DATA LIFECYCLE FRAMEWORK
& GLOBAL GEOSPATIAL IDENTIFICATION SYSTEM

TASK 6:

UTILITY PILOT: TECHNICAL NOTES

CONTENTS

The scope of this pilot as set forth in the agreement states: "For two (2) pilot municipalities, ESRC will perform a feasibility study relative to obtaining public utility data from Municipal Computer Assisted Mass Appraisal ("CAMA"[1]), as well as directly from the utility(ies) serving those communities. For the latter, the data solicitation will include a request for sewer/water lines in digital format. These potential data streams will be evaluated and integrated with the New Hampshire Mosaic Parcel Map. The task results will include an assessment of the feasibility of collecting similar data statewide."

The purpose of this task is to identify properties which are not served by public water and thereby would be eligible for NH Department of Environmental Services (NHDES) water quality sampling, either related to an identified potential contaminate in the area or as part of routine well water sampling activity. Identifying those properties not served by public water is most easily accomplished by first determining those properties that are served.

This pilot evaluated several potential methods for identifying properties served by public water by using three distinct data sources: (1) CAMA records, (2) Geospatial network files provided by the utility, and (3) Service address lists. To accomplish the task, the project team worked with NHDES to identify target utility providers and municipalities who could provide the required data. The data sets, municipalities and utility providers used for the pilot are listed in Table 1. Rye and Londonderry were selected for CAMA exports as Rye has both a village water district and a private water provider, and Londonderry has only a private water provider. Pennichuck Corporation provided their geospatial network files and Portsmouth Municipal Water Department provided both geospatial network data and a service address list.

Data Set	Municipality	Utility Providers
CAMA	Rye	Rye Village Water, Aquarion Water
CAMA	Londonderry	Pennichuck
Geospatial Network	Entire Service Area	Pennichuck
Service Addresses	Portsmouth, Rye, Newington, Greenland, New Castle, Madbury	Portsmouth Municipal Water Department

Table 1: Pilot Study Data

The source data used for this pilot and the data sets generated by the analyses described below have been provided under separate cover. A complete listing of these data sets is provided in Appendix I.

TOWN OF RYE

The town of Rye has two water utility providers, Rye Village Water District (RVWD) and Aquarion. The former is an incorporated municipal entity with the right to levy property taxes to support its water purification and distribution operations while the latter is a private company. Records on properties served by public water systems generally are maintained within the municipality's CAMA system. The town of Rye provided an export from their CAMA system containing the following fields:

- Map, Block, Lot, Unit
- Address
- Property Owners Name
- Land Use Code
- Use Description
- Water Identifier (Public, Well, NULL)
- Sewer Identifier (Septic, Public Sewer, NULL)

The data provided by RVWD was correlated with the 2018 NH Department of Revenue Administration (DRA) Parcel Mosaic with the results shown in Table 2 and Figure 1. These provide a tabular and geospatial listing of properties that are served by public water. Properties not shaded in Figure 1 are likely using a private well.

Data Set	Total Records	Public Water
CAMA	3,200	2,351
Parcels	2,900	2,279

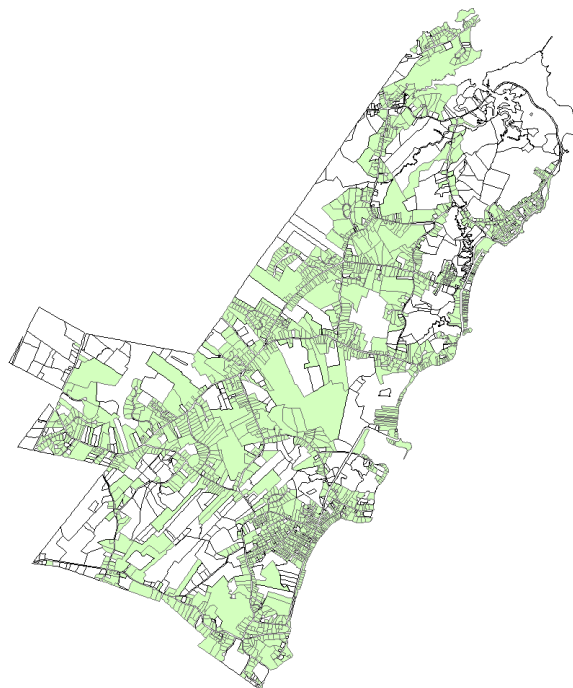**Table 2: Rye CAMA utility export correlation with 2018 DRA Parcel Mosaic Data**

Figure 1: Graphical Representation of Parcels Served by Public Water in Rye

The CAMA export provided by the town of Londonderry from the 2018 Parcel Mosaic contained utility information indicating if a property is served by public utilities, including water. Unlike Rye, Londonderry does not have an incorporated village water district. While the public water attribute is tracked within the CAMA database, it is not explicitly necessary for property tax billing and therefore may have varying accuracy. It should be noted that there are two public water providers in Londonderry, Pennichuck and the City of Manchester Water Department. The results of the 2018 CAMA data are provided in Table 3 and Figure 2.

Data Set	Records	Public Water
CAMA	10,253	2,123
Parcels	10,238	2,112

Table 3: Londonderry CAMA Export Correlation with 2018 DRA Parcel Mosaic Data
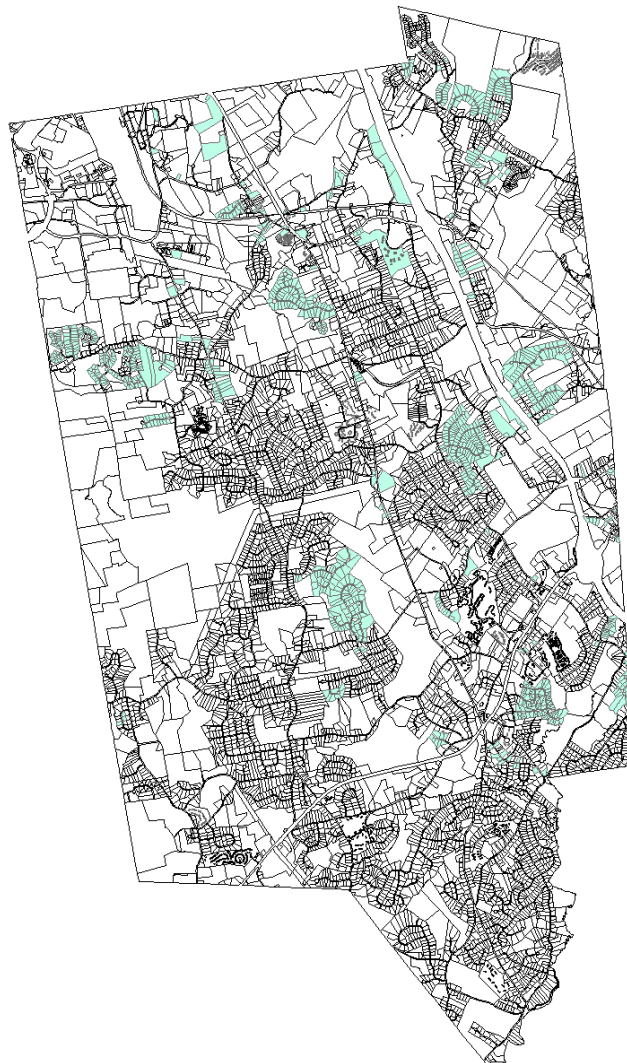

Figure 2: Graphical Representation of Londonderry CAMA Public Water

UTILITY SERVICE PROVIDER GEOSPATIAL DATA

PENNICHUCK

Geospatial utility network data from Pennichuck, a private water utility, was provided for their entire New Hampshire service area (see Figure 3). A list of the geospatial files provided is shown in Table 4.

| Name	Type	Description
Districts	Polygon	Defines service district boundary
Plan Area	Polygon	Areas where they have plan drawings
wStructure	Points	Locations of water structures
WaterMain	Polyline	Water Main
Lateral Line	Polyline	Lateral Line
wAbandonedLine	Polyline	Abandoned Line
wFitting	Points	Water Fitting
wHydrant	Points	Water Hydrant
wNetworkStructure	Points	Water Network Structure
wSampling Station	Points	Water Sampling Station
wSystem Valve	Points	Water System Valve
eCurbStopValve	Points	Water Curb Stop Valve

Table 4: Pennichuck Geospatial files provided
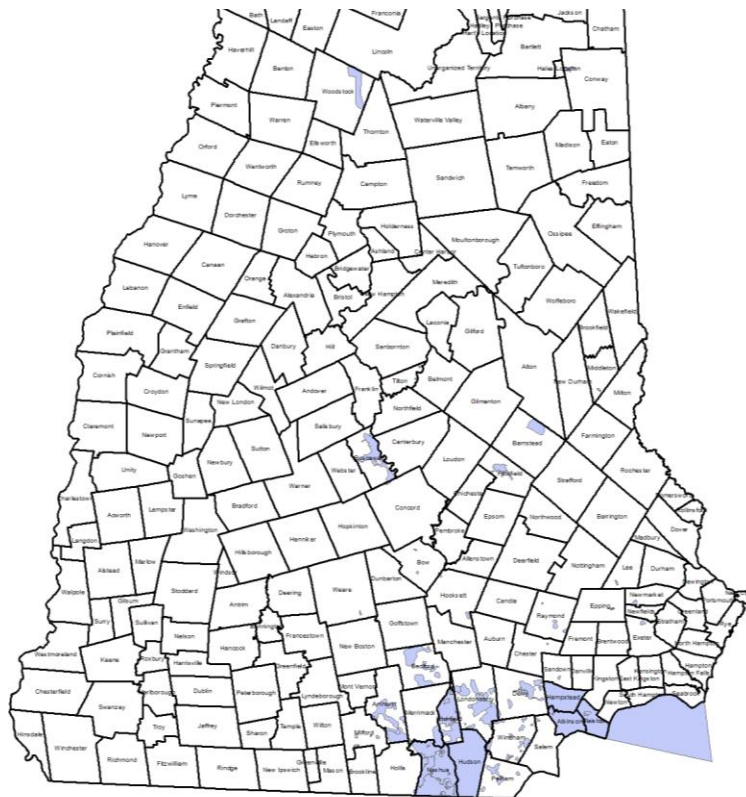
Figure 3: Pennichuck Service Territory

To facilitate a comparative analysis, Pennichuck data for the town of Londonderry was extracted and analyzed. The geospatial data and service area in Londonderry are shown in Figure 4. The primary files used in the analysis were the districts (polygon), plan areas (polygon), curb stop valves (points), and laterals (lines).

Figure 4: Pennichuck Data for Londonderry

To identify parcels that may be served using the geospatial data, a spatial analysis was performed using the following criteria:

- Parcel is within the district boundaries
- Parcel is within the plan area boundaries
- Parcel has a curb stop valve on it
- Parcel has a curb stop valve within 20'

The data provided by Pennichuck for Londonderry was correlated with the 2018 DRA Parcel Mosaic with the results shown in Table 5.

Data Set	Total Records	Within District	Within Plan Area	Curb Stop Valve on Parcel	Curb Stop Valve within 20' of Parcel
Parcels	10,238	3,431	3,477	723	1,486

Table 5: Pennichuck spatial analysis

PORTSMOUTH WATER DEPARTMENT

Geospatial utility network data was provided by the Portsmouth Water Department (PWD) which is managed by the City of Portsmouth. A list of the geospatial files provided is shown in Table 6.

Name	Type	Desc
WaterPipe	Line	Primary pipe file
WaterConnect1	Points	Primary service points
WaterConnect2	Points	Couples, unions, reducers
WaterConnect3	Points	Saddle and tee connections
WaterConnect4	Points	Locations of water structures
WaterNet_Junctions	Points	Junctions
PeripheralPoint	Points	Tie points
PeripheralLine	Line	Tie lines
PeripheralPolygon	Polygon	Unknown

Table 6: Portsmouth Water Geospatial Files

Based on spatial and visual analysis it was determined that the geospatial data provided by PWD would not correlate well to the DRA Parcel Mosaic. The discrepancies are illustrated in Figure 5 which shows Market Square in downtown Portsmouth. The purple points and lines are the PWD network data and the peach polygons are the Parcel Mosaic parcels. The PWD assets do not predictably exist on or near parcels serviced by public water and as such they cannot be reliably associated with individual parcels. For this reason, the correlation was not conducted with the spatial data. PWD also provided a list of their service addresses to the project team, and these addresses provided the basis for the PWD analysis outlined in the following section.
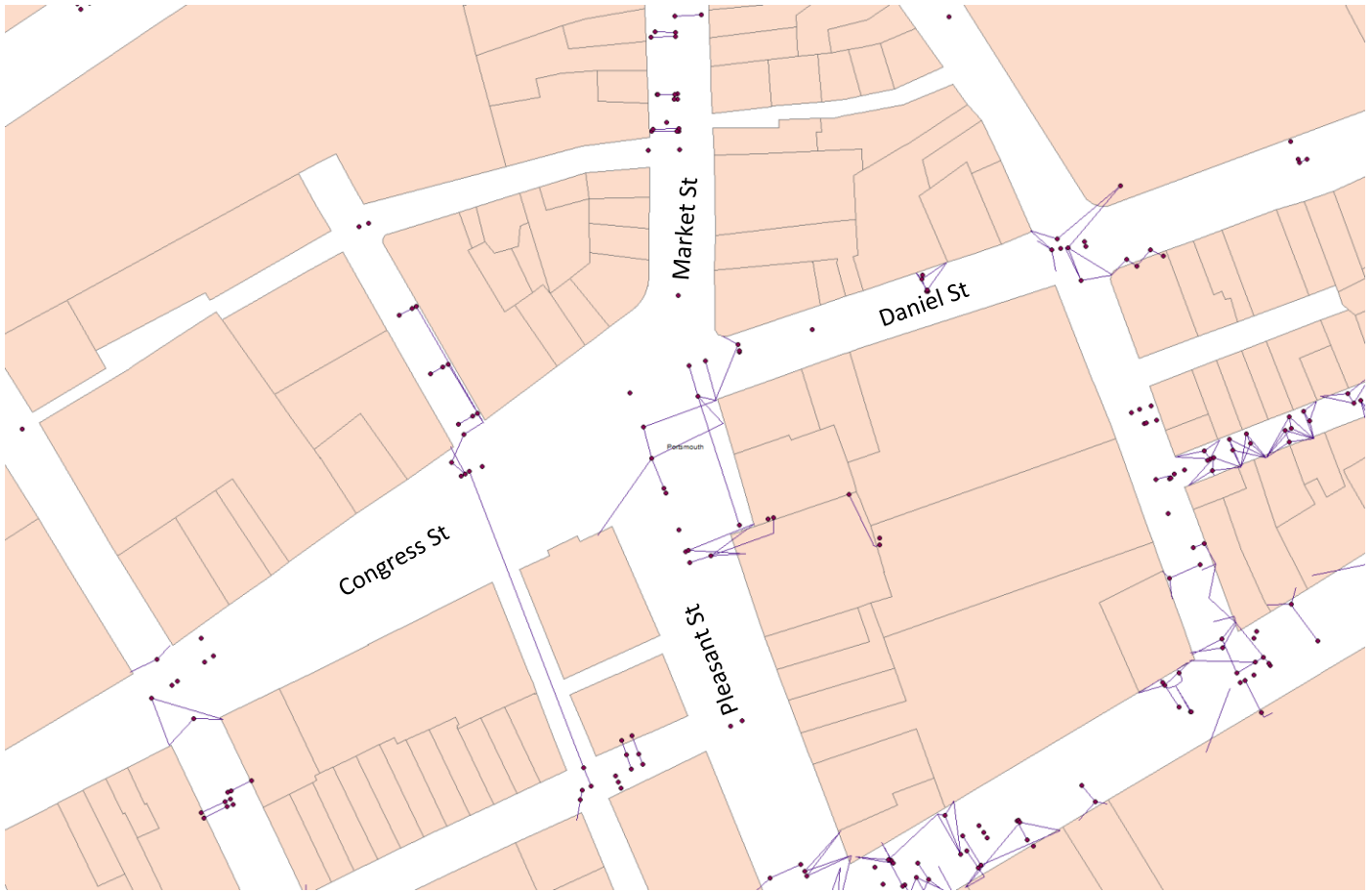
Figure 5: Portsmouth network data compared with geospatial parcels

UTILITY SERVICE PROVIDER ADDRESS INFORMATION

As noted above, service address data was also provided by the PWD. The file contained a street address (without City, State or Zip) which proved to be problematic as PWD serves locations in multiple municipalities (including Greenland, Portsmouth, Madbury, Newington, New Castle and Rye).

To correlate the service address information with the DRA Parcel Mosaic it was necessary to first geocode the tabular data. The first step was to create a composite geocoder using DRA Parcel Mosaic and related CAMA Data and US Census TIGER road centerline files. The composite geocoder was set to default to parcel addresses for primary geocoding and to use the TIGER centerline if a match could not be found. This method provides reliable results when geocoding single family homes, but is less reliable for apartments, condos and other types of multi-unit residences.

To provide more accurate geocoding, advanced postprocessing of Parcel Mosaic data in the PWD service region was conducted to develop parcel range addresses. Parcel range addresses enable geolocation of unit-based addresses like

Figure 6.

Figure 6: Range Address Parcels

To create range addresses, the existing single address field needed to be broken into its appropriate address pieces (street number, street name, street suffix, etc.). This was done by parsing the single address field based on spacing as shown in Figure 7.

Figure 7: Parsing single address field

Following the parsing of data, standardization of abbreviations was performed (ex. 'Street'→'ST','Terrace'→'TER', etc.). Once the address information was parsed and standardized, the data was spatially processed to determine the appropriate range addressing for each parcel.

To geocode the water service points the file provided by PWD was also parsed and standardized. Once this was performed the data was formatted in a similar fashion to the parcel data as shown in Figure 8.

Figure 8: Water service point data parsed and cleaned

Following the cleaning of the water service points, the data was geocoded to the range-addressed parcels using a custom algorithm as follows:

- Accounts are matched to parcels on exact matching of StName and StType, and the water accounts' AddressNumber falling within the parcels' AddressLow and AddressHigh range, and on the same side of the street.
- Accounts are matched to unmatched parcels if the water Address falls in the parcel Address range (and is on the same side of street), and the fields Address2, Address3, and Address4 all match.
- Accounts are matched to unmatched parcels if the water Address falls in the parcel Address range, is on the same side of street, and the fields Address2 and Address3 also match.
- Accounts are geocoded based on street centerlines and assigned to parcels that the geocoded point falls within.
- Accounts are then matched to the nearest parcel with a building on it (defined as BLD_G_AREA>0 from CAMA), *IF* that parcel is within 100' of the geocoded location (proximal).

The results of the analysis when correlated with the 2018 Parcel Mosaic are shown in Table 7.

Data Set	Records	Address 1	Address 2	Address 3	Centerline	Proximal
Service Address	8,221	7,668	7	69	137	80
		93.27%	.09%	.84%	1.67%	.97%

Table 7: Custom Geocoding Results

CAMA: Using municipal CAMA data to determine if a parcel is served by a public utility is inconsistent in its accuracy and availability across the state. For both Rye and Londonderry, it was unclear when parcels are indicated as having water service, who the service provider is. For example, in Rye, water service can be from Rye Village Water District or from Aquarion, and in Londonderry it can be Pennichuck or the City of Manchester Water Department. Utility data is currently provided in CAMA exports by a limited number of municipalities. Capturing and integrating utility attribution for municipalities that are not currently providing it would require substantial modification to the existing DRA Mosaic data conversion process which transforms local CAMA exports into the standardized statewide Mosaic schema.

Network Data: Using geospatial network data to determine if parcels are served by public utilities is inconsistent due to the variability in accuracy and completeness the data. Utilizing network data is first problematic due to the fact that most water network assets are in the road right-of-way. This results in most assets not directly coinciding with a parcel. Second, data can vary greatly between providers. For example, Pennichuck's data was sufficient to correlate parcels to curb stop valves with no buffer, and to curb stop valves with a buffer of 20 feet. Utility network data should be obtained by NHDES when it is available, but it is not suitable for use in statewide identification of public water service.

Service Addresses: Using service address data to determine if parcels are served by public utilities is the most ubiquitous and easily deployed statewide. The service address data provided by PWD posed issues only because it did not indicate the municipality the service address was located in. When dealing with a multi-jurisdictional provider that can be problematic. The parcel range address method will prove reliable in all NH towns currently except Clarksville, Dalton, Lancaster, Milton, Sandwich and Wilmot due to their native parcel data format (PDF). The ability to supplement parcel range data with road centerline data line from TIGER and other sources allows for future accuracy improvements. Collecting service address data presents a low barrier for the utility providers and will enable NHDES to compile a geospatially enabled database of known service addresses statewide.

APPENDIX A. GEOSPATIAL FEATURE INFORMATION

Listing of geodatabase contents (Production.gdb) provided under separate cover:

Feature Dataset	Feature	Type	Description
Londonderry	Districts	Polygon	Pennichuck district boundary
	Lateral Line	Line	Pennichuck lateral lines
	Plan Area	Polygon	Pennichuck area drawing boundaries
	Water_Main	Line	Pennichuck water mains
	wCurbStopValve	Point	Pennichuck curb stop valves
	wCurbStopValve20buffer	Polygon	Pennichuck curb stop valves 20' buffer area
	Londonderry	Polygon	Mosaic parcels with a Yes/No indicator for water service for each of the following fields: -District -Plan -Valve (direct) -Valve (20' buffer) -CAMA

Portsmouth	RangeParcels	Polygon	Mosaic parcels with addresses ranges created for geocoding
	ServiceAddresses	Points	Geocoded PWD service points
Rye	Rye	Polygon	Mosaic parcels with a Yes/No indicator for water service.