contributions to the development of Giovanni. We extend our thanks to data set algorithm developers and many users for their feedback and suggestions. GES DISC is funded by NASA's Science Mission Directorate.

References

Acker, J. G., and G. Leptoukh (2007), Online analysis enhances use of NASA Earth science data, *Eos Trans. AGU, 88*(2), 14–17, https://doi.org/10.1029/2007EO020003.

Berrick, S. W., et al. (2009), Giovanni: A Web service workflow-based data visualization and analysis system, *IEEE Trans. Geosci. Remote Sens., 47*(1), 106–113, https://doi.org/10.1109/TGRS.2008.2003183.

Liu, Z., et al. (2007), Online visualization and analysis: A new avenue to use satellite data for weather, climate, and interdisciplinary research and applications, in *Measuring Precipitation from Space: EURAINSAT and the Future, Adv. Global Change Res. Ser.,* vol. 28, edited by V. Levizzani et al., pp. 549–558, Springer, New York, https://doi.org/10.1007/978-1-4020-5835-6.

Author Information

Zhong Liu (email: zhong.liu-1@nasa.gov), NASA Goddard Earth Sciences Data and Information Services Center, Greenbelt, Md.; also at George Mason University Center for Spatial Information Science and Systems, Fairfax, Va.; and **James Acker,** NASA Goddard Earth Sciences Data and Information Services Center, Greenbelt, Md.; also at Adnet Systems, Inc., Bethesda, Md.