Visualization of Origin-Destination Commuter Flow Using CTPP Data and ArcGIS

Research & Analysis Department
Southern California Association of Governments

2015 ESRI User Conference | July 23, 2015 | San Diego, CA

Tom Vo
JiSu Lee
Jung Seo

Dr. Frank Wen
Dr. Simon Choi
Contents

- What is SCAG?
- Objectives
- Methodology
- Results
- Conclusions
- Future Studies
6 counties and 191 cities

15 sub-regions

18.4 million people (2012)

38,000 square miles

16th largest economy in the world (GRP: $924 Billion in 2013)

Nation’s largest Metropolitan Planning Organization (MPO)
Objectives

- Identify work destinations for each jurisdiction in the SCAG region at the Census Tract level
  - 191 cities and 6 counties
- Visualize the spatial patterns of the major work destinations for each jurisdiction
  - To understand where residents of each jurisdiction are employed
- Provide informational data resources to local jurisdictions for planning purposes
  - E.g. 2013 Local Profiles: Planning data reports
The CTPP is a set of special tabulations designed by transportation planners using large sample surveys conducted by the Census Bureau.

Utilize continuous survey called American Community Survey (ACS)

Three components of CTPP:
- Part 1: Residence-based tabulations summarizing worker and household characteristics
- Part 2: Workplace-based tabulations summarizing worker characteristics
- Part 3: Worker flows between home and work, including travel mode

CTPP supports a wide variety of transportation planning tasks
CTPP Program
DATA Description

- Census 2010 shapefile:
  - Tiger/Line shapefile
  - Place and Tract
- CTPP 2006-2010 Census Tract Flows
  - Download the database file (.accdb)
    - Total of 4,156,426 records (151 MB)
Census 2010 Shapefiles

Census 2010
- Place & Tract
Methodology for Identifying Work Destination

- Census tract to census tract flows from the new CTPP 2006-2010, using **CTPP 2006 – 2010 Census Tract Flows Query**

- Statistical Analysis System (SAS) program with CTPP raw datasets
  - To select census tract-to-tract commuter data within SCAG’s region
  - To generate top 10 working destinations at the Census Tract level for each jurisdiction in SCAG region
Combination of SAS and CTPP Datasets

- CTPP – ACS 2006/10
  - Origin-Destination (OD)
  - Residence Area & Workplace Area
- Total worker counts
- Enumerated by 2010 Census Tracts
CTPP Variable

CTPP raw dataset description

<table>
<thead>
<tr>
<th>Pos</th>
<th>Variable</th>
<th>Type</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>State_R</td>
<td>Number</td>
<td>FIPS code for residence state</td>
</tr>
<tr>
<td>2</td>
<td>County_R</td>
<td>Number</td>
<td>FIPS code for residence county</td>
</tr>
<tr>
<td>3</td>
<td>Tract_R</td>
<td>Number</td>
<td>FIPS code for residence tract</td>
</tr>
<tr>
<td>4</td>
<td>StateFP_W</td>
<td>Number</td>
<td>FIPS code for workplace state</td>
</tr>
<tr>
<td>5</td>
<td>CountyFP_W</td>
<td>Number</td>
<td>FIPS code for workplace county</td>
</tr>
<tr>
<td>6</td>
<td>TractFP_W</td>
<td>Number</td>
<td>FIPS code for workplace tract</td>
</tr>
<tr>
<td>7</td>
<td>EST</td>
<td>Number</td>
<td>Estimate of commuter</td>
</tr>
<tr>
<td>8</td>
<td>MOE</td>
<td>Number</td>
<td>Its associated margins of error</td>
</tr>
</tbody>
</table>
1. Import the CTPP main OD file and import correspondence table
2. Select SCAG's region from the OD file
3. Merge CTPP OD file with the correspondence table
4. Estimate new commuter (EST_NEW) number by Area Weighted Interpolation
   \[
   \text{EST\_New} = \frac{\text{tract\_acre}}{\text{Total Acre}} \times \text{EST}
   \]
   *tract\_acre: split tract after intersecting place and tract shapefile to create the correspondence table
   *Total Acre: original tract acre
   *EST: number of commuter
5. Rank EST\_New
6. Determine top 10 OD
## Top 10 OD Table

<table>
<thead>
<tr>
<th>Rank</th>
<th>Location</th>
<th>Distance (miles)</th>
<th>TRIP (in millions)</th>
<th>ZCTA</th>
<th>Total Area (acres)</th>
<th>Tract Area (acres)</th>
<th>Population</th>
<th>Propensity</th>
<th>EST Prop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>2</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>3</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>4</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>5</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>6</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>7</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>8</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>9</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
<tr>
<td>10</td>
<td>Aiken Co</td>
<td>10.2</td>
<td>3.1</td>
<td>102</td>
<td>341,620</td>
<td>35,669</td>
<td>2,461,680</td>
<td>34.6</td>
<td>0.0592</td>
</tr>
</tbody>
</table>
## Advantage and Disadvantage

<table>
<thead>
<tr>
<th></th>
<th>CTPP</th>
<th>Programming Scripts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td>▪ Easy-to-use interface and able to visualize data instantly</td>
<td>▪ Efficient to manipulate data for multiple areas at once</td>
</tr>
<tr>
<td></td>
<td>▪ Easy to perform analyses at multiple levels of geography</td>
<td>▪ Improve efficiency of managing and processing big data</td>
</tr>
<tr>
<td><strong>Cons</strong></td>
<td>▪ Inefficient to search/download for multiple locations</td>
<td>▪ Requires programming skills</td>
</tr>
<tr>
<td></td>
<td>▪ Requires GIS skills</td>
<td>▪ Requires GIS skills</td>
</tr>
</tbody>
</table>
Identifying Major Work Destinations

- O-D Analysis Steps

1. Correspondence between Census Tract codes and Place codes
2. Merging O-D dataset and Place codes by Census Tract codes
3. Sorting dataset in descending order
4. Visualization: CTPP OD Commuter Map
5. City to Tract OD Table
6. Rank Top 10 Work Destination
Visualizing Major Work Destinations

- Using ESRI ArcGIS application
  - *Data Driven Pages* – To create a multi-page map series from a single map document
- Using *Python* programming language
  - *Python* – Interpreted, object-oriented, high-level general-purpose programming language
  - Widely utilize in ArcGIS environment for automation process
Visualizing Major Work Destinations

- Utilizing ArcGIS and Python programming language to create a series of work destination maps for all jurisdictions in SCAG region
  - Automation process
  - Consistent and efficient process
Python and ArcGIS

- Importing ArcPy sitepackage in Python to automatically create work destination maps at the Census Tract level
  - Data Driven Page, Definition Query, Spatial Analyst, and Exporting functions in ArcGIS
    - Loop function (for row in cursor:)
    - Search Cursor (arcpy.da.SearchCursor)
    - Data Driven Page (mxd.dataDrivenPages.currentPageID)
    - Definition Query (lyr.definitionQuery = sqlExp)
    - Exporting (arcpy.mapping.ExportToJPEG)
O-D Commuter Flow Maps (Cont.)
Conclusions

- Significantly efficient in processing O-D analysis with SAS for numerous jurisdictions
- Python and ArcGIS improve the processing time and accuracy of map production as a set of maps or individual map
- Creating OD commuting flow at the Census Tract level helps to understand the travel pattern better
Future Studies

- Work destinations at different geographical level
- Relationship between demographic groups and estimated travel distance
- Include land use data
- Land use-weighted interpolation method to increase accuracy of the place of residence and work