Rapid Deployment of EPA Enterprise Geospatial Data Warehouse in Katrina's Aftermath

ESRI 2006 User Conference

May 18th, 2006

Torrin Hultgren & Qi Dai

EPA GIS Support Team

Research Triangle Park, North Carolina

Abstract

In the aftermath of Hurricane Katrina, the EPA had an urgent need for current geographic data on everything from known hazardous waste sites, flood conditions, streets and building footprints, and high-resolution satellite imagery. To respond to this need and make the data accessible to EPA personnel in dispersed locations, a high-availability ArcSDE/Oracle data warehouse and ArcIMS server were rapidly deployed in the EPA's National Computing Center in Research Triangle Park, North Carolina. Triage was performed on the terabytes of commercial imagery the NGA made available to federal agencies, and ArcGIS Spatial Analyst and ArcToolbox's Modelbuilder proved equal to the task of rapidly projecting, stretching and mosaicking large volumes of imagery.

EPA Emergency Response Role

- Field Testing – measuring levels of toxins and contaminants in floodwaters
- Mitigate risk to critical facilities and sites
- Contain larger spills when possible
- Coordinate safe debris removal and disposal
- Monitor air quality levels at debris burn sites
- GIS task: rapidly create maps and perform spatial analyses to support field engineers and decision-makers

EPA GIS Community

- Diverse Locations:
  - Region 5 – Louisiana
  - Region 4 – Mississippi and Alabama
  - Headquarters EOC – Washington, DC
- Mix of EPA Staff and Contractor Support
- Mix of GIS Experts and EOC GIS Novices
- Most located on the secure, high-speed EPA internal network

EPA Data Needs

- EPA Regulated Facilities Data
  - Already in SDE database
- Basemap data
  - As feature services
  - From various sources
- High-resolution post-Katrina imagery
  - Demand for imagery was intense
  - Response was mandate to NGA to purchase and disseminate as much imagery as possible to responders

Example Feature Layers

Airports
Area Codes
Buildings
Census Blocks
Cemeteries
Curblines
Debris Burn Locations
Drainage Lines
Evacuation Routes
Evacuation Zones
Federal Courts
Federal Lands
FEMA Zones
Fire Districts
Fire Stations
Fire Zones
Flood Extents
Golf Courses
Hospitals
Hurricane Swath
Hydrography
Interstates
Address Points
Lakes
Land Use
Landfills
Lots
Major Institutions
Major Roads
Neighborhoods
Parcels
Parish Boundaries
Parks
Police Dept Zones
Police Stations
Populated Places
Railroads
Recreation Areas
Schools
Sewer Lines
Sewer Nodes
State Boundaries
Street Centerlines
Topo Map Indices
Tracts
Truck Routes
Urban Areas
Water Bodies
Zip Codes

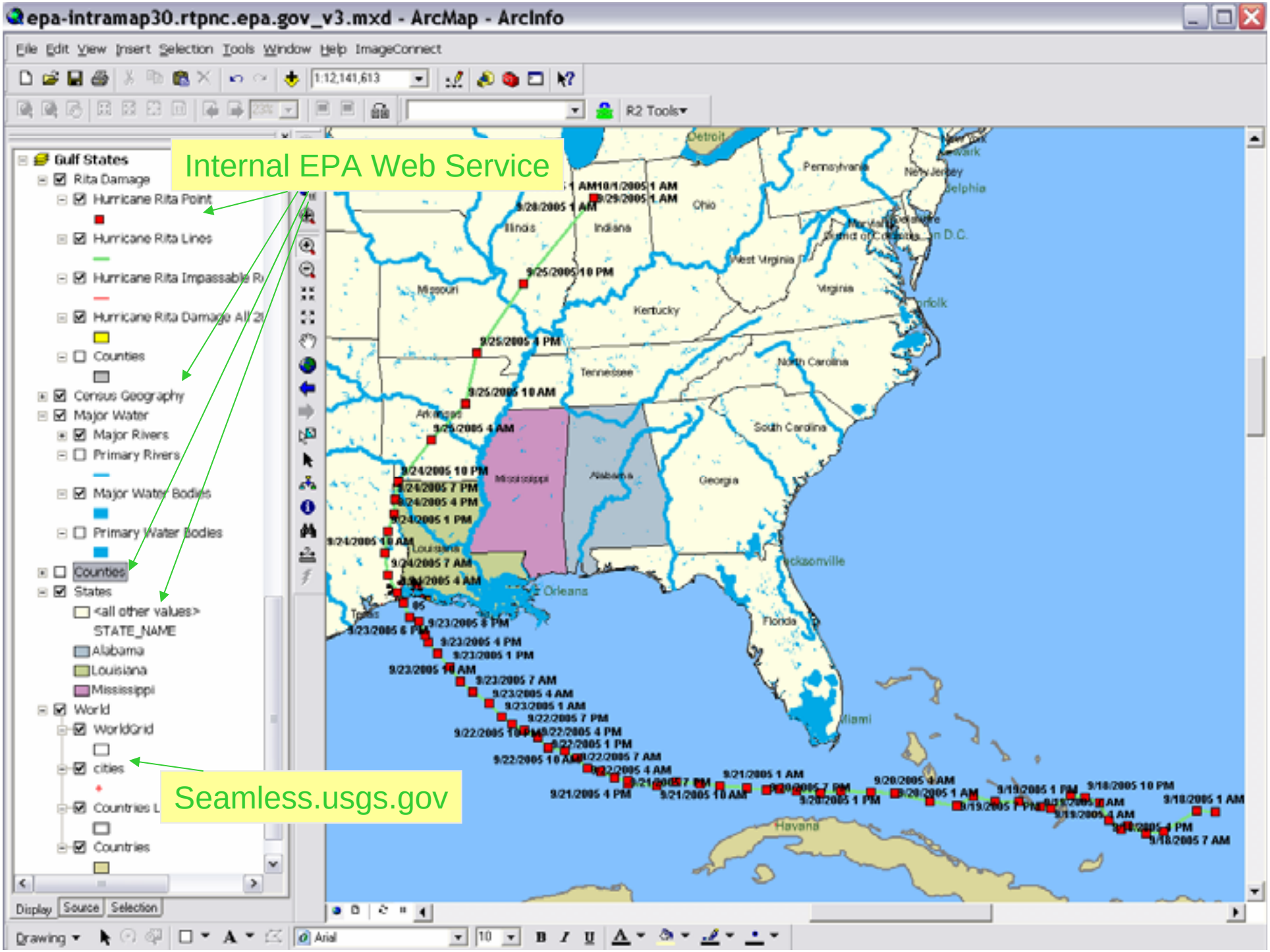
EPA National Computing Center

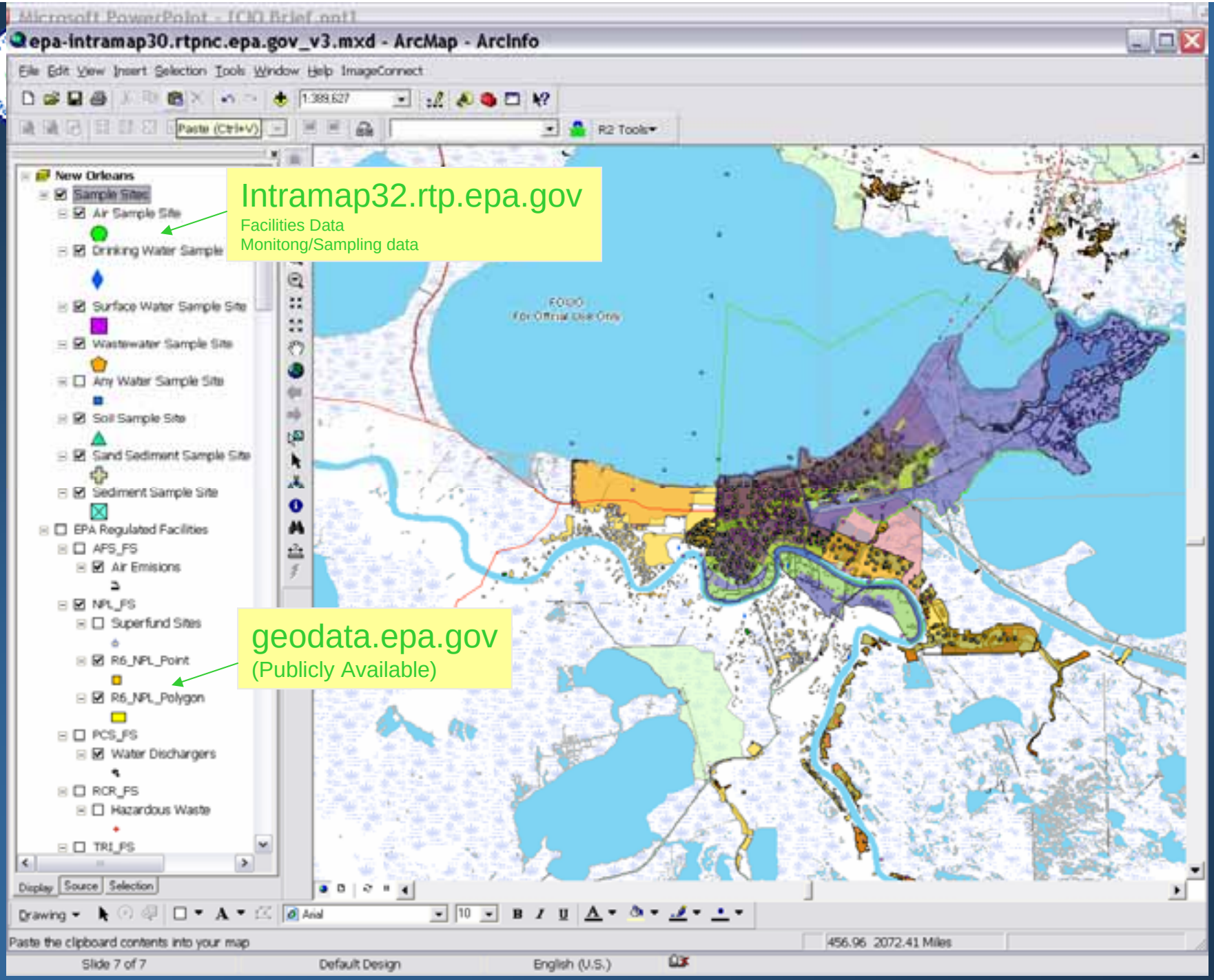
- Located in Research Triangle Park, NC
- Industrial strength data center
- Enterprise computing resources available for all EPA program offices, Research labs, and Regions
- Official host of all EPA GIS applications and databases for public access
- Official host of most EPA national GIS applications and databases for intranet access
- Agency-wide network resource hub and the gateway to the Internet
- 7x24 operation, backup and restore services, and emergency response
- EPA GIS group – responsible for public web mapping applications, internal geospatial databases, and general GIS and ESRI desktop support

Katrina Response Timeline

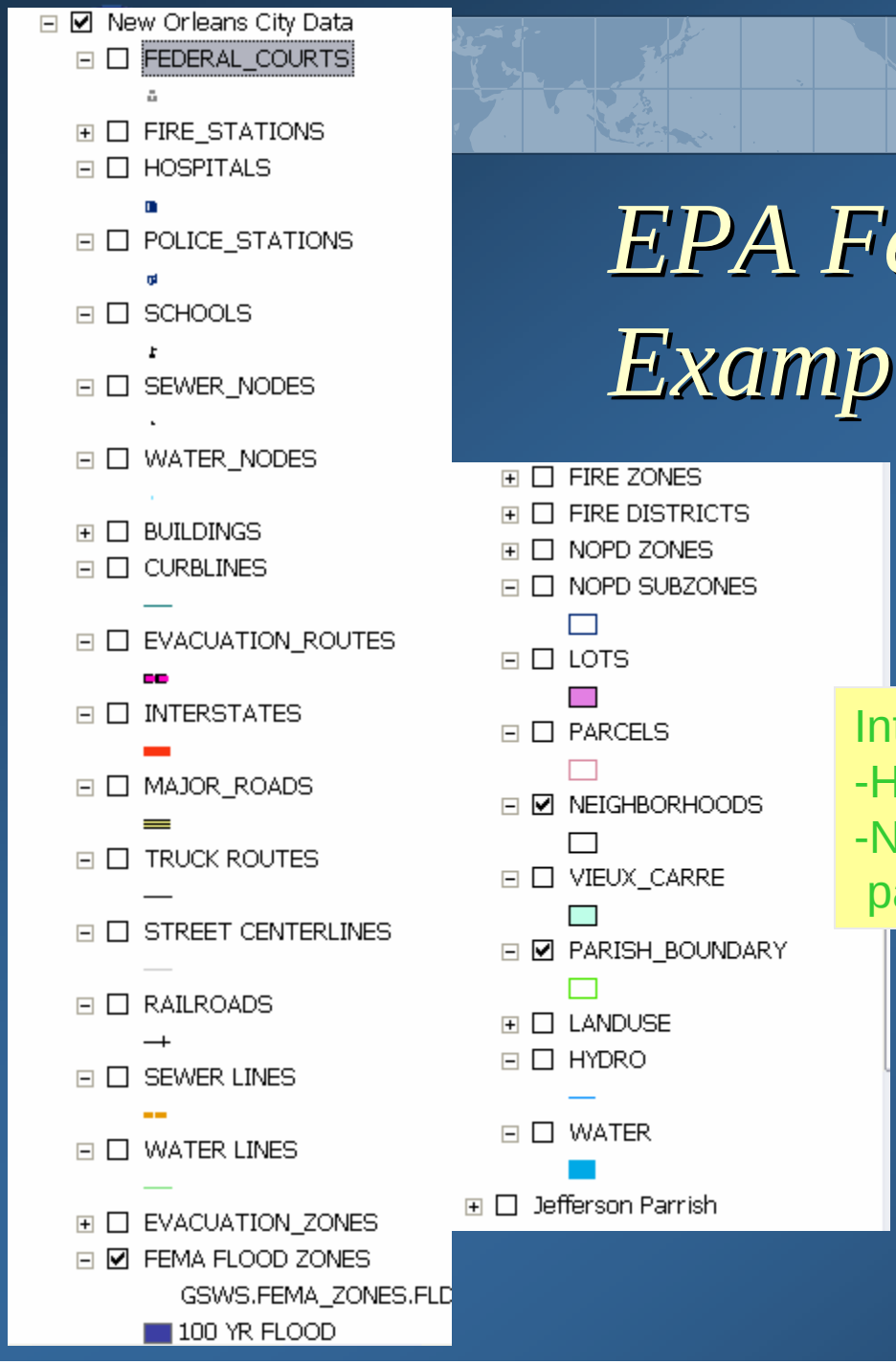
- Created new SDE instance on existing Oracle DB server on day of request
- Created load-balanced ArcIMS application server configuration the next day

8/29	9/01	9/09	9/16	9/23	10/04
R4 data request		R4/R6/EOC data request	R6/EOC data request	R6/EOC data request	R6/EOC data request
Katrina Hits Levees Break	5 Image services 5 feature services Added to the public access geodata.epa.gov	6 Image services 9 feature services 5 high res images Set up FTP site offering download of same data	5 feature services containing 35 data layers added to the intranet geo server Additional high res images added	2 feature services For Parcel Data containing 35 data layers added to Intranet geo server Additional high res images added	Rita Hits 4 Image services 4 feature services for Rita damage 5 SDE layers for Operational Facilities Data NWI feature service Additional high res images added


epa-intramap30.rtpnc.epa.gov_v3.mxd - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help ImageConnect
1:12,141,613
R2 Tools
Internal EPA Web Service
Gulf States
Rita Damage
Hurricane Rita Point
Hurricane Rita Lines
Hurricane Rita Impassable R
Hurricane Rita Damage All Zi
Counties
Census Geography
Major Water
Major Rivers
Primary Rivers
Major Water Bodies
Primary Water Bodies
Counties
States
<all other values>
STATE_NAME
Alabama
Louisiana
Mississippi
World
WorldGrid
cities
Countries L
Countries
Seamless.usgs.gov
Display Source Selection
Drawing
Arial
10
Detroit
Pennsylvania
New Jersey
Ohio
Illinois
Indiana
West Virginia
Virginia
Missouri
Kentucky
Tennessee
North Carolina
South Carolina
Arkansas
Mississippi
Alabama
Georgia
Louisiana
Texas
Florida
Orleans
Miami
Havana
9/25/2005 10 PM
9/25/2005 4 PM
9/25/2005 10 AM
9/25/2005 4 AM
9/24/2005 10 PM
9/24/2005 7 PM
9/24/2005 4 PM
9/24/2005 1 PM
9/24/2005 7 AM
9/23/2005 8 PM
9/23/2005 4 PM
9/23/2005 1 PM
9/23/2005 7 AM
9/23/2005 4 AM
9/23/2005 1 AM
9/22/2005 7 PM
9/22/2005 4 PM
9/22/2005 1 PM
9/22/2005 7 AM
9/22/2005 4 AM
9/21/2005 1 AM
9/21/2005 4 PM
9/21/2005 10 AM
9/20/2005 4 AM
9/20/2005 1 AM
9/20/2005 1 PM
9/19/2005 1 PM
9/18/2005 10 PM
9/18/2005 1 AM
9/18/2005 7 AM


epa-intramap30.rtpnc.epa.gov_v3.mxd - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help ImageConnect
New Orleans
Sample Sites
Air Sample Site
Drinking Water Sample
Surface Water Sample Site
Wastewater Sample Site
Any Water Sample Site
Soil Sample Site
Sand Sediment Sample Site
Sediment Sample Site
EPA Regulated Facilities
AFS_FS
Air Emissions
NPL_FS
Superfund Sites
R6_NPL_Point
R6_NPL_Polygon
PCS_FS
Water Dischargers
RCR_FS
Hazardous Waste
TRI_FS
Intramap32.rtp.epa.gov
Facilities Data
Monitong/Sampling data
geodata.epa.gov
(Publicly Available)
Paste the clipboard contents into your map
Slide 7 of 7
Default Design
English (U.S.)

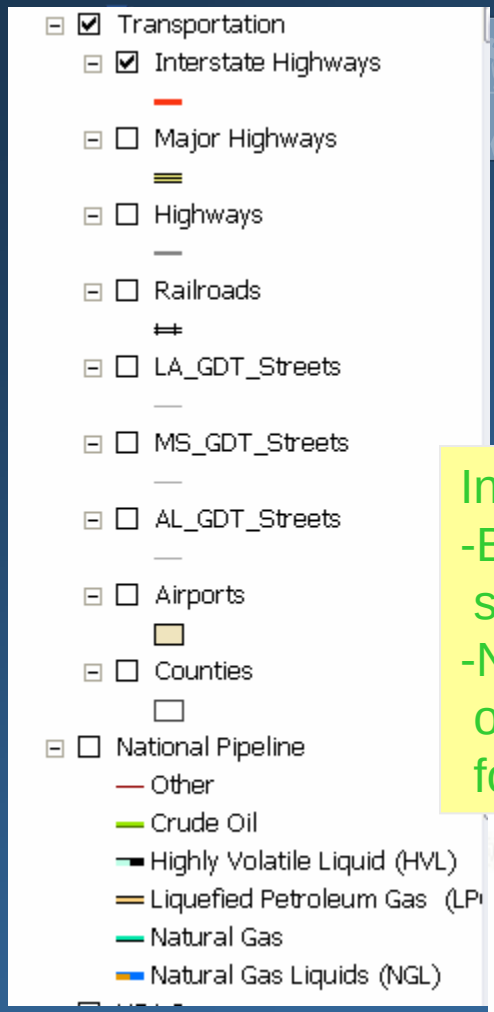
EPA Feature Service Examples – Local Data



Internal Web Services
-Hi-resolution local data
-New Orleans detailed transportation, parcel, administrative information

EPA Feature Service Examples:

Nationally Consistent Data Services



Internal EPA Web Services
-Broad range of data from GDT national subscription
-National Pipeline Data – acquired from Office of Pipeline Safety by GIO, critical layer for emergency response, for internal use only

Imagery: Need for Triage

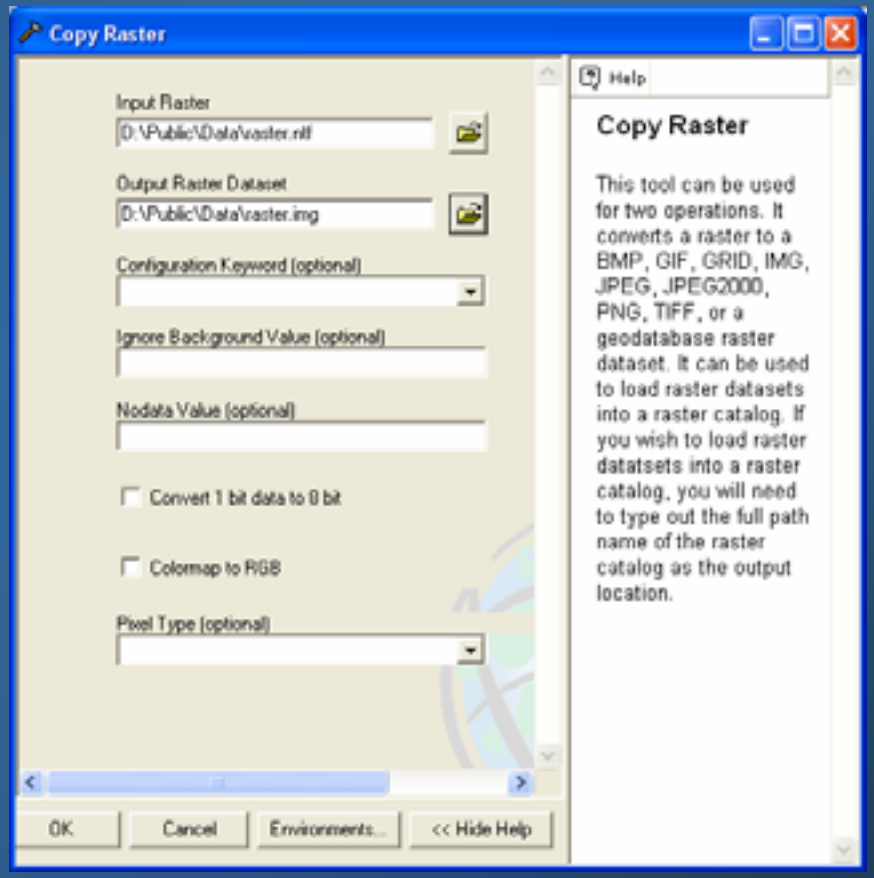
- The volume of imagery was in terabytes
- Imagery was provided by many sources – USGS, FEMA, NOAA, USACE, NGA, commercial providers, various universities
- The same images were posted online in many places, and most sites were updated daily
- Many websites had no ability to preview images prior to download
- Field GIS personnel responsible for making maps "five minutes ago"
  - Lacked bandwidth to handle "firehose" of downloadable imagery
  - Requested triage assistance from EPA RTP GIS Group

Response to Need for Triage

- RTP Staff monitored all imagery sites and maintained a list of potentially useful cloud-free images
- RTP Staff processed and republished restricted NGA imagery as ArcIMS Image Services
- Kept responders informed of updates to imagery via daily conference calls

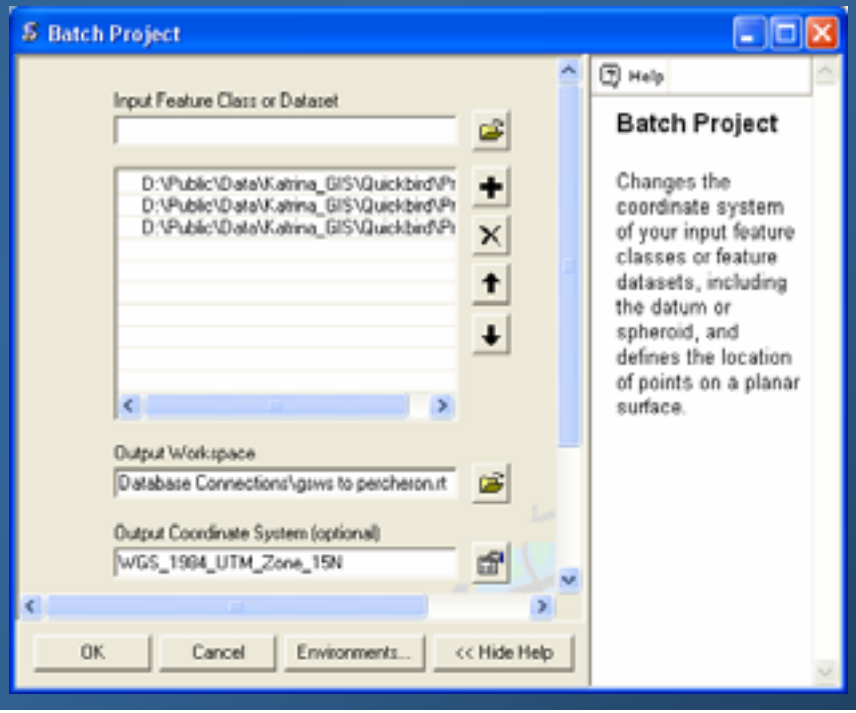
Geoprocessing: Format Conversion

- Raw, unrectified imagery posted as National Imagery Transmission Format (NITF) files
- Support in ENVI 4 and GeoXpress 4.0 was poor to nonexistent, but ArcGIS can recognize and process NITF reliably
- Copy Raster tool used for loading into SDE and conversion to MrSID compressible format



Geoprocessing: Reprojection

- New Orleans is divided by two UTM zones, imagery arrived in both projections
- ArcGIS can project images on-the-fly but all images in a SDE mosaic or Raster Catalog* must have the same projection

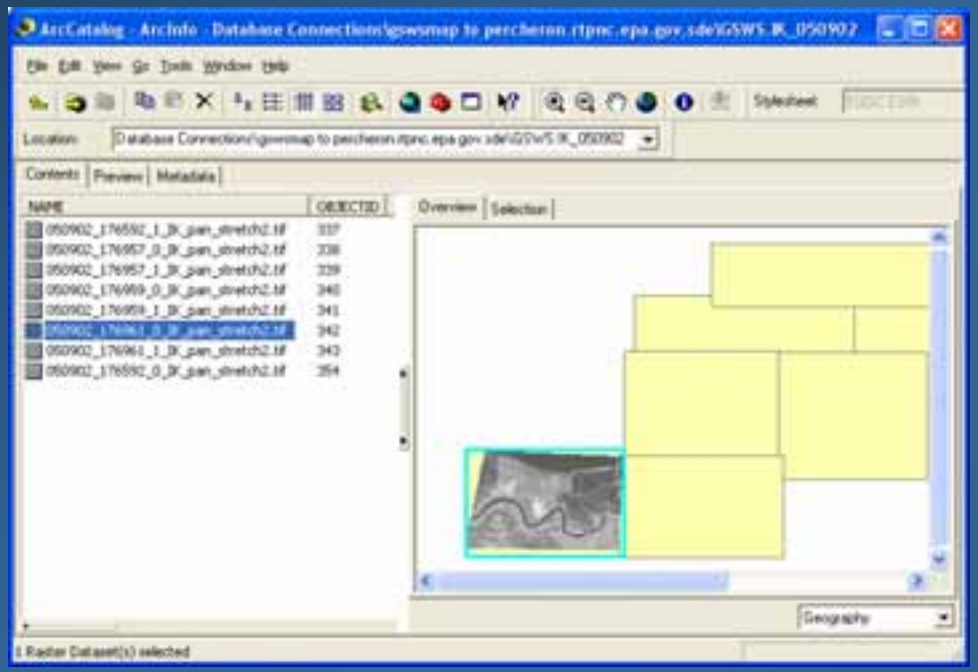


Geoprocessing: Stretching

- When viewing a standalone or SDE Raster with ArcCatalog or ArcMap, a 2 standard deviation stretch is applied by default
- When viewing an ArcIMS image service, no stretch can be applied
  - Raster must be saved with stretched values
- Stretch image in ArcMap, then use Reclassify from Spatial Analyst Toolbar (not ArcToolbox)
  - Current display lookup table loads by default
  - Can save table for manual editing or repeat use
  - Also useful for converting background values to nodata
  - Reduces file size by decreasing bit-depth

Geoprocessing: Mosaicking

- ArcSDE has two options for raster storage:



  - Single mosaicked raster
    - One set of pyramids, leading to faster display at smallest scales
  - Raster catalog
    - Eases management as individual rasters can be added and removed without affecting others

- Raster Catalogs are improving in 9.2!

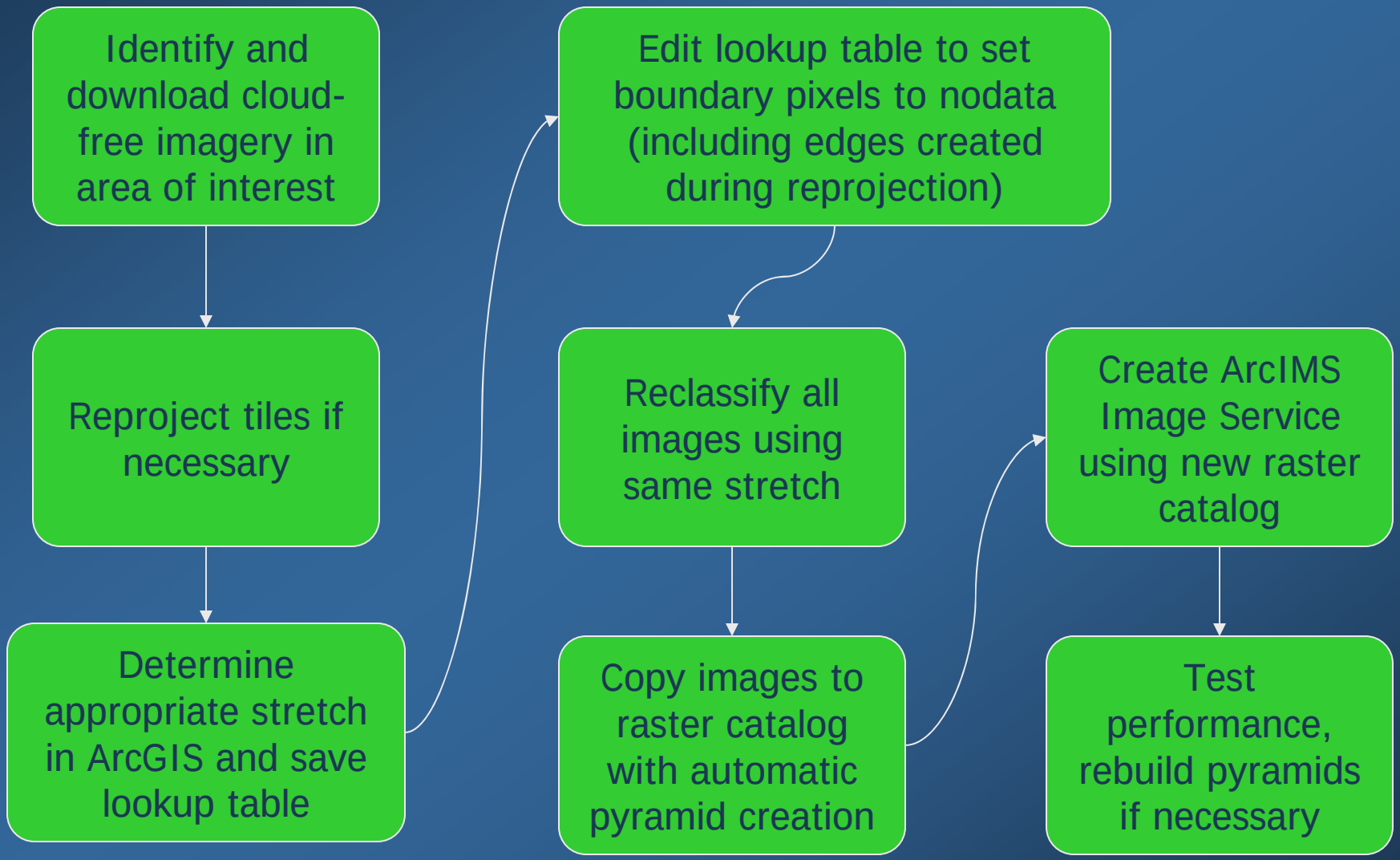
Geoprocessing: Pyramids

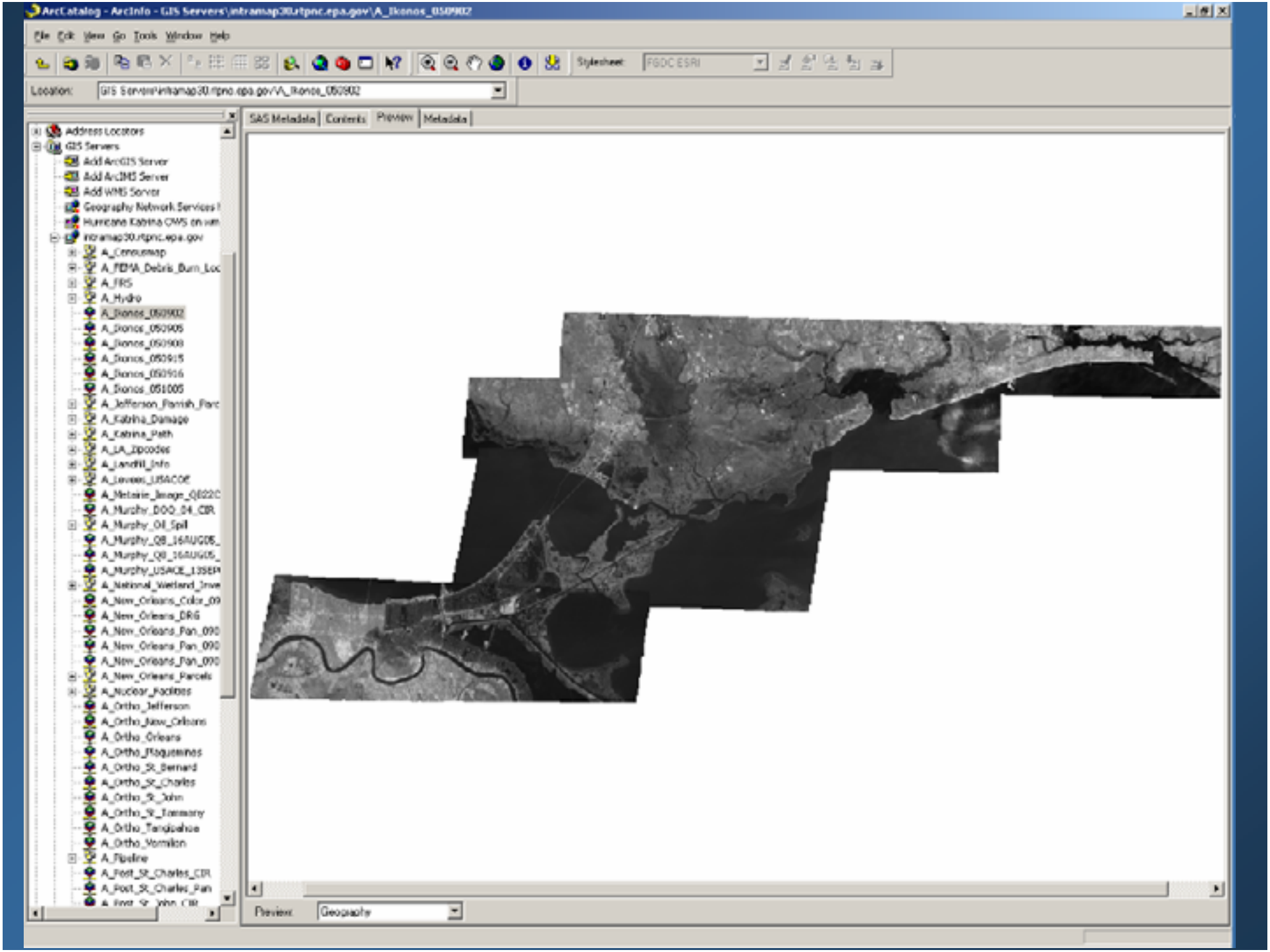
- Can be built automatically during copy/upload process
  - Specify in the geoprocessing environment settings
  - Pay attention to pyramid levels, not just resampling technique
- Pyramids will occasionally be corrupted – stripes or missing levels
  - Deleting and rebuilding won't always fix
  - Must choose different resampling technique

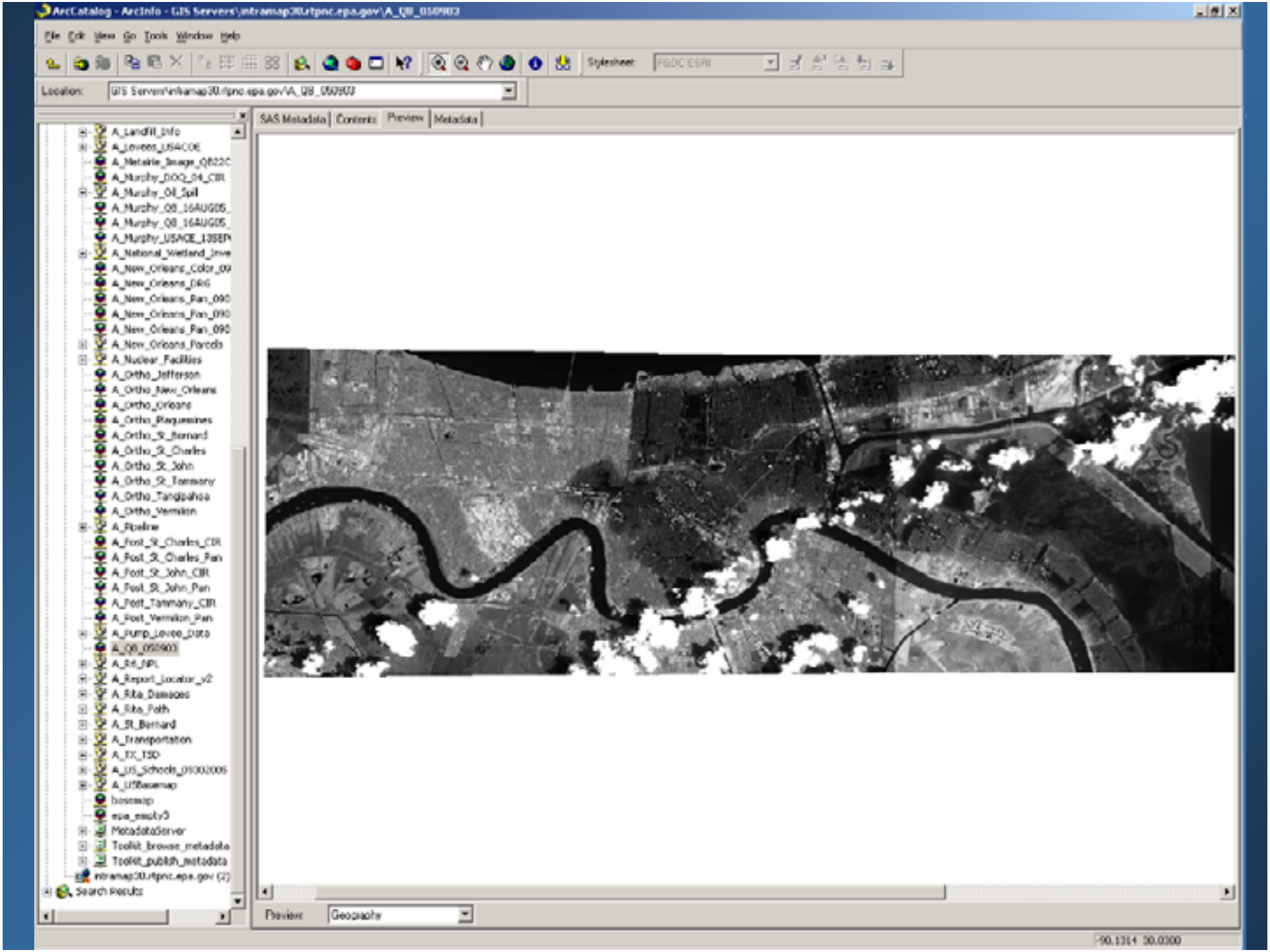
Geoprocessing: Batch Processes

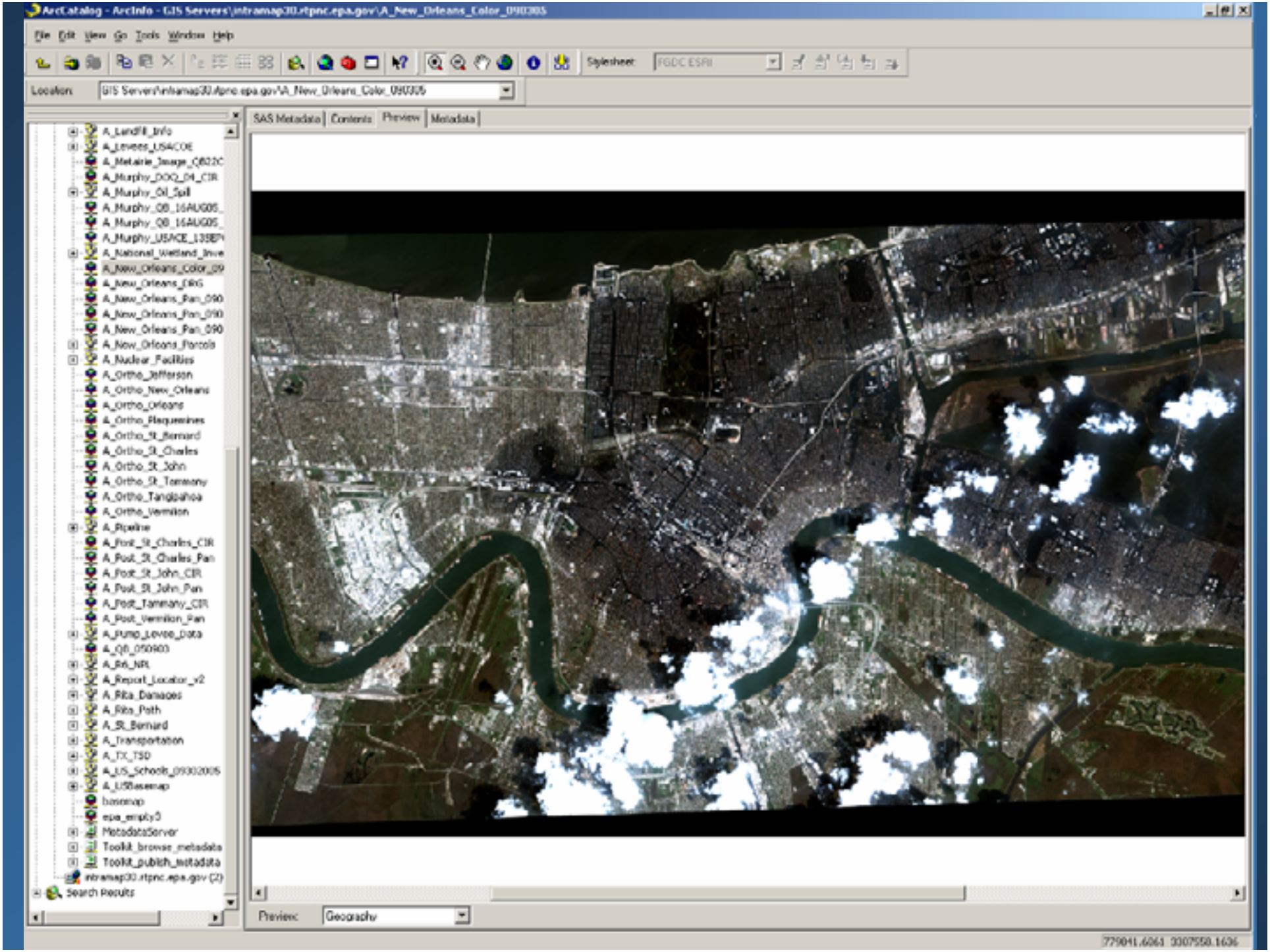
- Quick and dirty batch processing is easiest using the ArcCatalog Command Line
  - Copy and paste commands from text file
- ModelBuilder better for polished processes
  - Allows a nice GUI dialog box for setting input parameters
- When dealing with very large images in crisis response mode, quick and dirty is fine

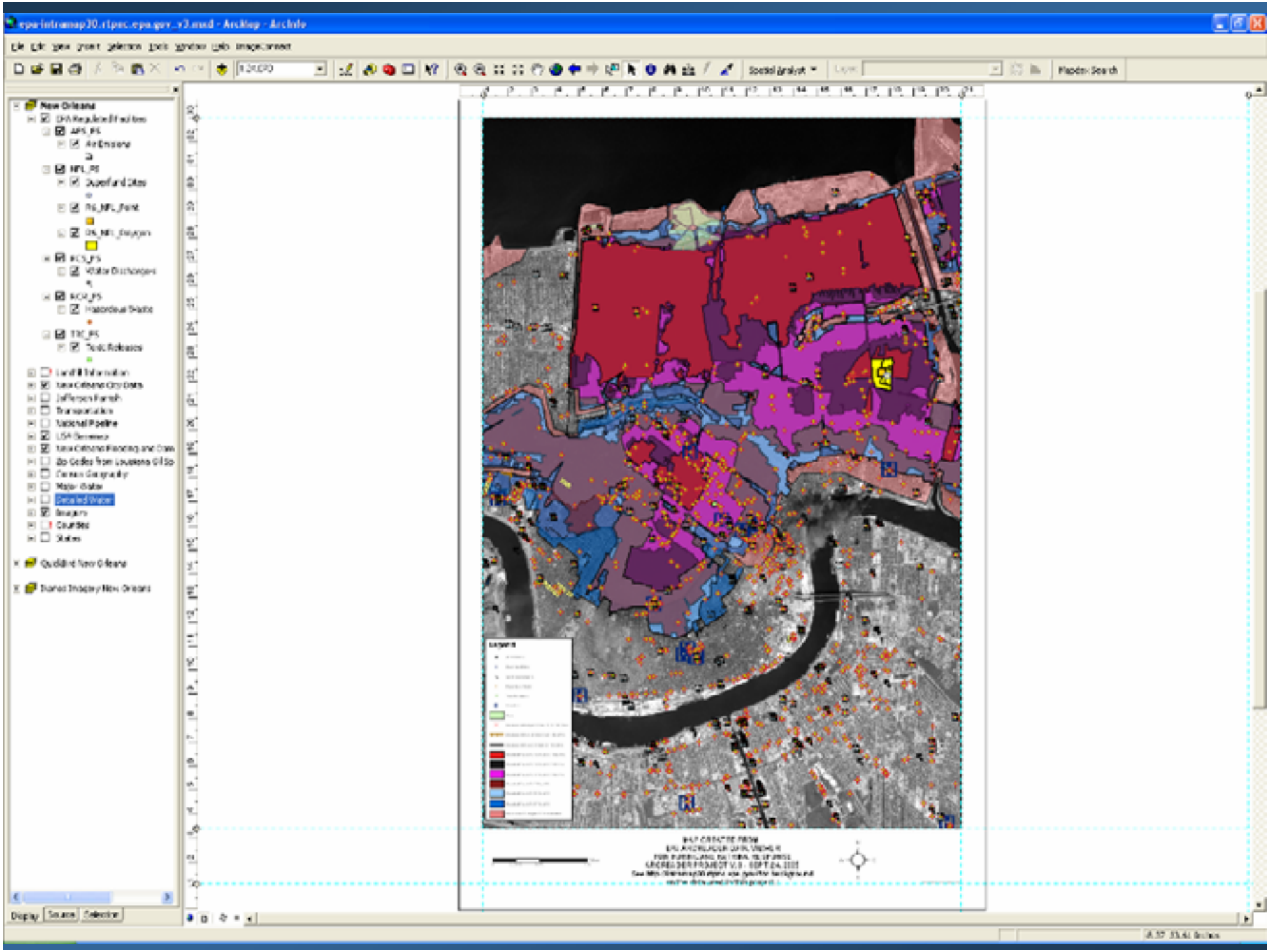
Workflow: Download to Service

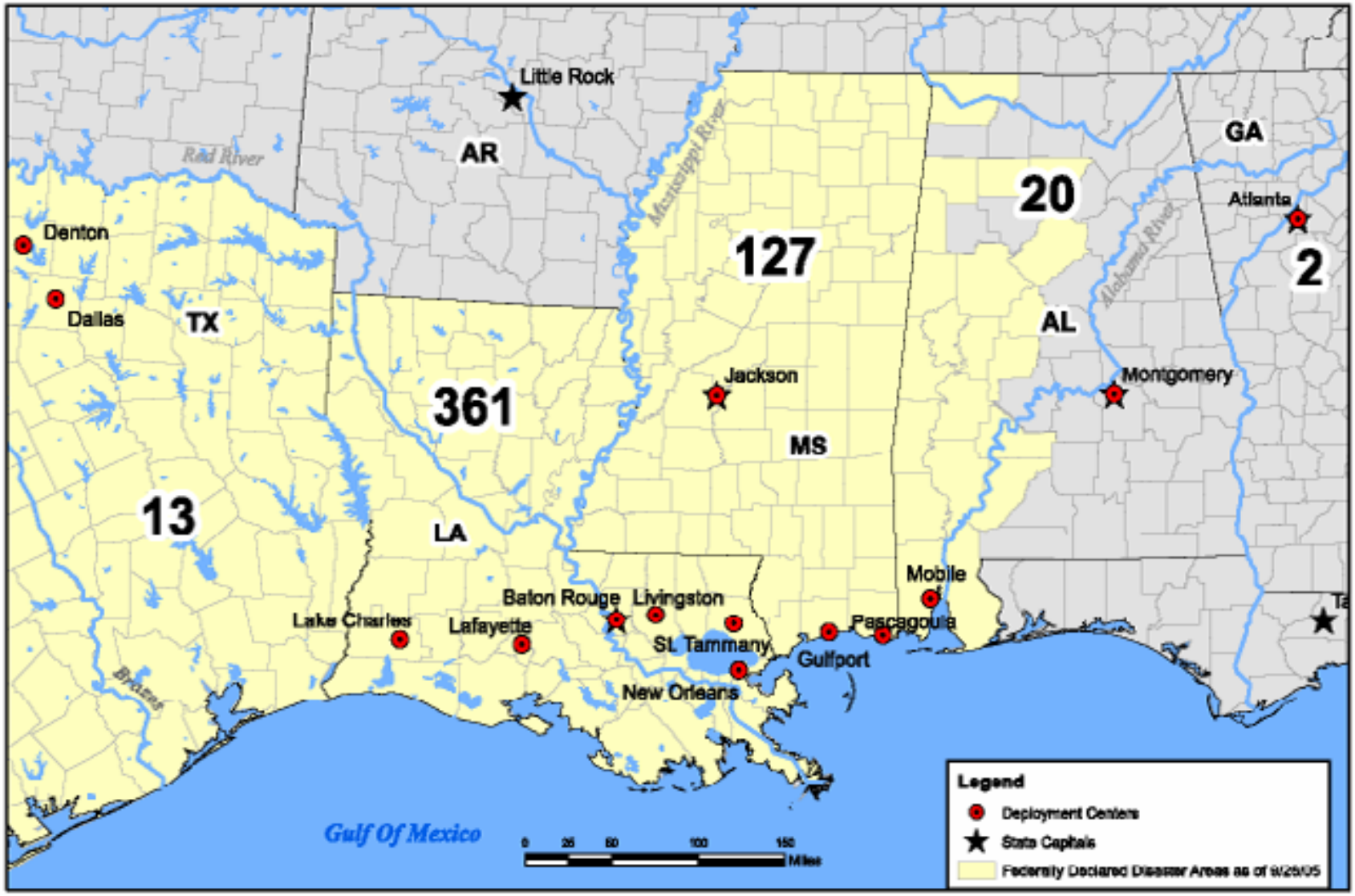



ArcCatalog - ArcInfo - GIS Servers\intramap30.rtpnc.epa.gov\A_Ikonos_050902
File Edit View Go Tools Window Help
Stylesheet:
FGDC ESRI
Location:
GIS Servers\intramap30.rtpnc.epa.gov\A_Ikonos_050902
Address Locators
GIS Servers
Add ArcGIS Server
Add ArcIMS Server
Add WMS Server
Geography Network Services
Hurricane Katrina OWS on vm
intramap30.rtpnc.epa.gov
A_Censusmap
A_FEMA_Debris_Burn_Loc
A_FRS
A_Hydro
A_Ikonos_050902
A_Ikonos_050905
A_Ikonos_050908
A_Ikonos_050915
A_Ikonos_050916
A_Ikonos_051005
A_Jefferson_Parrish_Parc
A_Katrina_Damage
A_Katrina_Path
A_LA_Zipcodes
A_Landfill_Info
A_Levees_USACOE
A_Metairie_Image_QB22C
A_Murphy_DOQ_04_CIR
A_Murphy_Oil_Spill
A_Murphy_QB_16AUG05_
A_Murphy_QB_16AUG05_
A_Murphy_USACE_13SEP
A_National_Wetland_Inve
A_New_Orleans_Color_09
A_New_Orleans_DRG
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Parcels
A_Nuclear_Facilities
A_Ortho_Jefferson
A_Ortho_New_Orleans
A_Ortho_Orleans
A_Ortho_Plaquemines
A_Ortho_St_Bernard
A_Ortho_St_Charles
A_Ortho_St_John
A_Ortho_St_Tammany
A_Ortho_Tangipahoa
A_Ortho_Vermilion
A_Pipeline
A_Post_St_Charles_CIR
A_Post_St_Charles_Pan
S&S Metadata
Contents
Preview
Metadata
Preview:
Geography


ArcCatalog - ArcInfo - GIS Servers\intramap30.rtpnc.epa.gov\A_QB_050903
File Edit View Go Tools Window Help
Stylesheet:
FGDC ESRI
Location:
GIS Servers\intramap30.rtpnc.epa.gov\A_QB_050903
SAS Metadata
Contents
Preview
Metadata
A_Landfill_Info
A_Levees_USACOE
A_Metairie_Image_QB22C
A_Murphy_DOQ_04_CIR
A_Murphy_Oil_Spill
A_Murphy_QB_16AUG05_
A_Murphy_QB_16AUG05_
A_Murphy_USACE_13SEP
A_National_Wetland_Inve
A_New_Orleans_Color_09
A_New_Orleans_ORG
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Parcels
A_Nuclear_Facilities
A_Ortho_Jefferson
A_Ortho_New_Orleans
A_Ortho_Orleans
A_Ortho_Plaquemines
A_Ortho_St_Bernard
A_Ortho_St_Charles
A_Ortho_St_John
A_Ortho_St_Tammany
A_Ortho_Tangipahoa
A_Ortho_Vermilon
A_Pipeline
A_Post_St_Charles_CIR
A_Post_St_Charles_Pan
A_Post_St_John_CIR
A_Post_St_John_Pan
A_Post_Tammany_CIR
A_Post_Vermilon_Pan
A_Pump_Levee_Data
A_QB_050903
A_RR_NPL
A_Report_Locator_v2
A_Rita_Damages
A_Rita_Path
A_St_Bernard
A_Transportation
A_TX_TSD
A_US_Schools_09302005
A_USBasemap
basemap
epa_study3
MetadataServer
Toolkit_browse_metadata
Toolkit_publish_metadata
intramap30.rtpnc.epa.gov (2)
Search Results
Preview:
Geography
-90.1314 30.0300

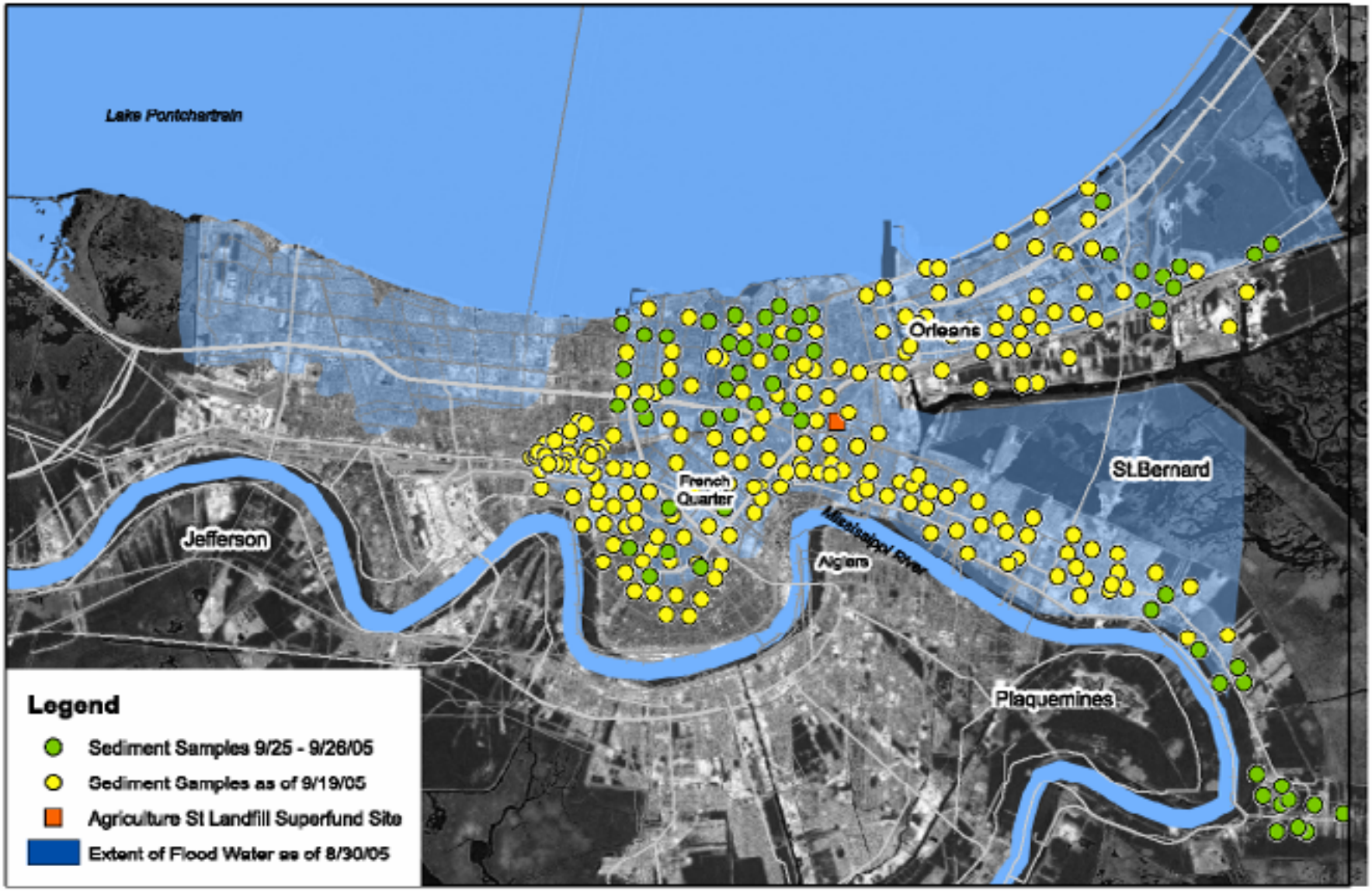

ArcCatalog - ArcInfo - GIS Servers\intramap30.rtpnc.epa.gov\A_New_Orleans_Color_090305
File Edit View Go Tools Window Help
Stylesheet
FGDC ESRI
Location:
GIS Servers\intramap30.rtpnc.epa.gov\A_New_Orleans_Color_090305
SAS Metadata
Contents
Preview
Metadata
A_Landfill_Info
A_Levees_USACOE
A_Metairie_Image_QB22C
A_Murphy_DOQ_04_CIR
A_Murphy_Oil_Spill
A_Murphy_QB_16AUG05_
A_Murphy_QB_16AUG05_
A_Murphy_USACE_13SEP
A_National_Wetland_Inve
A_New_Orleans_Color_09
A_New_Orleans_DRG
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Pan_090
A_New_Orleans_Parcels
A_Nuclear_Facilities
A_Ortho_Jefferson
A_Ortho_New_Orleans
A_Ortho_Orleans
A_Ortho_Plaquemines
A_Ortho_St_Bernard
A_Ortho_St_Charles
A_Ortho_St_John
A_Ortho_St_Tammany
A_Ortho_Tangipahoa
A_Ortho_Vermilion
A_Pipeline
A_Post_St_Charles_CIR
A_Post_St_Charles_Pan
A_Post_St_John_CIR
A_Post_St_John_Pan
A_Post_Tammany_CIR
A_Post_Vermilion_Pan
A_Pump_Levee_Data
A_QB_050903
A_R6_NPL
A_Report_Locator_v2
A_Rita_Damages
A_Rita_Path
A_St_Bernard
A_Transportation
A_TX_TSD
A_US_Schools_09302005
A_USBasemap
basemap
epa_empty3
MetadataServer
Toolkit_browse_metadata
Toolkit_publish_metadata
intramap30.rtpnc.epa.gov (2)
Search Results
Preview:
Geography
779941.6061 3307550.1636


epa-intramap30.rtpnc.epa.gov_v3.mxd - ArcMap - ArcInfo
New Orleans
Superfund Sites
Water Discharges
Hazardous Waste
Toxic Releases
Landfill Information
Transportation
Pipeline
Counties
States
Quickbird New Orleans
Display
Source
Selection



EPA Deployment For Hurricanes Katrina and Rita as of 9/28/05

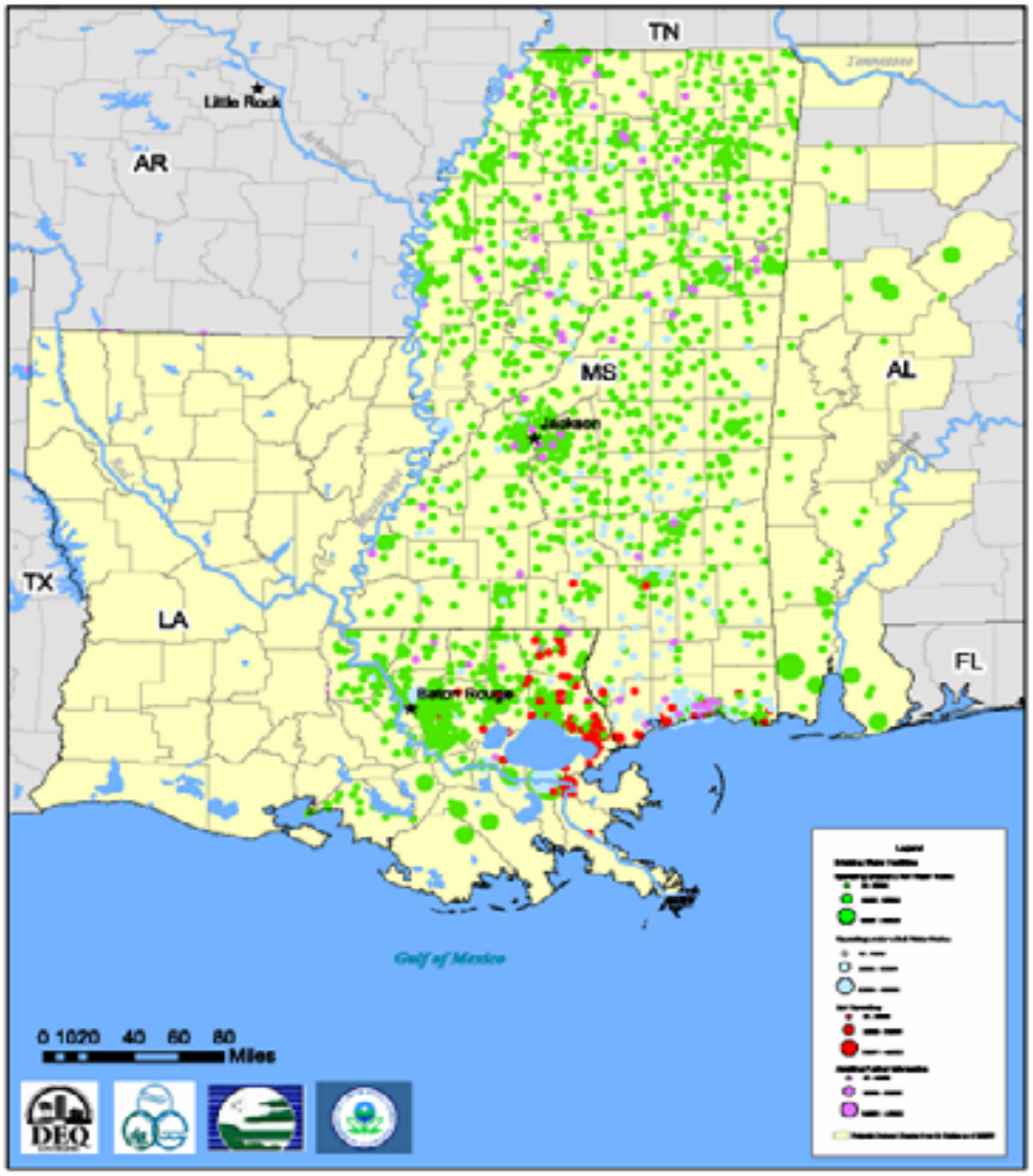
For Official Use Only
Source: EPA, ESRI



New Orleans Sediment Sampling as of 9/27/05

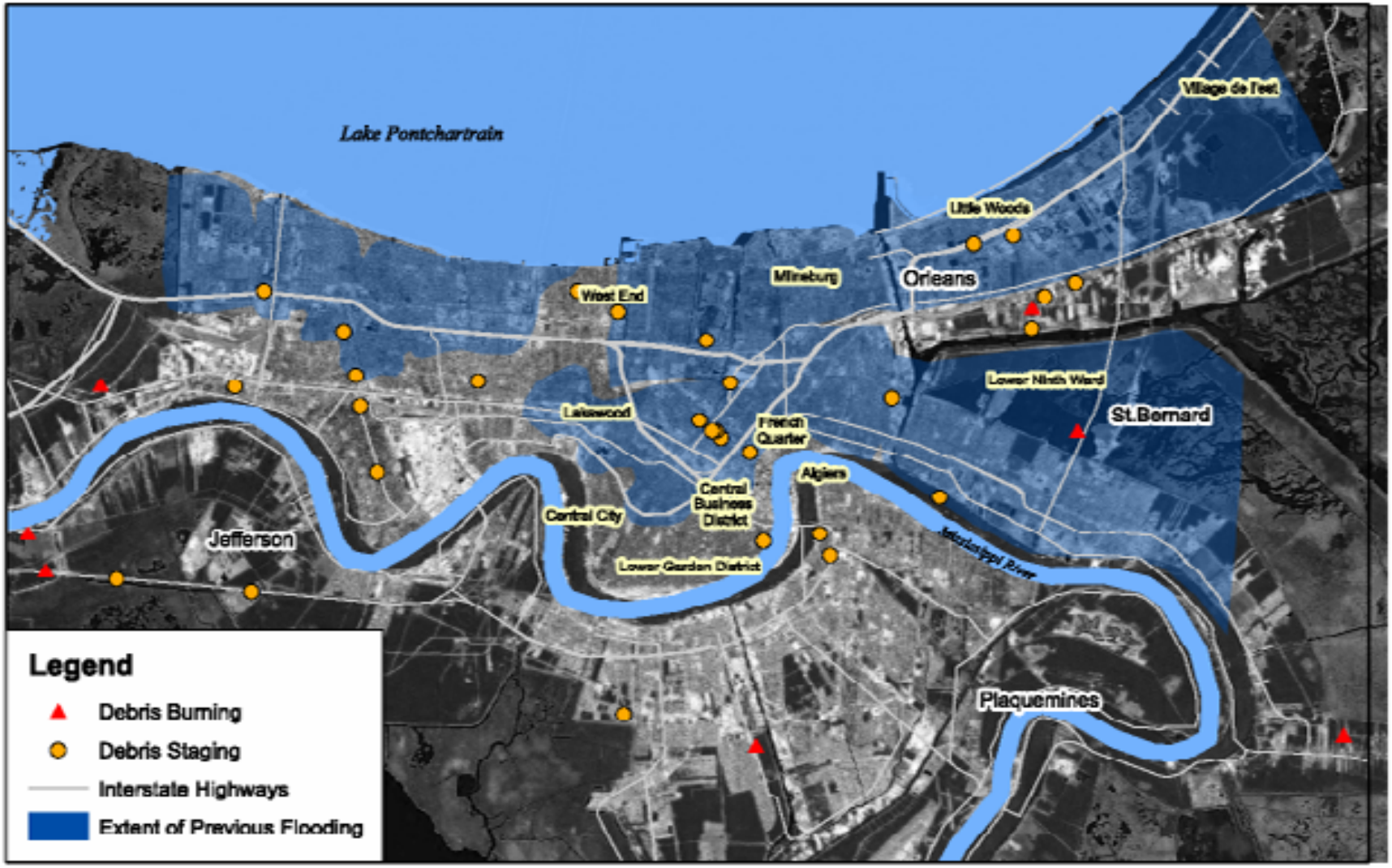
0 0.8 1.6 2.4 3.2 4 Miles

For Official Use Only
Source: EPA, NGA, USGS and ESRI Data



Drinking Water Facility Status
in Federally Declared Disaster Areas in MS and AL
and Impacted Areas in LA for Hurricane Katrina

Source: EPA, ESRI
For Official Use Only

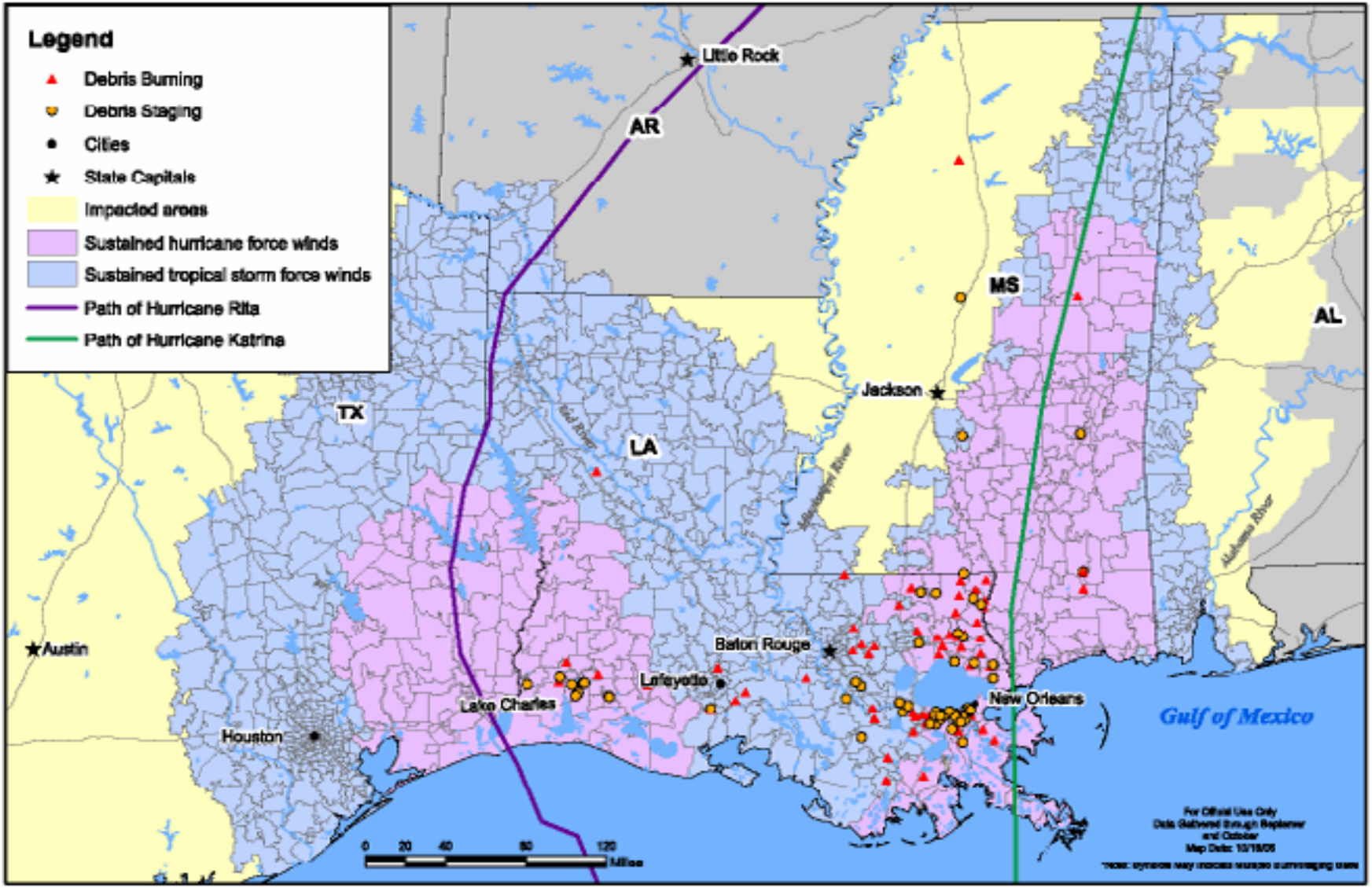


New Orleans Debris Burning and Staging Sites as of 10/14/05

N W E S

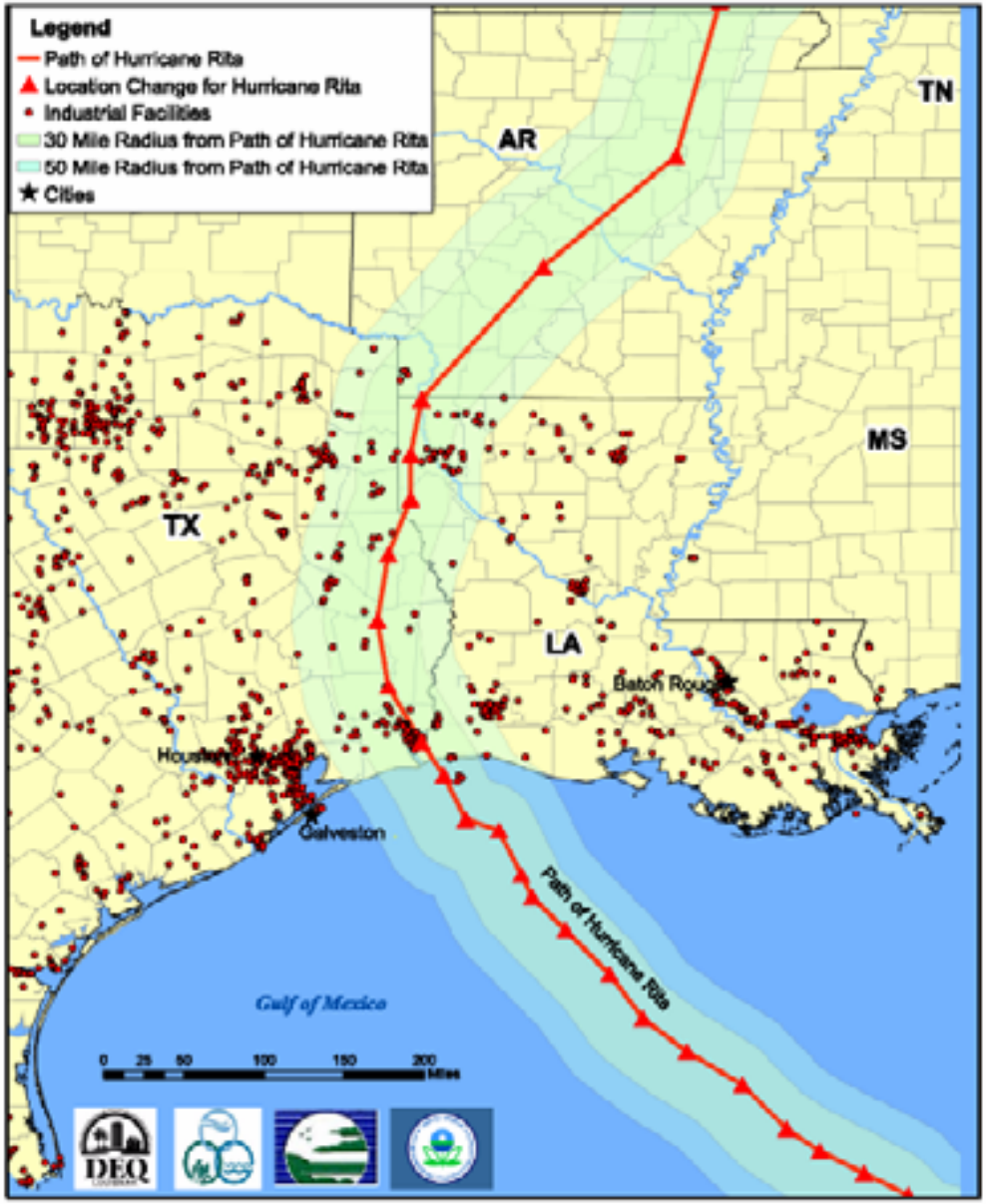
0 0.8 1.6 2.4 3.2 4 Miles

For Official Use Only
map date: 10/14/05



N
W E
S

Wind Speed by Zip Code and Debris Burning/Staging Sites in Federally DeclaredDisaster Areas for Hurricanes Katrina & Rita



Industrial Sites In Path of Hurricane Rita as of 9/27/05

For Official Use Only

ArcIMS Web Service Reality Check

- Network connectivity is unreliable in the field
- Responders prefer local copies of all data to live web services
- Feature services can be exported/clipped but image services cannot
- EPA HQ EOC utilized image services heavily
- Gulf coast responders preferred to download compressed imagery via ftp

Lessons Learned/Implemented

- Formalized process for EPA central GIS support to perform imagery triage and post compressed images on ftp site for download
- Added many intranet feature services, starting with GDT Streets, then census boundaries and EPA facilities
- Took the time to tune and optimize the performance of the services
- Set up ESRI Data Delivery Extension for large extracts

Performance Optimization

- Optimize the OS system of the database server
- Optimize the RDMS (Oracle or SQL server)
- Optimize the ArcSDE server environment
  - The dbtune table
  - The giomgr.defs
- Optimize the individual SDE data layers
  - The Spatial Index
    - Oracle Spatial and IBM DB2 Spatial extender
    - ArcSDE Ge0database
  - The Attribute Index
  - Minimize on-the-fly operation with big dataset
- Optimize the SDE clients

Optimize the ArcSDE Server Environment

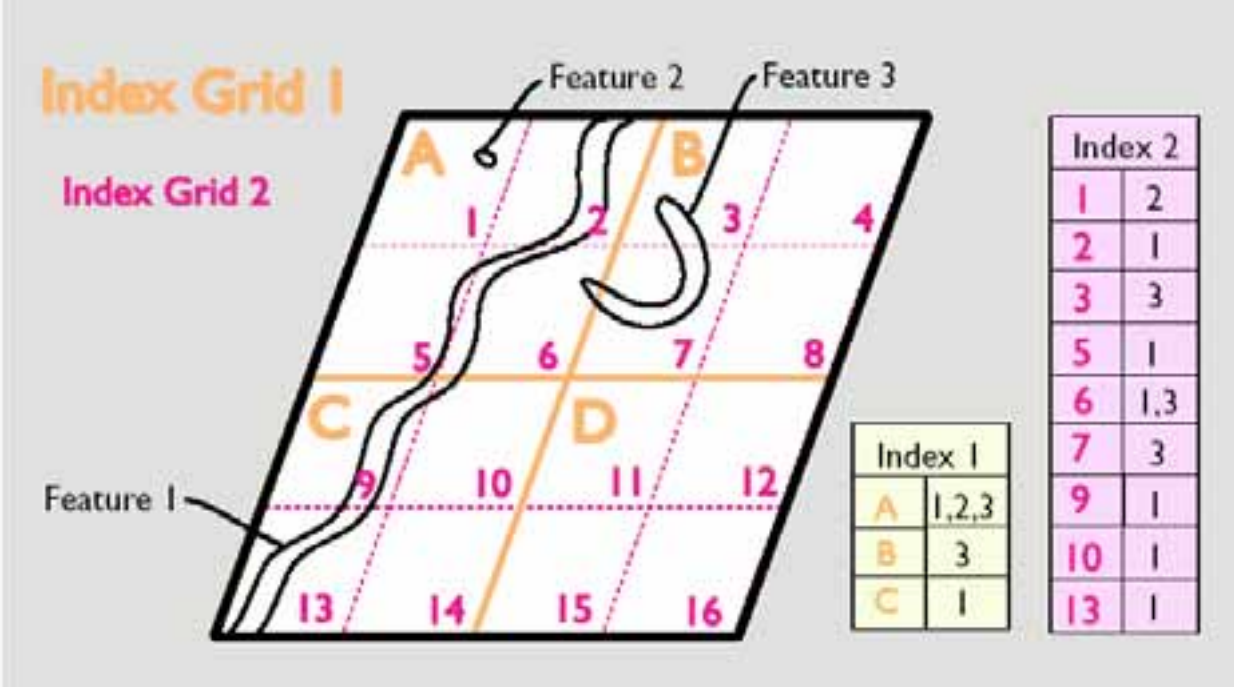
- The “dbtune” table
  - Control how ArcSDE clients create objects.
  - Determine which tablespace a table or index is created in.
  - Define the extent size of the data objects
  - Specify 1 of 4 available storage formats
  - sdedbtune command to export / import dbtune table
    - Sdedbtune –o export – export the current dbtune table into a text file format
    - Edit the text file
    - Sdedbtune –o import – re-load the update dbtune file into the database table
- The “giomgr.defs” file
  - Maximum SDE connection
  - TCPKEEPALIVE

Optimize the Spatial Index

- What is the spatial index?
- How do you know if the spatial index need optimized?
- How to change the spatial index?
  - ArcSDE command line
  - ArcGIS desktop GUI

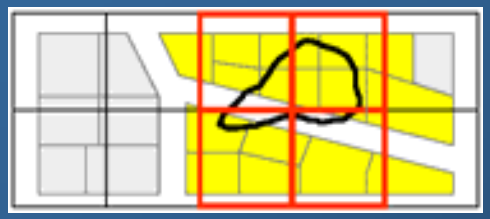
What Is The Spatial Index?

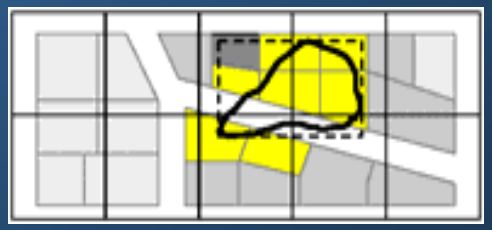
ArcSDE builds a spatial index by applying a grid to the feature class. It records features that fall within each grid cell in an index table (the S table of the feature class schema).



How Do Spatial Index Grids Work?

- Client submits a spatially-constrained query
  "retrieve all land parcels that overlap a ruptured storage tank's contamination plume"
- Server determines which parcels share an index grid with the plume
- Server finds shapes with overlapping envelopes
- Feature shapes compared to plume shape
- The key: only look at the data that meet your query criteria and ignore all other data





How To Check The Spatial Index Grid?

- ArcSDE command: sdelayer –o si_stats

```
Layer 5 Spatial Index Statistics:
Level 1,    Grid Size 0.1
|-------------------------------------------------------------------------|
| Grid Records: 212781                                                    |
| Feature Records: 199991                                                 |
| Grids/Feature Ratio:  1.06                                              |
| Avg. Features per Grid: 196.47                                          |
| Max. Features per Grid: 3887                                            |
| % of Features Wholly Inside 1 Grid: 93.96                               |
|-------------------------------------------------------------------------|
|                Spatial Index Record Count By Group                      |
| Grids:       <=4     >4     >10     >25     >50     >100    >250    >500 |
|---------- ------ ------ ------ ------ ------ ------ ------ ------ |
| Features: 199991       0                                0       0       0 |
| % Total:     100%      0%                               0%      0%      0%|
|-------------------------------------------------------------------------|
```

This value should be < 2. If it exceeds 4, the spatial index need to be modified

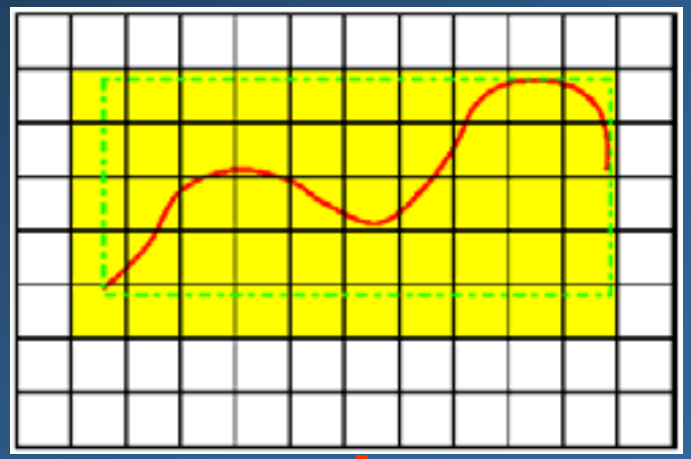
This value should be > 100 & < 200

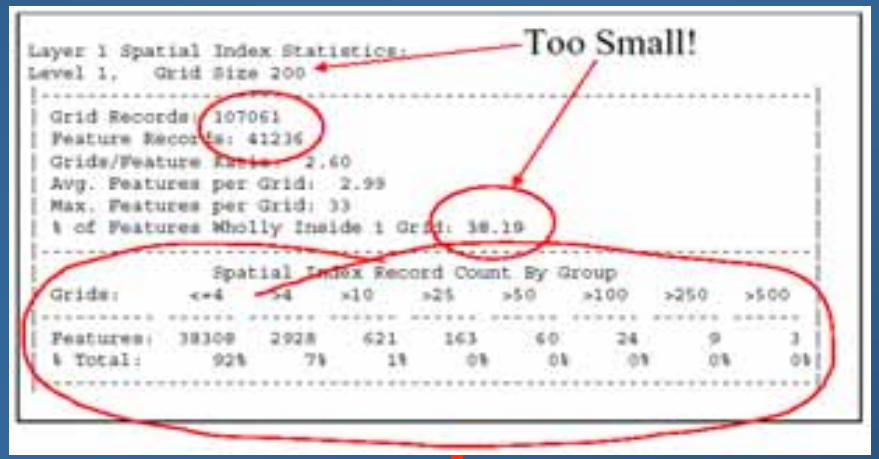
This value should be < 4000

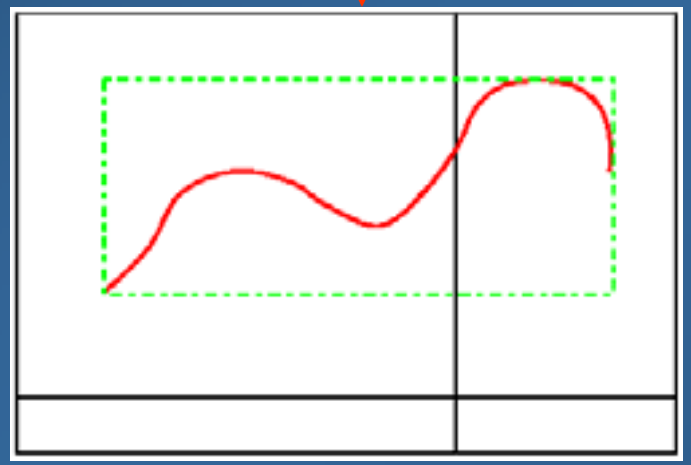
This value should be > 80

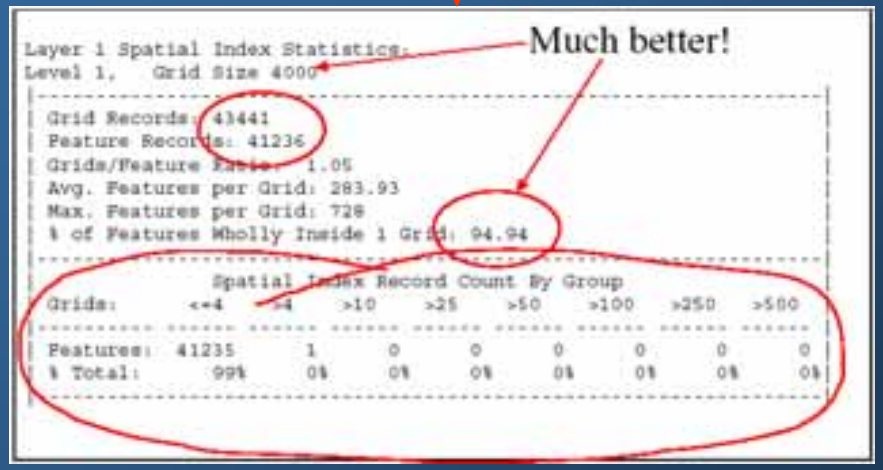
- The majority of the features should have fewer than four records.

Examples of Improving The Spatial Index







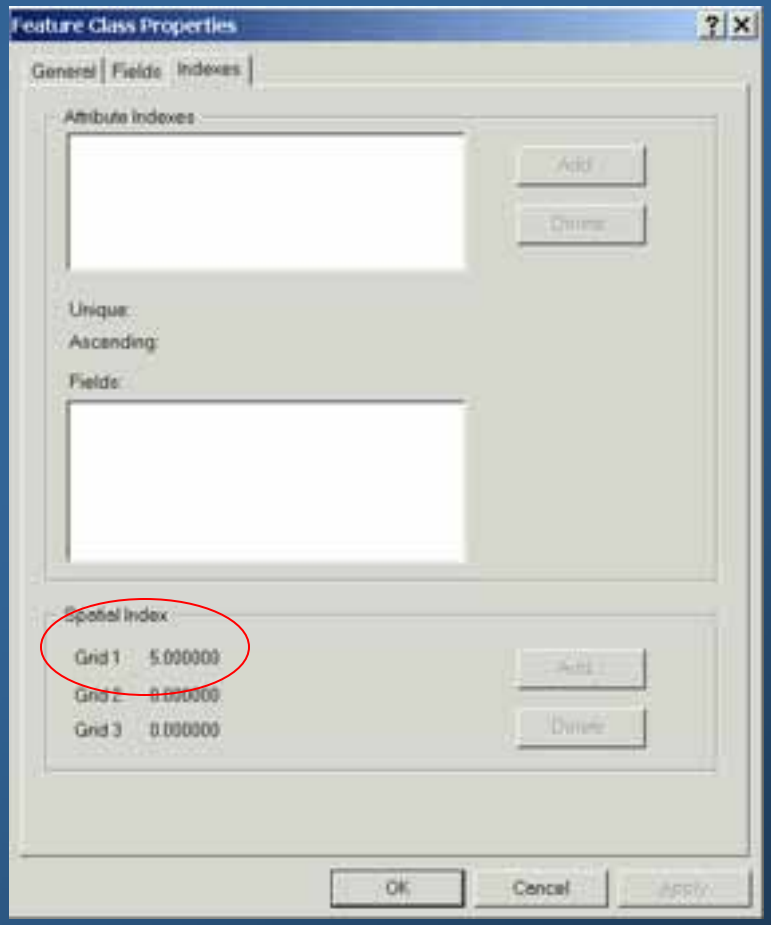


How To Change The Spatial Index Grid?

ArcSDE Command Line:

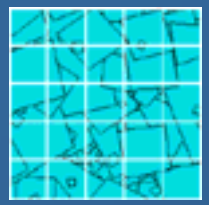
- Use "sdelayer -o load_only_io" to drop the spatial index. No spatial queries or data loading is allowed on this layer while it is in load-only mode.
- Use "sdelayer -o alter -g n,n,n" to specify new grid sizes. (Specify zero for the second or third grid size if not used.)
- Use "sdelayer -o normal_io" to rebuild the spatial index and make the layer accessible again.

ArcCatalog GUI



Spatial Index Recommendations

- Experiment with grid settings. Can change grid size any time.
- Start with 3X to the LARGER of:
  - Your layers average feature size or
  - Your average querying area
- Never make grids smaller than your features
- Most of the time **one grid level** is enough
- Keep average number of features per grid between 100 and 300
- Try not to let the max number of features per grid exceed 2000 to 4000
- Maintain a high % of features wholly in 1 grid
  - >90%=great; >80%=good; <80%=change
- Tuning grids to a “general” setting works well for most queries
- Automation of the optimization routine



1x of average feature extent.

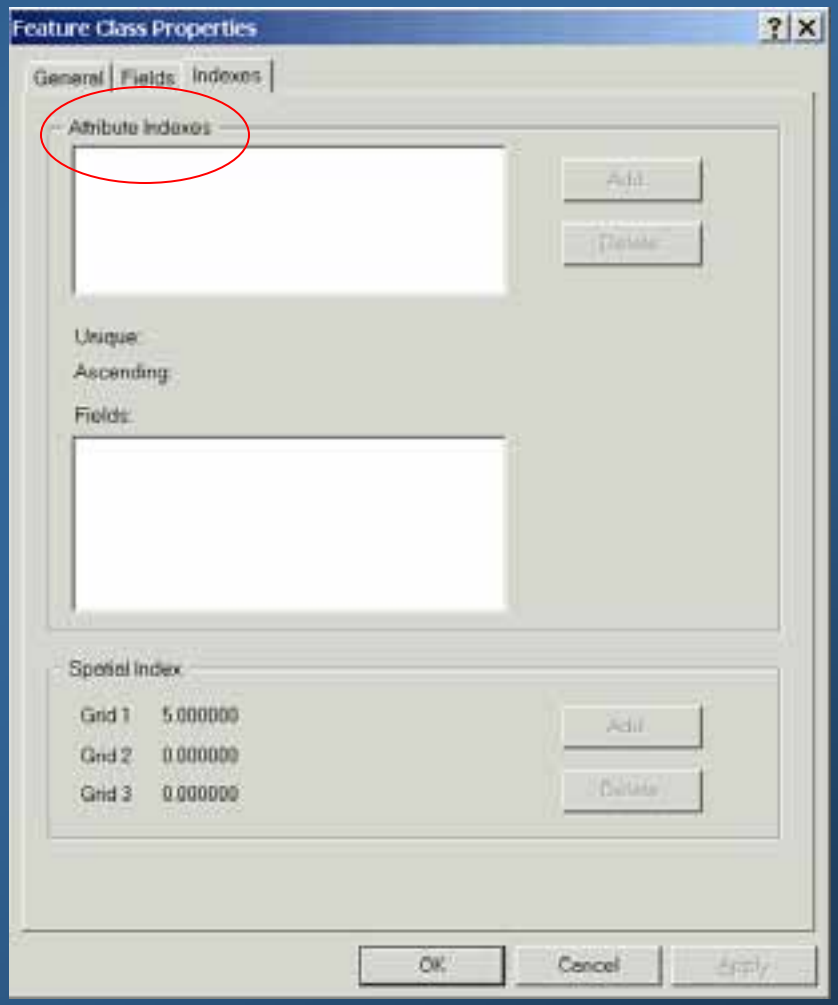
3x of average feature extent.

What is the Attribute Index?

- Two columns: indexed field and key
- Index table has the same number of records as the attribute table.
- The indexed field will be sorted with different indexing technology
- The index table could be very big

How to Build Attribute Indexes

- ArcSDE Command Line:
  - Use "sdetable“ SDE command to drop the attribute index.
  - Use "sdetable“ SDE command to build new attribute index
- ArcGIS Toolbox Wizard Interface
- ArcCatalog GUI
- Oracle Tools



Attribute Index Recommendations

- Index attribute fields you query often on
- Indexing both the “where” and “select” columns helps business table queries in the same way as it helps the S table queries.
- Different indexing technologies
  - unique; bitmap; unsorted; reverse
- Check application logic
  - Avoid unnecessary multiple criteria in “where” clause

Minimize On-The-Fly Operations

- The name system of the GDT street data contains 3 separated columns:
  - Prefix: e.g. South; Name: e.g. Miami; Street type: e.g. Drive
- To use the full name (South Miami Drive) for labeling, you have to join these 3 columns on-the-fly.
- To avoid on-the-fly operations, you can add a new attribute column to the database that contains all three parts of the name.
- If the dataset is small, the improvement may not be obvious. If the dataset is big, the improvement could be quite significant.

Performance Comparison Tests

- The same database server (SDE/Oracle)
- Same dataset: GDT street data
- Same SDE client: ArcGIS Desktop with connection to the SDE/Oracle database directly
- The before optimization test:
  - Default ESRI spatial index
  - No attribute index
  - No new attribute columns with all 3 name components
- The after optimization test:
  - Rebuilt the spatial index 8 times
  - Build attribute indexes for the FCC code and name label columns
  - Add a new attribute columns with all 3 name components

Results: Before vs After Optimization

- On average, the “after optimization” test is **30% faster** than the “before optimization” test
- In several cases, the performance improvement is over 50%
- For desktop operation, the improvement is quite obvious with big dataset
- For web application, the improvement is quite obvious even with small and mid size dataset

Web Service Considerations

- ArcIMS Image/Feature Services
  - Setting up proper number of spatial servers
  - Setting up proper number of instances
- ArcIMS OGC WMS/WFS connectors
- XML/SOAP based Geo Web Services
- Other Considerations

Conclusion/Summary

- EPA GIS Emergency Response is a team effort supporting the field personnel and decision-makers
- Imagery triage is a perfect task to offload onto centralized support
- ESRI geoprocessing and server tools met the challenge
- Taking stock of lessons learned is a fruitful interagency exercise that should help us be better prepared for the next crisis

Acknowledgements

RTP GIS Support Team

GIS/Remote Sensing Project Manager:
Dick Dulaney

GIS Analysts:
Mark S. Murphy
Catherine Harness
Binfei Chen

Senior System Administrator:
Janet Braxton

Customer Support Representative:
Claudia Benesch

EPA GIS Personnel

Geospatial Information Officer:
Brenda Smith

National Hosting Operations Coordinator:
Tim Richards

Region 2 GIS Coordinator:
Harvey Simon

Region 10 GIS Coordinator:
Scott Augustine

Contact Information

Qi Dai
GIS Technical Lead
dai.qi@epa.gov
919-767-7285

Torrin Hultgren
GIS Analyst
hultgren.torrin@epa.gov
919-767-7312

79 T.W. Alexander Drive, Building 4201
Research Triangle Park, NC 27709