San Jose Microsoft Footprint to Parcel Analysis

Pacific Northwest
NATIONAL LABORATORY

Kevin Keene – kevin.keene@pnnl.gov

Nora Wang – nora.wang@pnnl.gov

Match Footprint (MS) to Parcel

Overlap Threshold	Matches
0.00	344,485
0.01	315,452
0.02	303,820
0.03	294,884
0.04	287,357
0.05	**281,160**
0.06	276,115
0.07	272.083
0.08	268,494
0.09	265,546
0.10	262,935
0.15	254,025

GIS matching to create ground truth of parcel-building relationships

1. INTERSECT MS buildings to SJ parcels
2. Calculate percent overlap
   - Intersect area / building footprint area
3. Filter out overlaps less than 5% (slight overlap is likely mistake)

Number of parcels: 247,600
Number of MS footprints: 241,433
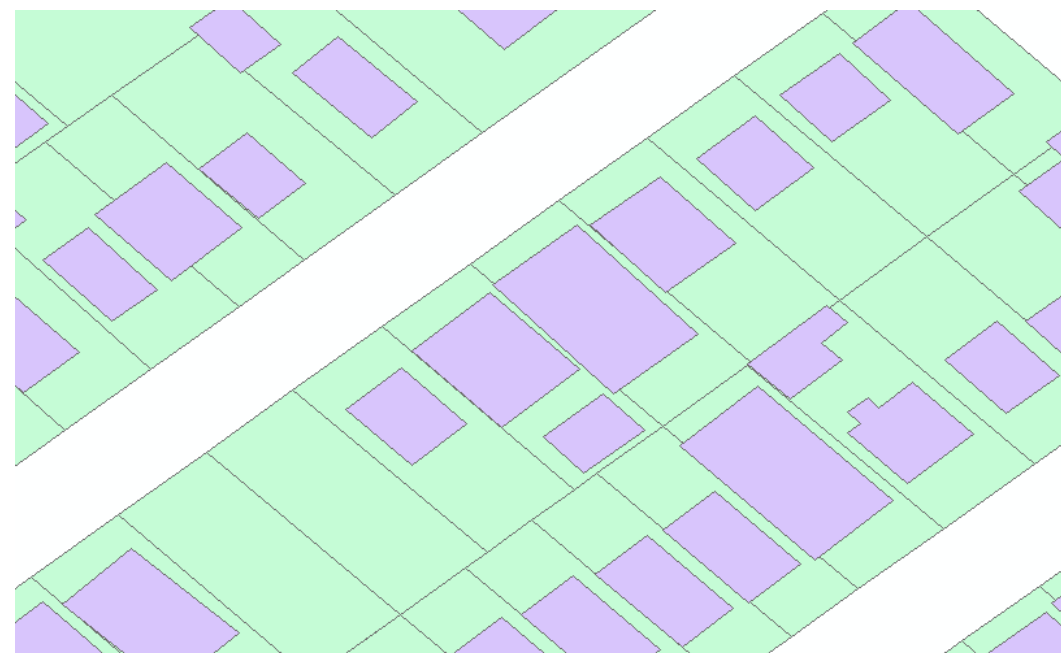
Relationships

- GIS, 0.05 overlap

Building-Parcel Relationship	# Intersections
One-to-one	147,732
Many-to-one	16,581
One-to-many	60,429
Many-to-many	56,418

UBID Cross Reference - % Footprint Overlap

- This is compared GIS intersection threshold of 5%
- This uses the area of intersection of bounding box divided by the area of the **building** bounding box as the criteria for UBID threshold

Overlap Threshold	# Matches to GIS	Total Intersections	Success Rate
0.0	281,160	997,385	0.28
0.1	280,968	826,908	0.34
0.2	278,306	699,105	0.40
0.3	274,222	623,937	0.43
0.4	270,985	563,520	0.47
0.5	267,071	498,747	0.52
0.6	261,451	430,555	0.58
0.7	255,240	383,817	0.62
0.8	249,442	353,900	0.65
0.9	239,115	328,000	0.65
1.0	237,954	325,501	0.65

Relationships

- UBID, overlap = 1.0

Building-Parcel Relationship	# Intersections
One-to-one	131,653
Many-to-one	41,091
One-to-many	39,117
Many-to-many	113,640

- GIS, overlap = 0.05

Building-Parcel Relationship	# Intersections
One-to-one	147,732
Many-to-one	16,581
One-to-many	60,429
Many-to-many	56,418

- Parcels like these account for some of the false matches…

- We can delete parcels with high area increases

Filtered Parcels, % Footprint Overlap

- Deleted parcels with percent area increase > 4 -> show different values at say 0.8 overlap
- Note that I didn't change the GIS parcels, so the success rate will be lowered some because the deleted parcels won't match with anything, but the reduction in false matches outweighs it

Overlap Threshold	# Matches	Total Intersections	Success Rate
0.0	276,769	934,562	0.29
0.1	276,584	767,898	0.36
0.2	274,019	643,113	0.42
0.3	269,984	570,343	0.46
0.4	266,806	512,053	0.51
0.5	262,962	449,317	0.56
0.6	257,457	383,443	0.63
0.7	251,409	338,712	0.68
0.8	245,775	310,687	0.71
0.9	235,703	287,021	0.71
1.0	234,572	284,845	0.71

Relationships

- UBID, overlap = 1.0

Building-Parcel Relationship	# Intersections
One-to-one	131,653
Many-to-one	41,091
One-to-many	39,117
Many-to-many	113,640

- UBID, 1.0 overlap, <4x area increase

Building-Parcel Relationship	# Intersections
One-to-one	146,789
Many-to-one	25,499
One-to-many	43,432
Many-to-many	69,125

- GIS, 0.05 overlap

Building-Parcel Relationship	# Intersections
One-to-one	147,732
Many-to-one	16,581
One-to-many	60,429
Many-to-many	56,418

- Problem: parcels that are have high area increase and high number of intersected buildings

- Indicators
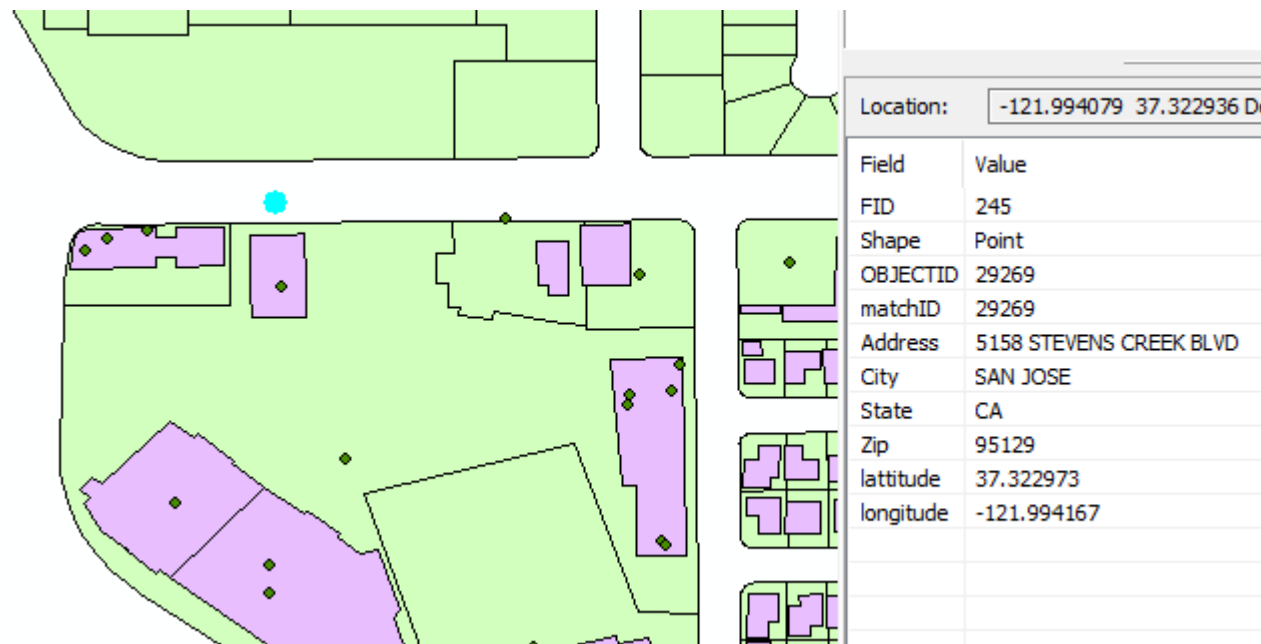 - Angle of rotation, % area increase, Overlap, Number of intersected buildings

# of buildings	BB % Increase	Angle Rotation	Overlap	Correct Matches	False Matches	Success
All	All	All	All	281,160	716,225	
All	< 4	All	All	276,769	657,793	

- UBID cross reference matches to parcel with nearest centroid
    - Must be within parcel BB to match

Jacob's Data
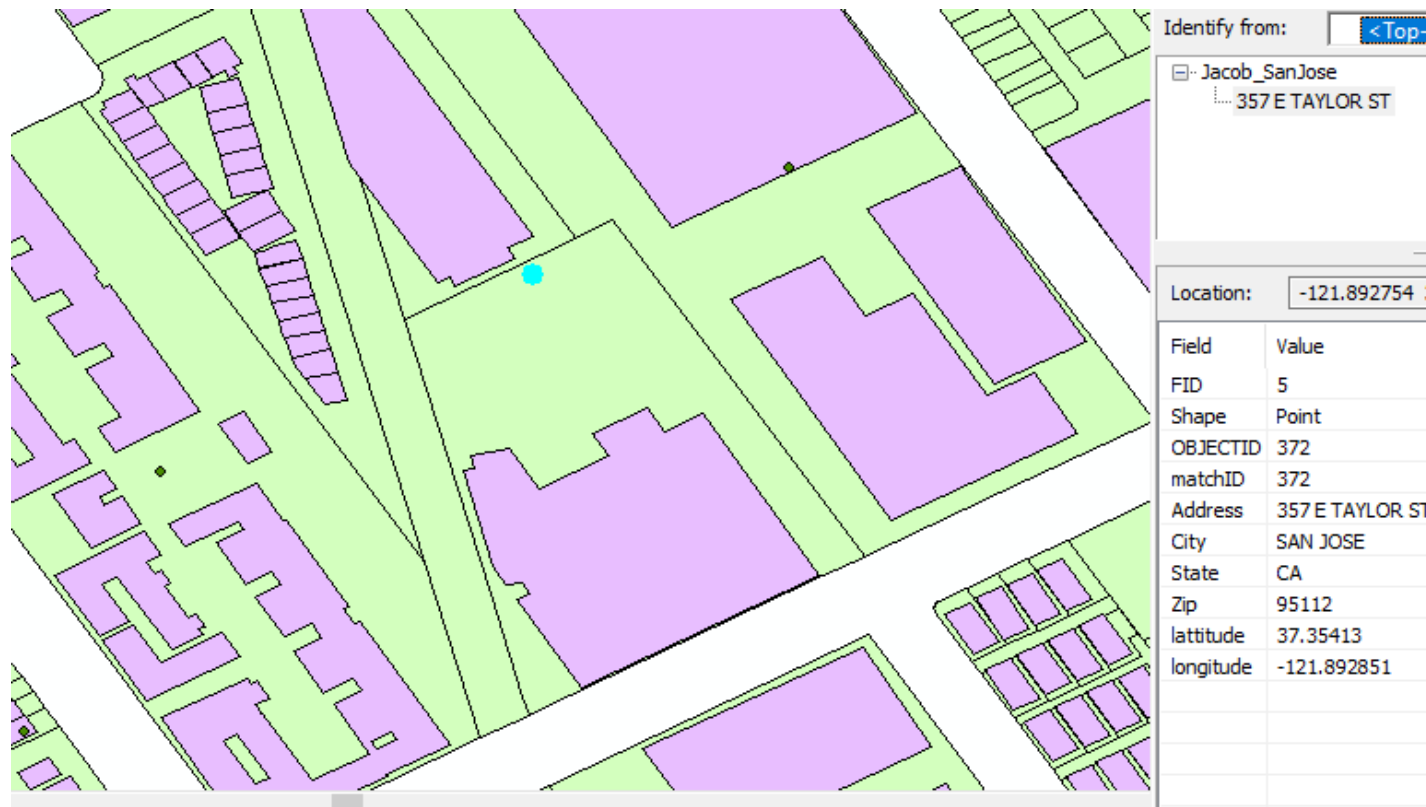
Jacob's Data

- 1764 addresses in San Jose from Jacob

  - 1753 are within a parcel footprint in GIS (99.2%)
    - ✓ 1761 are within a UBID parcel bounding box (99.8%)
      - See example below-left for address outside BB
    - ✓ UBID-GIS comparison
      - 1655/1753 matched correctly (94%)
        - See example below-right for incorrectly matched
        - Put UBID centroid on map

Blue address was matched to this parcel

Location:	-121.994079 37.322936 De
Field	Value
FID	245
Shape	Point
OBJECTID	29269
matchID	29269
Address	5158 STEVENS CREEK BLVD
City	SAN JOSE
State	CA
Zip	95129
lattitude	37.322973
longitude	-121.994167

Location:	-121.911299 37.382
Field	Value
FID	4091
Shape	Polygon ZM
OBJECTID_1	4092
PARCELID	534
INTID	534
APN	23729002
LOTNUM	
PARCELTYPE	Tax
PARC_AREA	455549.747123

- 1764 addresses in San Jose from Jacob

  - 1264 are within a building footprint in GIS (71.7%)
    - ✓ 1619 are within a UBID building bounding box
    - ✓ UBID-GIS comparison
      - With GIS inside footprint only – 1194/1264 (94%)
      - With GIS nearest match - 1483/1764 (84%)
        - Not perfect ground truth
        - See example below of building matched to nearest GIS but not in footprint

Jacob's Data

Thank you