# Web Statistics

**ESRL Physical Sciences Division**
**Science Review**
**May 12-14, 2015**

# PSD Web and FTP Traffic

PSD maintains an active Web presence, and provides dataset access via FTP, THREDDS and openDAP.

Services provided include:

- **Data Access:** Access to datasets including reanalyses, reforecast, Arctic observatories, AR observations, cruise, and satellite.
- **Data Analysis:** Online tools allow plotting, analysis and comparison of available datasets.
- **Forecast Products:** Access to experimental forecast products, including SST, winds, tropical atmosphere, hurricanes, decadal temperature.
- **Pre-Generated Climate/Weather Products:** Including our Maproom, Atmospheric River animations, BAO Tower Observatory.
- **Field Project Sites:** Access to data from a variety of field projects, including CalWater, WFIP, HMT Testbed, Texas Air Quality Study 2006.
- **Research Products:** Specialized products such as MJO indices, Climate Assessment Reports.

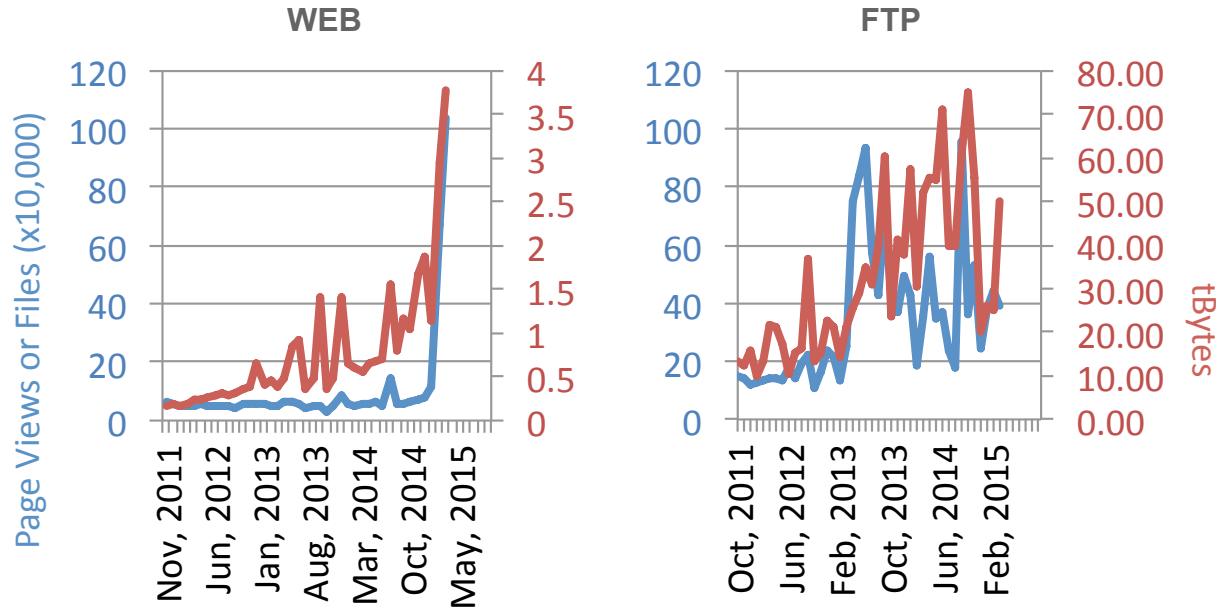
## WEB VISITORS – COUNTRY OF ORIGIN

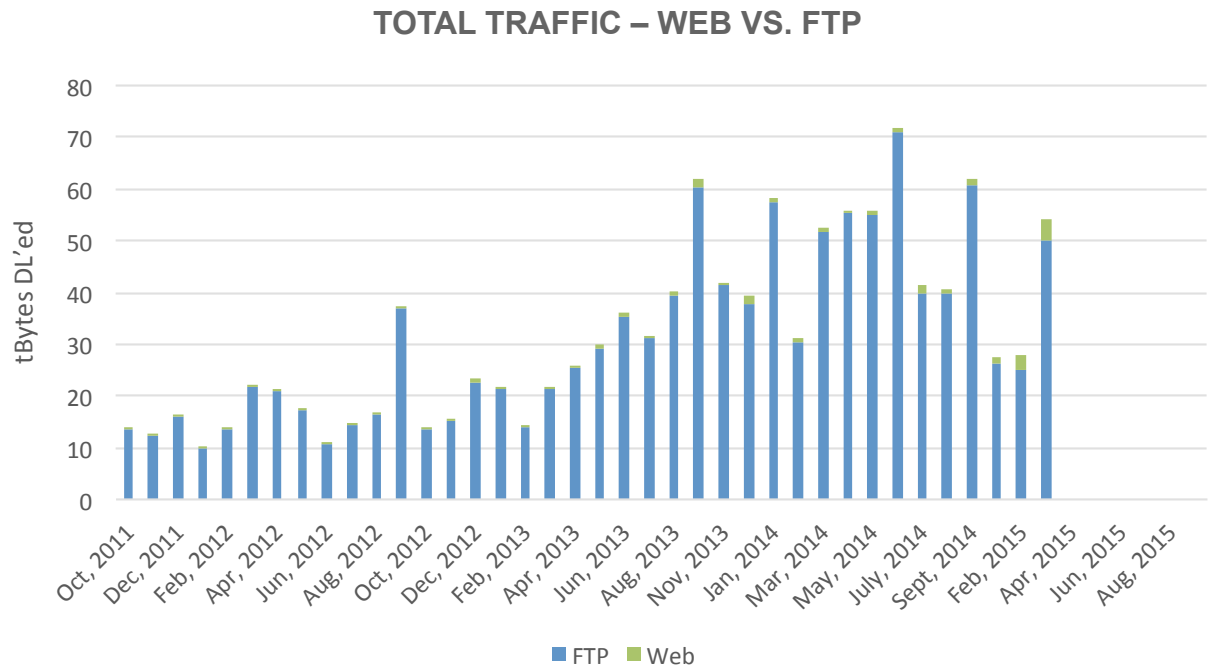
**PSD Web visits – FY 2012-15**

PSD's web presence is global, with page views from 225 countries, virtually every major country. Significant audiences include Australia, Canada, China, France, Germany, India, Japan, Mexico, Spain, the UK, and the USA.

# Total Traffic

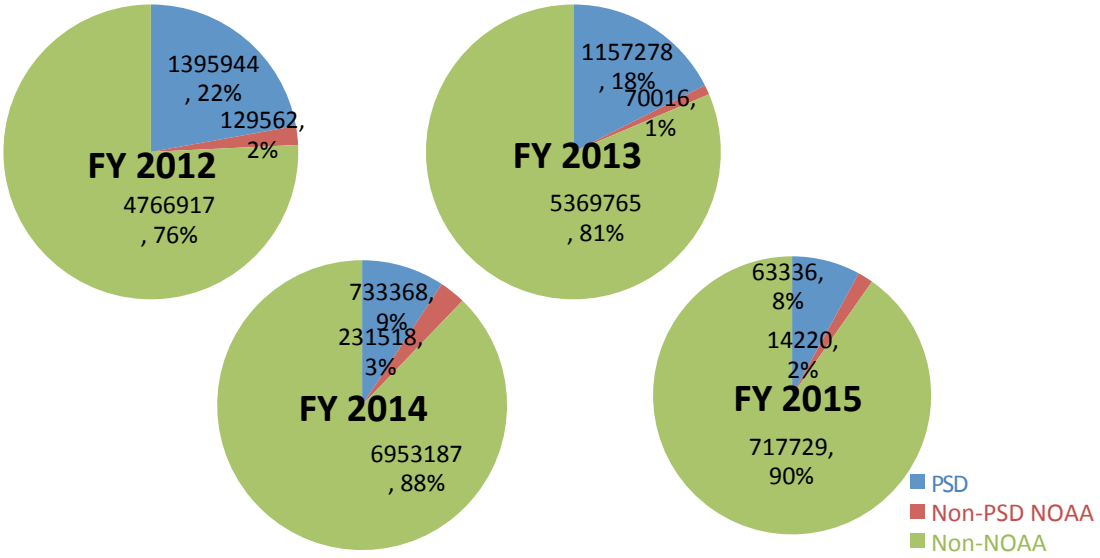
- PSD's Web & FTP Traffic has continued to increase over the past 4 years
- Close to 8 million web page views, nearly 50 million file downloads for FY14.
- 4.3 million plots produced over the last year.



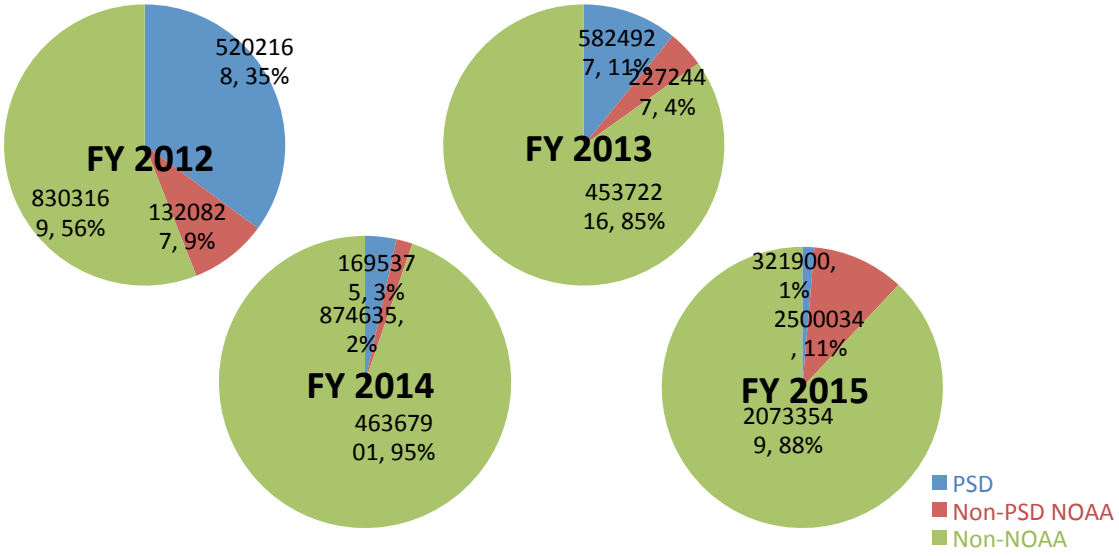

FTP continues to constitute the bulk of the data traffic, although web-based protocols are increasing in importance.

## WEB VISITOR BREAKDOWN BY ORGANIZATION (PAGE VIEWS)



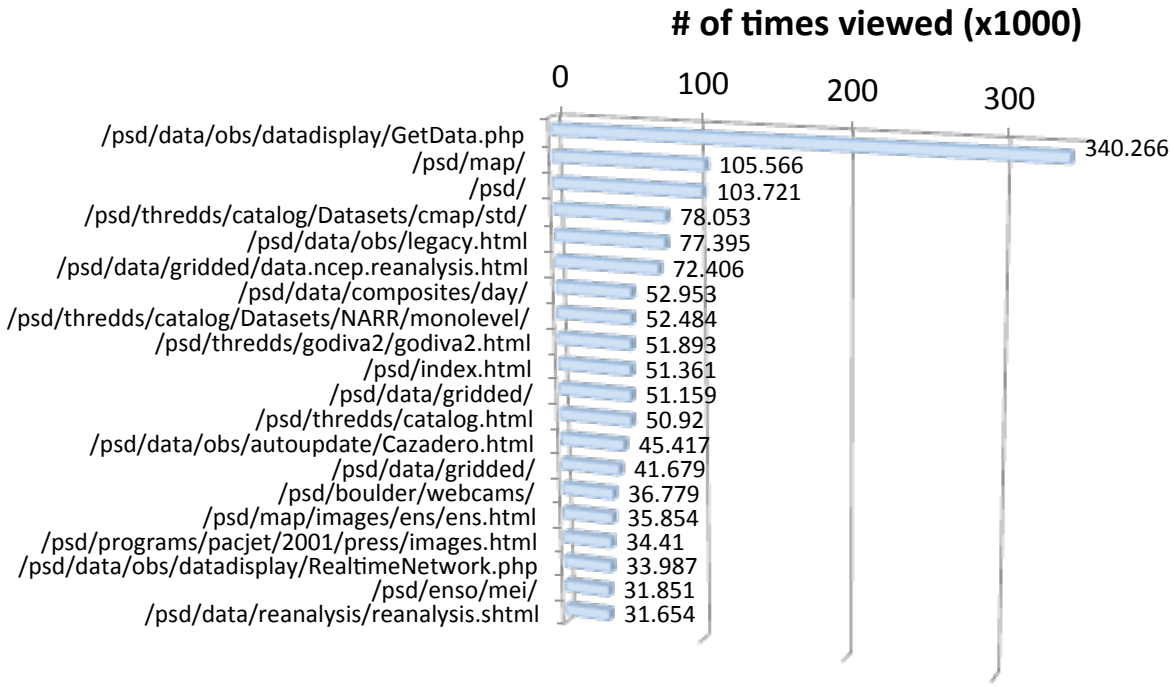

The majority of PSD's web visitors are from outside of NOAA. That fraction has been increasing of the past 4 years.

## FTP VISITOR BREAKDOWN BY ORGANIZATION (FILE DOWNLOADS)

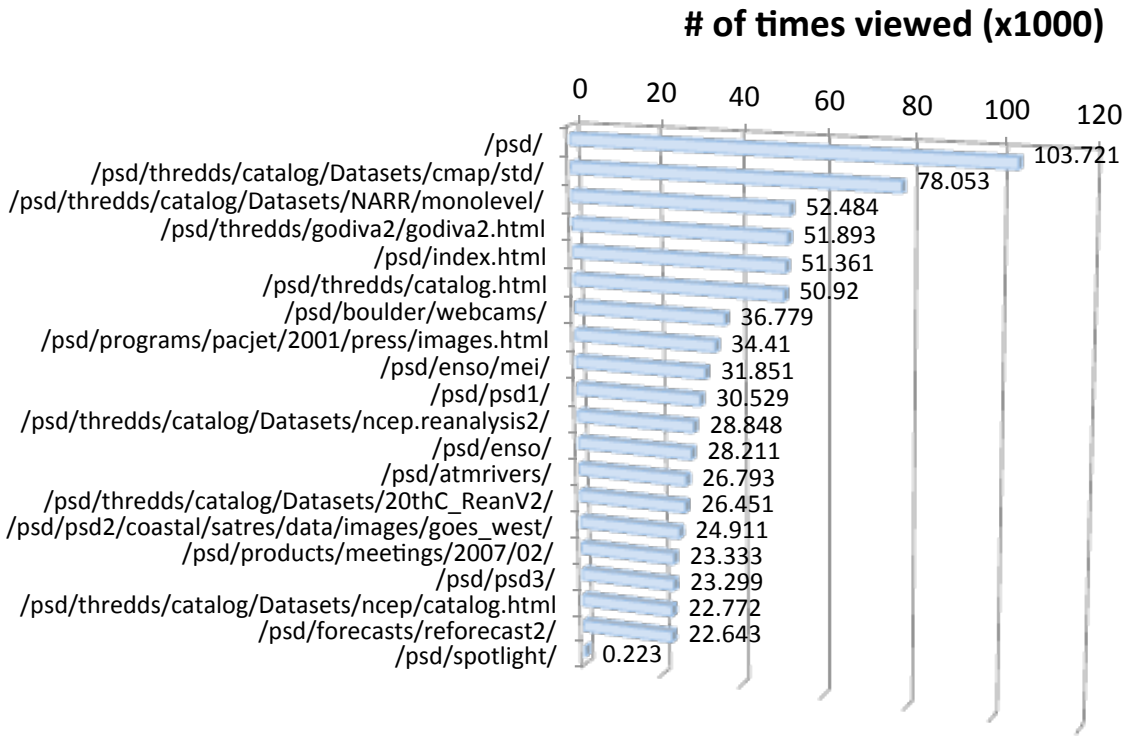


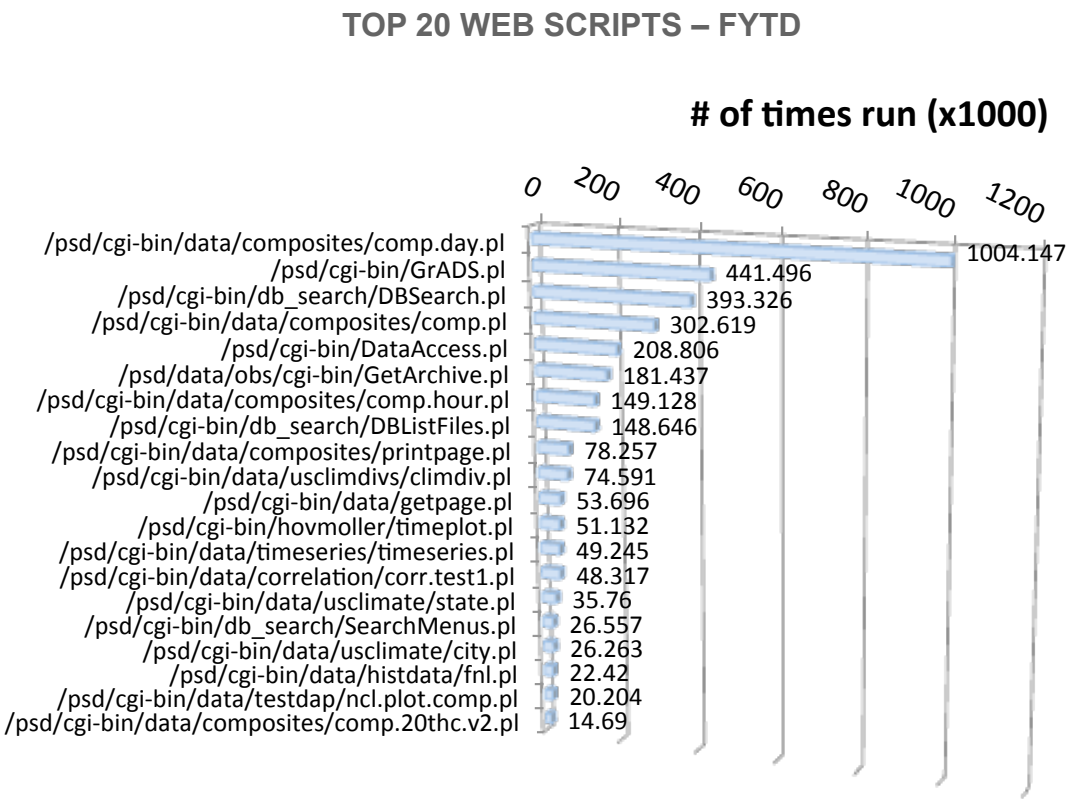

The majority of PSD's FTP downloads are also from outside of NOAA.

## TOP 20 WEB PAGES – FYTD




## TOP 20 NON-DATA WEB PAGES – FYTD






# Selected Kudos

*I just wanted to provide a quick note on the usefulness of the Daily Mean Composites (…) page that you provide through your web site. What an awesome tool! (…) "Job well done!" (Mike Huston, NWS Pocatello, 2011)*

*I use this page to make monthly, yearly and water year composites for presentations to a number of water resource groups. I also use this for monthly internal reports. (Aldis Strautins. Service Hydrologist NWS GJ, 2015)*

*I use your pages in support of classroom instruction and for student projects (…) Bottom line: Great resource! (Lance Bosart, Professor, U of Albany, 2012)*

*No other website widely available is as good as the one you maintain there at NOAA. To lose it would be a big blow to energy meteorologists around the world (Jess Torpey Senior Meteorologist, E.ON Global Commodities Düsseldorf, Germany, 2014)*

*This online tool is fantastic! I am a PhD student at the University of Alaska Fairbanks, and this resource is an amazing way to show synoptic overviews for the case study days in my research without laborious data manipulations. (Derek Starkenburg, University of Alaska Fairbanks student, 2014)*

*I have used ESRL PSD online plotting and analysis tool extensively. This tool is of great help and highly time saving one. The way you have created and designed it is simply superb. It gives access to huge*

*and diverse amount of data without actually downloading it. (Chinmay Khadke Research Fellow India Meteorological Department Pune, India, 2013).*

*We used the Monthly/Seasonal Climate Composites page in my Biogeochemical Cycles class today (taught by Dr. Oliver Wingenter), to model the Pacific Decadal Oscillation. Being able to visualize T, P and wind speed variations associated with phase changes in PDO was a great help. ( Annie Riggins Kottlowski Graduate Fellow NM Institute of Mining and Technology 2014)*

# Glossary

**Web traffic** – Network traffic (viewing of html pages, file downloads, running scripts, etc) associated with the HTTP or HTTPS (Hypertext Transfer Protocol, or Hypertext Transfer Secure Protocol). Http(s) is the protocal of web browsers and typically was it used to view web pages. It can also be used to download files although for this purpose FTP is much more robust and efficient.

**FTP traffic –** Network traffic (file downloads) associated with the FTP (File Transfer Protocol) – a standard network protocol often used to transfer files from one host to another over the Internet.

**Web Page –** A web page consists of single URL a visitor might bookmark and all the associated elements (images, sub-parts, etc) that might be download by the visitor in the course of viewing the page. A web page often consists of multiple elements each of which must be retrieved by the visitor's browser.

**Non-Data Web Page –** a web page in the PSD hierarchy that does not have /psd/data/ or /psd/map in the URL.

**Scripts –** defined as a url that has *.pl in it. Typically, a program or other action that a visitor causes to happen on our server by accessing a web page, or clicking on a "submit" button on one of our pages. The action involved might generate an image, provide a specific subset of a data page, or accomplish a number of other things for that visitor on our web server.

**Page Views –** a page view consists of single visitor accessing a page on the PSD website and retrieving all the associated elements. Each element thus retrieved consists of a "hit" on the website.

**gBytes -** in reference to web traffic, the total amount of data downloaded by visitors to our site over a given period. This may include web pages and associated elements, data generated by running scripts, or downloading data files via the web browser directly.

**Hits –** any request to the web server from a visitors web browser. This might include accessing a page and each of its associated elements (each access is a hit), running a script, or accessing a specific data file. Viewing a single web page may involve tens, if not hundreds, of hits, depending upon the complexity of the page.

# Technical Notes & Caveats

- Information for countries comes from a lookup against the GeoIP database from MaxMind.
- Top Level domains are determined by doing nslookups on the IP addresses.
- A Page View is defined as a log entry that had a result of 200 or 304 and had *.*htm* in the URL or the URL ends in a "/".
- A Non Data Page is a Page View that does not have "/psd/data*" or "/psd/map*" in the URL.
- A Script is defined as having "*.pl" in the URL. Note: all our scripts are perl.
- A Hit is defined as any log entry with a status of 200 or 304.
- PSD is defined as domain matching "*psd*noaa.gov"
- NOAA is defined as domain matching "*noaa.gov"
- A Visitor is defined as a unique IP Address within the time period. This number is approximate because or proxies and firewalls. As multiple visitors may come from a single IP.
- A Page Visitor is a Visitor who viewed a Page.
- Robot hits are not counted in the statistics.

## NOTES ON URL COUNTS:

- /boulder/index.html and /boulder are the identical page. Same for /psd/ and /psd/index.html.
- All the /obs/img/ are navigation images and while they should be counted overall, they probably shouldn't be listed in top 20/other hits