

Big Data and Decision Making – In Practice

Al Karlin, Ph.D. GISP

Mapping & GIS

Southwest Florida Water Management District



Big Data at the SWFWMD

- WMIS -** Watershed Management Information System
- NexRAD Rainfall Data
 - ePermitting Data
- GWIS -** Geographic Watershed Information System
- ArchHydro-Surface Water (GeoICPR)
 - ArchHydro-Groundwater
- TIPS -** Topographic Information & Publication System
- Light Detection and Ranging (LiDAR)
 - Sound Navigation and Ranging (Sonar)
 - Digital Elevation Models (DEMs)
- LABINS-** Natural Color and False Color-InfraRed Imagery

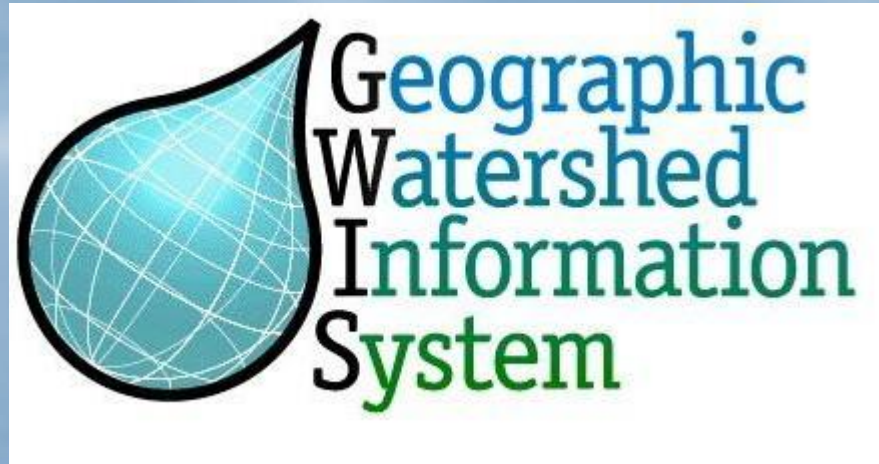
Big Data at the SWFWMD



WMIS contains different types of scientific data, including hydrological (groundwater and surface-water levels), atmospheric (rainfall and evapotranspiration), and water quality (e.g. nutrients, pesticides, organics, metals) measurements. The data available in WMIS represent District-collected values only and account for about 10TB of data.

The district maintains approximately 2300 groundwater monitoring wells (water quantity) and approximately 200 water quality stations monitored through the SCADA system

Big Data at the SWFWMD

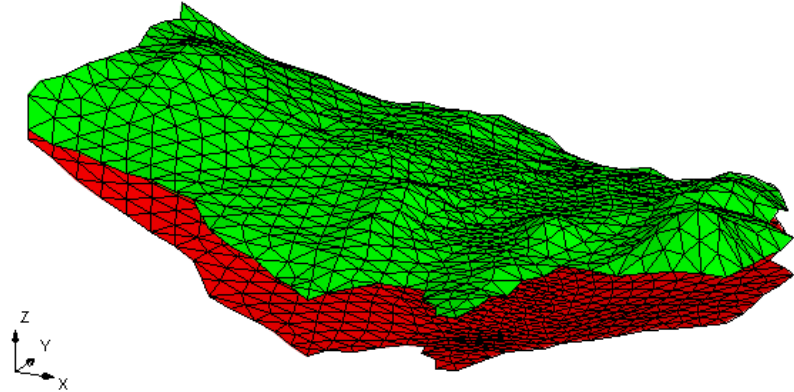


GWIS contains the surface ArchHydro model data that is used as the basis for Hydrological and Hydraulic modeling. The data are maintained for each of the District's 277 Watersheds and 370 Planning Units. The ArchHydro/GWIS database accounts for approximately 1.5 TB of data with the H&H model data making up another 8 TB of data.

The District is currently in the process of assembling the Regional Observation and Monitoring well Program (ROMP) physical data into an ArchHydr-Groundwater database. Preliminary estimates indicate that these data, and their derived ArchHydro products will account for another 2 TB of data.

Big Data at the SWFWMD

Topographic
Information
&
Publication
System



TIPS contains the LiDAR and Bathymetric/Topographic data for the District. The data are distributed into the ArcGIS environment either as Digital Elevation Models (DEM) or as Laser Range Files (LAS) through custom-build toolbars. The DEMs are offered as either tiled elevation models, county-wide DEMs, and/or as Hillshades. Derived products include, for most areas, LiDAR-derived 1' and 2' contours.

Data distributed to the public are delivered in a compressed LAS format (zLAS or LAZ). The TIPS dataset accounts for approximately 6 TB of data (and is rapidly growing)!

Big Data at the SWFWMD

Aerial Imagery



The District has been collecting aerial imagery since the early 1970. Categories include:

- Ortho-imagery (RGB & fCIR)
- Flood Imagery
- Seagrass Imagery

Internally data are accessed through custom toolbars; Public access is through the Land Boundary Information System (LABINS/ www.LABINS.org)

This database contains ~ 4 TB of MrSID compressed Imagery

SWFWMD “Big Data” issues

Perhaps our biggest issues are:

- Staff (engineers, hydrologists, environmental scientists) don't have the basic IT skills (programming in R/SAS/SQL/Python, etc.) to easily access and use the tools
- Staff don't have a good understanding of what data are available to them,
- We also don't have all of the data well organized.



“Big Data” Challenges for the SWFWMD:

- Modeling data
- Time series data collected by contractors and/or staff working outside of our Data Collection Bureau
- A lot of our unstructured content is not in repositories or not well organized in our document management system
- Land Management/Acquisition – related information.
- Hydrogeological data (well logs, etc.)
- Procurement/contractual documents.

