

An ArcGIS Tool for Geographic Address Verification

Brittany Meunier, Derek Sutton – Citizens Energy Group

Mehmet Ercan – Arcadis

ESRI

User Conference

June 27-July 1, 2016 | San Diego, CA

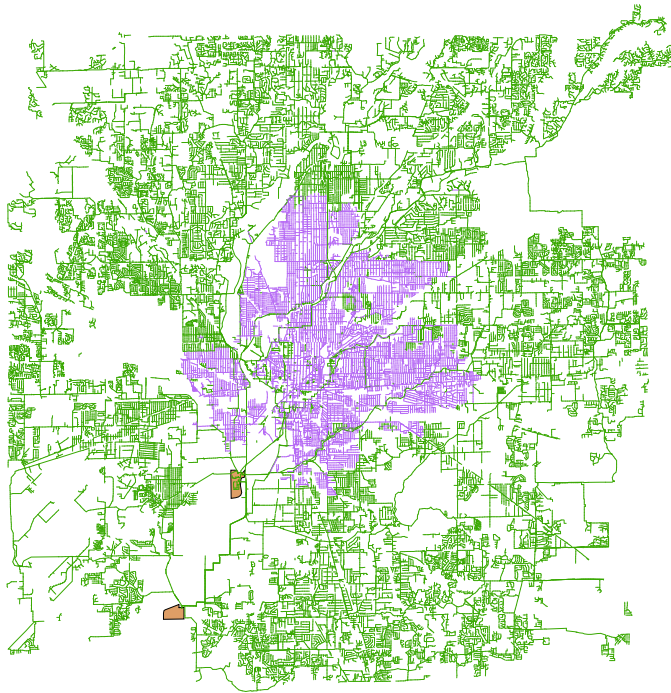
Outline

- Motivation
 - Importance and Challenges of managing the GIS network for utilities
 - Spatial Issues in Customer Addresses
- Methodology
 - Use of More Confident GIS Layers
 - Building the Address Comparison Tool
- Discussions
 - Applications
- Conclusions

Outline

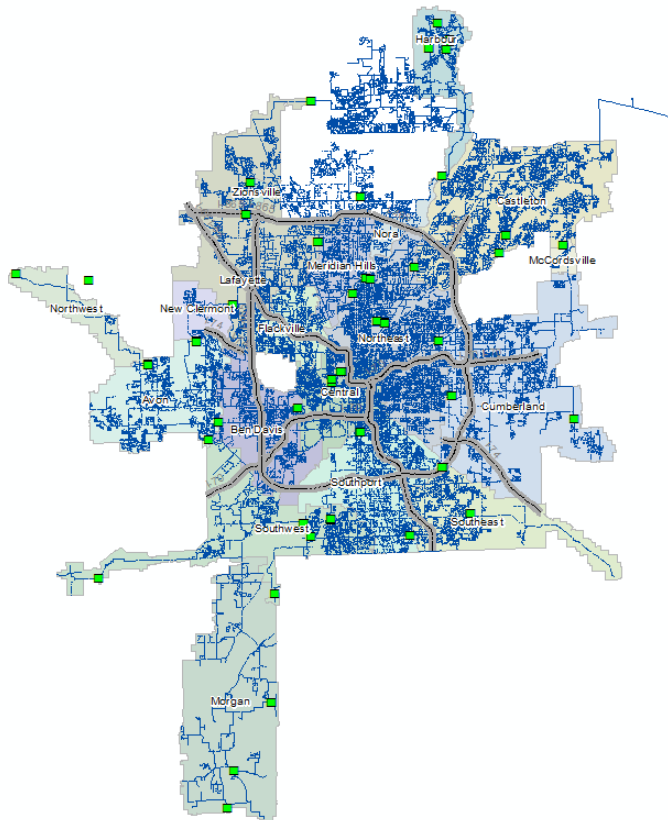
- Motivation
 - Importance and Challenges of managing the GIS network for utilities
 - Spatial Issues in Customer Addresses
- Methodology
 - Use of More Confident GIS Layers
 - Building the Address Comparison Tool
- Discussions
 - Applications
- Conclusions

Managing the GIS Network for Citizens - Sewer



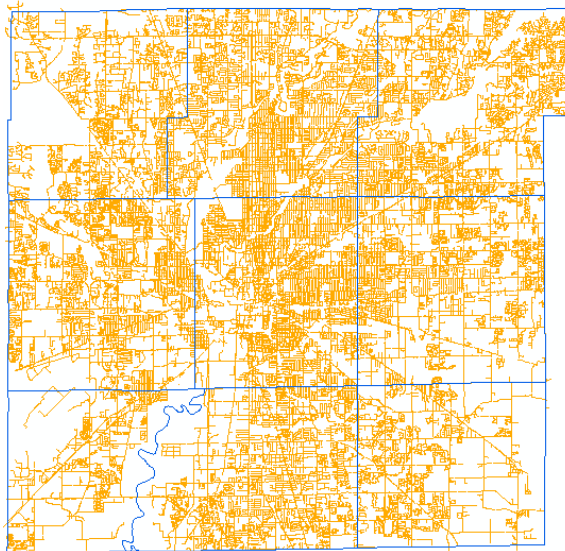
Feature Statistics for Citizens Sewer Assets	
Item	Number
Number of Structures	>75,000
Pipes in Miles	3,175
Pipes in Segments	75,000
Lift Stations	269
CSOs	130
Permanent Meter Locations	170

Managing the GIS Network for Citizens - Water



Feature Statistics for Citizens Water Assets	
Item	Number
Pressure Districts	19
Treatment Plants	9
Booster Stations	21
Storage Tanks	16
Pipes in Miles	>4,300

Managing the GIS Network for Citizens - Gas



Feature Statistics for Citizens Gas Assets

Item	Number
Number of Services	272,706
Miles of Distribution	4,157
Miles of Transmission	268
Regulating Stations	20

Spatial Issues in Customer Addresses

- New customers added into the billing system and GIS every month
- Manual entry introduces possible errors
- Spatial errors causing problems
 - Locating customer for maintenance
 - Calculating consumption for an area
 - Locating non-billed users
- Some layers like Indiana Department of Homeland Security (IDHS) address points are more accurate which can be used to address spatial issues by using coordinates of matching addresses
 - Using current ArcGIS tools does not achieve above 60% match rate

Outline

- Motivation
 - Importance and Challenges of managing the GIS network for utilities
 - Spatial Issues in Consumer Addresses
- Methodology
 - Use of More Confident GIS Layers
 - Building the Address Comparison Tool
- Discussions
 - Applications
- Conclusions

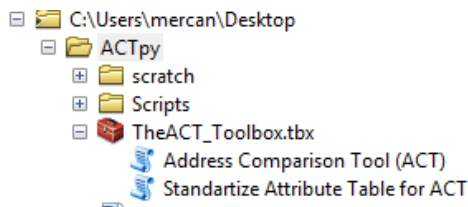
Use of more Confident GIS Layers

- Citizens is more confident in spatial accuracy of some GIS layers including
 - IDHS address points
 - Property owner (parcel) polygons maintained by the City or County
 - Unit addresses maintained for apartments, townhomes, etc.
- The spatial information extracted from these layers could be used to verify manually entered or missing customer coordinates
- Since manual comparison could bring extra errors and would take incredible amount of time, an automated process was necessary

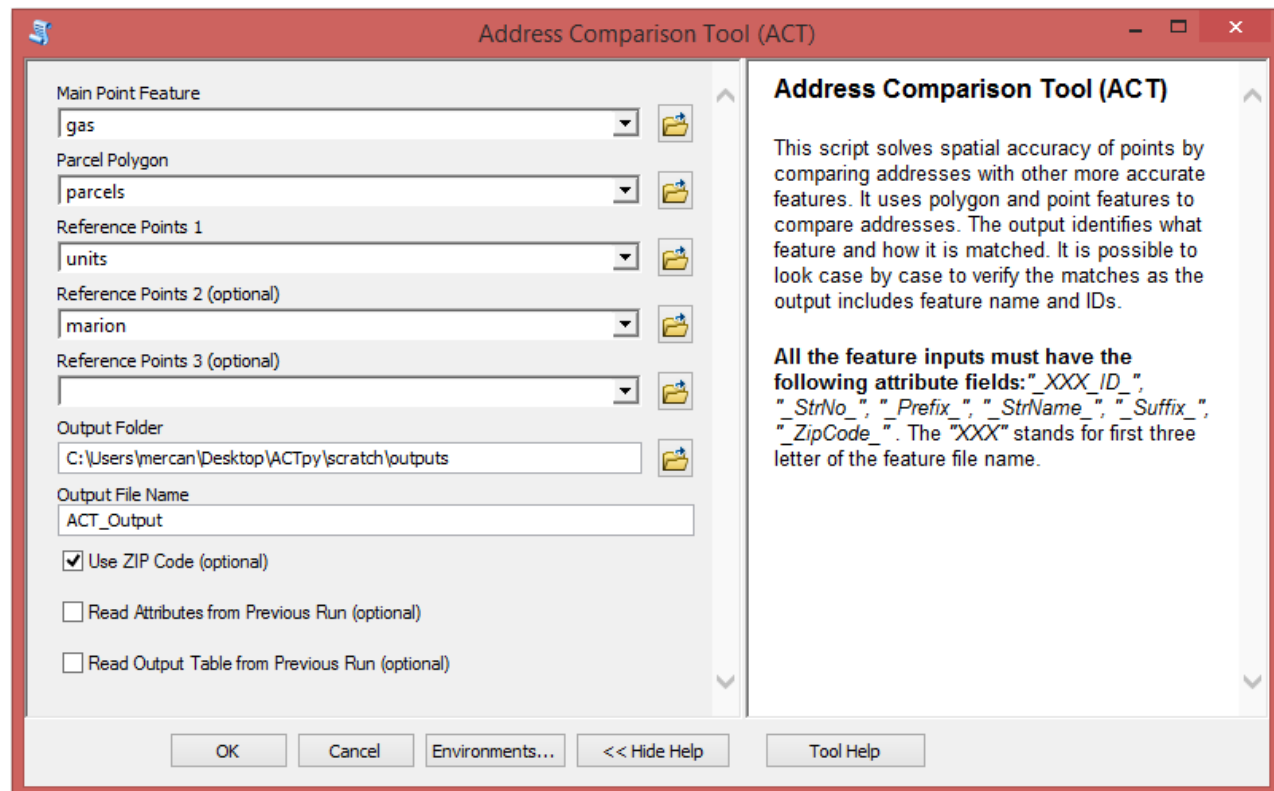
Building the Address Comparison Tool

- Started addressing spatial inaccuracies in new & existing sewer customers
 - Realized existing ArcGIS tools are not customized to Citizens input layers
 - The rate of match using existing tools could not achieve more than 60%
 - Considering total of 270,000 customers, there would be about 108,000 to manually check
- Decided to create customized ArcGIS tool leveraging ArcPy libraries
- Later adjusted the script to be used for comparing Gas customer points
- When it was decided to be used for water as well, it was generalized in a Python library called ACTpy
- Finally, an ArcGIS toolbox was created using ACTpy library

Building the Address Comparison Tool

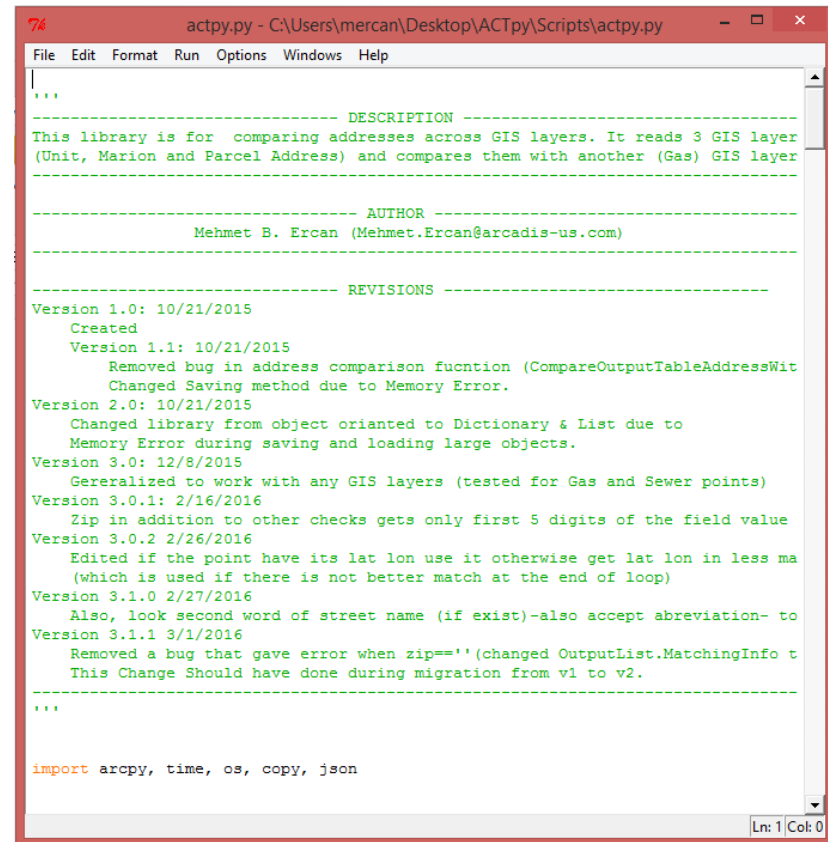


- User friendly and same as regular ArcGIS tools
- Each tool input has explanation



Building the Address Comparison Tool

- The script has improved a great deal since it was created
- Some of the major revisions:
 - Inefficient calculations addressed
 - Added ability to re-start broken run
 - Data was stored in object oriented way and changed to dictionary and list orientation because of memory problems

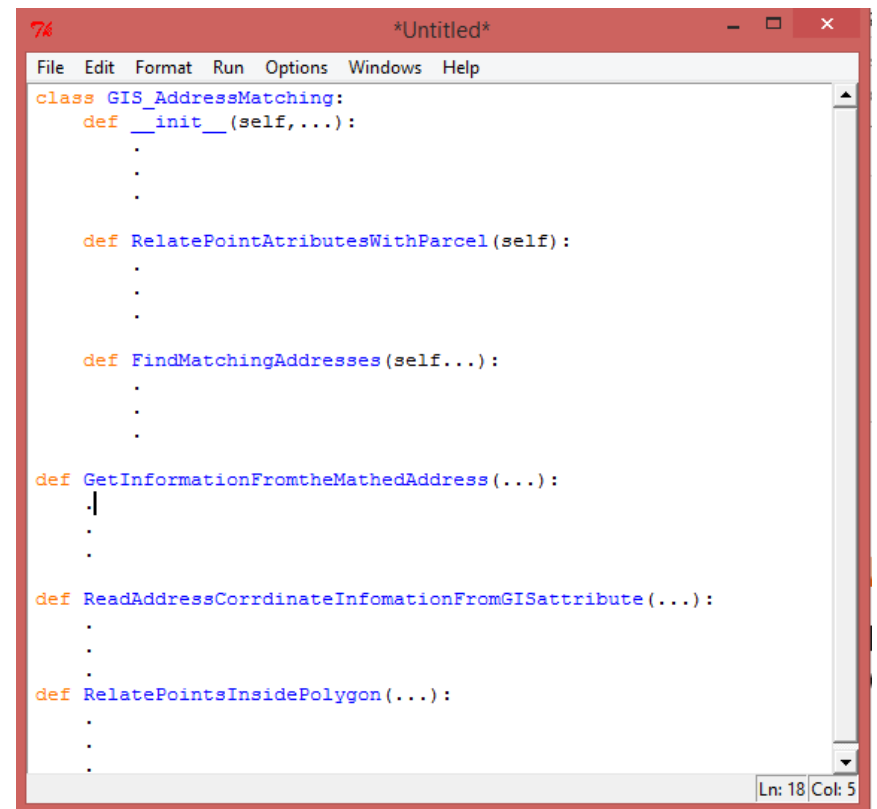


```
actpy.py - C:\Users\mercan\Desktop\ACTpy\Scripts\actpy.py
File Edit Format Run Options Windows Help
'''
----- DESCRIPTION -----
This library is for comparing addresses across GIS layers. It reads 3 GIS layer
(Unit, Marion and Parcel Address) and compares them with another (Gas) GIS layer
-----
----- AUTHOR -----
Mehmet B. Ercan (Mehmet.Ercan@arcadis-us.com)
-----
----- REVISIONS -----
Version 1.0: 10/21/2015
Created
Version 1.1: 10/21/2015
Removed bug in address comparison fuction (CompareOutputTableAddressWith
Changed Saving method due to Memory Error.
Version 2.0: 10/21/2015
Changed library from object orianted to Dictionary & List due to
Memory Error during saving and loading large objects.
Version 3.0: 12/8/2015
Generalized to work with any GIS layers (tested for Gas and Sewer points)
Version 3.0.1: 2/16/2016
Zip in addition to other checks gets only first 5 digits of the field value
Version 3.0.2 2/26/2016
Edited if the point have its lat lon use it otherwise get lat lon in less ma
(which is used if there is not better match at the end of loop)
Version 3.1.0 2/27/2016
Also, look second word of street name (if exist)-also accept abbreviation- to
Version 3.1.1 3/1/2016
Removed a bug that gave error when zip='' (changed OutputList.MatchingInfo t
This Change Should have done during migration from v1 to v2.
'''

import arcpy, time, os, copy, json
```

Building the Address Comparison Tool

- The Address Matching class has two functions
 - Relate Point Attributes With Parcel (self)
 - Find Matching Addresses (self)
- The three external functions used
 - Read Address Coordinate Information From GIS attribute
 - Relate Points Inside Polygon
 - Get Information From the Matched Address



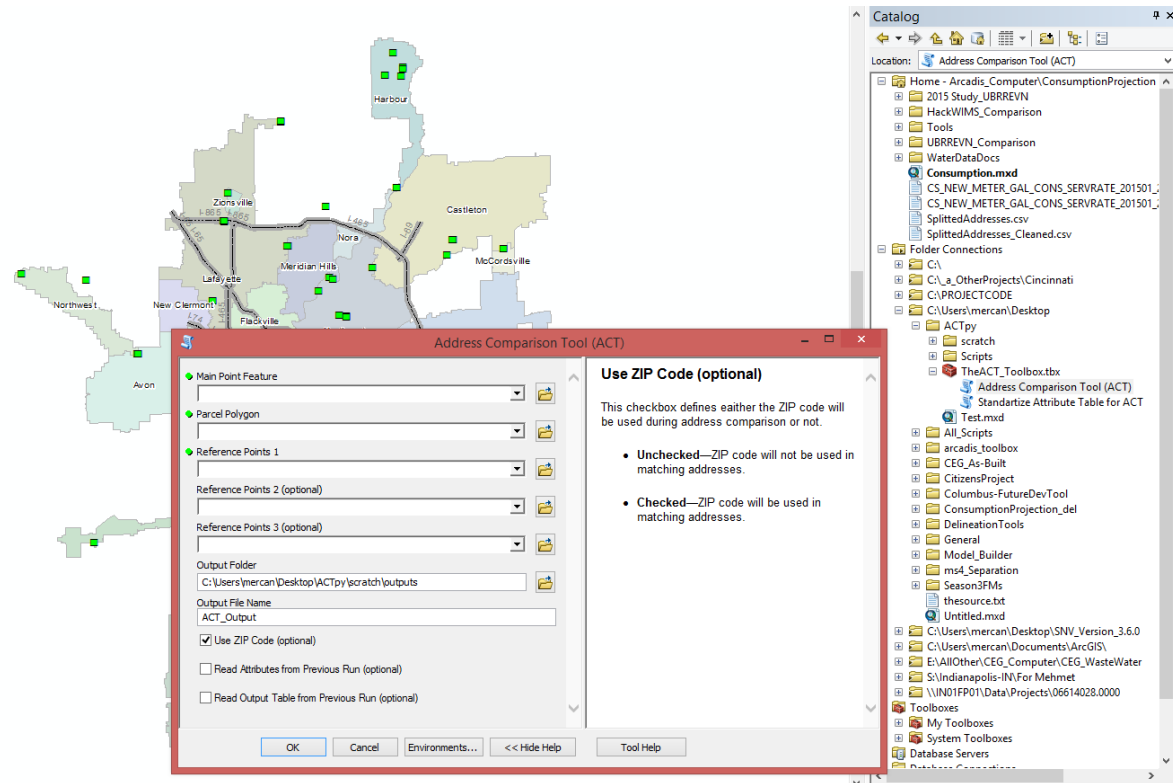
```
74 *Untitled*
File Edit Format Run Options Windows Help
class GIS_AddressMatching:
    def __init__(self,...):
        .
        .
        .
    def RelatePointAttributesWithParcel(self):
        .
        .
        .
    def FindMatchingAddresses(self...):
        .
        .
        .
def GetInformationFromtheMathedAddress(...):
    .|
    .
    .
def ReadAddressCorrdinateInfomationFromGISattribute(...):
    .
    .
    .
def RelatePointsInsidePolygon(...):
    .
    .
    .
Ln: 18 Col: 5
```

Outline

- Motivation
 - Importance and Challenges of managing the GIS network for utilities
 - Spatial Issues in Consumer Addresses
- Methodology
 - Use of More Confident GIS Layers
 - Building the Address Comparison Tool
- Discussions
 - Applications
- Conclusions

Applications

- Customizable
- Select layers
- User options
- Output files
- CSV
- Json files
- (to rerun where it left off)



Applications

- Successfully used in Sewer, Gas, and Water providing structured csv output
- Improved address matching rate from ~60% to ~90%
- Extra scripts written to process inputs and outputs

The screenshot shows a spreadsheet application with a data table and a file explorer window. The spreadsheet has columns A through I, with headers: FID, InParcel, MatchingInfo, MatchdUniqueID, IsInSameParcel, Flag, IsStNoMatch, IsPrfxMatch, IsStNmMatch, and IsSf:object. Row 6 is highlighted, showing FID 15, InParcel marion, MatchdUniqueID 10, IsInSameParcel No, Flag Match: Yes, IsStNoMatch Yes, IsPrfxMatch Yes, IsStNmMatch Yes, and IsSf:object No. The file explorer window shows a directory structure: ACTpy > Scripts > HelpScripts, with files: 1_SplitStreetNumber&Prefix.py, 2_SplitStreetName&Suffix.py, 2b_StandarizeAttributes4ArcGIStoolbox.py, 3_CreatPointshpFrmXYpoints_ReadCSV.py, 4_CreatPointshpFrmXYpoints_ReadJson.py, and info.txt.

	A	B	C	D	E	F	G	H	I	
1	FID	InParcel	MatchingInfo	MatchdUniqueID	IsInSameParcel	Flag	IsStNoMatch	IsPrfxMatch	IsStNmMatch	IsSf:object
2	11	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	Yes
3	10	299461	Parcel	93246	No	Match: Yes	Yes	Yes	Yes	Yes
4	13	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	Yes
5	12	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	Yes
6	15		marion	10	No	Match: Yes	Yes	Yes	Yes	No
7	14	245388	Parcel	245388	Yes	Match: Yes	Yes	Yes	Yes	Yes
8	1	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	Yes
9	0	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
10	3	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
11	2	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
12	5	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
13	4	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
14	7	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
15	6	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
16	9	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
17	8	299461	units	0	Yes	Match: Yes	Yes	Yes	Yes	No
18										

Outline

- Motivation
 - Importance and Challenges of managing the GIS network for utilities
 - Spatial Issues in Consumer Addresses
- Methodology
 - Use of More Confident GIS Layers
 - Building the Address Comparison Tool
- Discussions
 - Applications
- Conclusions

Conclusions

- The ACT tool was created to
 - Address spatial inaccuracies in customer points
 - Reduce the amount of time spent on these inaccuracies
 - Work smarter not harder
- The tool developed and improved using inputs from different Citizens GIS professionals
- It is designed to work for any GIS network (e.g. sewer, water, and gas)
- The ACT improved results compare to existing tools from ~60% to ~90% matching rate

Questions?



Mehmet Ercan
Mehmet.Ercan@Arcadis.com



© Arcadis 2015

Brittany Meunier
BMeunier@citizensenergygroup.com

Derek Sutton
DSutton@citizensenergygroup.com

