EXPLORING FIRE RELATED DATA THROUGH THE WEB

Stefan Falke
Environmental Engineering Science and the
Center for Air Pollution Impact and Trend Analysis
Washington University, St. Louis, Missouri

1. SUMMARY

A web browser provides access to a multitude of fire related data sets and visualization applications. Web sites delivering these data and applications are independent of one another and therefore require substantial effort by the user in order to combine data from multiple sites. A more dynamic approach that networks these providers of data and tools would better serve decision support systems.

Information technologies are making feasible the concept of a cyberinfrastructure of fire related data and tools. In particular, the modular and self-describing nature of web services allows them to be dynamically combined to create tools customized for specific end use applications.

Two fire detection data analysis tools developed using DataFed web services are presented as examples of the type of work that could be done through a web-based infrastructure. The presented tools spatially and temporally aggregate and compare fire detection data from different sensors.

The Federation of Earth Science Information Partners (ESIP) offers a framework for improving the information flow between data providers and data users. It is suggested that the ESIP infrastructure could bring together the data and information technologies needed to foster the development of a network of fire related data and applications that, in turn, could support decision support systems for fire, smoke, and air quality management.

2. FIRE-RELATED DATA ON THE WEB

A variety of government, academic, commercial, and non-profit web sites disseminate an impressive collection of fire related data. The data available on these http or ftp sites include fire location, fire characteristics, vegetation, fire weather, modeled smoke patterns, and air pollution concentrations. The prevalence of fire related data and applications on the web is due to many factors including attention given to recent severe wildland fire seasons, the wide range of organizations involved in managing fire and air quality, the multiple uses of the data, a variety of sensors collecting fire data, and the relative simplicity in communicating and understanding the data (e.g., fire locations, satellite image of a large smoke plume). Many of these datasets, particularly those derived from satellite imagery, are available at global scales. A single inventory of the web resources does not exist but a good starting point is the list presented in Appendix V of GAO (2003).

Web sites offering fire related data generally focus on data delivery and visualization within a single application. Some of these applications allow users to explore and, to some extent, integrate data. For instance, BlueskyRAINS (www.blueskyrains.org) uses a WebGIS to overlay fire location data with forecasted smoke plumes and meteorological data. More powerful applications would be available if the data and applications from these distributed sites were connected and shared.

A recent workshop sponsored by NASA, EPA, and NOAA that examined potential collaborative research among the agencies concluded that two primary issues in using fire data in air quality applications were knowing what data were available and understanding the relationships among similar fire data products. (EPA/NASA/NOAA, 2004)

3. CYBERINFRASTRUCTURE

Cyberinfrastructure is defined as the information sciences and technologies, including distributed computer, information and communication technologies, used to build new types of scientific and engineering knowledge environments with the goal of pursuing research and management more effectively and efficiently. *"Contemporary projects require effective federation of both distributed resources (data and facilities) and distributed, multidisciplinary expertise and cyberinfrastructure is a key to making this possible."* (Atkins et al., 2003) Aided by new information science tools and computer science networks, fire and air quality researchers and managers are poised to exchange information more effectively, integrate data and analyses more efficiently, and interact more actively (NSF, 2002; CyRDAS, 2004).

One set of technologies that could lead to shared processing, analyzing, and visualizing of data are web services; self-contained software modules using XML-based standards for describing themselves and communicating with other web resources (Alameh, 2002; Thakkaret al., 2002). Web services are designed to be independent of any particular database or application platform, and are therefore ideally suited for building cyberinfrastructure.

An example of the capabilities attainable through web services is the Open Geospatial Consortium (OGC) specifications (Buehler and McKee, 1998). The OGC specifications have been adopted by a wide range of applications in enabling exchange of GIS data layers and maps.

DataFed (www.datafed.net) is a web infrastructure that provides the foundation for accessing distributed air quality data and for processing and visualizing these data through web services (Husar et al., 2004). DataFed provides mediator software for creating "views" of data, including maps, time series, and tables, that are distributed among multiple web servers. The views are created using web services thereby allowing them to be used and reused in custom applications with standard web programming languages.

4. WEB TOOLS FOR DATA ANALYSIS

Data analysis tools represent an important component of a fire, smoke, and air quality cyberinfrastructure. Two fire detection data analysis tools developed using DataFed web services are presented as examples of the type of work that could be done through a web-based infrastructure.

Fire location data are typically presented as maps for a specific day or time period. However, exploring the temporal dimension of the fire activity is also useful. A web service was created using the DataFed environment that interactively defines a set of grid nodes at which the neighboring fire detections counted. The spatially aggregated fire count lends itself to spatial pattern analysis and. temporal trend analysis because the fire counts at each fixed location node create a record over time.

Figure 1 illustrates an example using fire pixel data from the NOAA-NESDIS Hazard Mapping System (http://www.ssd.noaa.gov/PS/FIRE/). During the summer of 2004, a number of large fires burned in central Alaska. The map in Figure 1 shows the aggregated fire pixel count on August 18, 2004 at nodes centered on 50km^2 radius grid cells. The size of each red square in the map is proportional to the number of fires counted at a node. The companion time series shows the temporal pattern of the fires for a particular node in the map. The time plot indicates an outbreak of fires in late July followed by a lull until substantial fire activity is detected again in mid-August. Clicking on another location in the map or a different date in the time plot will update the display for the respective location or date thereby allowing simple exploration of the dataset.

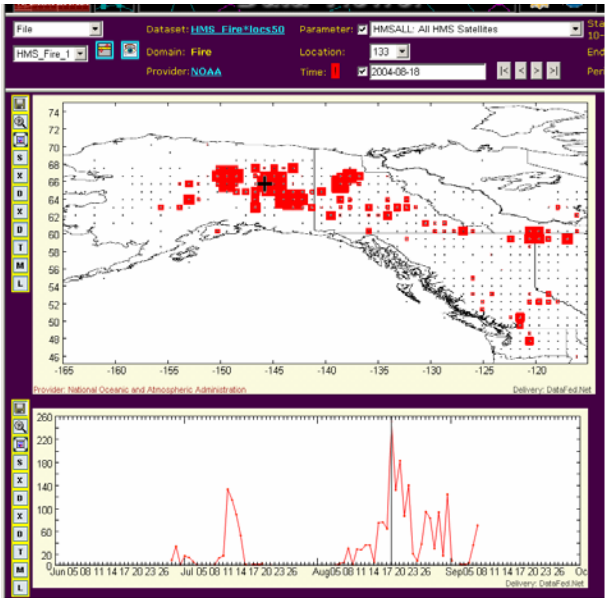

Figure 1. Spatial-temporal analysis of fire counts.
http://webapps.datafed.net/dvoy_services/datafed.aspx?page=Fire_Pixel_Count_AK

An outstanding issue in using fire detection data in air quality applications is the relationship among various fire location datasets (Boschetti, 2004). For example, comparisons are needed to understand the differences in satellite fire detections due to a sensor's spatial resolution, sampling time, or the algorithm used in detecting fires. A web application was developed to conduct exploratory spatial-temporal analysis of fire location data among satellites and field observations. The tool works by finding the nearest fire detection point in dataset B for each fire detection point in dataset A. The distance and number of days between a fire detection point in dataset A and its nearest neighbor dataset B are calculated and displayed on a map.

Figure 2 presents a comparison between GOES and MODIS fire pixels from the NOAA-NESDIS Hazard Mapping System where colored squares indicate the spatial and temporal correspondence between the two satellite derived fire pixel datasets. A red *shaded* square indicates the distance separating the MODIS and GOES pixels was small while a blue *shaded* square indicates the nearest neighbor between the datasets were far apart. A red *outlined* square indicates the nearest neighbor was detected on the same day while a blue *outlined* square indicates a longer time separation. Gray shaded and/or outlined squares indicate that a nearest neighbor was not found between the two datasets given the search parameters (in this example case, 100 km and 2 days). On July 17, 2004, both GOES and MODIS detected fires in the Central Plains. The distance (both in space and time) is small between GOES and MODIS pixels and are therefore colored red in the maps. The top map displays GOES fire locations colored by their distance to the nearest MODIS pixel. The bottom map displays MODIS fire locations colored by their distance to the nearest GOES pixel. While there is good agreement in Kansas, the fires detected by MODIS in northern Texas were not detected by GOES (and are therefore colored gray).

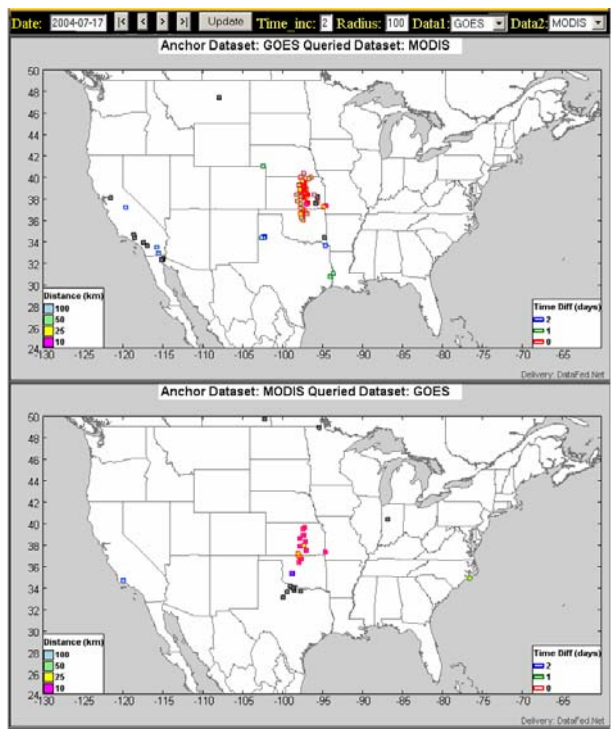

Figure 2. Comparison of fire detection datasets.
http://www.datafed.net/WebApps/MiscApps/ModisGoes/FireLocationComparison.htm

This type of comparative analysis is merely the beginning of a more involved analysis to understand the influence of sensor characteristics and algorithms in differences between fire detection datasets. The tools presented are designed to be flexible and adaptive so that they can import other datasets, be extended with more advanced analysis algorithms, or be used in a "chain" of processing services for other applications.

5. ESIP AIR QUALITY CLUSTER

The datasets and information technologies needed for a fire/air quality cyberinfrastructure exist but a framework is needed that allows these pieces to connect. An available framework in which to develop such partnerships is the Federation of Earth Science Information Partners (ESIP). ESIP brings together government agencies, universities, and businesses with the goal of facilitating information flow between data providers and end users (www.esipfed.org). Its Air Quality Cluster focuses on applying information science and technology to improve access, integration and analysis of data through air quality case studies.

The Air Quality Cluster is considering as a case study the integration of distributed surface data, satellite imagery, and model results in improving smoke emission estimates used by forest fire managers and air quality planners. A key activity in developing case studies is building the community of data providers, information science, and data users and the Cluster invites any interest and participation among the fire and air quality communities.

REFERENCES

Alameh, N., 2002: The Frontier of GIS: GIS Web Services, In Proceedings of Next Generation of Community Statistical Systems Conference, Tampa, Florida.

Atkins, D.E.; Droegemeier, K.K; Feldman, S.I.; Garcia-Molina, H.; Klein, M.L.; Messerschmitt, D.G.; Messina, P.; Ostriker, J.P.; Wright, M.H., 2003: Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. Technical report, National Science Foundation, http://www.cise.nsf.gov/sci/reports/toc.cfm

Boschetti, L., 2004: Lessons to be learned from the comparison of three satellite-derived biomass burning products. *Geophysical Research Letters*, 31, L21501.

Buehler, K. and L. McKee, 1998: The OpenGIS Guide: Introduction to Interoperable Geoprocessing and the OpenGIS Specification, Waltham, MA.

EPA/NASA/NOAA, 2004: Joint EPA/NASA/NOAA Workshop on Air Quality and Related Climate Change Issues, Research Triangle Park, http://www.cleanairinfo.com/EPA_NASA_NOAA/

General Accounting Office (GAO), 2003: Geospatial Information: Technologies Hold Promise for Wildland Fire Management, but Challenges Remain, Report to Congressional Requestors, GAO-0301047 http://www.gao.gov/new.items/d031047.pdf

Husar, R.B; Hoijarvi, K.; and Falke, S.R., 2004: DataFed at a Glance, online document, http://www.datafed.net/datafed_Q.asp.

National Science Foundation, 2002: Cyberinfrastructure for Environmental Research and Education, Report from a workshop held at the National Center for Atmospheric Research.

Steering Committee for Cyberinfrastructure Research and Development in the Atmospheric Sciences (CyRDAS), 2004: Cyberinfrastructure for the Atmospheric Sciences in the 21st Century: A report from the Ad Hoc Committee for Cyberinfrastructure Research, Development and Education in the Atmospheric Sciences.

Thakkar, S., C.A. Knoblock, J. L. Ambite, C. Shahabi, 2002: Dynamically Composing Web Services from On-line Sources, In Proceedings of AAAI-02 Workshop on Intelligent Service Integration, Edmonton, Canada.

BIOGRAPHY

Stefan Falke is a research assistant professor of Environmental Engineering Science and in the Center for Air Pollution Impact and Trend Analysis (CAPITA) at Washington University in St. Louis. His research involves spatial-temporal data analysis and the development of web-based tools to support air quality research and management, particularly where integration of surface, satellite and model data is needed. He is a coordinator of the Earth Science Information Partners (ESIP) Air Quality Cluster and can be contacted at stefan@wustl.edu. .