

GROUNDWATER SECURITY

DELIVERING **BIG DATA** AND THE CHANGING LANDSCAPE OF WEB-BASED **TECHNOLOGIES**

Sachin Shah

Hydrogeologist, International Lead

USGS Geospatial Science + Cyber Innovation Branch



GROUNDWATER SECURITY: BIG DATA AND WEB TECHNOLOGIES

GROUNDWATER SECURITY



BIG DATA

↑ VOLUME

Millions of data points

↑ VARIETY

More than just groundwater data

↑ VELOCITY

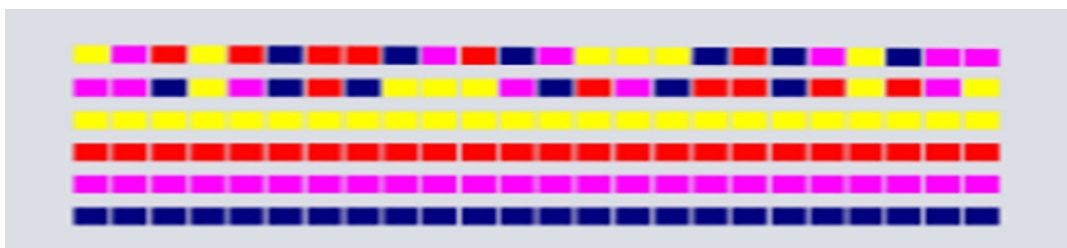
Sense of urgency

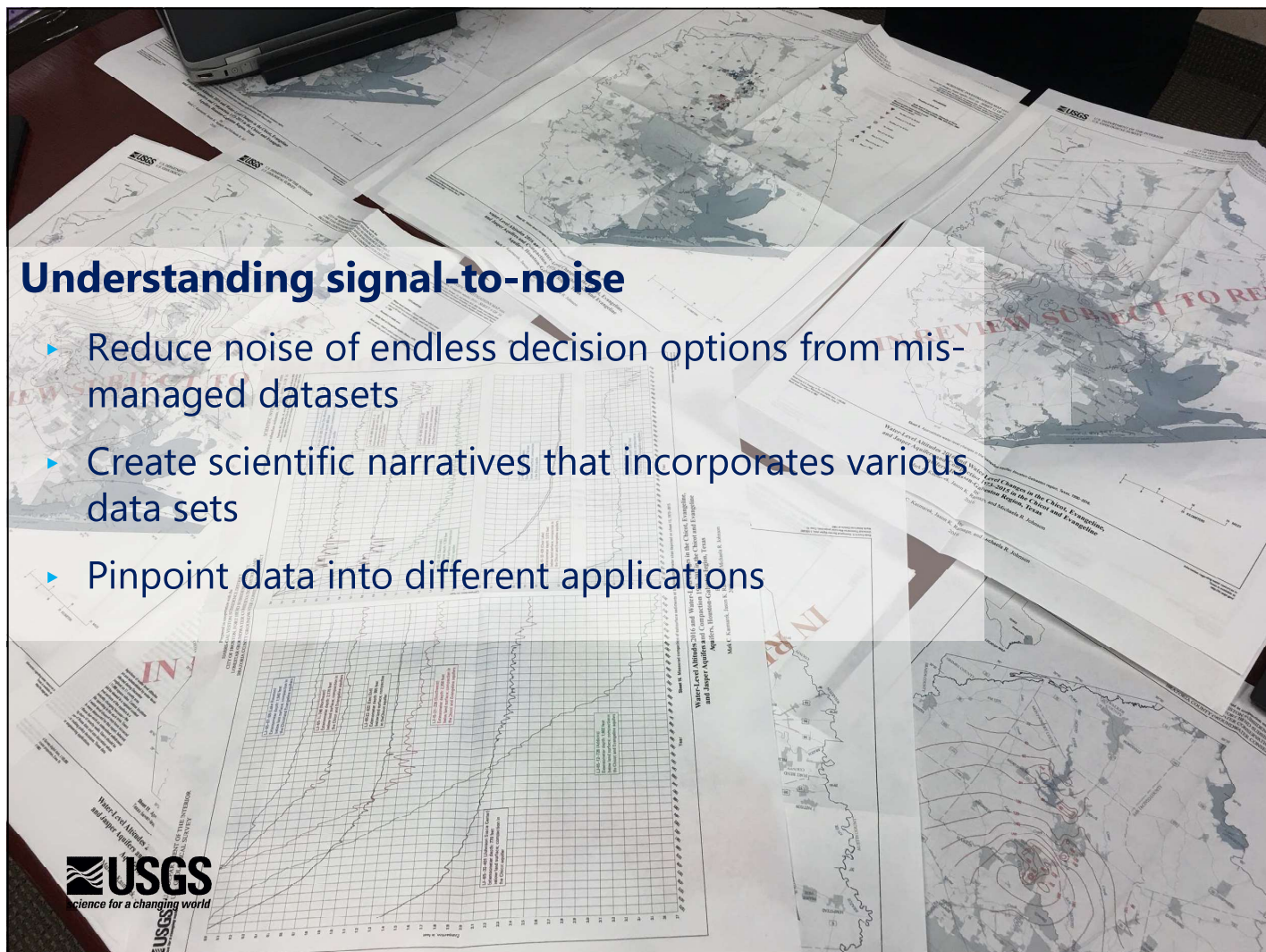


GROUNDWATER DATA → DECISION SUPPORT

Can these applications help meet development objectives?

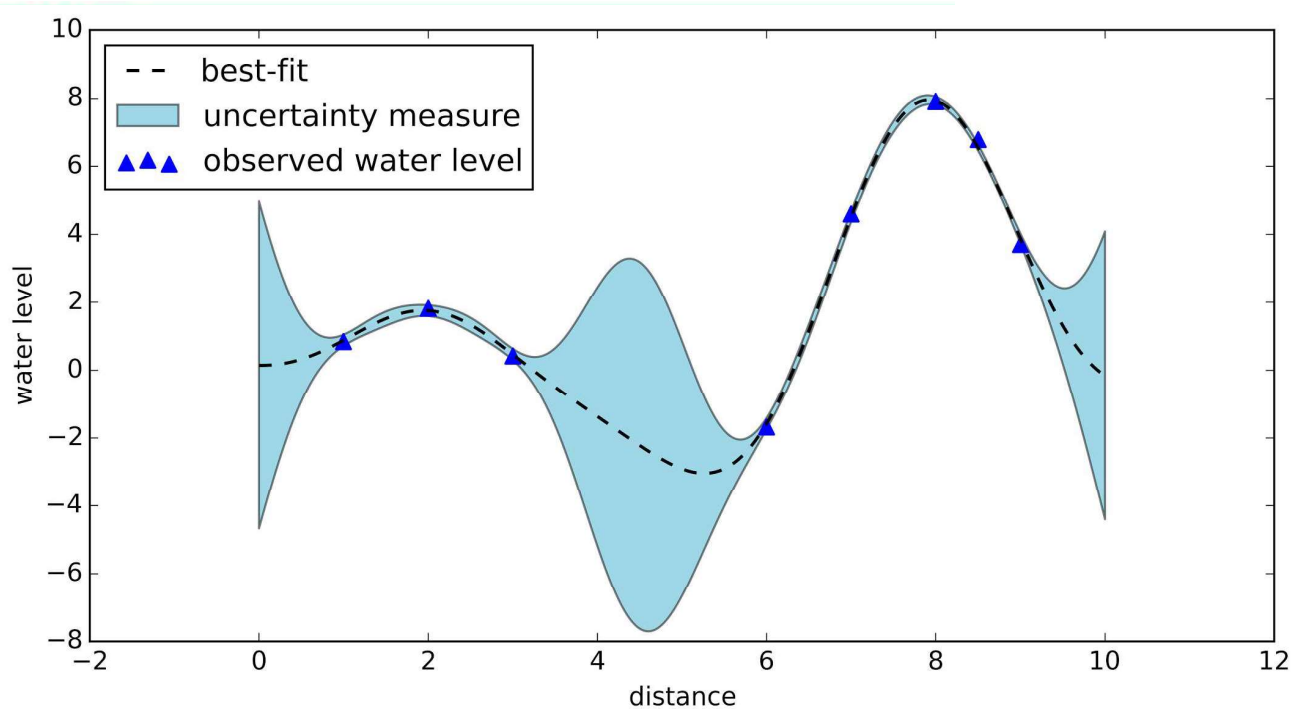
BIG DATA DECISIONS ANALYSIS



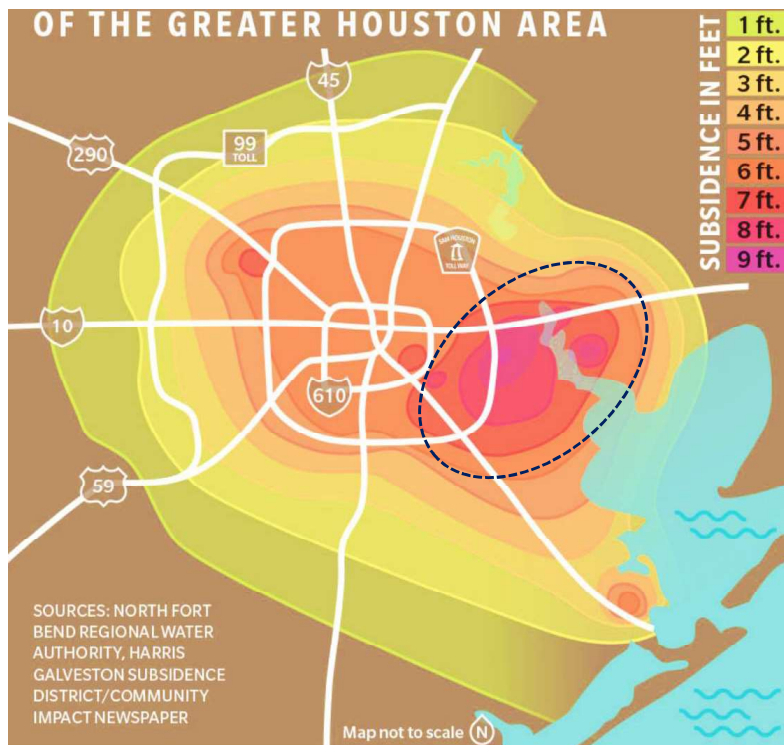


GROUNDWATER SECURITY: BIG DATA AND WEB TECHNOLOGIES

NOISE REDUCTION

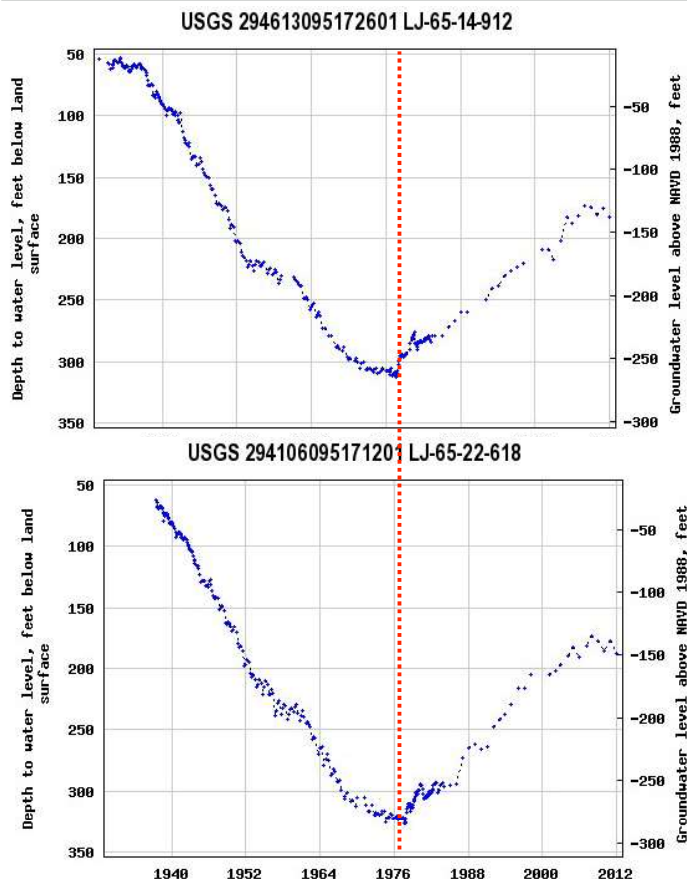


HOUSTON, TEXAS, USA GULF COAST REGION



- ▶ 3 meters of subsidence in southeast Houston
- ▶ seawater intrusion
- ▶ flooding

HOUSTON
population 3 million

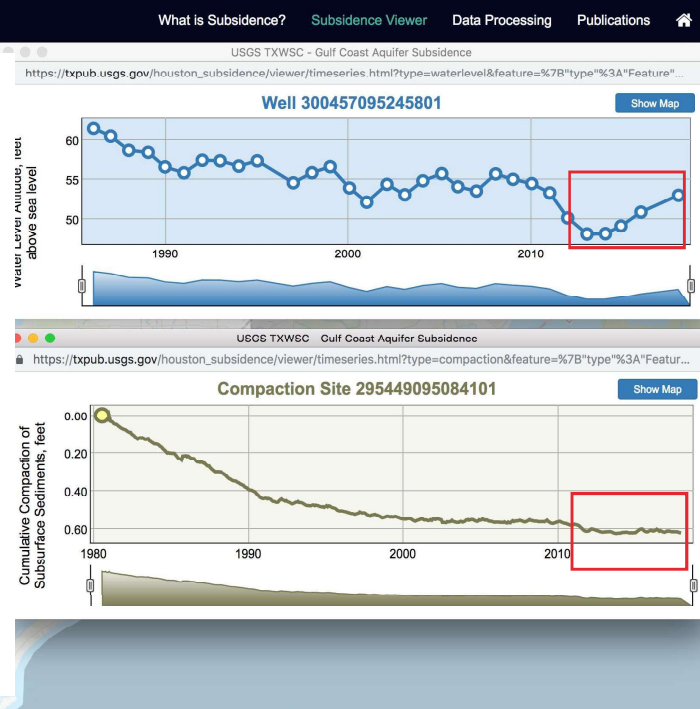


- ▶ Subsidence agency established in 1975
- ▶ USGS data collection started in 1976
- ▶ 750 public water supply wells measured annually
- ▶ 3,000 data points generated every year

SUBSIDENCE APPLICATION → GROUNDWATER POLICY



- ▶ Over 42 years of groundwater and subsidence data
- ▶ Comparisons of groundwater levels versus compaction of areas in Houston, Texas, USA.
- ▶ Changes in regulations and management over time
- ▶ Change in land-surface elevation



SUBSIDENCE APPLICATION → GROUNDWATER POLICY



SDG6 Indicators

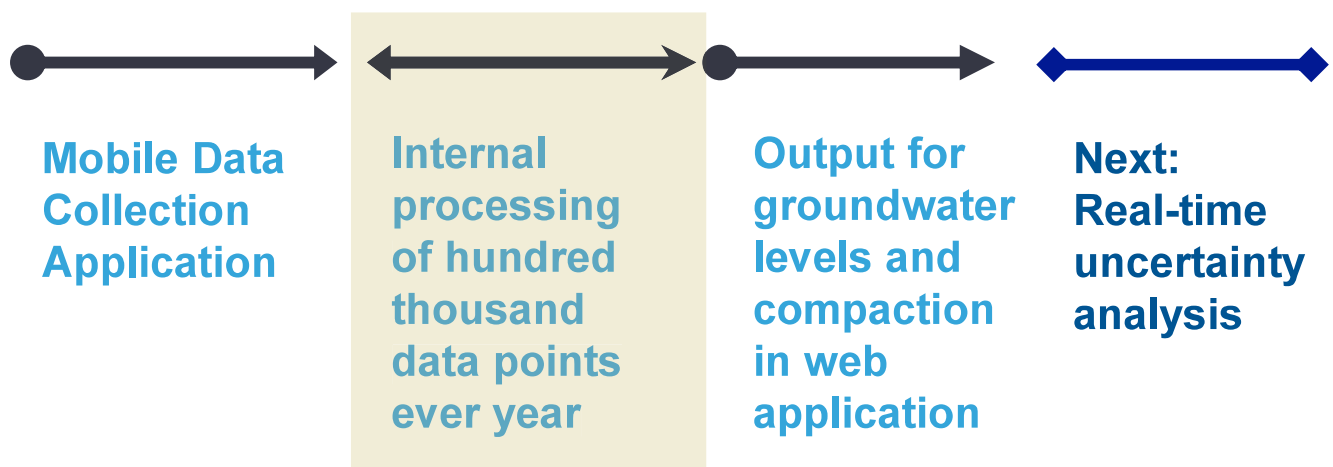
- ▶ Change in *Water Use Efficiency* over time
- ▶ Level of *Water Stress* (groundwater availability)
- ▶ Degree of *Integrated Water Resources Management* Implementation

Data Opportunities

- ▶ Develop *internal monitoring processes* for more updated and timely data
- ▶ *Disaggregate data* if possible to develop science narratives (noise reduction)
- ▶ Incorporate *decision tools and data* to measure success

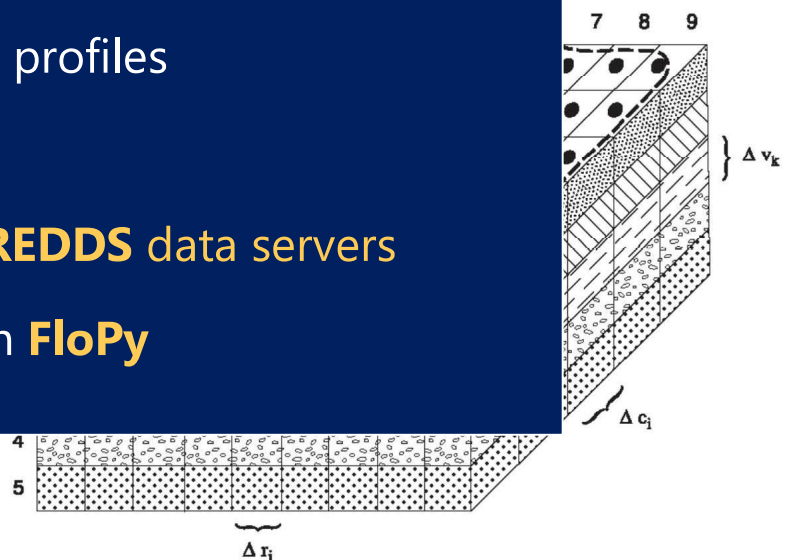


SUBSIDENCE APPLICATION MODERNIZATION PROCESS



GW WEB FLOW → BIG DATA PROCESSING

- ▶ Web-based user-interface for **MODFLOW** models
- ▶ Visual inputs and outputs
 - Maps, time series, depth profiles
- ▶ Scenario analysis
- ▶ Uses **netCDF** files and **THREDDS** data servers
 - Programmatic support in **FloPy**





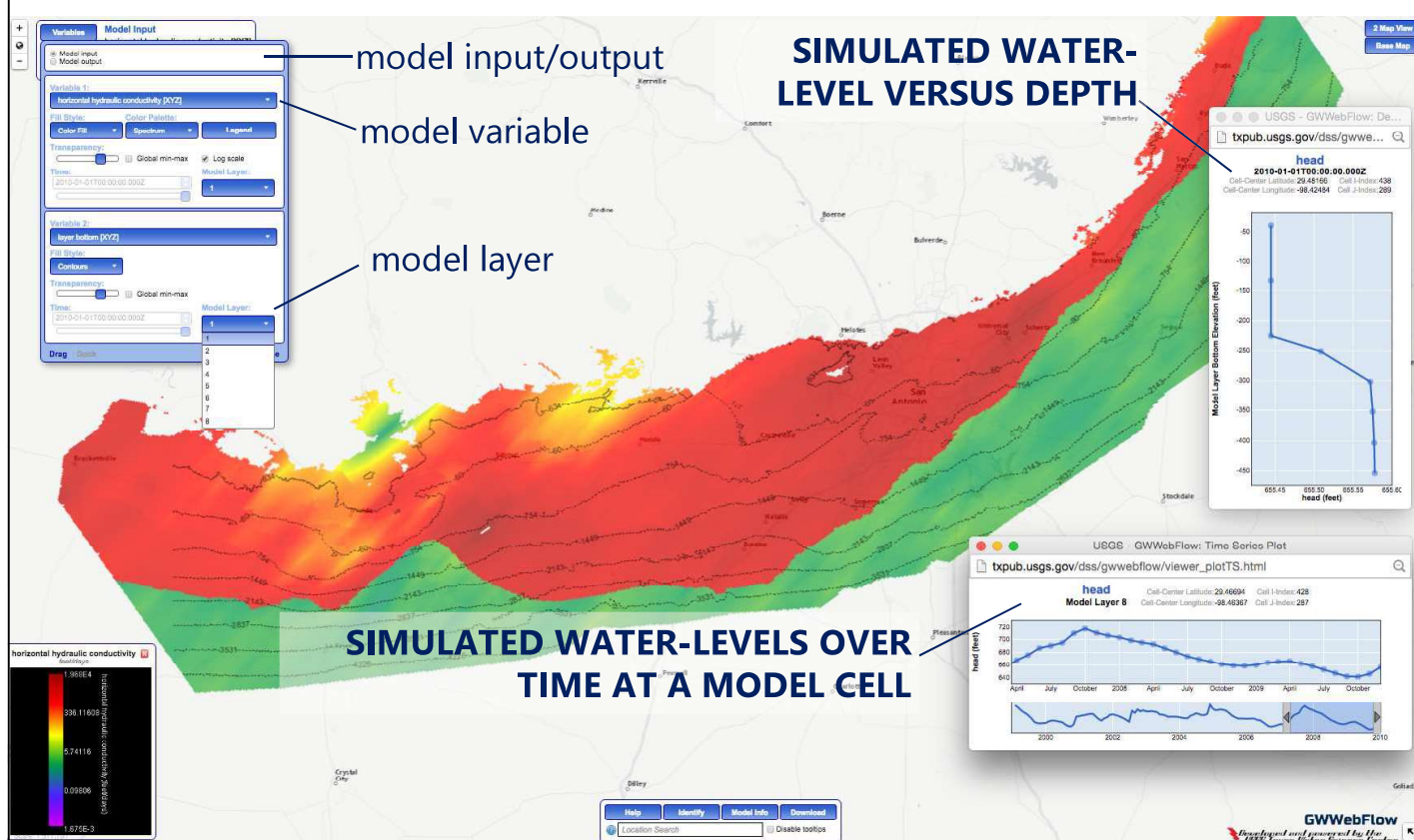
```

1 def cf_model(model, layer, model_name, base):
2     model.rem
3     lrcQ = {1
4     wel = fpm
5     wel.write
6     model.run
7     hfile = f
8     cfile = f
9     step_peri
10    h = hfile
11    cap_frac
12    for tstep
13        if h[
14            c
15        else:
16            v
17
18        v
19        v
20        v
21        full3D=True)
22        cap_frac[tstep] = ((v1[0].sum() + v2[0].sum() +
23        v3[0].sum() - base) / (-Q)
24    return cap_frac

```

- A programmatic modeling approach facilitates analyses that can be difficult to complete
- Many model runs with slightly different inputs that are required for:
 - parameter estimation
 - uncertainty analysis
 - drawdown analysis
 - capture-fraction analysis

GW WEB FLOW



GW WEB FLOW

GWWebFlow

Data Output

- ▶ Provides **status** of the model run based on inputs/outputs
- ▶ Important: provides a **traceability analysis** (metadata)
- ▶ Able to **reproduce and trace** large datasets to develop complex applications

Data Opportunities

- ▶ Develop a consistent internal approach to model file outputs for more efficiency and timely processing
- ▶ Incorporate various modeling packages to encourage data aggregation

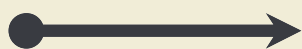
Scale: 1:2,311,162

Location Search

Disable tooltips

GW WEB FLOW ORDER OF OPERATIONS: FLO-PY

MODFLOW



FloPy:
Make output
processing of
millions of cells
simpler



Output grid

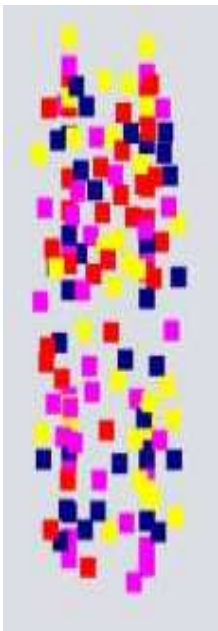


Traceability
analysis and
summary

USE FLO-PY TO MAKE DATA EASIER TO CURATE FOR PROCESSING

SOLVING THE INFORMATION CHALLENGE

CHALLENGES



Data sources, quality, and reliability vary across regions.

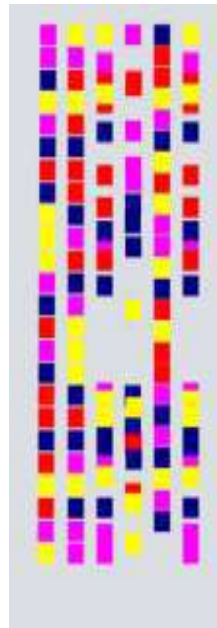
Data availability

Frequency and coverage vary widely.

Data curation

Data storage

OPPORTUNITIES



Standardization of monitoring to improve comparability of data.

Leverage open source statistical packages

Build capacity for new analytical skills

3

GOOD HEALTH AND WELL-BEING

Mapping the movement of mobile phone users can help predict the spread of infectious diseases

6

CLEAN WATER AND SANITATION

Sensors connected to water pumps can track access to clean water

Common themes

- ▶ Need vast amounts of different data to download, aggregate, and manage
- ▶ Data needs to be consumable scientific narrative
- ▶ Translate data into decision support

DISTRIBUTION OF WATER LEVEL VALUES

