Three-Dimensional Air Quality System (3D-AQS)
Initial Benchmark Report

by

Erica Zell and Jill Engel-Cox

Battelle

2101 Wilson Boulevard, Suite 800

Arlington, VA 22201-3008

Subaward Agreement 0000003276
Prime Award NNS06AA02A

To the University of Maryland, Baltimore County
Joint Center for Earth Systems Technology

April 12, 2007

Three-Dimensional Air Quality System (3D-AQS)
Initial Benchmark Report

The Three-Dimensional Air Quality System (3D-AQS) project incorporates satellite, lidar, and other Earth science sensor data into decision support systems used by EPA, NOAA, and tribal, state, and local governments for air quality management, forecasting, and public health tracking. The resulting 3D-AQS will greatly enhance the understanding of the geospatial extent and transboundary transport of air pollutants, particularly fine particulate matter. The following five performance metrics will be used to benchmark progress of the 3D-AQS project. For each metric, a baseline measurement has been collected at the start of the project, and a final measurement will be collected at the conclusion of the project.

METRIC #1- End User Group Input

Description: Because this project focuses on enhancing decision support systems that are used by the air quality forecasting, regulatory, and public health communities, a key aspect of this project is the involvement of an End User Group to provide feedback on project development and interim project outputs in order to maximize the usefulness of the project results. The 3D-AQS End User Group consists of approximately 21 members and is representative of potential stakeholders including Federal, regional, state, and local air quality managers and forecasters and university researchers. The End User Group survey includes general questions on use of satellite air-quality data, temporal and spatial resolution, data format and delivery, and visualizations. A similar survey will be conducted at the end of the project, and a summary comparison of the initial and final surveys will be prepared.

Baseline Measurement: Sixteen of the 21 members provided written responses to the baseline survey. Fourteen out of the 16 respondents already use or work with satellite remote sensing data for air-quality purposes, although 6 of those 16 reported very limited use of satellite data. Eleven respondents use the data for historical analysis such as examining transport or comparing it against modeling simulations. Five of the respondents use data for air-quality forecasting on a regular basis. Respondents expressed a wide range of priorities in terms of interest in data products and imagery. Although trends were evident (e.g., many seem to prefer images), the main message is that different products are important to different users. Seven respondents listed AOD images as their number one priority (4 requested AOD alone, and 3 requested hybrid images of AOD and True Color). Another six respondents listed their interest in AOD images (either alone or hybrid) as one of several high priorities. Five respondents noted that True Color images interfere with event identification due to overlapping color schemes with AOD. Finally, all the respondents selected at least two options for preferred spatial delivery. The most popular was single AOD values associated with ground-based monitors, identified as of interest by 14 respondents. Complete survey results are listed in Appendix A.

METRIC #2- Online Questionnaire Linked to Smog Blog and IDEA Websites

Description: Similar to the End User Group Survey, a multiple-choice format questionnaire was conducted through an online web site. This questionnaire covered

general use of satellite data for air quality, usability and frequency of existing data and imagery sources, and preferred source enhancements such as format and spatial and temporal resolution. The survey was accessed through an advertised link on both the UMBC Smog Blog and IDEA websites. A similar survey will be conducted at the end of the project, and a summary comparison of the initial and final surveys will be prepared.

Baseline Measurement: The baseline survey ran for a period of about five weeks, and was completed by 15 respondents. Twelve out of the 15 respondents already use or work with satellite remote sensing data for air-quality purposes on a regular basis. The respondents were roughly split in primary interest between air-quality forecasting and air-quality modeling and historical event analysis. In terms of factors that limit the respondents use of satellite data for air-quality, seven respondents listed the need to know more about the relation to ground-based monitoring sites, and seven respondents listed the need to know more about vertical distribution of pollution. The other most popular response (five respondents) was a lack of time/resources/management support. The respondents gave a wide range of responses in terms of the desired option for receiving AOD data. Complete survey results are listed in Appendix B.

METRIC #3- Review and Comparison of Smog Blog and IDEA Website User Logs

Description: The user logs of the Smog Blog and IDEA websites will be tracked throughout the project with benchmarks at the start and end of the project. Results from 2006 will be used as a baseline year. The analysis includes the monthly total of following parameters:

Table 1. Web site Parameters Tracked.

Parameter	Description
Hits	Total http requests, including html files and images
Kilobytes	Data transferred from server to users
Visits	Series of requests for pages in set time period
Sites	Unique IP addresses making request to server, rough indicator of unique users
Pages	URLs .htm, .html, or .cgi (page views)
US Educational Users	.edu
US Government Users	.gov
Network Users	.net
US Commercial Users	.com
Unresolved/Unknown Users	Unknown
United States Users	.us

Baseline Measurement: As shown by Figure 1, the number of monthly visits to the Smog Blog between January 2006 and February 2007 has varied between about 25,000 and 38,000. As shown by Figure 2, the number of monthly visits to the IDEA web site between January 2006 and February 2007 has varied between about 2,900 and 5,000. Complete weblog results including the parameters listed in the table above are tabulated in Appendix C.

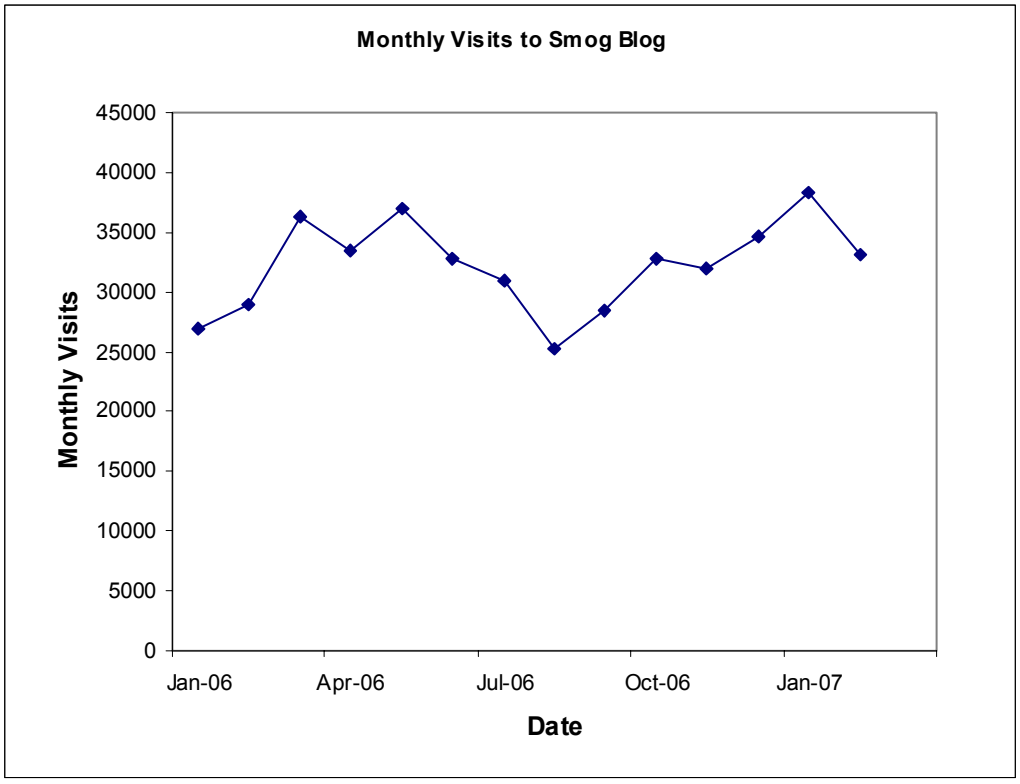


Figure 1. Monthly Visits to Smog Blog.

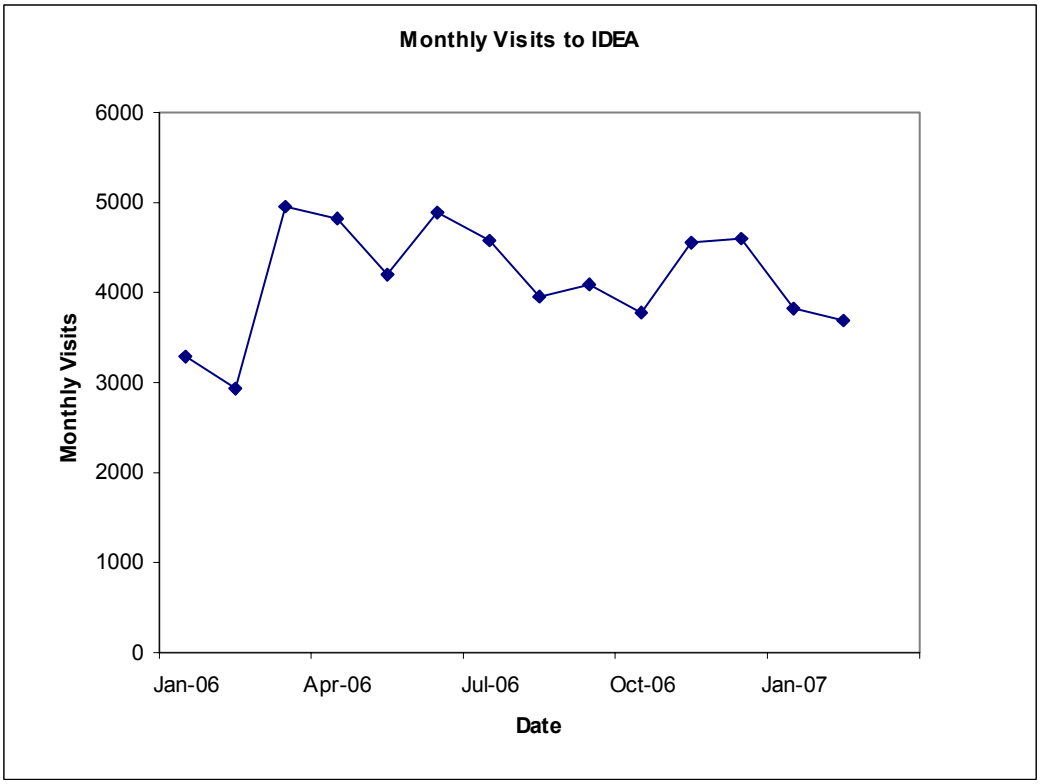


Figure 2. Monthly Visits to IDEA Website.

METRIC #4- Data Accessibility

Description: The major activities of the 3D-AQS project will improve data availability. The key activities that will serve as metrics are:

- Transition of IDEA to a NOAA operational environment
- Integration of NASA and UMBC sensor data into EPA's AirQuest system, specifically the number of discreet datasets integrated and the volume of data downloaded by users.

Baseline Measurement: To demonstrate NOAA commitment to achieving transition of IDEA to a NOAA operational environment, Dr. Alfred Powell, Director of the Center for Satellite Applications and Research (STAR) of NOAA NESDIS has sent a letter to Lawrence Friedl and Teresa Fryberger of NASA Applications Directorate stating his support for the transition. EPA has demonstrated commitment to integration of NASA sensor data into the AirQuest system by participating actively in the 3D-AQS project and 3D-AQS End User Group. Specifically, End User Group member Phil Dickerson and others at EPA are responsible for maintaining IT systems such as AirQuest and AirNow, and stated in the End User Group Survey that they are interested in integrating ground-based observations with satellite data. The intent of this metric is to ensure long-term access to three dimensional air quality data through NOAA and EPA decision-support systems which are already in place.

METRIC #5- Improved Air Quality Forecast Accuracy

Description: This quantitative metric measures the increased value that the satellite data and accessibility provided by this project brings to end users by tracking the accuracy of an experimental pilot-scale air-quality forecast for the eastern US generated by the New York State Department of Environmental Conservation (NYSDEC). Dr. Gopal Sistla, Director of the NYSDEC Bureau of Air Quality Analysis and Research, is participating in the 3D-AQS End User Group. The modeling group at NYSDEC generates air-quality forecasts based on the Community Multiscale Air Quality (CMAQ) model, and does not currently incorporate satellite air-quality data. Due to their interest in incorporating satellite air-quality data that may help improve the accuracy of the air quality forecast, especially for $PM_{2.5}$, Dr. Sistla has agreed to provide the 3D-AQS team with statistics on forecast accuracy throughout the life of the project. The metric will compare the pre- and post-project ability of the CMAQ-based forecast to predict the onset, severity, and duration of an air quality event, utilizing a measure such as accuracy or bias that is already maintained by NYSDEC.

Baseline Measurement: A paper accepted for publication in the Journal of Applied Meteorology[1] documents the results of CMAQ-based air quality forecasts generated by researchers at NYSDEC. The paper compares the CMAQ-based air-quality forecasts to actual observed values and to forecasts generated by more traditional methods of combining weather predictions, statistical analysis, and expert judgment. For the 3D-

[1] C. Hogrefe, W. Hao, K. Civerolo, J.-Y. Ku, G. Sistla, R.S. Gaza, L. Sedefian, K. Schere, A. Gilliland, and R. Mathur, Daily Simulation of Ozone and Fine Particulates Over New York State: Findings and Challenges, Accepted for publication in the Journal of Applied Meteorology

AQS project, we are most interested in tracking any improvement in NYSDEC CMAQ-based air-quality forecasts in comparison to actual observed values, based on incorporation of satellite air-quality data. The time periods studied in the paper are July through September 2004, January through March 2005, and June to October 2005. The paper includes correlation coefficients between observations and model predictions at each monitoring station. For 24-hr average $PM_{2.5}$ concentrations, correlations (r) are greater than 0.65 for all but two locations in upstate New York and range from 0.45 to 0.75 in the New York City metropolitan area. As the 3-D AQS project proceeds, similar statistics will be calculated and analyzed to assess improvements or trends.

The results of the CMAQ-based forecasts were also presented at the 2007 National Air Quality Conference.[2] This presentation noted that there is a strong correlation between 24-hour average $PM_{2.5}$ values predicted by CMAQ and AIRNow observations, but that there were a number of missed alarms. As shown in the Figure 3, there is a tendency for CMAQ to overestimate low and underestimate high observed $PM_{2.5}$ concentrations.

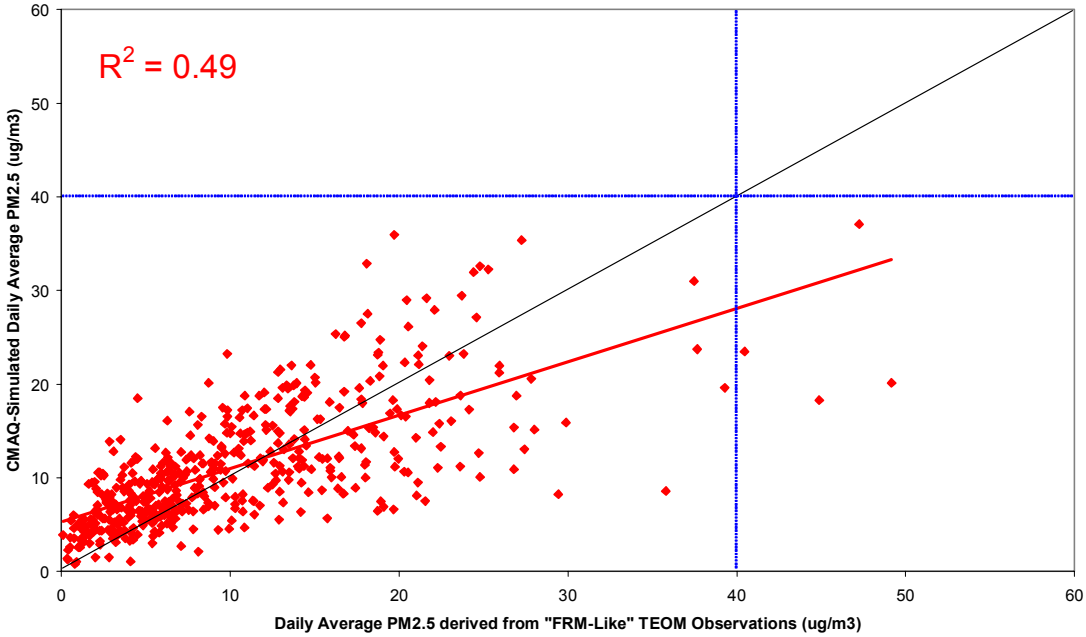


Figure 3. 24-Hour Average $PM_{2.5}$ June 2005-October 2006, Upper Hudson Valley; CMAQ vs. AIRNow Observations.

As shown in the Figure 4, for the New York City metro area, there is also a strong correlation but a large number of false alarms. CMAQ systematically overestimates observed $PM_{2.5}$ concentrations in this region, probably due to a combination of errors in emission inventory and temporal profiles, and shallow predicted boundary layer depths.[1,2] On some days, overprediction of "region concentration" may be driven by a few grid cells.

[2] C. Hogrefe, W. Hao, P. Doraiswamy, K. Civerolo, J.-Y. Ku, and G. Sistla, Two Years of Near-Real-time Simulations of Ozone and $PM_{2.5}$ Over New York State: Experiences and Future Directions, National Air Quality Conference, February 13, 2007, Orlando, Florida.

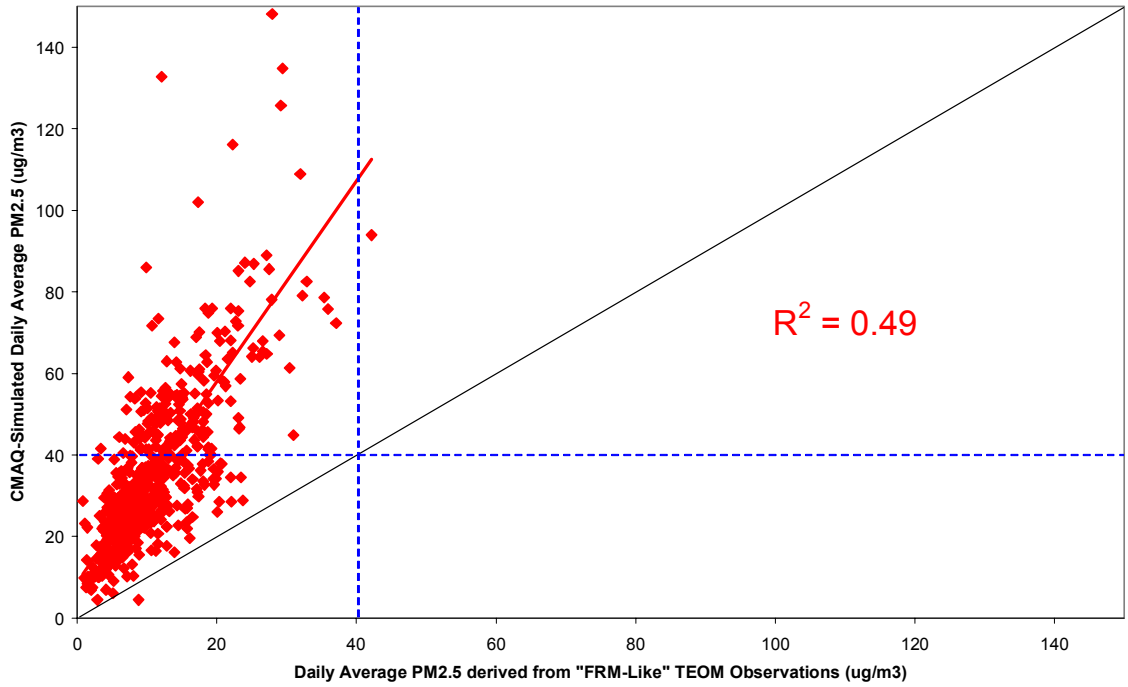


Figure 4. 24-Hour Average $PM_{2.5}$ June 2005-October 2006, NYC Metro; CMAQ vs. AIRNow Observations.

At present, evaluation of forecasts is conducted on a batch by batch basis. The metrics analyzed include bias, root mean square error and correlation coefficients. In addition, comparisons against measured air quality index values are conducted, and categorical evaluation metrics such as false alarm ratio, probability of detection and critical success index are evaluated.[1]

APPENDIX A: End User Group Survey Results (Metric #1)

Three-Dimensional Air Quality System (3D-AQS)
End User Group Kickoff Call - Data Requirements Survey Response Summary

The following summarizes the 13 End User Group kickoff call survey responses (representing 16 End User Group members).

General Questions

How and to what extent do you currently use satellite remote sensing data for understanding air pollution?

Fourteen out of 16 of the respondents already use or work with satellite remote sensing data for air-quality purposes, although 6 of those 16 reported very limited use of satellite data. Eleven respondents use the data for historical analysis such as examining transport or comparing it against modeling simulations. Five of the respondents use data for air-quality forecasting on a regular basis. (Note that some respondents use it for both historical analysis and forecasting, as reflected in the above numbers). One respondent (Phil Dickerson, EPA) notes that his group is responsible for maintaining IT systems such as AIRNow and AIRQuest and is interested in integrating ground-based observations with satellite data.

What satellite-based air quality information (images and/or data) do you use routinely and where do you currently obtain the information?

Three respondents listed only one source for satellite-based air quality information (one said IDEA, one said MODIS data by working with NASA and academic scientists and one said NASA's web page). The other respondents that use satellite data listed a variety of sources. The following tallies the number of times each source was mentioned by any respondent:

<u>For AOD:</u>
IDEA – 4
GASP – 5
NRL Monterey Aerosol - 1

<u>For True Color Images:</u>
MODIS processed images (source not specified) – 7
DataFed (CATT tool) – 2

<u>For Aura/OMI:</u>
NASA Giovanni website – 1

<u>For Fire & Smoke Products:</u>
Hazard Mapping System (HMS) – 2
NESDIS fire page- 1
AVHRR (for fire M3 Hotspots and Fire & Smoke) – 1
ATSR (for fire M3 Hotspots) – 1
DMSP/OLS (for Fire & Smoke) – 1

For Combined Data Sources:
UMBC Smog Blog – 2
FX-Net - 1
Google Earth - 1

Rank your level of interest in the following:

satellite-based true-color images
aerosol optical depth (AOD) images
hybrid/layered images of true-color and AOD
AOD data
fire count or other fire related data
interpreted products (such as fire/smoke detection)
other (please list)

Respondents expressed a wide range of priorities. Although trends were evident (e.g., many seem to prefer images), the main message is that different products are important to different users. Seven respondents listed AOD images as their number one priority (4 requested AOD alone, and 3 requested hybrid images of AOD and TrueColor). Another six respondents listed their interest in AOD images (either alone or hybrid) as one of several high priorities. Five respondents noted that True Color images interfere with event identification due to overlapping color schemes with AOD. The complete tabulation of responses is shown below. (Some respondents provided rankings of 1-6, which are reflected directly in the table below. Other respondents indicated items of high or medium priority, which were translated into a rank of 2 and 3, respectively. Items that were not ranked by a respondent were translated into a rank of 6. An average across all participants for each item is shown in the far right column, and the items are ordered from most to least important based on this average.)

Respondent #	1	2	3	4	5	6	7	8	9	10	11	12	Average
aerosol optical depth (AOD) images	2	1	2	2	2	4	1	2	1	1	2	2	**1.8**
hybrid/layered images of true-color and AOD	1	2	2	3	2	6	3	1	6	2	1	6	**2.9**
AOD data	2	3	2	1	6	6	4	4	2	1	3	2	**3.0**
satellite-based true-color images	4	4	2	4	2	3	2	3	6	6	4	6	**3.8**
fire count or other fire related data	5	6	3	5	2	2	5	6	6	4	5	2	**4.3**
interpreted products (such as fire/smoke detection)	6	5	3	6	5	6	6	5	3	5	6	2	**4.8**
other (please list)	3	6	6	6	6	1	6	6	6	3	6	2	**4.8**
Assume High = 2. medium = 3, not ranked = low = 6	tropospheric/ total column ozone & trace gases					aerosol profiles				data for PM chem. comp., gases such as CO, NO2, HCHO		spatial analysis (isopleths) of AOD (highlight ~0.56)	

Temporal Questions

As an end-user of air quality data, are you more interested in forecasting (requiring real-time data) or retrospective analysis (requiring historical data)? What are the interests of your organization as a whole?

Eight respondents expressed interest in both forecasting and retrospective analysis. Seven expressed interest only in retrospective analysis (including near real-time data) and one expressed interest solely in forecasting.

If interested in forecasting, how rapidly after real-time do you need the data/imagery? What precisely do you need (you can reference data format questions below)?

Six respondents answered this question. One noted to ask MARAMA forecasters, and another said that it depends on the deposition rates, reaction rates, and dynamics of the variable in question. Another said that they need at least hourly data between 12Z and 17Z (and the forecast is at 1830Z). Another said that 14:00 EST is the latest they would be useful for next day forecasts. Finally one respondent said that one to two hours after collection would be best, but definitely available by 15 to 17 UTC. One asked for AOD daily so that it can be used potentially to adjust model-based forecasts.

As the Terra and Aqua satellites pass from east to west, the data on the East Coast is available first. Are you interested in viewing the East Coast data/imagery as quickly as possible, or waiting for a full set of data for the entire continental U.S.?

Seven respondents requested East Coast data as it becomes available, six preferred to wait for all data or at least adjacent regional data, and three had no preference. Several respondents asked about the costs and resources necessary to view data early.

For retrospective analysis, how rapidly do you need data (e.g., 1 day, 1 week, monthly)?

Most respondents requested data on at least a monthly basis. One requested data the day after for exceptional events, and two said that data up to one year old would be useful. One also noted that quality assured data may take longer but raw data could be useful in the interim.

GOES data can be available in as small as 15 minute increments. Do you need data at that fine a temporal resolution? Or some other scale (e.g., 30 minutes, 1 hour)?

All but two respondents said that hourly data was fine. The other two respondents (from states, with a strong interest in forecasting) would like 15 minute increment data. Several respondents noted the need to align with ground-based increments which can be as small as one minute, and another requested daily or episode aggregations.

Spatial Questions

There are a variety of options for gridding (or not gridding) the data. MODIS AOD is available on roughly a 10 kilometer average, although due to the shifting path of the satellite overpass each day, the grid is not the same from day to day. Which of the following options are you interested in for receiving AOD data?

The number of respondents that expressed interest in each of the following options is shown by the corresponding number below.

- 10 km AOD matched with latitude and longitude - 8
- 12 km standard CMAQ grids (recalculated from 10km AOD) - 9
- County level or ZIP code level data - 7
- Selected single AOD values for points associated with ground-based monitors - 14
- Others? – 1 (raw data for self aggregation)

All the respondents selected at least two options for preferred spatial delivery. The most popular was single AOD values associated with ground-based monitors, identified as of interest by 14 respondents.

Data Format and Delivery

Do you currently use the IDEA site (http://idea.ssec.wisc.edu/) and/or the University of Maryland at Baltimore County Smog Blog (http://alg.umbc.edu/usaq/) to access satellite data? (Or another site?) Which of these sites should we focus on for improvement first?

Five respondents said they use the IDEA site and four respondents said they use the Smog Blog. Six respondents requested an initial focus on improving IDEA (4 of those 6 respondents requested an initial focus on improving AirQuest, followed by IDEA). One respondent requested an initial focus on improving the Smog Blog, specifically with respect to lidar explanations and the ability to select and customize lidar images.

Would you like to access satellite data through EPA's AirQuest (expanded relational database including monitoring/compliance data)? Or would you prefer accessing satellite data through a meteorological site such as the National Weather Service Advanced Weather Information Processing System (AWIPS) (http://www.weather.gov/geodata/)? Or another database?

Nine respondents suggested AirQuest (or possibly AQS or one of its derivatives such as Datamart) for accessing satellite data. Six respondents did not express a preference between AirQuest and AWIPS, but offered other related comments. Three respondents suggested an improvement to FX-net to reduce redundancy. Other respondents requested maximum flexibility and good integration with ground-based values. One respondent preferred AWIPS.

What data formats (e.g., ASCII, HDF, GeoTiff, jpg) do you need or prefer?

Most respondents focused on the need for flexibility and compatibility over one specific format. Three respondents said that data needs to be compatible with Microsoft Access or Excel, and five respondents specifically mentioned ASCII. Two respondents mentioned SAS, and four respondents mentioned the use of ArcMap GIS and the need for GeoTiffs or .shp files. Two respondents also suggested .jpgs for quick and easy image creation. One respondent suggested .png files for imagery, and another suggested NetCDF. Another respondent suggested translation tools to allow users to switch between formats.

As we add lidar data, we are planning new data visualizations in both 2- and 3-dimensions. In the future we will have some suggested prototypes for you to view and comment on. In the meantime, what suggestions or requests do you have on data visualization?

About half the respondents replied to this question. The individual replies are listed below to capture the diversity of suggestions. Several respondents requested the ability to show time series data, and also made suggestions regarding lidar data and display.

- I would prefer to have access to the raw data with a sophisticated query tool for subsetting and selecting lidar data to integrate with other data types. In terms of the visualization I think being able to integrate visually the time evolution of the observation with a static background image will be useful. In addition being able to show multiple of these on a map simultaneously (across sample sites or along satellite track) will be very useful.

- Is it possible to make sure that any data visualizations have accompanying data files that can be downloaded? Problem is finding a neat picture and then having no idea how that picture was generated or being able to use the information from that picture for further analysis. We are interested in creating and incorporating quantitative indicators for analysis and decision making.

- Animated 2-D visualizations of the wind fields at various altitudes/layers over the area of interest with ground-based monitoring data superimposed. The animations should run over the course of an episode (1 week or more). It would be great to be able to select the pollutant to be displayed at ground stations while the animation displays the transport pattern/wind field over time. Animated 3-D visualizations at specific locations, like major cities (especially Baltimore), would be interesting to see.

- A plotting capability to choose plot options for time, height, and scaling would be nice. The Advanced Display Options used by the MADIS profiling site is a good example of plotting options: http://www.madis-fsl.org/cap/profiler.jsp?options=full

- Time series and vertical x-sections would be great. I'm assuming backscatter or extinction will be used. Getting a handle on clouds would be useful, because their uncertainty. We get temporary jumps (decreases) in PM2.5 concentrations in the SE in summer during isolated convection events. Anything that can get a better

handle on clouds would be great. LIDAR data will be especially useful to understand residual O3 layer over urban areas during summer regional episodes, as well as transported PM2.5 into area.

- Separate images which allow visualization of the very lowest boundary and then the entire boundary layer might be helpful. In the past when I needed to inspect boundary layer information (i.e. profiler data, forecast wind or temp fields) often there are images which are good for one or the other but not both. Because processes can be happening over such thin layers, particularly in the nighttime, more than one scale in the vertical would be helpful so that one can look at different processes.

- One thing not always clear with lidar is the depth of the different range gates (if that is the right term), the lovely images give the illusion of continuous data but as I understand it the measurements are averaged over a discrete layer, or am I wrong? Anyway, knowing that information in a intro page would be nice. The lidar images in Smog Blog were informative.

- Both 2- and 3-dimensions.

- If data are provided in NetCDF, visualization is not expected to be a problem.

Ray Hoff provided the following images and requested End User interpretation.

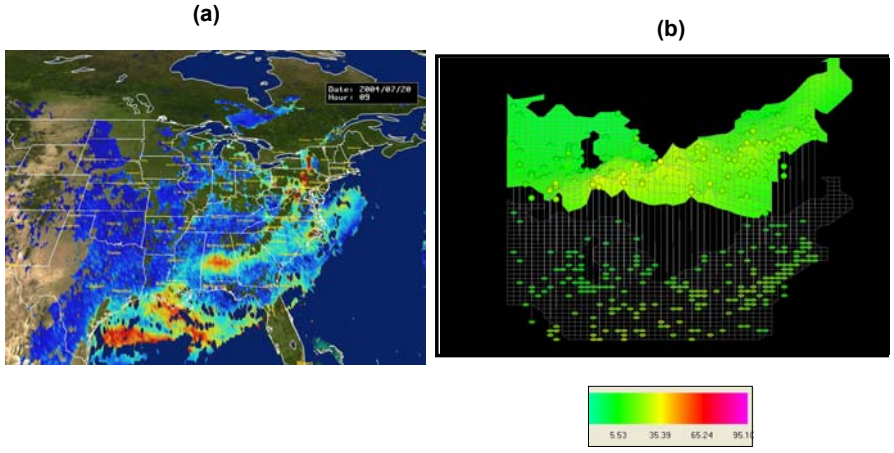


Figure A is a TrueColor image overlaid with AOD showing transport of an air mass during a fire event. Figure B is an overlay of the combination (merger) of CMAQ forecast PM2.5 and the ground based AIRNOW PM2.5 data (from a Bayesian analysis done by EPA), plotted as a contour map. (Figure B was included as an example of something that is not visually self-explaining.)
Only two respondents ventured a guess:
One on left shows enhanced AOD over Alabama, while one on right shows 3D backscatter plot.

My guess is that a) is a true color image overlaid with AOD showing transport of air mass of some fire event and b) is a temperature depiction of Europe (with the large green area over water).

APPENDIX B: Online Survey Results (Metric #2)

1. How often do you access satellite remote sensing data or imagery for understanding air pollution?		
Answer	**Count**	**Percentage**
No answer	0	0.00%
Daily (Q101)	10	66.67%
Weekly (Q102)	2	13.33%
Monthly (Q103)	0	0.00%
During major air pollution events (several times per year) (Q104)	1	6.67%
Never (Q105)	2	13.33%
2. What are your primary areas of interest: (pick all that apply)?		
Answer	**Count**	**Percentage**
Air quality forecasting (Q201)	13	86.67%
Air quality modeling and historical event analysis (Q202)	12	80.00%
Academic Interest (Q203)	4	26.67%
Personal health interest (Q204)	0	0.00%
Personal interest in general (Q205)	1	6.67%
Other (see responses below)	2	13.33%
(Lidar data analysis, air pollution measurement) (Environmental policy & program assessment)		
3. What satellite-based air quality information do you use? (pick all that apply)		
Answer	**Count**	**Percentage**
TrueColor images (like a photograph of the Earth) (Q301)	8	53.33%
Aerosol optical depth (AOD) images (Q302)	11	73.33%
Aerosol optical depth (AOD) data (Q303)	6	40.00%
Fire count or other fire-related data (Q304)	10	66.67%
Interpreted fire/smoke products (Q305)	9	60.00%
Ground-based lidar (in combination with satellite data) (Q306)	4	26.67%
None (Q307)	3	20.00%
Other (see response below)	1	6.67%
(trajectories)		
4. Where do you currently obtain satellite information? (pick all that apply)		
Answer	**Count**	**Percentage**
NASA MODIS Rapidfire Browse (Q401)	9	60.00%
University of Wisconsin MODIS Direct (Q402)	8	53.33%
NASA/EPA/NOAA/UW Infusing Satellite Data Into Environmental Applications (IDEA) (Q403)	9	60.00%
NOAA NESDIS GOES Aerosol/Smoke Product (GASP) (Q404)	8	53.33%
University of Maryland Baltimore County US Air Quality Weblog (Smog Blog) (Q405)	8	53.33%

NASA Level 1 and Atmospheric Archive and Distribution Service (LAADS) (Q406)	3	20.00%
NASA OMI Ozone & Aerosol (Q407)	2	13.33%
NASA OMI Realtime Aerosol Index (Q408)	4	26.67%
NOAA Hazard Mapping System Fire and Smoke Product (Q409)	6	40.00%
NASA Earth Observatory (Q410)	3	20.00%
NASA's Visible Earth (Q411)	0	0.00%
USGS Earth Explorer (Q412)	1	6.67%
NOAA Operational Significant Event Imagery (Q413)	3	20.00%
NASA Goddard Redhook EOS Data Gateway (Q414)	2	13.33%
NASA AERONET Sunphotometer (Q415)	3	20.00%
NOAA Hysplit Model Daily U.S. Trajectories (Q416)	6	40.00%
National Research Laboratory Monterey Aerosol Page (Q417)	6	40.00%
I do not use satellite data (Q418)	2	13.33%
Other (see responses below)	3	20.00%
(Langley DAAC Calipso, NWS Sites, NOAA GOESS and AVHRR)		

5. Do any of the following websites have applications that exceed your system or network requirements? If so, please describe the problems.

Answer	Count	Percentage
NASA MODIS Rapidfire Browse (Q501)	0	0.00%
University of Wisconsin MODIS Direct (Q502)	0	0.00%
NASA/EPA/NOAA/UW Infusing Satellite Data Into Environmental Applications (IDEA) (Q503)	0	0.00%
NOAA NESDIS GOES Aerosol/Smoke Product (GASP) (Q504)	0	0.00%
University of Maryland Baltimore County US Air Quality Weblog (Smog Blog) (Q505)	0	0.00%
NASA Level 1 and Atmospheric Archive and Distribution Service (LAADS) (Q506)	0	0.00%
NASA OMI Ozone & Aerosol (Q507)	0	0.00%
NASA OMI Realtime Aerosol Index (Q508)	0	0.00%
NOAA Hazard Mapping System Fire and Smoke Product (Q509)	0	0.00%
NASA Earth Observatory (Q510)	0	0.00%
NASA's Visible Earth (Q511)	0	0.00%
USGS Earth Explorer (Q512)	0	0.00%
NOAA Operational Significant Event Imagery (Q513)	0	0.00%
NASA Goddard Redhook EOS Data Gateway (Q514)	0	0.00%
NASA AERONET Sunphotometer (Q515)	0	0.00%

NOAA Hysplit Model Daily U.S. Trajectories (Q516)	0	0.00%
National Research Laboratory Monterey Aerosol Page (Q517)	0	0.00%
I do not use satellite data (Q518)	2	13.33%

6. If you use the IDEA (Infusing Satellite Data Into Environmental Applications) website, how important is the daily forecast discussion (which is updated daily from April 1 to September 30)?

Answer	Count	Percentage
No answer	5	33.33%
Very Important (Q601)	6	40.00%
Somewhat Important (Q602)	1	6.67%
Occasionally Important (Q603)	3	20.00%
Not Important (Q604)	0	0.00%

7. What factors limit your use of satellite data for air quality? (pick up to 3)

Answer	Count	Percentage
Lack of general understanding/explanation on available data and sources (Q701)	2	13.33%
Lack of simple, user friendly comprehensive data source (Q702)	1	6.67%
Need to know more about relation to ground-based monitoring sites (Q703)	7	46.67%
Need to know more about vertical distribution of pollution (Q704)	7	46.67%
Inadequate spatial resolution (Q705)	2	13.33%
Inadequate temporal resolution (Q706)	1	6.67%
Need customizable image (Q707)	1	6.67%
Need different data file type (Q708)	0	0.00%
Need information on different pollutants (beyond particulate matter) (Q709)	4	26.67%
Lack of software/technical capability to work with large datasets (Q710)	1	6.67%
Lack of time/resources/management support (Q711)	5	33.33%
None (generally satisfied with satellite air quality data and products) (Q712)	3	20.00%
Other (see response below)	1	6.67%
(Anything for nightime air quality/AOD?)		

8. Which of the following web sites should NASA-funded researchers focus on for improvement first? (choose single most pressing need)

Answer	Count	Percentage
No answer	5	33.33%
IDEA site (http://idea.ssec.wisc.edu/) (Q801)	5	33.33%
University of Maryland, Baltimore County U.S. Air Quality (The Smog Blog - http://alg.umbc.edu/usaq/) (Q802)	4	26.67%
Other	1	6.67%

9. Part of this project aims to incorporate satellite remote sensing data into a system such as EPA's Air Quality System (AQS). For this, there are a variety of options for gridding (or not gridding) the AOD data. MODIS AOD is available on roughly a 10-km average, although due to the shifting path of the satellite overpass each day, the grid is not the same from day to day. Which of the following options are you interested in for receiving AOD data? [Ranking 1]

Answer	Count	Percentage
10 km AOD matched with latitude and longitude (Q901)	5	33.33%
12 km standard CMAQ grids (recalculated from 10km AOD) (Q902)	3	20.00%
County level or ZIP code level data (Q903)	0	0.00%
Selected single AOD values for points associated with ground-based monitors (Q904)	4	26.67%
Format that allows multiple options listed above (customizable) (Q905)	2	13.33%

[Ranking 2]

Answer	Count	Percentage
10 km AOD matched with latitude and longitude (Q901)	4	26.67%
12 km standard CMAQ grids (recalculated from 10km AOD) (Q902)	0	0.00%
County level or ZIP code level data (Q903)	5	33.33%
Selected single AOD values for points associated with ground-based monitors (Q904)	2	13.33%
Format that allows multiple options listed above (customizable) (Q905)	2	13.33%

[Ranking 3]

Answer	Count	Percentage
10 km AOD matched with latitude and longitude (Q901)	1	6.67%
12 km standard CMAQ grids (recalculated from 10km AOD) (Q902)	3	20.00%
County level or ZIP code level data (Q903)	3	20.00%
Selected single AOD values for points associated with ground-based monitors (Q904)	2	13.33%
Format that allows multiple options listed above (customizable) (Q905)	4	26.67%

[Ranking 4]

Answer	Count	Percentage
10 km AOD matched with latitude and longitude (Q901)	3	20.00%
12 km standard CMAQ grids (recalculated from 10km AOD) (Q902)	5	33.33%
County level or ZIP code level data (Q903)	0	0.00%

Selected single AOD values for points associated with ground-based monitors (Q904)	4	26.67%
Format that allows multiple options listed above (customizable) (Q905)	1	6.67%
[Ranking 5]		
Answer	**Count**	**Percentage**
10 km AOD matched with latitude and longitude (Q901)	1	6.67%
12 km standard CMAQ grids (recalculated from 10km AOD) (Q902)	2	13.33%
County level or ZIP code level data (Q903)	5	33.33%
Selected single AOD values for points associated with ground-based monitors (Q904)	1	6.67%
Format that allows multiple options listed above (customizable) (Q905)	4	26.67%
10. For incorporation into EPA system, what requirements do you have for data and/or imagery compatibility? (pick all that apply)		
Answer	**Count**	**Percentage**
Compatible with Microsoft Access or Excel (Q1001)	8	53.33%
Compatible with GIS (e.g. ArcMap) such as GeoTiffs or .ship files (Q1002)	8	53.33%
Imagery files (e.g., .jpgs) for quick and easy image creation. (Q1003)	10	66.67%
Translation tools to allow users to switch between formats. (Q1004)	5	33.33%
Compatible with data analysis tool such as SAS or MatLab (Q1005)	4	26.67%
I do not know (Q1006)	2	13.33%
Other	0	0.00%
11. If you are involved in air quality forecasting, what time of day do you finalize the forecast for the following day? Choices: (pick one)		
Answer	**Count**	**Percentage**
No answer	8	53.33%
12 noon (Q1101)	0	0.00%
2 p.m. (Q1102)	5	33.33%
4 p.m. (Q1103)	1	6.67%
6 p.m. (Q1104)	0	0.00%
8 p.m. (Q1105)	0	0.00%
Other	1	6.67%
12. For incorporation into EPA system, NOAA GOES aerosol optical depth data can be available in as small as 15 minute increments. At what temporal resolution do you prefer the data?		
Answer	**Count**	**Percentage**
No answer	4	26.67%
Every 15 minute reading (Q1201)	2	13.33%
Hourly average (Q1202)	8	53.33%
Twice daily readings (Q1203)	0	0.00%

Daily average (Q1204)	1	6.67%
Monthly average (Q1205)	0	0.00%
Seasonal average (Q1206)	0	0.00%
Prefer raw data files to manipulate as needed (Q1207)	0	0.00%
Other (-oth-)	0	0.00%

13. What type of organization do you work for?

Answer	Count	Percentage
No answer	0	0.00%
Federal Government (Q1301)	5	33.33%
State government (Q1302)	3	20.00%
Local government (Q1303)	4	26.67%
Regional organization (Q1304)	0	0.00%
Private Sector (Q1305)	0	0.00%
Non-governmental organization (Q1306)	0	0.00%
University/Academic (Q1307)	2	13.33%
Individual (access primarily for personal interest) (Q1308)	1	6.67%
Research Institute (Q1309)	0	0.00%
Other (-oth-)	0	0.00%

14. Please provide any other comments or suggestions for improving the utility of satellite remote sensing data for air quality.

Answer	Count	Percentage
Answer (see answers below)	3	20.00%
No answer	12	80.00%

Need more lidars.

The main focus should be to provide ground truthing of the AOD data with real-time PM monitoring and develop adjusted AOD data using correlation methods.

Expand IDEA Forecast Discussion to Year Round

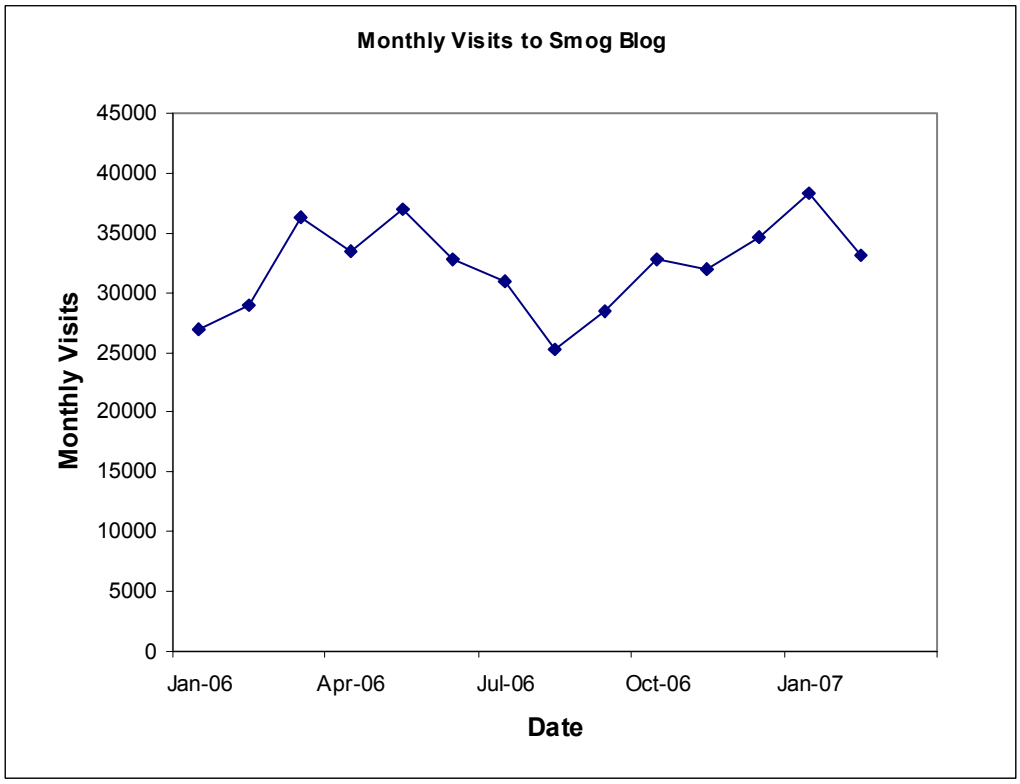
Survey Respondents (Who Provided Contact Info)

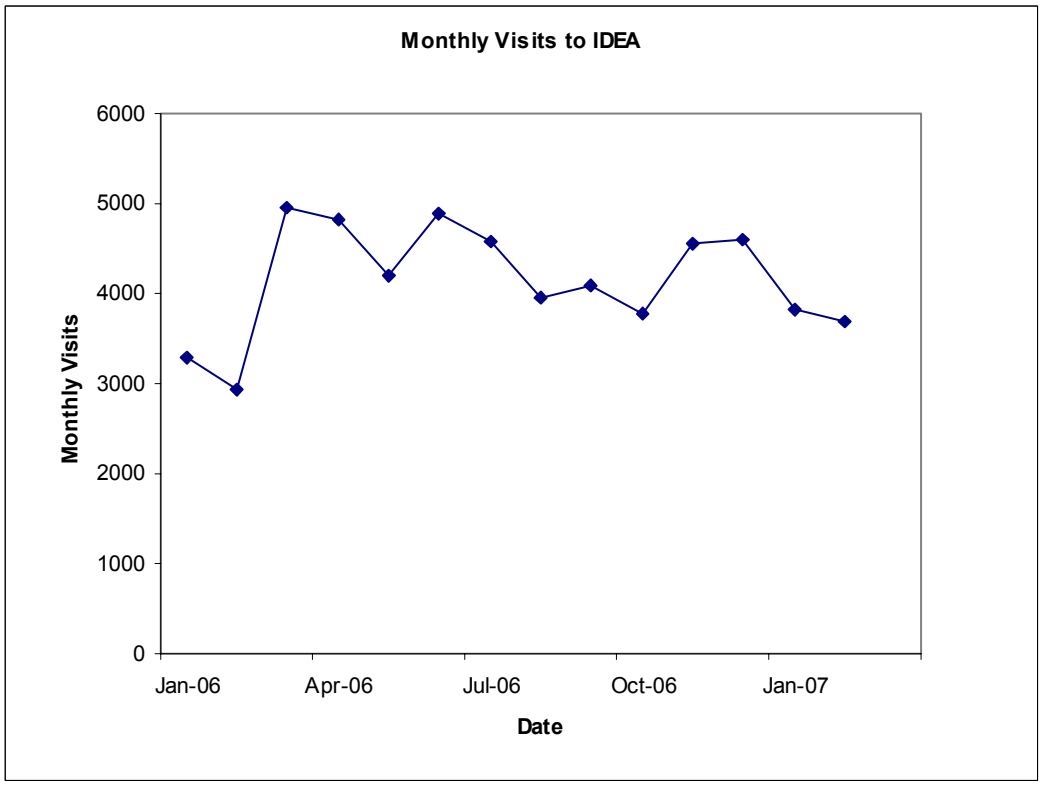
Name	Organization	Email
Ray Hoff	UMBC	hoff@umbc.edu
Dan Salkovitz	Virginia Department of Environmental Quality	ddsalkovitz@deq.virginia.gov
kamonayi mubenga	Hamilton County Dept. of Env. Services	kamo1@umbc.edu
Evan Shipp	San Joaquin Valley APCD	evan.shipp@valleyair.org
Bebhinn Do	NC DAQ	Bebhinn.Do@ncmail.net
Richard Haeuber	EPA/OAR/Clean Air Markets Division	
Brian Lee	EPA	lee.brian@epa.gov
Harry St. Clair	Hamilton County	harry.st.clair@hamilton-co.org

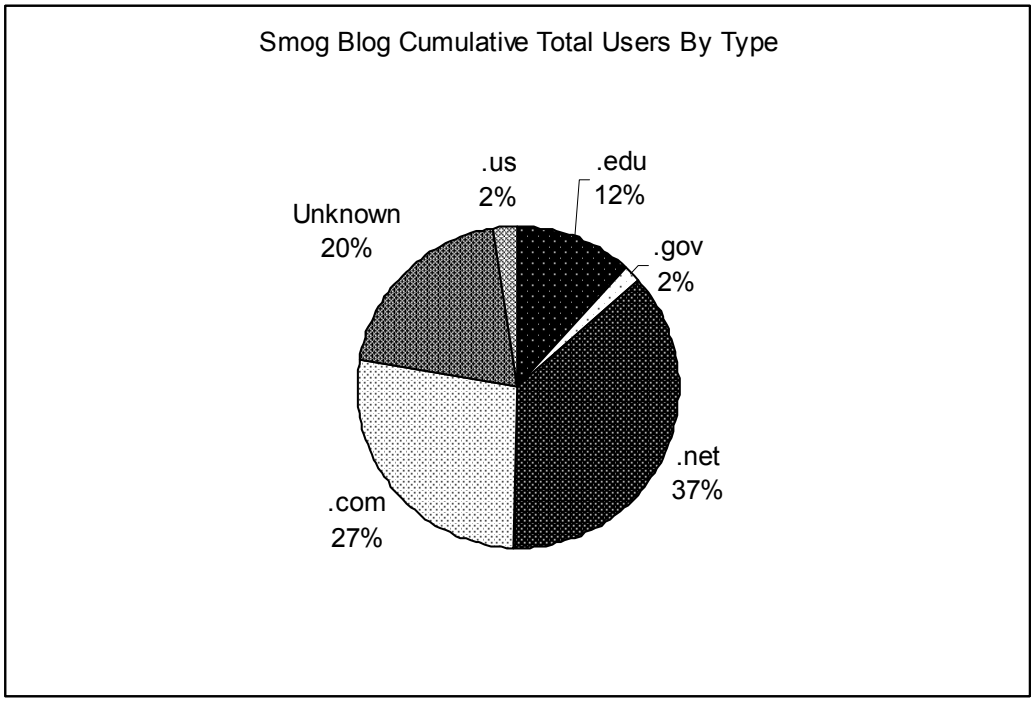
Phil	Wilderness Center Astronomy Club	pcreed4863@hotmail.com
Edward	Naval Research Laboratory	edward.hyer@nrlmry.navy.mil

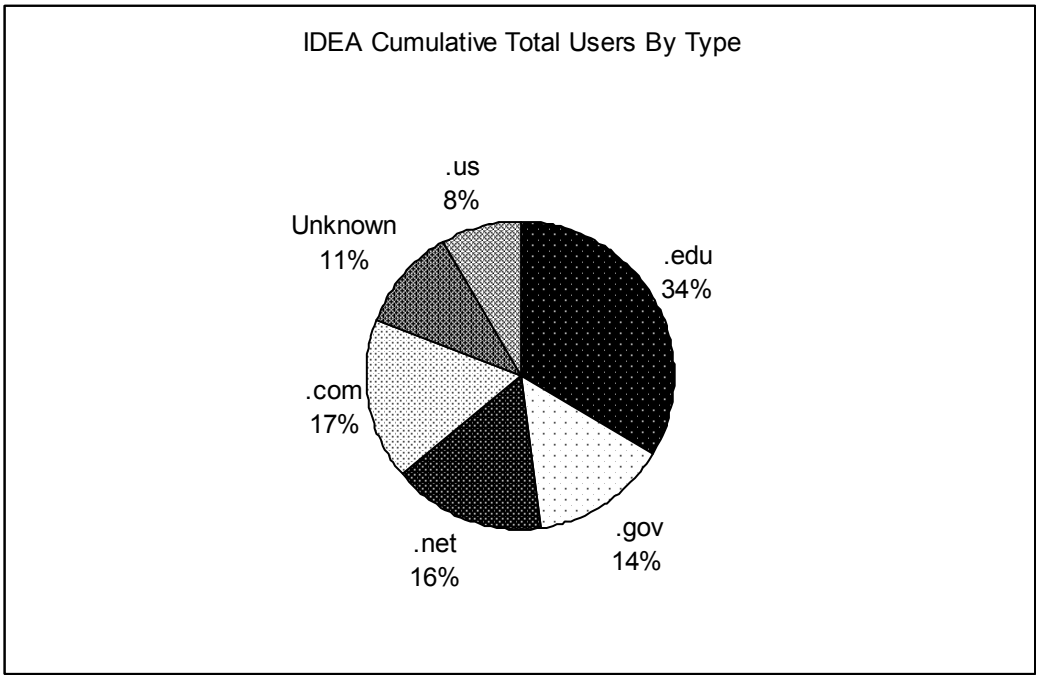
APPENDIX C: Smog Blog and IDEA Website User Logs Statistics (Metric #3)

Item	Explanation	Units	2006 Jan-06	Feb-06	Mar-06	Apr-06	May-06	Jun-06	Jul-06	Aug-06	Sep-06	Oct-06	Nov-06	Dec-06	2007 Jan-07	Feb-07
IDEA																
Activity																
Hits	Total http requests, including html files and images	Monthly Total	47289	52044	76676	124432	77798	95886	100786	97124	89466	64003	55209	48967	51123	41628
KBytes	Data transferred from server to users	Monthly Total	4302894	1883484	2534361	4554016	2476264	2847859	6101082	3800766	3434649	1961167	1632817	1391156	2223158	1331428
Visits	Series of requests for pages in set time period	Monthly Total	3282	2932	4948	4812	4203	4895	4568	3949	4086	3774	4557	4600	3813	3683
Sites	Unique IP addresses making request to server, rough indicator of unique users	Monthly Total	1508	1479	1965	2423	1919	2107	2016	1844	1710	1639	1880	1729	1703	1597
Pages	URLs .htm, .html, or .cgi (page views)	Monthly Total	9226	8457	18632	22862	19568	22198	22087	25337	22137	17192	20067	19587	20714	15652
User Type																
U.S.Educational	.edu	Monthly Total Hits		4098	18481	76699	18396	30849	29269	27675	21321	18718	16974	12704	17728	11980
U.S. Government	.gov	Monthly Total Hits		7146	11672	7219	10842	15190	12777	17120	17573	12680	4475	5884	3305	4705
Network	.net	Monthly Total Hits		9248	9516	11718	13027	12416	15822	16886	14374	8876	8507	10181	9355	6914
U.S. Commercial	.com	Monthly Total Hits		21718	17374	13917	12243	13570	16357	10444	10131	6276	6956	8549	8467	6443
Unresolved/Unknown	Unknown	Monthly Total Hits		3786	7757	5536	5705	6896	12195	13385	13716	5920	7446	4022	5873	6404
United States	.us	Monthly Total Hits		2068	5851	4794	13002	12225	7105	5851	6134	5058	3637	5093	3350	1909
Smog Blog																
Activity																
Hits	Number of files requested from server	Monthly Total	331,075	298105	360468	338486	359549	326388	325870	344718	427894	438571	438531	363701	459889	433225
KBytes	Data transferred from server to users	Monthly Total	38092591918	33548832	41007956	40689178	41608536	35681318	44747598	54955326	65572168	70753107	68951174	64621442	79860230	71914833
Visits	Approximate number of actual individual users	Monthly Total	27003	28920	36356	33478	37028	32857	30874	25235	28484	32722	31932	34558	38306	33170
Page Views	Number of web pages viewed by those visitors	Monthly Total	76559	67666	75211	72999	74797	76077	59341	51796	53898	57032	66608	55196	63361	59922
User Type																
U.S.Educational	.edu	Monthly Total Hits			42125	34907	23522	60218	45578	29485	33878	40282	40233	24259	27418	31445
U.S. Government	.gov	Monthly Total Hits			4192	5498	5699	7826	6655	6361	5584	5517	5690	4813	5161	5511
Network	.net	Monthly Total Hits			102956	99000	108912	85020	94416	96733	133110	138128	128133	117959	146662	129757
U.S. Commercial	.com	Monthly Total Hits			81540	77951	79133	75097	72481	81802	94412	87839	89236	80278	101989	91024
Unresolved/Unknown	Unknown	Monthly Total Hits			61318	54939	64498	42724	44543	57376	63445	69776	75097	54428	71727	78259
United States	.us	Monthly Total Hits			5604	5943	8210	4017	2918	4988	11056	11056	11806	7477	9460	9130


Monthly Visits to Smog Blog
Monthly Visits
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
Jan-06
Apr-06
Jul-06
Oct-06
Jan-07
Date


Monthly Visits to IDEA
Monthly Visits
6000
5000
4000
3000
2000
1000
0
Jan-06
Apr-06
Jul-06
Oct-06
Jan-07
Date


Smog Blog Cumulative Total Users By Type
.us
2%
.edu
12%
Unknown
20%
.gov
2%
.net
37%
.com
27%


IDEA Cumulative Total Users By Type
.us
8%
Unknown
11%
.edu
34%
.com
17%
.gov
14%
.net
16%