

Analysis of SSO data

- H-GAC has been obtaining and analyzing SSO (sanitary sewer overflow) data for the BIG for several years
- In the past two years, data have been obtained from the TCEQ Central Office rather than the regional office, in a format that greatly simplifies data management and analysis
- However, there is considerable duplication in the data – the same SSO may be entered multiple times
- Removing these duplicates relies upon unique key fields like the permit number, cause, and reported location
- “Location” data is not reported in a useable or consistent format

Region 12 raw data – 2018/2019

- TCEQ provides statewide SSO data to H-GAC annually
- 4997 records for SSO events in Region 12 were extracted from the 12,783 records provided for 2018 and 2019
- The “permit” field in the data includes every permit held by the permittee (waste water, sludge, storm, water, etc.) and the WWTF permit number must be extracted via text mining techniques
- The “permit” field for 177 region 12 records was blank, and the records were deleted from the analysis
- After H-GAC processing and data cleaning, there were 3463 unique Region 12 SSO events in the 2018/2019 data – a removal of about 30 percent

Location data in the TCEQ database is limited and inconsistent

- The “Source” field often includes partial address data, but without city names or zip codes needed for geocoding
- The table on the right is a rough count of the types of data entered in the “Source” field
- Raw 2018-2019 data

Type of information in “Source” field	Number of incidents
Partial Address	3664
Manhole	73
Material/Cause	6
Other	1056
Permittee (City)	162
WWTF	36

Region 12 – 2019 Data Only

- Raw data:
1858
records
- After
processing
: 1558

Cause Code	Events- Raw Data	Events-after H-GAC Processing
ACT OF GOD	177	145
CONTINUOUS RELEASE	3	3
CORROSION	1	.
EQUIPMENT FAILURE	180	149
GREASE BLOCKAGE	643	569
HUMAN ERROR	30	23
INFILTRATION AND INFLOW	328	273
INTENTIONAL DISCHARGE	1	.
LINE BLOCKAGE (NON-GREASE)	288	244
LINE BREAK	31	25
OTHER	122	96
POWER OUTAGE	34	26
UNKNOWN	19	14
VANDALISM	1	1

The value of accurate location data

- Accurate and unambiguous locations of the SSO events – ideally, latitude and longitude- are vital for creating a global view of SSO challenges in the BIG
 - It would allow better, and more “granular” mapping of individual events
 - Accurate location data would simplify removal of *true* duplicate records
 - Better location data could identify particular infrastructural trouble spots
 - SSO events could be linked to specific assessment units, allowing the application of statistical methods to determine the impact of SSOs
 - City of Houston staff have developed a mobile app for recording SSO information in the field that includes latitude and longitude of events