Federated Data System
DataFed

Rudolf Husar, Washington University, St. Louis
With may who participated and supported DataFed since 2004

2009 National Ambient Air Monitoring Conference
November 2-5, 2009, Gaylord Opryland Hote, Nashville, TN
Summary: http://datafed.net

DataFed: Agile, Federated Data System
Mediator for Accessing 100+ AQ Datasets
Incorporates Tools for AQ Data Analysis
Applied to Research, EE Analysis
Users Welcome

but

Needs to be more friendly to Sate/Reg/Fed Analysts

Research/Development Supported by
Screencast Links

DataFed Background

Canada Smoke
Georgia/Florida June 2007 Smoke Exceptional Event

CATT All stations one day

AQS PM2.5  http://webapps.datafed.net/datafed.aspx?dataset_abbr=AQS_D

VIEWS AerChem  http://webapps.datafed.net/datafed.aspx?dataset_abbr=VIEWS

AirNOW  http://webapps.datafed.net/datafed.aspx?dataset_abbr=Airnow
Current info systems are **project/program oriented** and provide end-to-end solutions.

Part of the data resources of any project can be **shared for re-use** through DataFed.

Through the Federation, the **data are homogenized** into multi-dimensional cubes.

**Data processing** and rendering can then be performed **through web services**.

Each project/program can be **augmented by Federation data and services**.
Main tools of DataFed

**Viewer:** General purpose spatio-temporal data browser and view editor applicable for all DataFed datasets

**CATT:** Combined Aerosol Trajectory Tool for the browsing backtrajectories for specified chemical conditions

**Consoles:** Data from diverse sources are displayed to create a rich context for exploration and analysis.
Example Datasets in DataFed

- Data are accessed from autonomous, distributed providers
- DataFed ‘wrappers’ provide uniform geo-time referencing
- Tools allow space/time overlay, comparisons and fusion

**Near Real Time Data Integration**

**Delayed Data Integration**

**Surface Air Quality**
- AIRNOW: O3, PM25
- ASOS_STI: Visibility, 300 sites
- METAR: Visibility, 1200 sites
- VIEWS_OI: 40+ Aerosol Parameters

**Satellite**
- MODIS_AOT: AOT, Idea Project
- GASP: Reflectance, AOT
- TOMS: Absorption Index, Refl.
- SEAW_US: Reflectance, AOT

**Model Output**
- NAAPS: Dust, Smoke, Sulfate, AOT
- WRF: Sulfate

**Fire Data**
- HMS_Fire: Fire Pixels
- MODIS_Fire: Fire Pixels

**Surface Meteorology**
- RADAR NEXRAD
- SURF_MET: Temp, Dewp, Humidity...
- SURF_WIND: Wind vectors
- ATAD: Trajectory, VIEWS locs.
Data Consoles
A Sample of Datasets Accessible through ESIP – DataFed Mediation
Near Real Time (~ day)

MODIS Reflectance
MODIS AOT
TOMS Index
GOES AOT
GOES 1km Reflec
MODIS Fire Pix
NEXTRAD Radar
NWS Surf Wind, Bext
NRL MODEL
CATT: A Community Tool!
Part of an Analysis Value Chain

AEROSOL

- Collection IMP. EPA
- Integration VIEWS
- Aerosol Sensors → Aerosol Data → Integrated AerData → AerData Cube → CATT-In CAPITA → Aggreg. Aerosol → Next Process

TRANSPORT

- Assimilate NWS
- Trajectory ARL
Quebec Smoke Event, July 07, 2002
SeaWiFS Satellite, VIEWS Aerosol Organics
Back Trajectories for All IMPROVE Sites on 7/7/02 Unweighted (top left), & color-weighted for OC (top right), SO4 lower right & Cl (lower left)
Combined Aerosol Trajectory Tool

Resources/Discussion

Manual - pdf, ppt

Ensemble Trajectory Browsers

<table>
<thead>
<tr>
<th>Single Site, Single Day</th>
<th>Multi-Site, Single Day</th>
<th>All Sites, Single Day</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Single Site, Single Day" /></td>
<td><img src="image2" alt="Multi-Site, Single Day" /></td>
<td><img src="image3" alt="All Sites, Single Day" /></td>
</tr>
</tbody>
</table>

User-Defined Filter

![User-Defined Filter](image4)

Gridded Transport Metrics Browsers

<table>
<thead>
<tr>
<th>Inc. Prob. 'Poirot'</th>
<th>Src.Contr. 'Hopke'</th>
<th>Avg. Conc. 'Kenski'</th>
<th>Weighed Prob. 'Green'</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Inc. Prob. 'Poirot'" /></td>
<td><img src="image6" alt="Src.Contr. 'Hopke'" /></td>
<td><img src="image7" alt="Avg. Conc. 'Kenski'" /></td>
<td><img src="image8" alt="Weighed Prob. 'Green'" /></td>
</tr>
</tbody>
</table>
Sulfate Transport to BIBE, GRSM and LYBR

All Data

80-100 Percentile

0-20 Percentile

Big Bend, TX

Great Smoky, TN

Lynbrook, VT
Exceptional Event Rule:
Evidence Needed to Flag Data as Exceptional

• The event was not reasonably controllable or preventable

• Would be no exceedances or violation but for the event.

• The event is in excess of historical values.

• Clear casual relationship of data and the event
1. **The event** not reasonably controllable/preventable

Show that the cause is in category of **uncontrollable/preventable**

**Transported Pollution**
- Transported African, Asian Dust; Smoke from Mexican fires & Mining dust, Ag. Emissions

**Natural Events**
- Nat. Disasters.; High Wind Events; Wildland Fires; Stratospheric Ozone; Prescribed Fires

**Human Activities**
- Chemical Spills; Industrial Accidents; July 4th; Structural Fires; Terrorist Attack
2. No exceedance/violation **but for** the event.

Show that the exceedance is explicitly caused by the exceptional event

---

**Exceptional Event**

The 'exceptional' concentration raises the level above the standard. A valid EE to be flagged.

**NOT Exceptional Event**

Controllable sources are sufficient to cause exceedance. Not a 'but for', not an EE.

**NOT Exceptional Event**

No exceedance, hence, there is no justification for an EE flag.
3. The event is in excess of historical values.

Evidence from comparison of flagged data with historical values

Frequency Distribution

The 'exceptional' concentration is an outlier on the frequency.

Time Series Analysis

Event data deviate from the regular seasonal concentration pattern.
4. Clear support of event causality with data.

EE causality may come from multiple lines of observational evidence

**Chemical Signature**
The EE sample shows the fingerprints of 'exceptional' source.

**Source & Transport**
Clear evidence of transport from known source region.

**Spatial Pattern**
Unusual spatial pattern as evidence of Exceptional source.

**Temporal Pattern**
Unusual concentration spike as indication of an Exceptional Event.
Near-Real-Time Data for May 11, 07 GA Smoke

Pane 1,2: MODIS visible satellite images – smoke pattern
Pane 5,6: AirNOW Ozone, Surf. Wind – Ozone, transport pattern
Pane 7,8: OMI satellite Total, Tropospheric NO2 – NO2 column conc.
Pane 9,10: OMI satellite Aerosol Index, Fire P-xels – Smoke, Fire
Pane 11,12: GOCART, NAAPS Models of smoke – Smoke forecast

Console Links
May 07, 2007
May 08, 2007
May 09, 2007
May 10, 2007
May 11, 2007
May 12, 2007
May 13, 2007
May 14, 2007
May 15, 2007
May 2007 Georgia Fires

The fires in S. Georgia emitted intense smoke throughout May 07.

Google Earth Video (small 50MB, large 170mb)
The Exceedance would not Occur, But For the Exceptional Event
May 07 Georgia Fires: User-Supplied Qualitative Observations

Searching and pruning user-contributed Internet content yielded rich, but qualitative description of the May 07 Georgia Smoke Event.

Google and Technorati blog searches yielded entries on GA Smoke.

Videos of smoke were found on YouTube.

Smoke images, were also found searching Flickr and Google.

Visually pruned blogs, videos and images were bookmarked and tagged for later analysis.
Screencast Links

DataFed Background

Canada Smoke
Georgia/Florida June 2007 Smoke Exceptional Event

CATT All stations one day

AQS PM2.5 http://webapps.datafed.net/datafed.aspx?dataset_abbr=AQS_D

VIEWs AerChem http://webapps.datafed.net/datafed.aspx?dataset_abbr=VIEWs

AirNOW http://webapps.datafed.net/datafed.aspx?dataset_abbr=Airnow
Summary

- Major advances in fire detection (fire pixels, burn scars)
- Satellite and surface smoke detection is also advanced
- Still smoke quantification is elusive

Integrative Analysis ‘revolution’ (200? ++)

- Need to integrate most of the sensory-model info
- Propose an InfoTech – enabled collaborative approach
Summary: http://datafed.net

DataFed: Agile, Federated Data System
Mediator for Accessing 100+ AQ Datasets
Incorporates Tools for AQ Data Analysis
Applied to Research, EE Analysis
Users Welcome
but
Needs to be more User Friendly
Mystery Winter Haze:
Natural? Nitrate/Sulfate? Stagnation?

 CONTRIBUTED BY THE FASNET COMMUNITY, SEP. 2004

CORRESPONDENCE TO R HUSART, R POIROT

COORDINATION SUPPORT BY

INTER-RPO WG - FAST AEROSOL SENSING TOOLS FOR NATURAL EVENT TRACKING, FASTNET
NSF - COLLABORATION SUPPORT FOR AEROSOL EVENT ANALYSIS
NASA - REASON COOP
EPA - OAQPS
Secondary MP25 Peak in February-March

- The AIRNOW PM25 data are available real-time for 300+ stations since July 2002.
- The 30-day smoothing of the average hourly data shows the Eastern US PM25 seasonality.
- The seasonal pattern shows the summertime sulfate peak and a second Feb/Mar peak.
- The existence, characteristics, and origin of this regional peak is not known.
- The objective of effort is to characterize this ‘mysterious’ phenomenon over the EUS.
- The approach is to seek out the community as a resource for collaborative analysis.
Regional Average PM25 Concentration Pattern

Based on AIRNOW

From Rudy Husar

- **Time pattern of the 0502 Event**
  - The overall event lasted about 10 days, Jan 28-February 7
  - The Upper Midwest peaked first (Jan 31-Feb 2); Industrial MW later (Feb 3 – 6)
  - The Industrial MW region show more diurnal variation (lowest in the mid-afternoon)
Seasonal PM25 by Region

The 30-day smoothing average shows the seasonality by region.

The Feb/Mar PM25 peak is evident for the Northeast, Great Lakes and Great Plains.

This secondary peak is absent in the South and West.
Monthly average FRM PM25 are shown as circle and contour (Blue: 0; Red: 25 µg/m³)

- The Feb/Mar peak is clearly evident in the Midwest region; also in January
- Hence, there is some deviation in peak location and time among the networks