



# Halifax Water Pipe Segmentation Tool

David Blades

Luke Butler

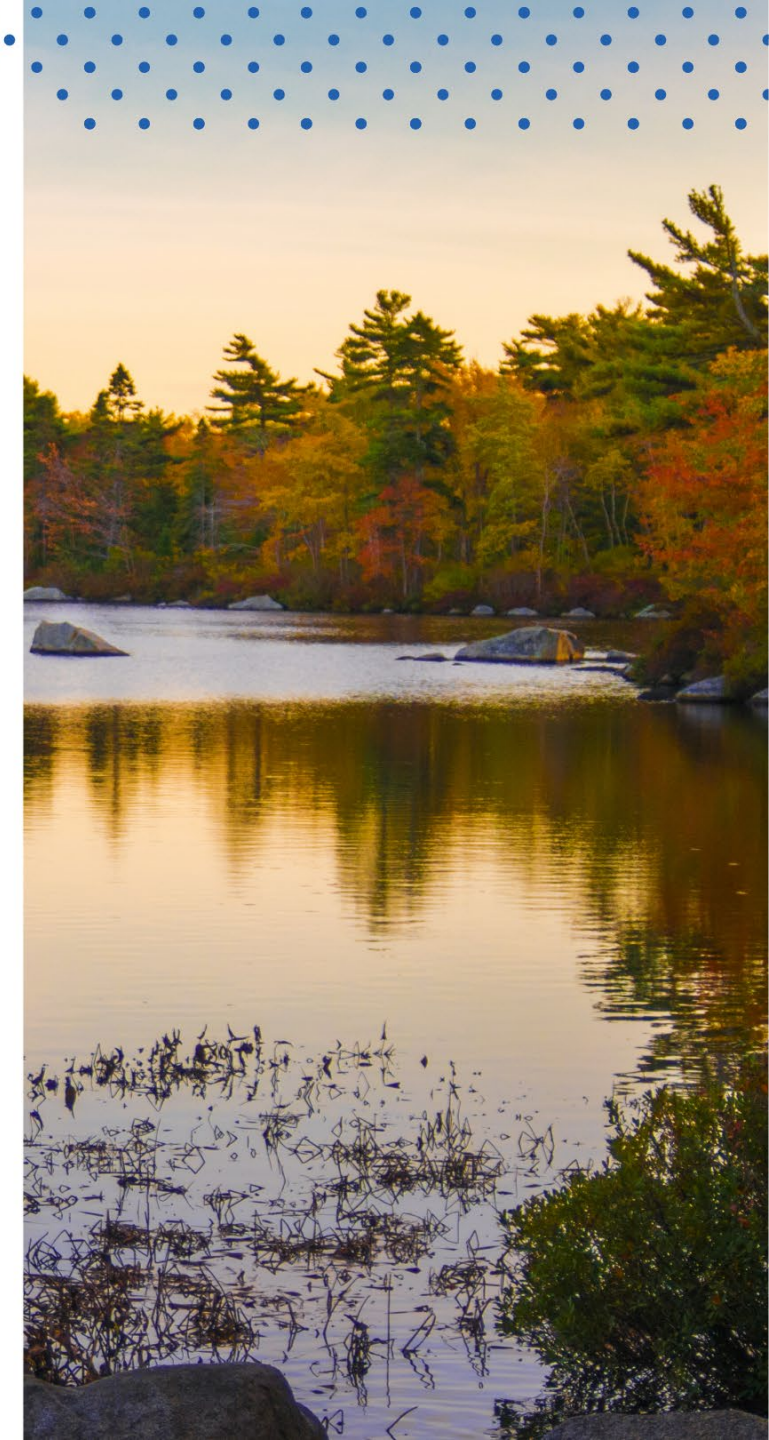
Michelle Scott

CWWA

NATIONAL WATER AND WASTEWATER CONFERENCE 2022

November 8, 2022

**STRAIGHT from  
the SOURCE**



# Developing the Digital Twin

There are some key distinctions between a conventional model and a Digital Twin.



The water industry has a long history of using models to leverage collected data to help design and manage water systems—and with great success; however, Digital Twins look to overcome conventional model limitations while helping realize the longstanding aspirations of these models—to help operate water systems more efficiently."

—Pusker Regmi,  
Wastewater Process Expert



## Conventional Models



## Digital Twin

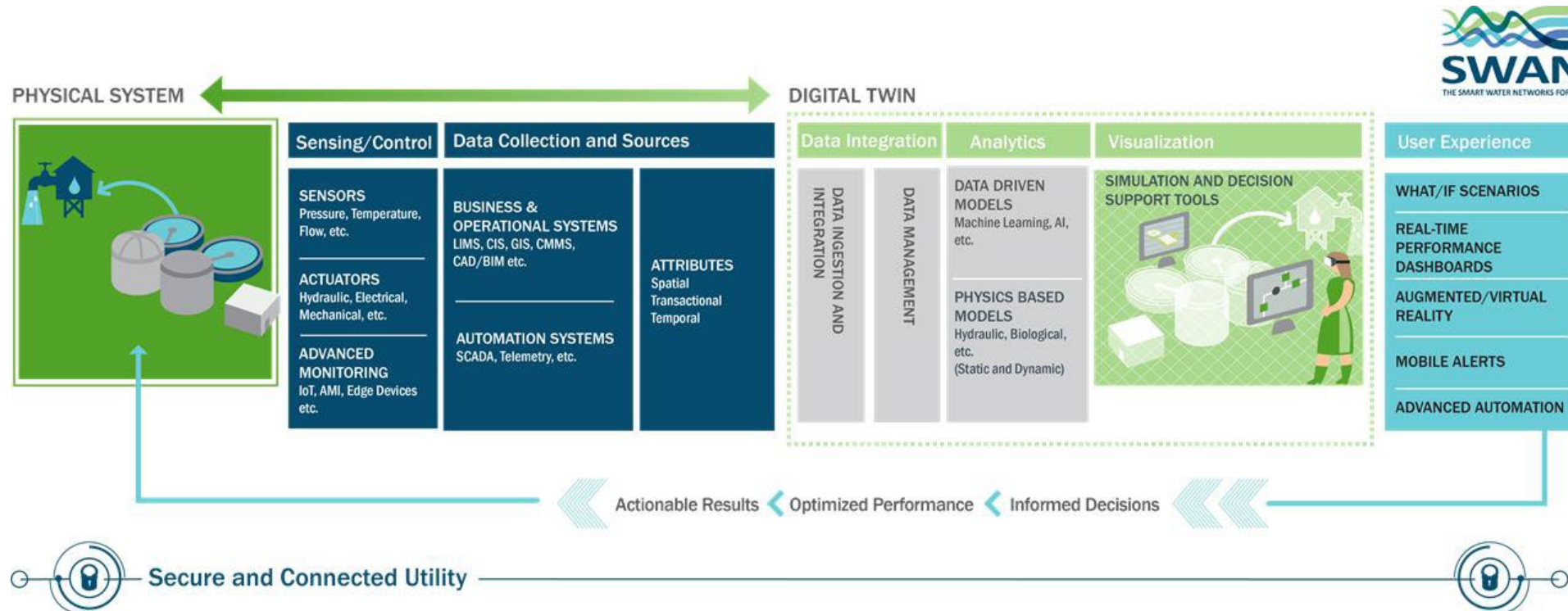
Based on static datasets	Diverse systemwide information that includes real-time and meta data
Data screening and cleansing	Includes advanced data analytics
Models updated periodically with static data	Models closely represent physical system and update continuously and dynamically
Slower decision support; model outputs are interpreted manually	Faster decision support; models outputs are interpreted automatically
Outputs and potential actions are interpreted manually	Outputs are interpreted automatically and provide real-time alerts, insights, and actions



# Digital Twin

Definition: A Digital Twin is a dynamic digital representation of real-world entities and their behaviors using models with static and dynamic data that enable insights and interactions to drive actionable and improved outcomes.

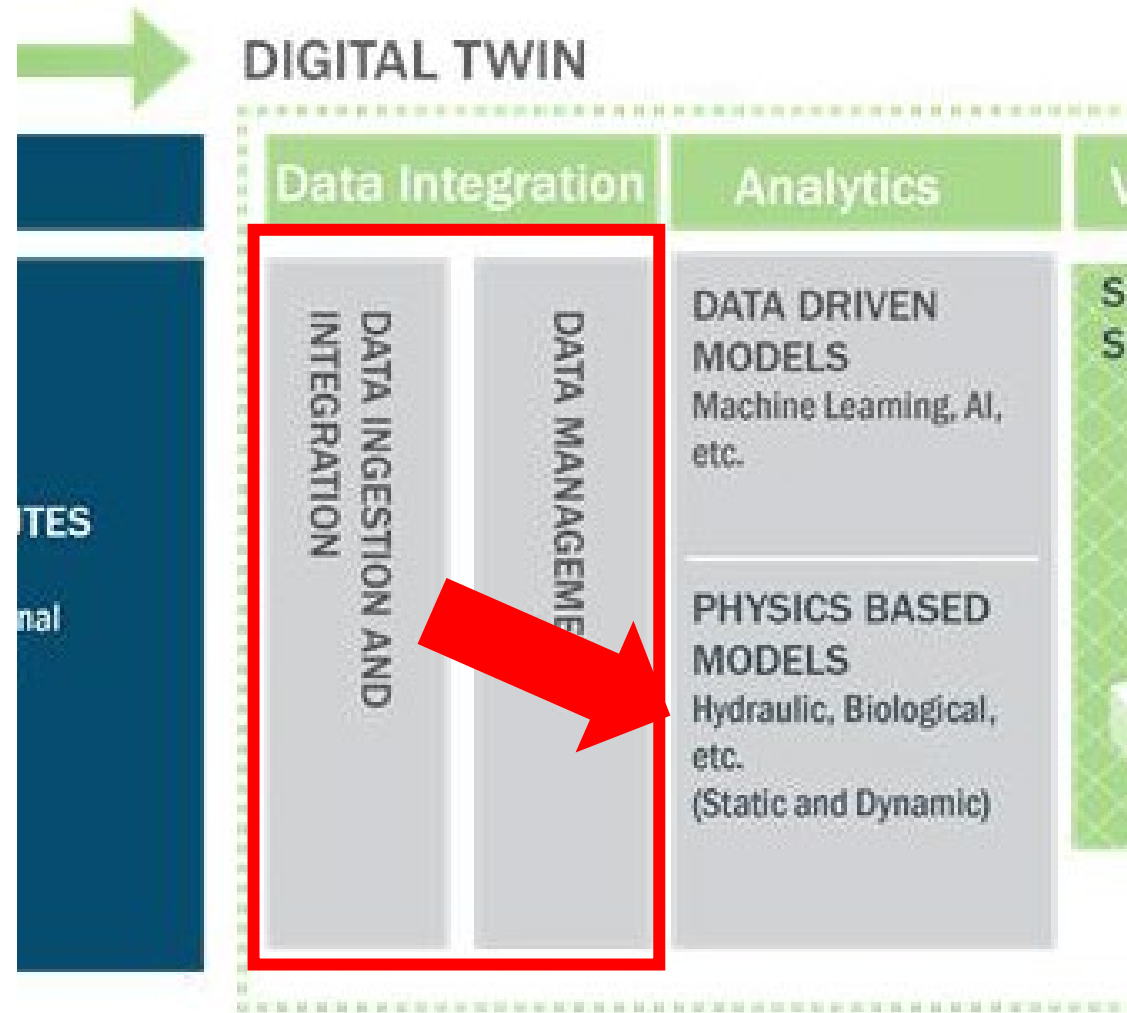
❖ Industry aligned definition led by SWAN Digital Twin Working Group and "American Water Works Association Digital Twins Committee



# Digital Twin

Definition: A Digital Twin is a dynamic digital representation of real-world entities and their behaviors using models with static and dynamic data that enable insights and interactions to drive actionable and improved outcomes.

❖ *Industry aligned definition led by SWAN Digital Twin Working Group and "American Water Works Association Digital Twins Committee*



# Modelling Program Touchstones

## Leverage Primary Data Sources

- Avoid duplication of existing datasets
- GIS is the source of infrastructure data

## Reduce Manual Data Cleaning/Manipulation

- Implement automated data import procedures where possible
- Reduce roadblocks to calibration

## Traceability and Repeatability of Modelling

- Develop modelling guidelines
- Implement version control

## Get Analysis in Hands of Decision Makers

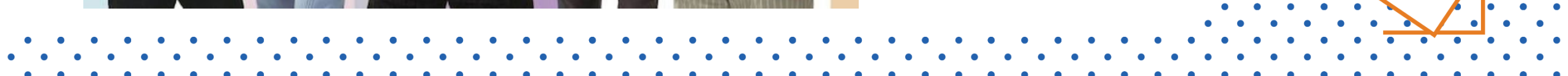
- Don't get caught in analysis paralysis
- Find ways to present analysis in an easy to digest format



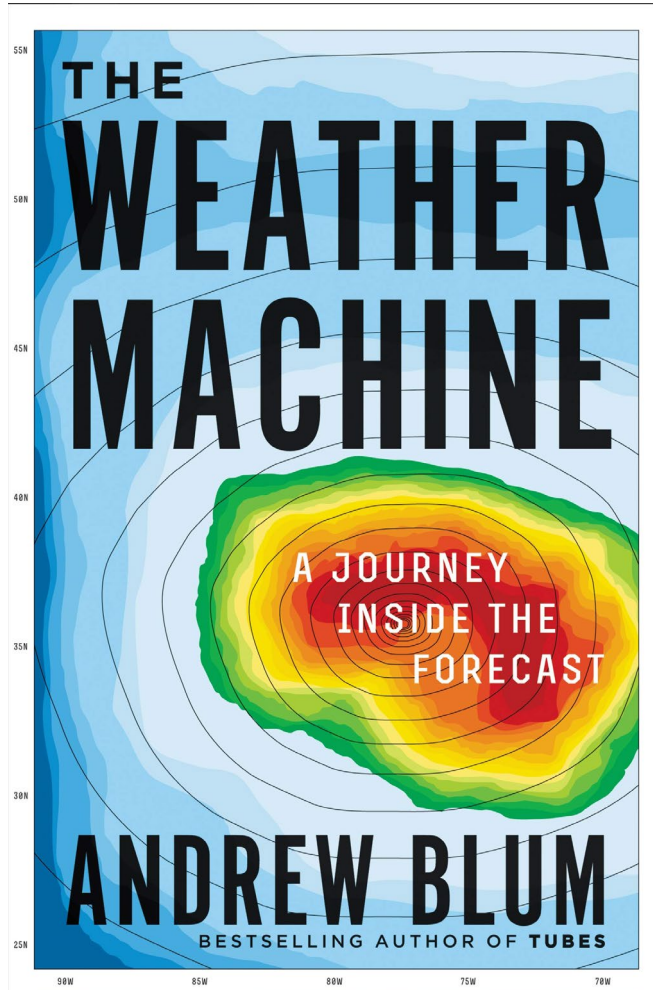


# Digital Twin Illustration

Don't worry, it's all been done before!



# Digital Twin Illustration



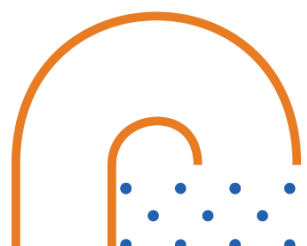
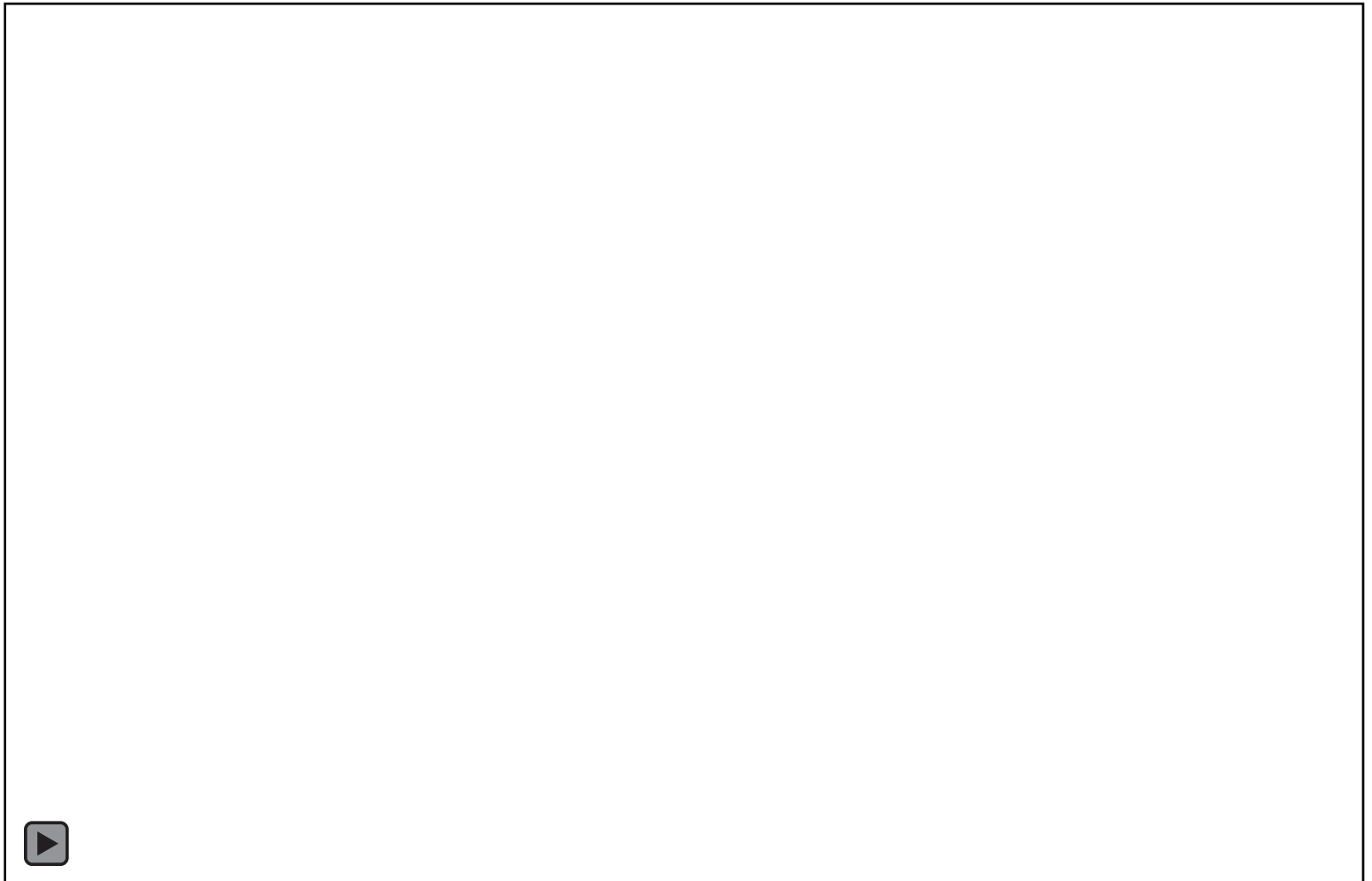
- Multiple sources of data
- Robust to failure
- Frequent simulations
- Visualization generated (automatic)
- Used for Decision Making





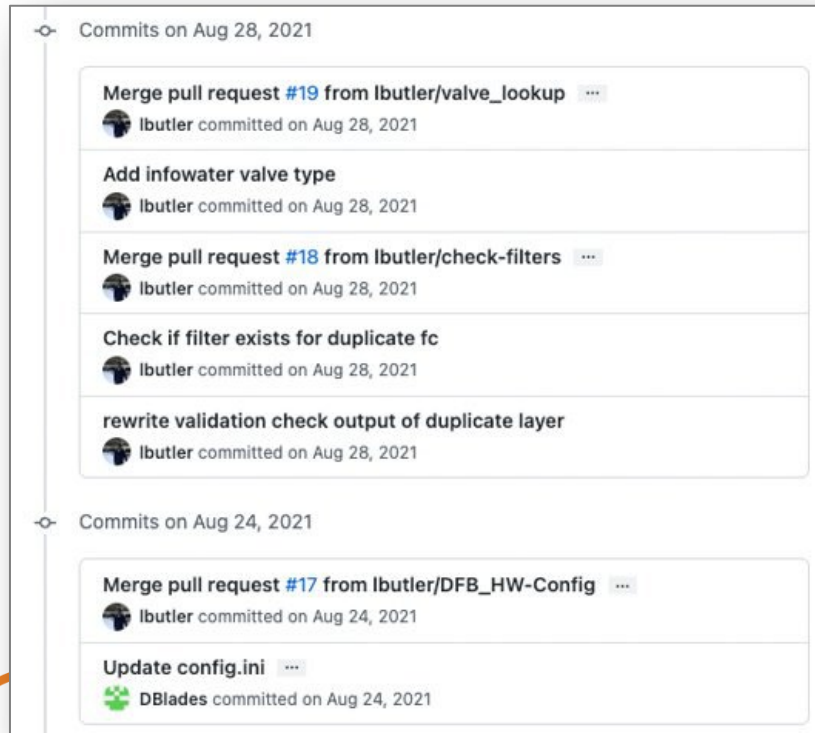
# Halifax Water Pipe Segmentation Tool

- Python script for ArcGIS
- Run as a background process
- Can be automated and run frequently
- Removes the previously manual GIS clean up tasks
- Provides a consistent output allowing easier merging of data



# Best Practice

- Source Control
- Modular
- Logging

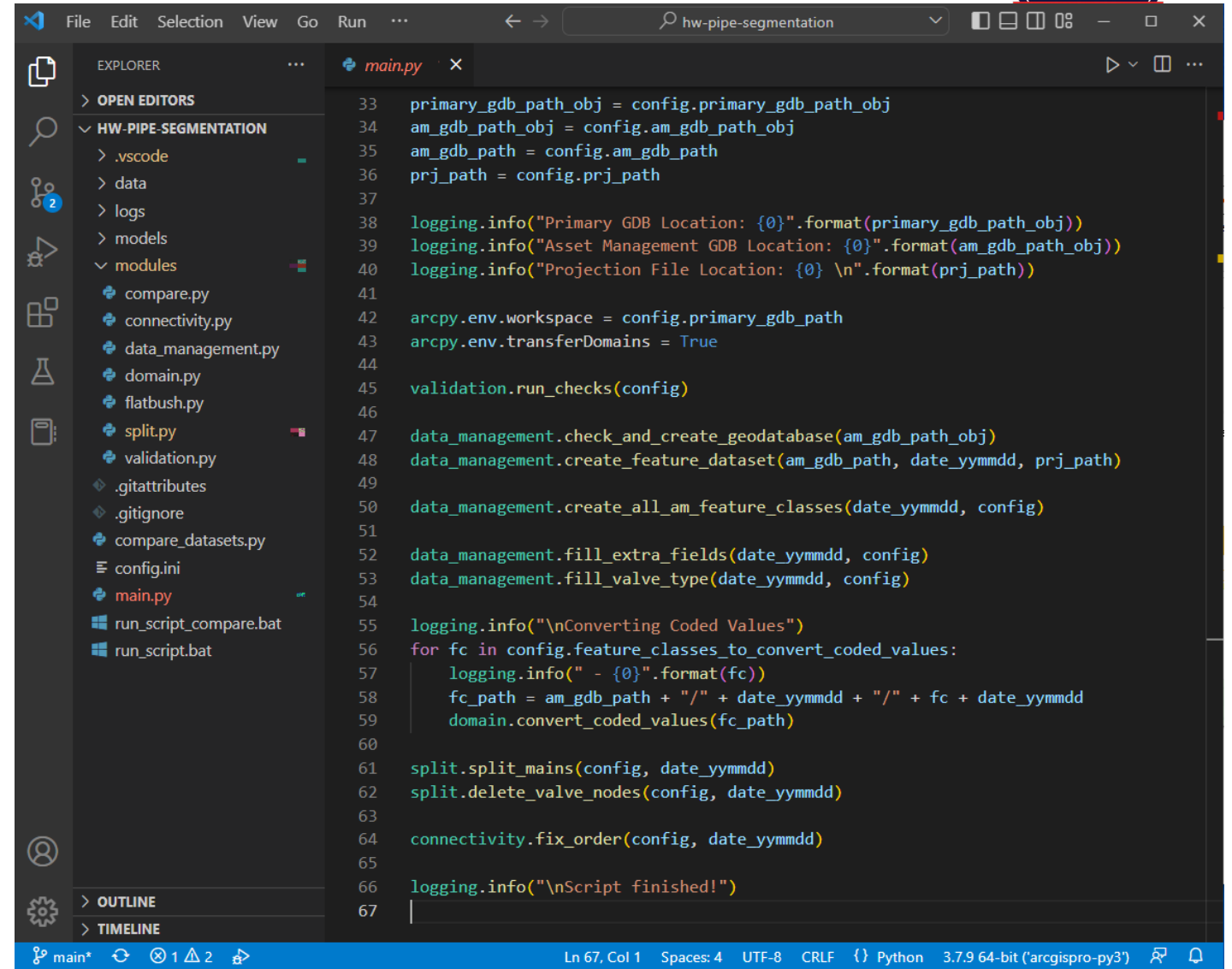


Commits on Aug 28, 2021

- Merge pull request #19 from lbutler/valve\_lookup ...  
lbutler committed on Aug 28, 2021
- Add infowater valve type  
lbutler committed on Aug 28, 2021
- Merge pull request #18 from lbutler/check-filters ...  
lbutler committed on Aug 28, 2021
- Check if filter exists for duplicate fc  
lbutler committed on Aug 28, 2021
- rewrite validation check output of duplicate layer  
lbutler committed on Aug 28, 2021

Commits on Aug 24, 2021

- Merge pull request #17 from lbutler/DFB\_HW-Config ...  
lbutler committed on Aug 24, 2021
- Update config.ini ...  
DBlades committed on Aug 24, 2021



hw-pipe-segmentation

```
33 primary_gdb_path_obj = config.primary_gdb_path_obj
34 am_gdb_path_obj = config.am_gdb_path_obj
35 am_gdb_path = config.am_gdb_path
36 prj_path = config.prj_path
37
38 logging.info("Primary GDB Location: {0}".format(primary_gdb_path_obj))
39 logging.info("Asset Management GDB Location: {0}".format(am_gdb_path_obj))
40 logging.info("Projection File Location: {0} \n".format(prj_path))
41
42 arcpy.env.workspace = config.primary_gdb_path
43 arcpy.env.transferDomains = True
44
45 validation.run_checks(config)
46
47 data_management.check_and_create_geodatabase(am_gdb_path_obj)
48 data_management.create_feature_dataset(am_gdb_path, date_yymmdd, prj_path)
49
50 data_management.create_all_am_feature_classes(date_yymmdd, config)
51
52 data_management.fill_extra_fields(date_yymmdd, config)
53 data_management.fill_valve_type(date_yymmdd, config)
54
55 logging.info("\nConverting Coded Values")
56 for fc in config.feature_classes_to_convert_coded_values:
57     logging.info(" - {0}".format(fc))
58     fc_path = am_gdb_path + "/" + date_yymmdd + "/" + fc + date_yymmdd
59     domain.convert_coded_values(fc_path)
60
61 split.split_mains(config, date_yymmdd)
62 split.delete_valve_nodes(config, date_yymmdd)
63
64 connectivity.fix_order(config, date_yymmdd)
65
66 logging.info("\nScript finished!")
67
```

Ln 67, Col 1 Spaces: 4 UTF-8 CRLF {} Python 3.7.9 64-bit ('arcgispro-py3')

# Best Practice

- Source Control
- Modular
- Logging
- Syntax checking

```

6
7 def split_mains(config, date_yymmdd):
8     logging.info(f"\nStarting splitting process")
9
10     stringNotUsed = " I'm not used in the code "
11
12     dataset_path = config.am_gdb_path + "/" + date_yymmdd
13
14     fc_target_split_pipe = dataset_path + "AM_water_pipe_s
15     fc_service_line = dataset_path + "AM_water_service_lin
16     fc_target_nodes = dataset_path + "AM_nodes" + date_yv

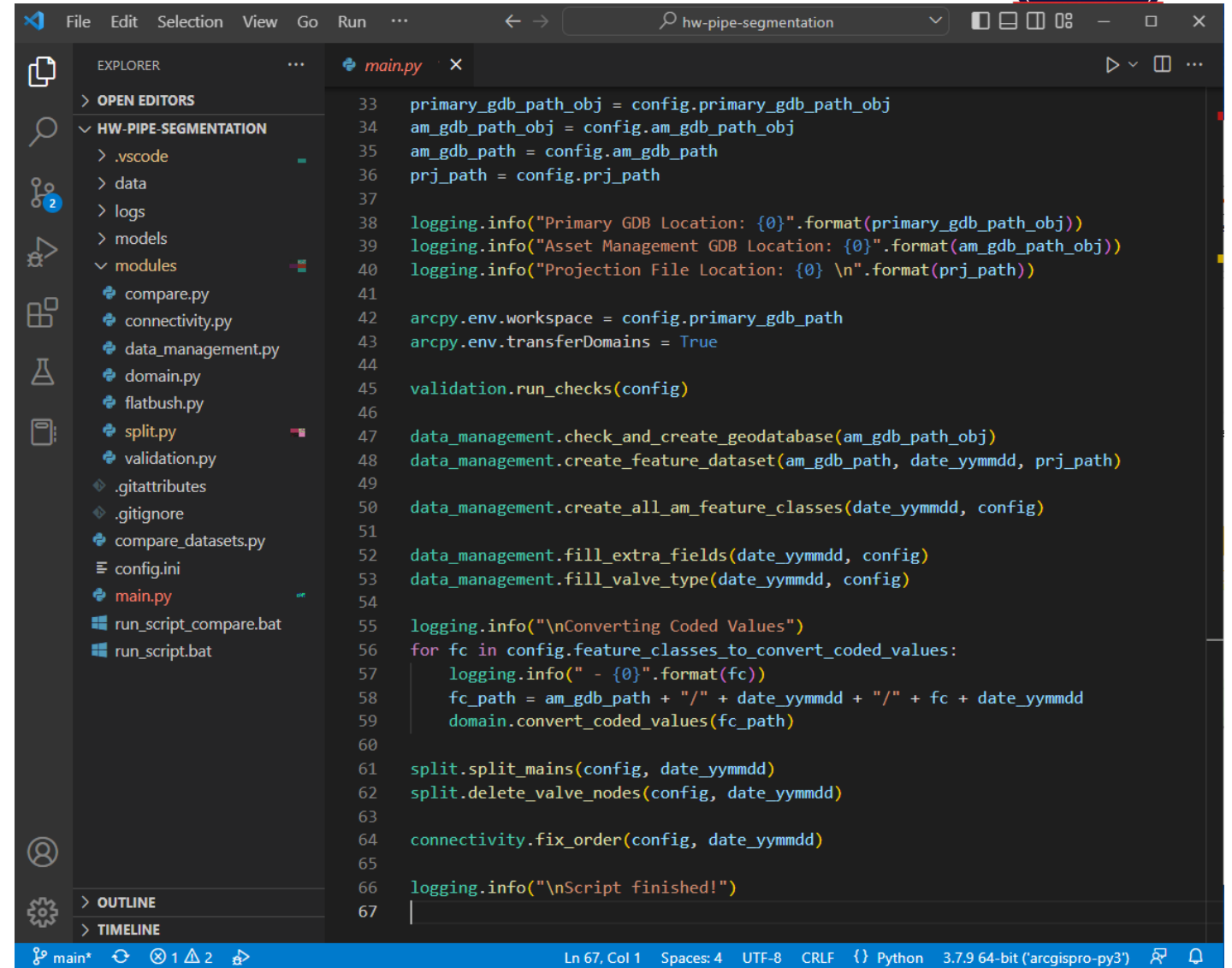
```

OUTPUT TERMINAL JUPYTER: VARIABLES PROBLEMS 4

Filter (e.g. text, \*\*/\*.ts, !\*\*/node\_modules/\*\*)

split.py modules 4

- ⊗ f-string is missing placeholders flake8(F541) [Ln 8, Col 18]
- ⊗ local variable 'stringNotUsed' is assigned to b... flake8(F841) [Ln 10, Col 5]
- ⊗ f-string is missing placeholders flake8(F541) [Ln 42, Col 18]
- ⊗ line too long (122 > 120 characters) flake8(E501) [Ln 80, Col 121]



```

33 primary_gdb_path_obj = config.primary_gdb_path_obj
34 am_gdb_path_obj = config.am_gdb_path_obj
35 am_gdb_path = config.am_gdb_path
36 prj_path = config.prj_path
37
38 logging.info("Primary GDB Location: {0}".format(primary_gdb_path_obj))
39 logging.info("Asset Management GDB Location: {0}".format(am_gdb_path_obj))
40 logging.info("Projection File Location: {0} \n".format(prj_path))
41
42 arcpy.env.workspace = config.primary_gdb_path
43 arcpy.env.transferDomains = True
44
45 validation.run_checks(config)
46
47 data_management.check_and_create_geodatabase(am_gdb_path_obj)
48 data_management.create_feature_dataset(am_gdb_path, date_yymmdd, prj_path)
49
50 data_management.create_all_am_feature_classes(date_yymmdd, config)
51
52 data_management.fill_extra_fields(date_yymmdd, config)
53 data_management.fill_valve_type(date_yymmdd, config)
54
55 logging.info("\nConverting Coded Values")
56 for fc in config.feature_classes_to_convert_coded_values:
57     logging.info(" - {0}".format(fc))
58     fc_path = am_gdb_path + "/" + date_yymmdd + "/" + fc + date_yymmdd
59     domain.convert_coded_values(fc_path)
60
61 split.split_mains(config, date_yymmdd)
62 split.delete_valve_nodes(config, date_yymmdd)
63
64 connectivity.fix_order(config, date_yymmdd)
65
66 logging.info("\nScript finished!")
67

```

main\* 1 2 Ln 67, Col 1 Spaces: 4 UTF-8 CRLF Python 3.7.9 64-bit (arcgispro-py3)

# Best Practice

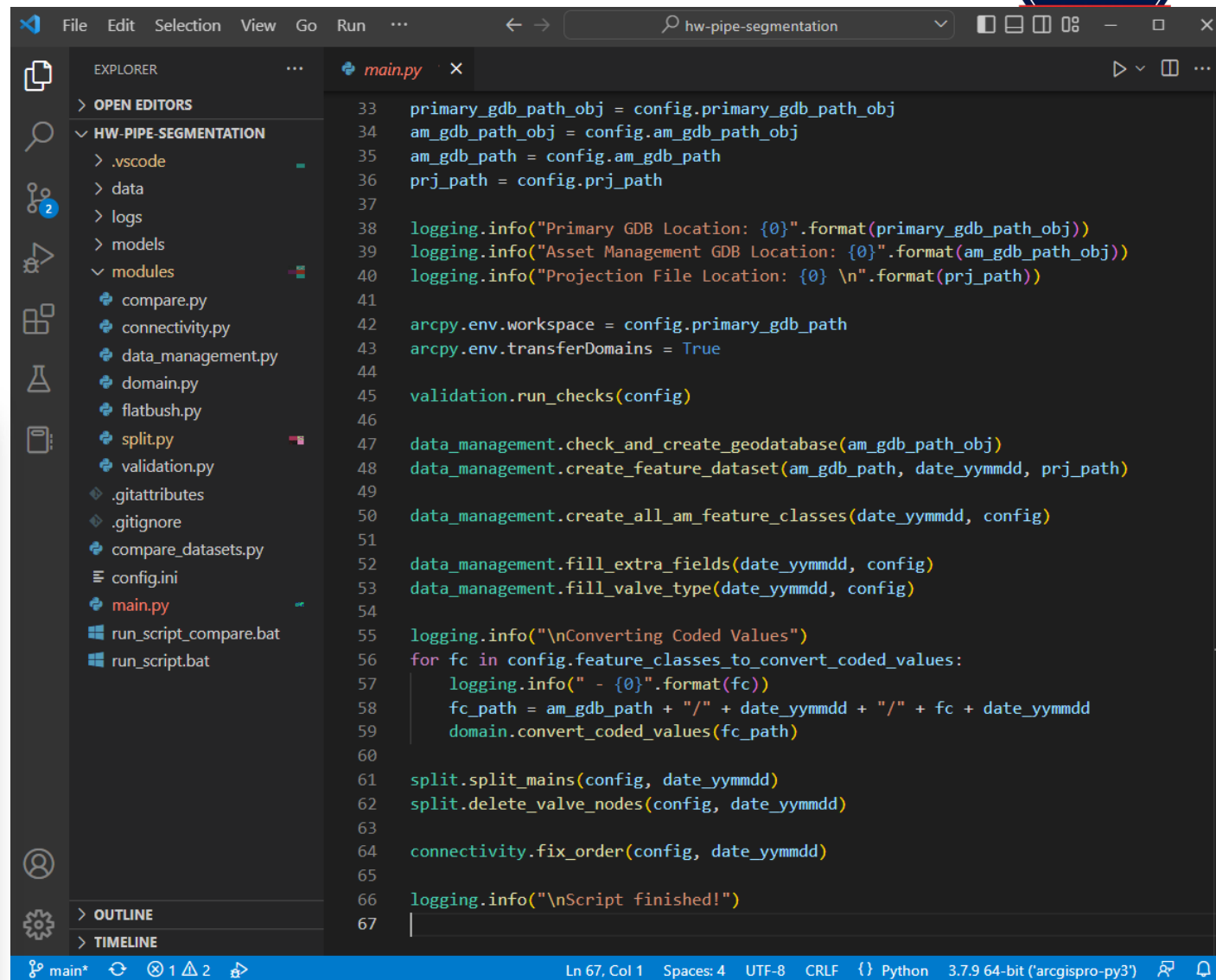
- Source Control
- Modular
- Logging
- Syntax checking
- Configuration

```
[Paths]
PrimaryGeodatabase=./data/C3_Export_V4.gdb
AMGeodatabase=./data/am_data.gdb
prj=./data/EPSG8083.prj

# Filters are applied to feature classes when AST DB is duplicated to AM DB
# The key must be the same as layer listed in the FeatureClasses duplicate array in config
[Filters]
AST_water_pipe = LIFECYCLESTATUS <= 3 And WATERTYPE = 1 And MAINTAINEDBY = 2
AST_water_fitting = FCODE IN ('WCCHFTCR', 'WCPTCR')
AST_water_hydrant = LIFECYCLESTATUS <= 3 And WATERTYPE = 1 And MAINTAINEDBY = 2
AST_water_service_line = FCODE IN ('WCHYLD')
AST_water_valve = (STATUS IN (4, 3) And FCODE IN ('WCVLBF', 'WCCHVLBF', 'WCCHVGLA', 'WCVLGA'))
or FCODE IN ('WCCHVLP', 'WCVLPR')

[FeatureClasses]
#Filter is applied if layer name exists in [FILTER]
duplicate = [
  ["AST_water_pipe", "AM_water_pipe"],
  ["AST_water_service_line", "AM_water_service_line"],
  ["AST_water_hydrant", "AM_water_hydrant"],
  ["AST_water_fitting", "AM_water_fitting"],
  ["AST_water_valve", "AM_water_valve"]
]
convert_coded_values = [
  "AM_water_pipe", "AM_water_service_line", "AM_water_hydrant",
  "AM_water_fitting", "AM_water_valve"
]
unique_id_check = [
  ["AST_water_pipe", "PIPEKEY"],
  ["AST_water_hydrant", "HYDRANTNUMBER"]
]

[DropFields]
AM_water_pipe = [
  "Enabled", "CapitalJobNumber", "CREATIONDATE", "STREETCODE",
  "EXTERIORPROTECTION", "JOINTTYPE", "PRIVATE", "RECORDED",
  "OPERATOR", "WARRANTYDATE", "LOCDESC", "MAINTAINEDBY"
]
AM_water_hydrant = [
  "ROTATION"
]
```



```
33 primary_gdb_path_obj = config.primary_gdb_path_obj
34 am_gdb_path_obj = config.am_gdb_path_obj
35 am_gdb_path = config.am_gdb_path
36 prj_path = config.prj_path
37
38 logging.info("Primary GDB Location: {0}".format(primary_gdb_path_obj))
39 logging.info("Asset Management GDB Location: {0}".format(am_gdb_path_obj))
40 logging.info("Projection File Location: {0} \n".format(prj_path))
41
42 arcpy.env.workspace = config.primary_gdb_path
43 arcpy.env.transferDomains = True
44
45 validation.run_checks(config)
46
47 data_management.check_and_create_geodatabase(am_gdb_path_obj)
48 data_management.create_feature_dataset(am_gdb_path, date_yymmdd, prj_path)
49
50 data_management.create_all_am_feature_classes(date_yymmdd, config)
51
52 data_management.fill_extra_fields(date_yymmdd, config)
53 data_management.fill_valve_type(date_yymmdd, config)
54
55 logging.info("\nConverting Coded Values")
56 for fc in config.feature_classes_to_convert_coded_values:
57     logging.info(" - {0}".format(fc))
58     fc_path = am_gdb_path + "/" + date_yymmdd + "/" + fc + date_yymmdd
59     domain.convert_coded_values(fc_path)
60
61 split.split_mains(config, date_yymmdd)
62 split.delete_valve_nodes(config, date_yymmdd)
63
64 connectivity.fix_order(config, date_yymmdd)
65
66 logging.info("\nScript finished!")
67
```

# Best Practice

- Source Control
- Modular
- Logging
- Syntax checking
- Configuration
- Validation

```
Command Prompt
Halifax Water Pipe Segmentation - Version 0.2
Primary GDB Location: C:\hw-pipe-segmentation\hw_data.gdb
Asset Management GDB Location: C:\hw-pipe-segmentation\hw_data.gdb
Projection File Location: C:\Luke-git\hw-pipe-segmentation\data\EPSG8083.prj
Running pre-script validations
✓ | ArcGIS for Desktop Advanced license available
✓ | Primary GeoDatabase Exists
✓ | Projection Definition Exists
✓ | Feature class to duplicate into AM DB exists: C:\hw-pipe-segmentation\hw_data.gdb/AST_water_pipe
✓ | Feature class to duplicate into AM DB exists: C:\hw-pipe-segmentation\hw_data.gdb/AST_water_service_line
✓ | Feature class to duplicate into AM DB exists: C:\hw-pipe-segmentation\hw_data.gdb/AST_water_hydrant
✓ | Feature class to duplicate into AM DB exists: C:\hw-pipe-segmentation\hw_data.gdb/AST_water_fitting
✓ | Feature class to duplicate into AM DB exists: C:\hw-pipe-segmentation\hw_data.gdb/AST_water_valve
✓ | Filter in config for feature class to duplicate into AM DB exists: AST_water_pipe
✓ | Filter in config for feature class to duplicate into AM DB exists: AST_water_service_line
✓ | Filter in config for feature class to duplicate into AM DB exists: AST_water_hydrant
✓ | Filter in config for feature class to duplicate into AM DB exists: AST_water_fitting
✓ | Filter in config for feature class to duplicate into AM DB exists: AST_water_valve
! | WARNING - Dataset already exists, overwrite? C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25
Please type YES to overwrite data: YES
✓ | No duplicates in PIPEKEY field in layer C:\hw-pipe-segmentation\hw_data.gdb/AST_water_pipe
✓ | No duplicates in HYDRANTNUMBER field in layer C:\hw-pipe-segmentation\hw_data.gdb/AST_water_hydrant
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_nodes_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_fitting_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_hydrant_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_pipe_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_pipe_split_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_service_line_2022_10_25
✓ | Can acquire schema lock on layer C:\hw-pipe-segmentation\hw_data.gdb/_2022_10_25/AM_water_valve_2022_10_25
- Validation Done
```

# Functionality Overview

Create a  
Filtered Copy

Populate  
Model Data

Re-segment  
Pipes and  
Create  
Junctions

Check and  
Flag Changes



Create a Filtered Copy

Populate Model Data

 Re-segment Pipes and  
 Create Junctions

Check and Flag Changes

 AST  
 Geodatabase

 AM  
 Geodatabase

- ☐ AST\_water\_fitting
- ☐ AST\_water\_hydrant
- ☐ AST\_water\_meter\_pressure\_zone
- ☐ AST\_water\_pipe
- ☐ AST\_water\_pump
- ☐ AST\_water\_service\_line
- ☐ AST\_water\_structure\_point
- ☐ AST\_water\_structure\_polygon
- ☐ AST\_water\_valve
- ☐ HW\_facility\_polygon
- ☐ HWDS\_flowmeter
- ☐ HWDS\_operational\_region
- ☐ HWDS\_tank



- ☐ am\_data.gdb
  - ☐ \_2019\_11\_01
    - ☐ AM\_water\_fitting\_2019\_11\_01
    - ☐ AM\_water\_hydrant\_2019\_11\_01
    - ☐ AM\_water\_pipe\_2019\_11\_01
    - ☐ AM\_water\_service\_line\_2019\_11\_01
    - ☐ AM\_water\_valve\_2019\_11\_01

- Halifax Water's AST corporate geodatabase is copied into new AM geodatabase
- Not all features in AST are required for the model build. Filters are used.
- Examples of filters:
  - Hydrants: Maintained by = Halifax Water
  - Valves: Type = PRV or Isolation
  - Pipes = water type = potable



Create a Filtered Copy

Populate Model Data

Re-segment Pipes and  
Create Junctions

Check and Flag Changes

- Not all features required for the model build that are included in the AST geodatabase feature classes.
- Attributes are added to AM feature classes in this step.
- Examples includes:
  - Copies of domain fields
  - Pipe size in metric & numeric format
  - Installation year in numeric format
  - C-factor
  - Unique Element ID
  - Valve Type

### AST Attributes

PipeSize	100 mm, 4"
InstallDate	1/1/1985
Owner	1



### Attributes Used for Model

Diameter	100
Year	1985
Owner	Halifax Water
Roughness	120





Create a Filtered Copy

Populate Model Data

Re-segment Pipes and  
Create Junctions

Check and Flag Changes

- Pipe must have upstream and downstream node and a node must be present at any location where a pipe is split.
- The script runs through the geometry of the model elements.
- Pipes are split within a specified tolerance (ex: 0.01 meters) of a node or the end of another pipe.
- Additional nodes created at the ends of pipes where one does not already exist.

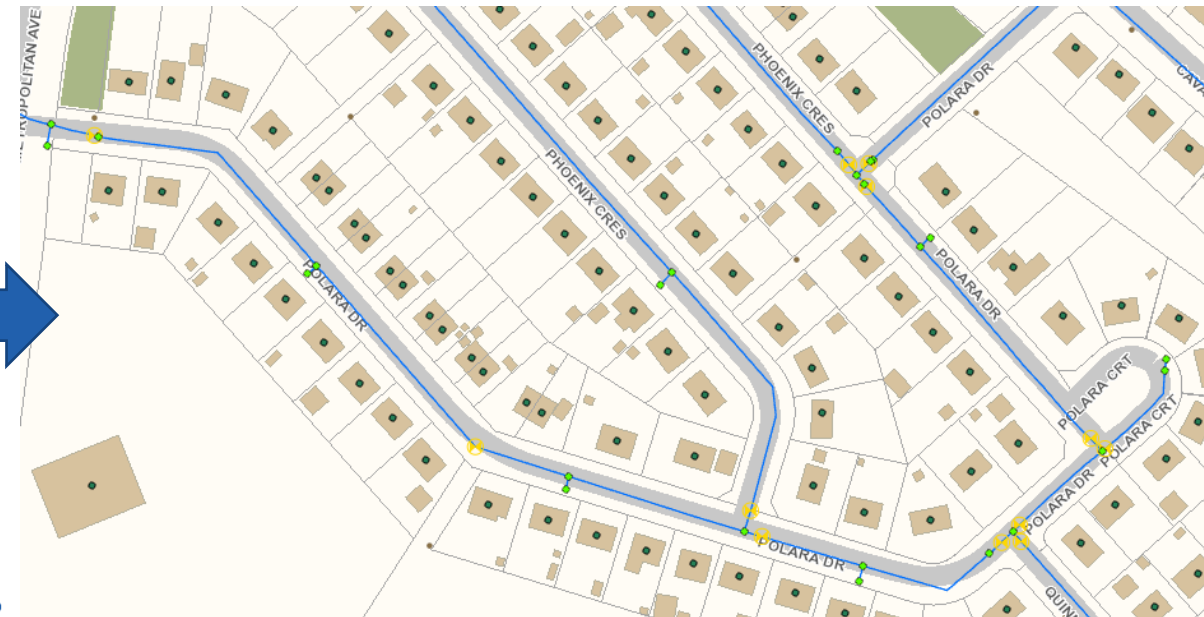
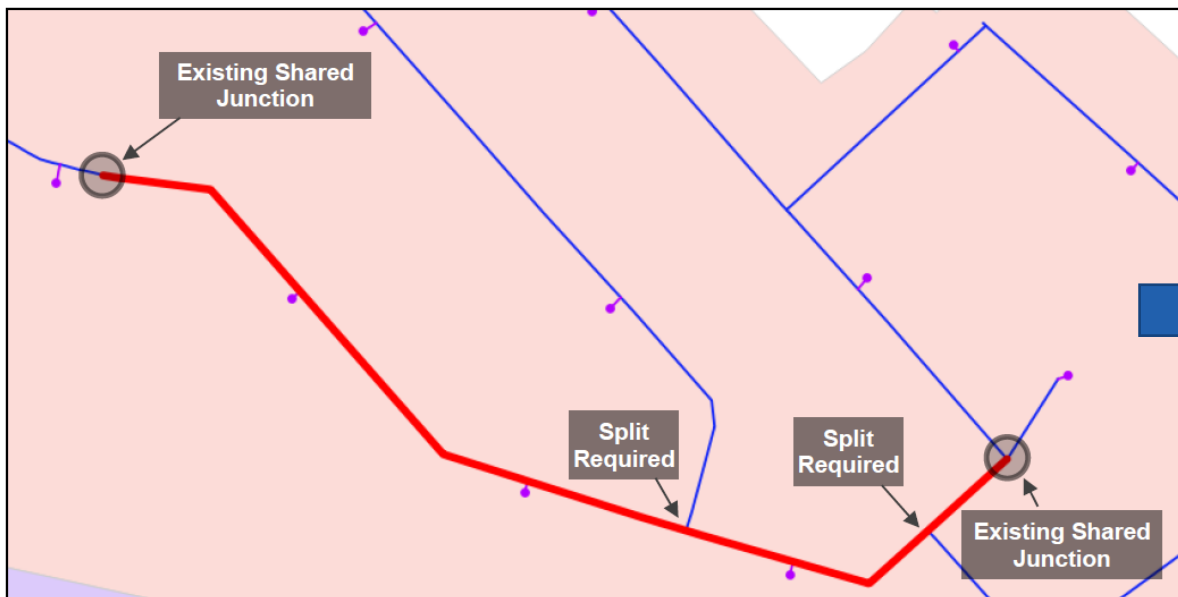
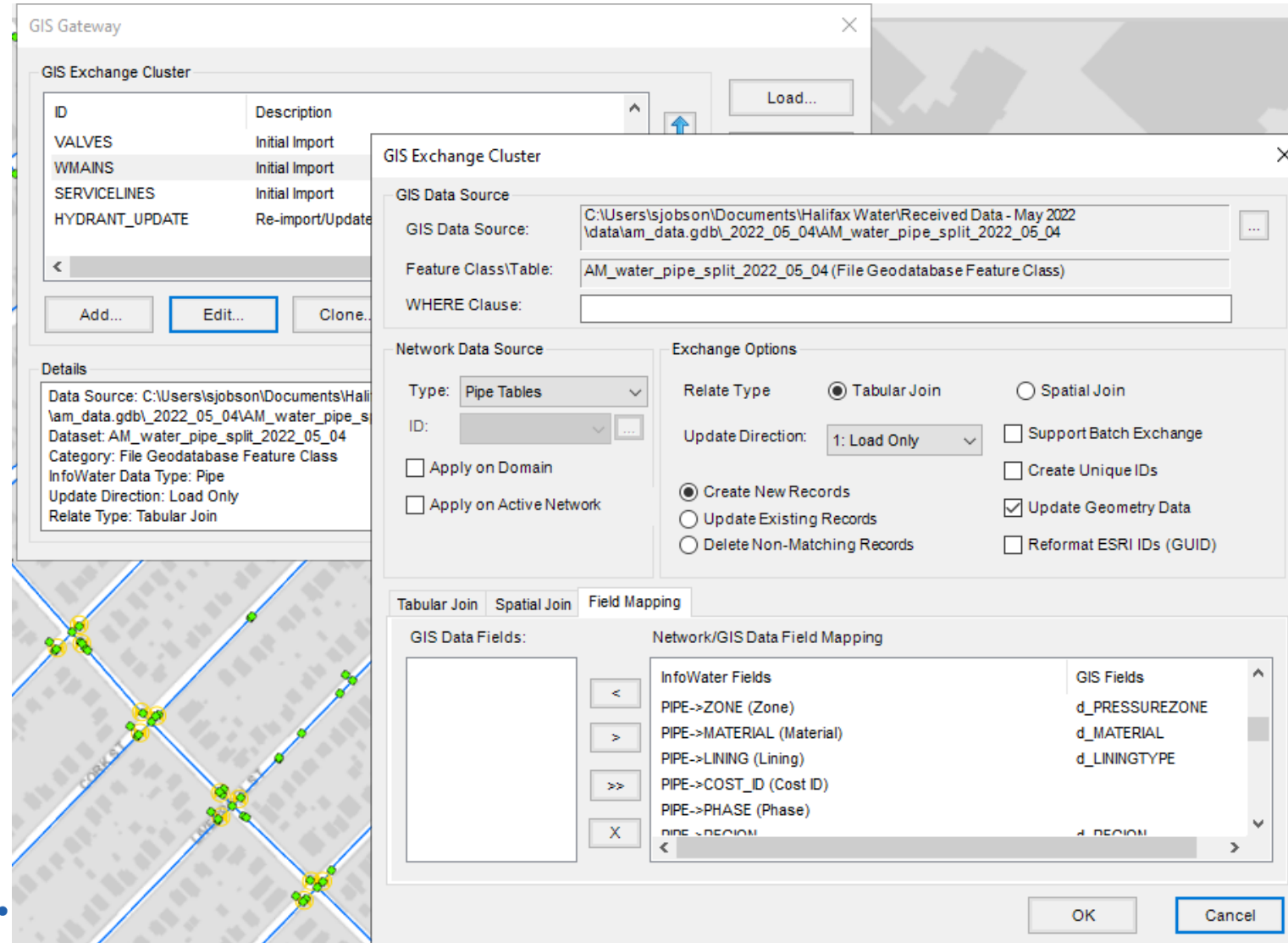


Figure 8-1 Splitting required on pipe W413862

# Import to InfoWater

- Once GIS has been formatted, it can be imported into InfoWater Pro using the GIS Gateway Tool
- Tabular joins can be used to update existing records or populate new elements



The screenshot displays the GIS Gateway tool interface, which is used for importing GIS data into InfoWater Pro. The interface is divided into several sections:

- GIS Exchange Cluster:** A table listing data sources with columns for ID and Description.
 

ID	Description
VALVES	Initial Import
WMAINS	Initial Import
SERVICELINES	Initial Import
HYDRANT_UPDATE	Re-import/Update
- Details:** A section providing specific information about the selected data source:
  - Data Source: C:\Users\sjobson\Documents\Halifax Water\Received Data - May 2022\am\_data.gdb\2022\_05\_04\AM\_water\_pipe\_split\_2022\_05\_04
  - Dataset: AM\_water\_pipe\_split\_2022\_05\_04
  - Category: File Geodatabase Feature Class
  - InfoWater Data Type: Pipe
  - Update Direction: Load Only
  - Relate Type: Tabular Join
- GIS Data Source:** A section for configuring the data source:
  - GIS Data Source: C:\Users\sjobson\Documents\Halifax Water\Received Data - May 2022\am\_data.gdb\2022\_05\_04\AM\_water\_pipe\_split\_2022\_05\_04
  - Feature Class/Table: AM\_water\_pipe\_split\_2022\_05\_04 (File Geodatabase Feature Class)
  - WHERE Clause: (Empty)
- Network Data Source:** A section for configuring the network data source:
  - Type: Pipe Tables
  - ID: (Empty)
  - Apply on Domain
  - Apply on Active Network
- Exchange Options:** A section for configuring the exchange options:
  - Relate Type:  Tabular Join,  Spatial Join
  - Update Direction: 1: Load Only
  - Support Batch Exchange
  - Create Unique IDs
  - Update Geometry Data
  - Reformat ESRI IDs (GUID)
- Field Mapping:** A section for mapping fields between the GIS data and the network:
  - GIS Data Fields: (Empty)
  - Network/GIS Data Field Mapping:
 

InfoWater Fields	GIS Fields
PIPE->ZONE (Zone)	d_PRESSUREZONE
PIPE->MATERIAL (Material)	d_MATERIAL
PIPE->LINING (Lining)	d_LININGTYPE
PIPE->COST_ID (Cost ID)	
PIPE->PHASE (Phase)	
PIPE->REGION	d_REGION

Create a Filtered Copy

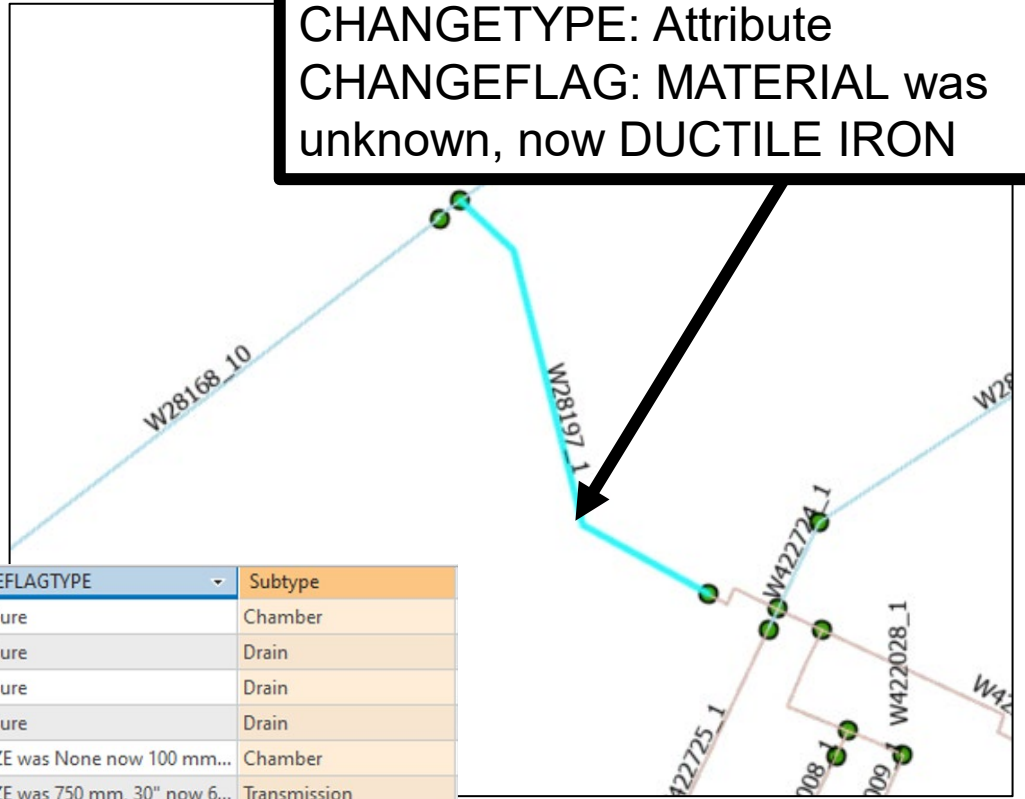
Populate Model Data

Re-segment Pipes and Create Junctions

Check and Flag Changes

- Model periodically updated.
- Useful to know changes made to the GIS feature classes since the last update.
- Tool compares to previous AM feature dataset and flags differences.
- The following change types checked:
  - New feature
  - Geographic shape
  - Attribute Change (ex: size, install date, owner, material)
  - Deleted feature
  - Pipe segmentation count change

**CHANGETYPE: Attribute**  
**CHANGEFLAG: MATERIAL was**  
**unknown, now DUCTILE IRON**



PipeKey	CHANGEFLAGTYPE	Subtype
W422837	New Feature	Chamber
W422871	New Feature	Drain
W422877	New Feature	Drain
W422878	New Feature	Drain
W7129	d_PIPESIZE was None now 100 mm...	Chamber
W417155	d_PIPESIZE was 750 mm, 30" now 6...	Transmission
W20508	d_PIPESIZE was 50 mm, 2" now 25...	Distribution
W10103	d_PIPESIZE was 50 mm, 2" now 25...	Distribution
W26605	d_PIPESIZE was 350 mm, 14" now 1...	Distribution
W25297	d_PIPESIZE was 150 mm, 6" now 20...	Distribution
W25609	d_PIPESIZE was 150 mm, 6" now 20...	Distribution
W417104	d_PIPECLASS was Unknown now C...	Transmission
W27615	d_PIPECLASS was Unknown now C...	Distribution
W417116	d_PIPECLASS was Unknown now C...	Transmission
W421941	d_PIPECLASS was CL-52 now CL-54	Distribution
W28561	d_MATERIAL was Unknown now PVC	Distribution





# Model Build (May 4<sup>th</sup> to November 1<sup>st</sup>) GIS Updates

```
You selected _2022_05_04 and it will be compared against _2022_11_01
Older FC: /_2022_05_04/AM_water_pipe_2022_05_04
Newer FC: /_2022_11_01/AM_water_pipe_2022_11_01
Start script: 2022-11-01 10:19:51
Key field is "PIPEKEY"

Comparing fields: ['PIPEKEY', 'INSTALLDATE', 'd_PIPESIZE', 'd_LININGTYPE', 'd_PIPECLASS', 'd_MATERIAL', 'd_LIFECYCLESTATUS', 'd_WATERTYPE']

Dictionary created: 2022-11-01 10:19:51
Total Time Elapsed: 0:00:00.280831

Checking new number 14839

Added to Latest = 114; Deleted from Older = 28; Attributes Changed = 22; Shape Changed = 0;
Older Count: 14750; Latest Count: 14839

Total Time Elapsed: 0:00:00.907371

Creating deleted feature class
- Creating Temporary Layer
- Making Filter Selection
- Copying Features
- Cleaning Temporary Features

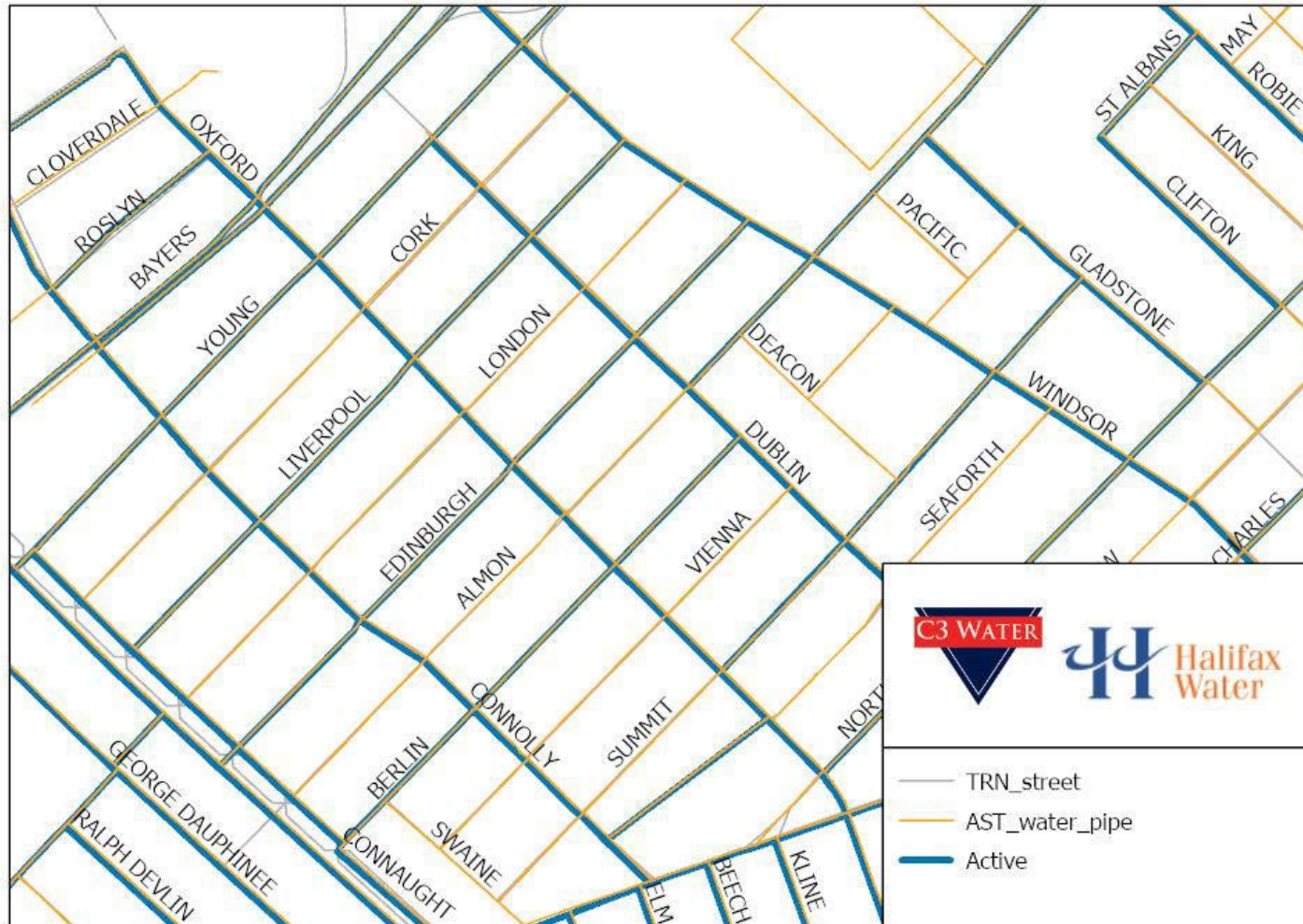
Filtering deleted feature

Comparing and counting splits on AM_water_pipe
- running through older count
- running through new count
- comparing the two datasets
- Pipe id W7641 had 3 pipe segments and now has 2
- Pipe id W8410 had 6 pipe segments and now has 4
- Pipe id W1965 had 3 pipe segments and now has 2
- Pipe id W2469 had 3 pipe segments and now has 2
- Pipe id W7641 had 3 pipe segments and now has 2
```

- 114 - New Pipes
- 28 - Removed Pipes
- 156 - Changed Segments
- 22 - Changed Attributes
- 0 - Changed Shape



# Next Steps...



2016:  
Total Pipes: 6,098  
Total Junctions: 4,613  
Valves (PRV): 136  
Valves (TCV): 0





# Next Steps...



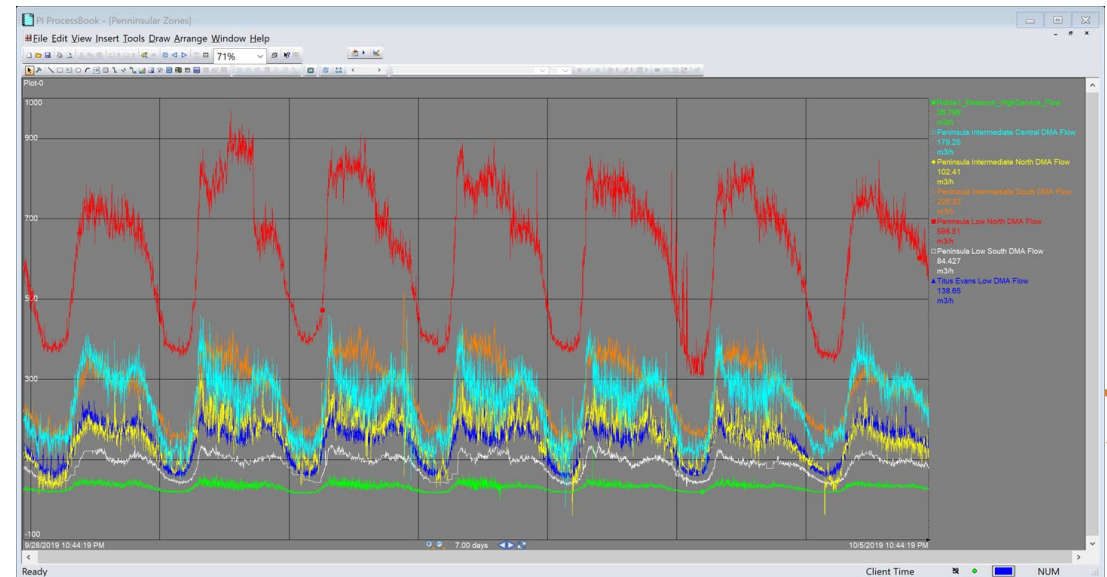
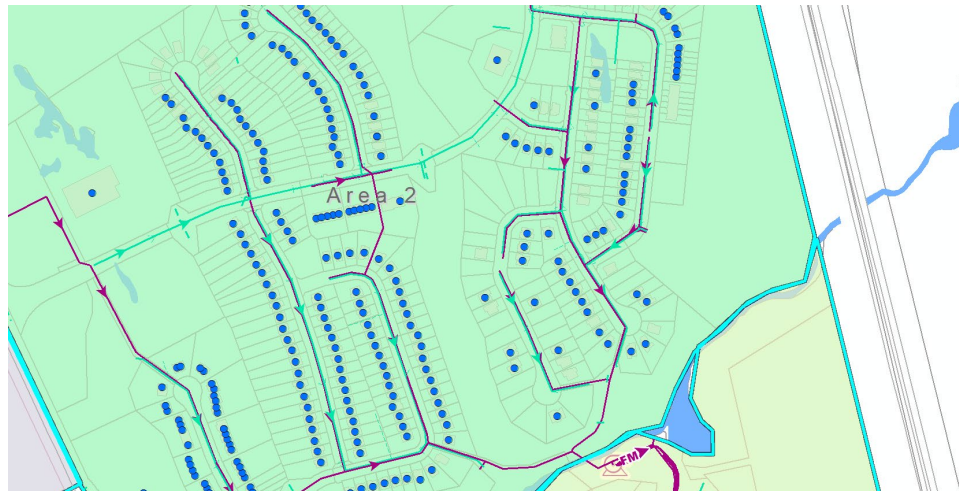
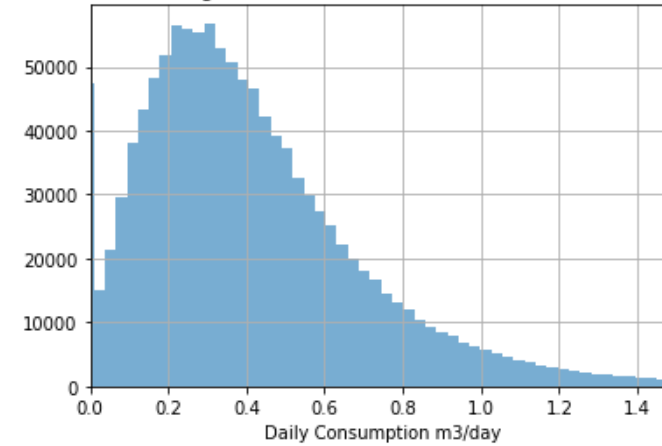
2022:  
Total Pipes: 52,384  
Total Junctions: 34,058  
Valves (PRV/TCV): 15,706



# Next Steps...

- Advanced Metering Infrastructure (AMI)
- C-Factors
- SCADA “Live Data” linkage
- Document
- Document
- Document

Histogram of 5/8" Meters that are Zoned R-1



# Questions?

[david.blades@halifaxwater.ca](mailto:david.blades@halifaxwater.ca)

[michelle.scott@c3water.ca](mailto:michelle.scott@c3water.ca)

[luke@matrado.ca](mailto:luke@matrado.ca)

