

Methodology and Automation in Estimating Area Type for Transportation Modeling

By Petya Maneva and Wang Zhang, Maricopa Association of Governments

Summary

For the purpose of this paper we define the economic intensity of an area as a function of certain socio-economic characteristics such as employment, population, and land use. High economic intensity is often associated with high traffic volumes whereas low economic intensity is associated with low traffic volumes. Hence economic intensity can be used to estimate both capacities and free flow travel speeds within a travel demand model.

This paper describes the Maricopa Association of Governments (MAG) new methodology to delineate area type polygons and the subsequent incorporation of the area type at the line segment level for the modeled transportation network. A seamless integration of ArcInfo, Python and GISDK allows the user to automatically assign an area type to roadway segments based on the land use and socio-economic characteristics of the surrounding territory.

The procedure determines the economic intensity associated with a centroid for the developed area of land for each transportation analysis zone. A raster type economic intensity surface is generated, reclassified and converted into area type polygons. Transportation network links are then automatically assigned area types based on their geographic location. The application is easy to customize and can leverage additional GIS data as appropriate.

Role of Area Type and Economic Intensity in Travel Demand Modeling

Area type has traditionally been used in travel demand modeling as an indicator of economic intensity. In many travel demand forecasting models area type affects a roadway's capacity, free-flow speed, and the parameters of the Volume-Delay Function (VDF) as seen in *Tables 1 and 2*. A more clearly defined area type can improve the fidelity of a travel demand model and provide for better validations in terms of speeds, flows and delays.

Table 1: *Volume-Delay Function Codes with Area Type and Functional Class*

Area Type	Functional Class									
	Freeways (1)	Expressways (2)	Collectors (3)	6 Leg Arterials (4)	Centroid Connectors (5)	Major Arterials (6)	Unmetered Ramps (7)	Metered Ramps (8)	C/D Roads (9)	HOV Lane (0)
(1) CBD	11	12	13	14	15	16	17	18	19	10
(2) Outlying CBD	21	22	23	24	25	26	27	28	29	20
(3) Mixed Urban	31	32	33	34	35	36	37	38	39	30
(4) Suburban	41	42	43	44	45	46	47	48	49	40
(5) Rural	51	52	53	54	55	56	57	58	59	50

Table 2: Free Flow Speed and Volume-Delay Function Parameters

Area Type	Code	Functional Class	Free-flow Speed (mph)	BPR	
				Alpha	Beta
1 CBD (Central Business District)	0	HOV Lane	65	0.74	4.60
	1	Freeway	62	0.87	5.00
	2	Expressway	47	0.96	2.30
	3	Collector	25	0.96	2.30
	4	6-Leg Arterial	36	0.96	2.30
	5	Centroid Connector	15	--	--
	6	Major Arterial	34	0.96	2.30
	7	Unmetered Ramp	39	0.96	2.30
	8	Metered Ramp	30	0.96	2.30
	9	C/D Road	50	0.87	5.00
2 Outlying Central Business District	0	HOV Lane	68	1.88	3.10
	1	Freeway	65	0.75	4.00
	2	Expressway	49	1.13	2.64
	3	Collector	28	1.13	2.64
	4	6-Leg Arterial	37	1.13	2.64
	5	Centroid Connector	20	--	--
	6	Major Arterial	39	1.13	2.64
	7	Unmetered Ramp	40	1.13	2.64
	8	Metered Ramp	30	1.13	2.64
	9	C/D Road	53	0.75	4.00
3 Mixed Urban	0	HOV Lane	75	1.88	3.10
	1	Freeway	71	0.71	3.47
	2	Expressway	55	0.73	2.36
	3	Collector	37	0.73	2.36
	4	6-Leg Arterial	39	0.73	2.36
	5	Centroid Connector	25	--	--
	6	Major Arterial	39	0.73	2.36
	7	Unmetered Ramp	44	0.73	2.36
	8	Metered Ramp	30	0.73	2.36
	9	C/D Road	57	0.71	3.47
4 Suburban	0	HOV Lane	75	1.88	3.10
	1	Freeway	73	0.71	3.47
	2	Expressway	55	0.33	3.00
	3	Collector	39	0.33	3.00
	4	6-Leg Arterial	40	0.33	3.00
	5	Centroid Connector	25	--	--
	6	Major Arterial	40	0.33	3.00
	7	Unmetered Ramp	44	0.33	3.00
	8	Metered Ramp	30	0.33	3.00
	9	C/D Road	59	0.71	3.47
5 Rural	0	HOV Lane	75	1.88	3.10
	1	Freeway	75	0.71	3.47
	2	Expressway	57	0.16	3.80
	3	Collector	40	0.16	3.80
	4	6-Leg Arterial	41	0.16	3.80
	5	Centroid Connector	30	--	--
	6	Major Arterial	44	0.16	3.80
	7	Unmetered Ramp	46	0.16	3.80
	8	Metered Ramp	30	0.16	3.80
	9	C/D Road	61	0.71	3.47

Legacy Approach

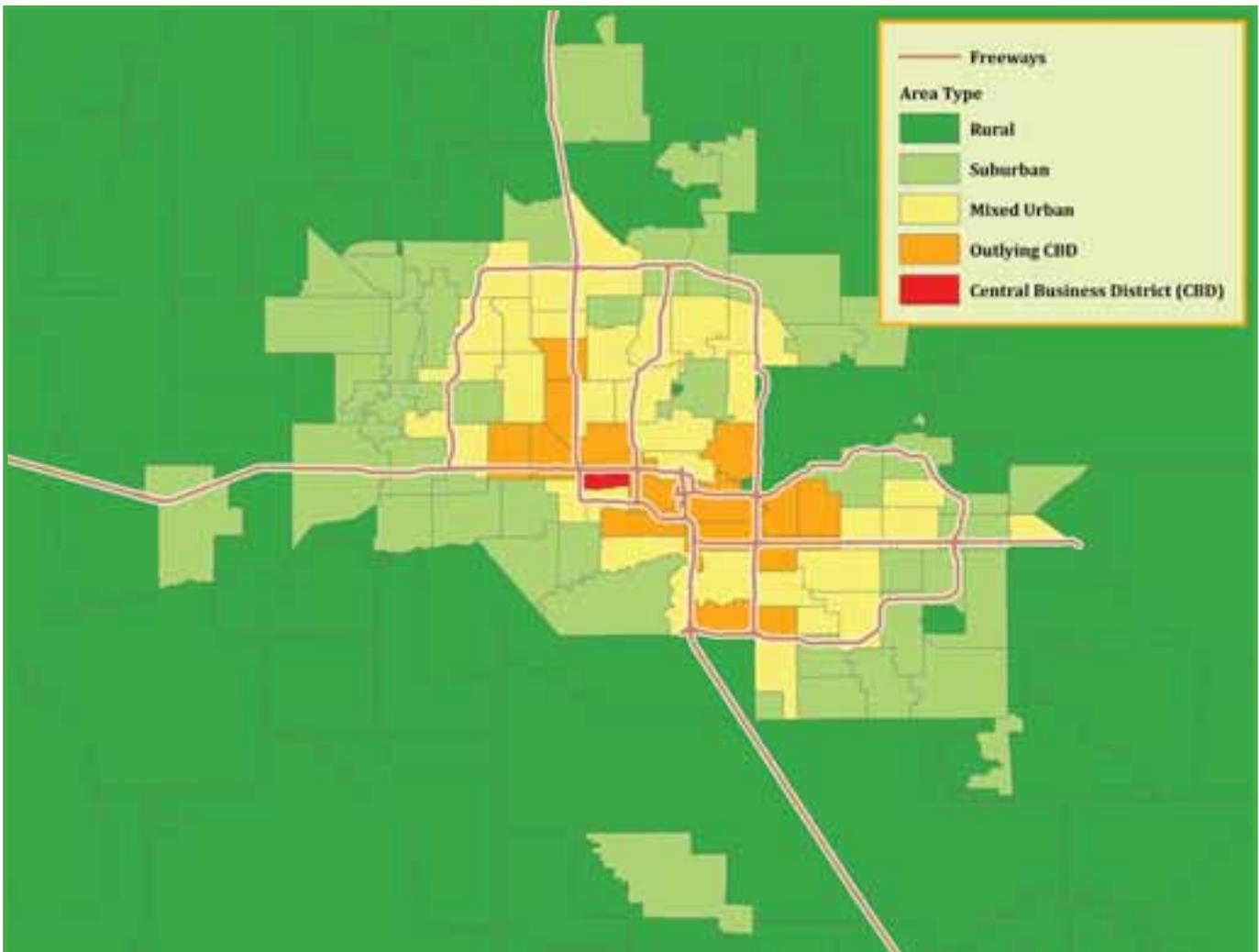
The economic intensity was previously calculated as a function of employment and population at the Regional Analysis Zone (RAZ) level.

$$EI = \frac{(E * 2 + P)}{A}$$

Where: **EI** is economic intensity for RAZ,
E is total employment in RAZ,
P is total population in RAZ,
A is RAZ area in square miles.

Figure 1 shows a map of area type polygons created with the legacy approach. Prior to implementing this new methodology, the area-type information was hard-wired in the roadway network links as an attribute. This made it difficult to update numerous network versions with changes in area type values. Also, it was not feasible to accurately assign area type attributes in the future years' model networks.

Figure 1 : *Area Type Polygons with Old Methodology*



New Approach

While employment and population will continue to be used as explanatory variables in the new approach, a few additional variables such as retail space and university enrollment have been added to the formulae to better model the economic intensity. Additionally the geographic level of analysis has been moved from Regional Analysis Zone (RAZ) to the significantly smaller Traffic Analysis Zone (TAZ).

A Kernel Interpolation with Barriers is performed based on the intensity at the centroid points of the developed parts of each TAZ.

A two-dimensional function of economic intensity is created and calibrated using traffic data collected from the field. Initially, the economic intensity is treated as a continuous surface. It is later reclassified in five categories of differing area types. With the new approach the area-type attributes are automatically assigned to the network links based on the relevant economic intensity. This can be accomplished for all horizon years and versions of the travel demand model network.

Input Datasets

The new procedure requires the user to have the socio-economic data for travel demand modeling (employment and population) at the TAZ level as well as land use data. Both datasets need to be in GIS format and available for the horizon years for which area type is being determined.

The modeling area of MAG includes the Arizona counties of Maricopa and Pinal as well as parts of Yavapai and Gila. It covers 16,080 square miles with a population of 4,198,925 people according to 2010 Census. The transportation network dataset consists of approximately 20,000 individual link segments (the number varies depending on the horizon year) and a total of 3009 TAZs.

Software Requirements

This methodology was developed using ArcInfo, Python and GISDK, along with software licenses for Geostatistical Analyst and Spatial Analyst. Additionally all datasets are stored in a Geodatabase. The use of GISDK is necessary as the MAG travel demand and forecasting model runs within a GISDK shell.

Note: although it is possible to develop this procedure in a different software environment, the scope of this paper is restricted to the software packages and licensing as stated above.

A New Methodology in Estimating Area Type

The following is a step-by-step procedure on how area type is determined in the MAG travel demand and forecasting model.

1. A dataset of natural barriers is compiled from the land use dataset.
2. A dataset of developed land is derived from the land use dataset.
3. The acreage of high density retail land use is calculated for each TAZ.
4. Initial intensity is calculated using the following formula:

$$II = \frac{(E * 2 + P)}{A}$$

Where: **II** is initial economic intensity for TAZ,
E is total employment in TAZ,
P is total population in TAZ,
A is area of TAZ in acres.

5. Adjustment factors (AF) are being determined.
6. Final intensity is calculated using the formula:

$$FI = \frac{II}{AF}$$

Where: **II** is initial economic intensity for TAZ,
FI is final economic intensity for TAZ,
AF is adjustment factor.

7. An economic intensity surface (raster format) is generated using Kernel Interpolation with Barriers among the centroids of the developed parts of TAZs.
8. The intensity raster is reclassified and converted into area type polygons.
9. Area type values are transferred from area type polygons to transportation network links.
10. Volume Delay Function (VDF) values at the link level are updated to reflect the new area type.

Adjustment Factor and Cases of Special Consideration

The spatial interpolation based on the new economic intensity formulae was executed multiple times and the resulting prediction surface was compared to traffic field data. For most zones the predicted economic intensity closely matched the traffic patterns from field data. There were however a few areas that experienced either a significant underprediction or overprediction of intensity. A variable adjustment factor is used to correct underprediction errors whereas a value of 1 is assigned for all other zones.

The Case of Underprediction of Economic Intensity

It was noticed that the interpolation surface underpredicted the level of economic intensity in the vicinities of large retail centers and the campuses of Arizona State University (ASU). To account for this, an adjustment factor is estimated that is based upon land use and ASU enrollment as shown in *Table 3*.

Table 3: *Cases of Very High Economic Intensity*

TAZ Characteristics	Adjustment factor
30 acres < Area of Retail Center <= 60 acres	0.8
60 acres < Area of Retail Center <= 100 acres	0.6
100 acres < Area of Retail Center	0.3
5,000 < ASU Enrollment <= 10,000	0.5
10,000 < ASU Enrollment <= 20,000	0.4
20,000 < ASU Enrollment <= 30,000	0.3
30,000 < ASU Enrollment	0.2

The Case of Overprediction of Economic Intensity

In the greater Phoenix area there are some large and very sparsely populated TAZs. The intensity surface generated through interpolation overpredicted the economic intensity in these areas. The following business rules were introduced to help compensate for the model overprediction: If the total area of a TAZ is over 5,000 acres and the total employment is less than 1,000, the initial intensity is assumed to be zero. Similarly, if the total population in a zone is no more than 1,000 residents and the total retail area does not exceed 30 acres, a value of zero is assigned as the initial intensity of that zone. The zones with no developed land in them are also given an initial intensity of zero along with an adjustment factor of one.

As already noted, all zones for which the modeled economic intensity matches the observed traffic patterns have an adjustment factor of 1.

Notes on the Spatial Interpolation

The spatial interpolation is based on the final intensity calculated for each TAZ and is assigned to the centroid of the developed portions of the TAZ. The resulting economic intensity surface is shown on *Figure 2*.

Figure 2: Continuous Economic Intensity Surface



The bright red colors represent areas of very high economic intensity; yellow colors denote average intensity and the greens stand for low economic intensity. The considerations behind this approach are presented below.

The type of spatial interpolation is Kernel with Barriers using the Epanechnikov function, which is a discontinuous parabolic function:

$$1 - \left(\frac{r}{h}\right)^2 \quad \text{For} \quad \frac{r}{h} < 1$$

Where: r is the radius,

h is the bandwidth (search radius).

This particular method was chosen for its capability to reflect the presence of ‘islands’ of low economic intensity inside MAG modeling area. The technique allows the exclusion of places with little traffic such as parks, mountains, Native American communities, etc. from the interpolation procedure. Through trial and error, the bandwidth was set at 4 miles and the kernel function was chosen

as the Epanechnikov function as this combination created the interpolation surface that most closely match field collected traffic data.

Another important aspect in running the interpolation procedure was the decision to generate centroid points that are based upon the developed parts of each TAZ as opposed to the entire zone.

Going From Continuous to Categorical Variable

The threshold values that define the following five categories of area type were determined by applying natural breaks (Jenks) as the initial classification method. The minimum and maximum values of each category were incrementally adjusted until the outlines of the area types closely matched those of speed and traffic count field data.

The next step was to convert the continuous economic intensity surface into five categories of polygons shown in *Table 4*. Generalization was applied—specifically polygons with area under 640 acres or 1 square mile were merged with adjacent polygons that share the largest common border. The area type polygons delineated at this step of the process are shown in *Figure 3*.

Figure 3: *Discrete Economic Intensity: Area Type Polygons with New Methodology*

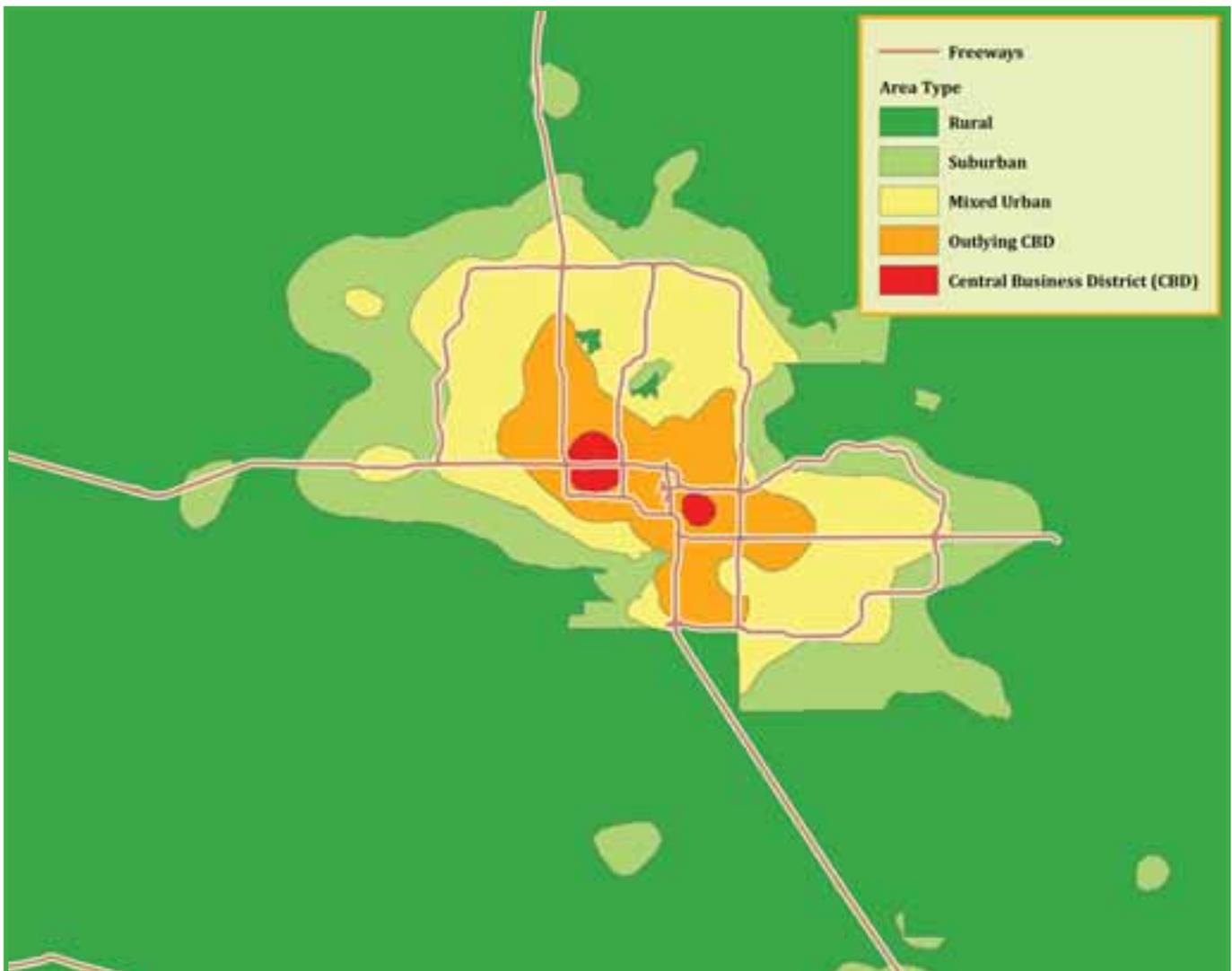


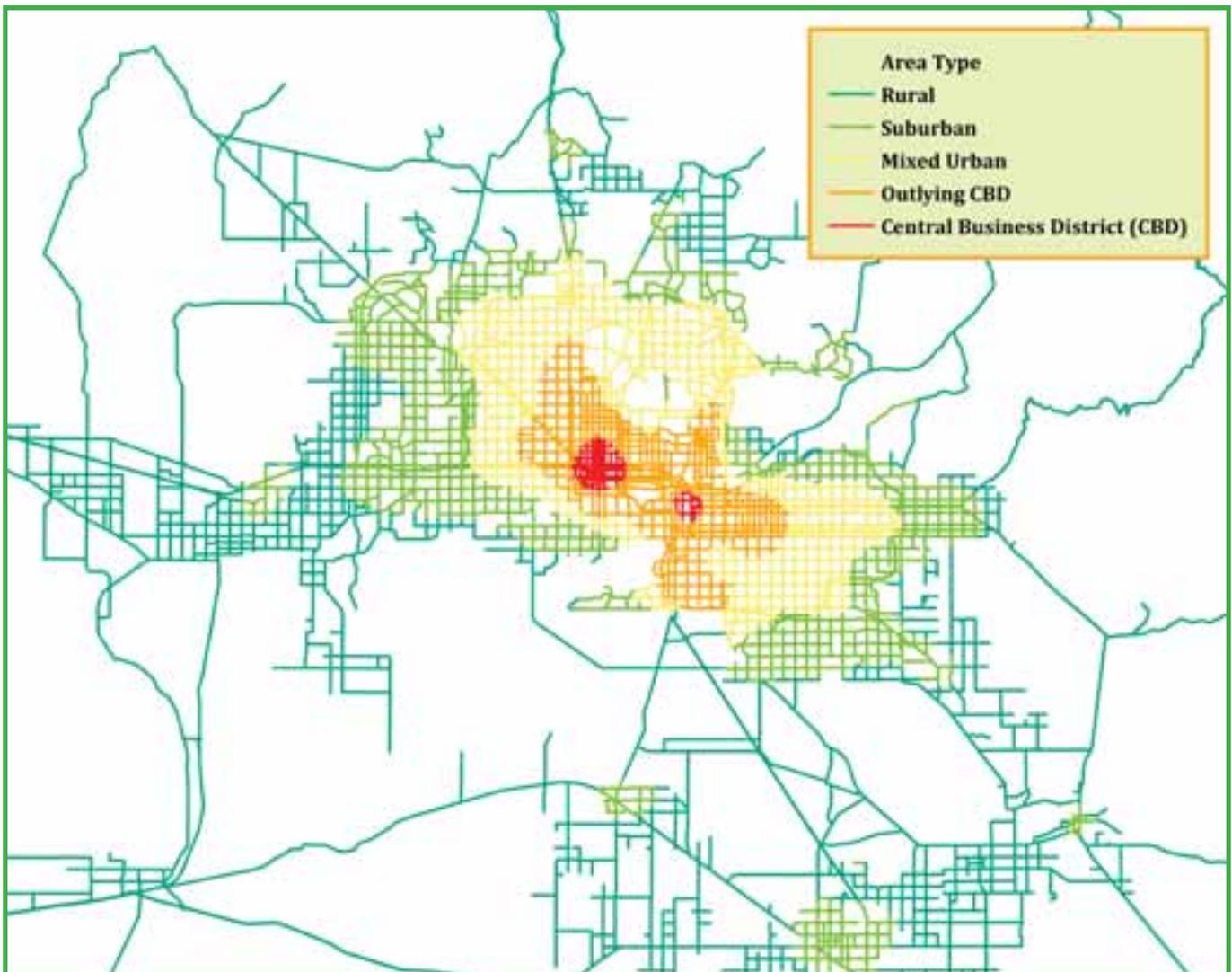
Table 4: List of Area Type Values and their Corresponding Economic Intensity

Area Type Value	Description	Level of Economic Intensity
1	Central Business District (CBD)	Very high economic intensity
2	Outlying CBD	High economic intensity
3	Mixed Urban	Moderate economic intensity
4	Suburban	Low economic intensity
5	Rural	Very low economic intensity

Transfer of Area Type Values from Polygons to the Transportation Network Links

A single area type value was assigned to each link within the network that is dependent upon the location of the link mid-point. The use of other criteria, such as assigning the link to the area type that contains its longest part were also considered but not used, as they were found to provide only negligible improvements. *Figure 4* shows a map of the transportation network links after having an area type assigned to them.

Figure 4: Area Type at the Level of Network Links

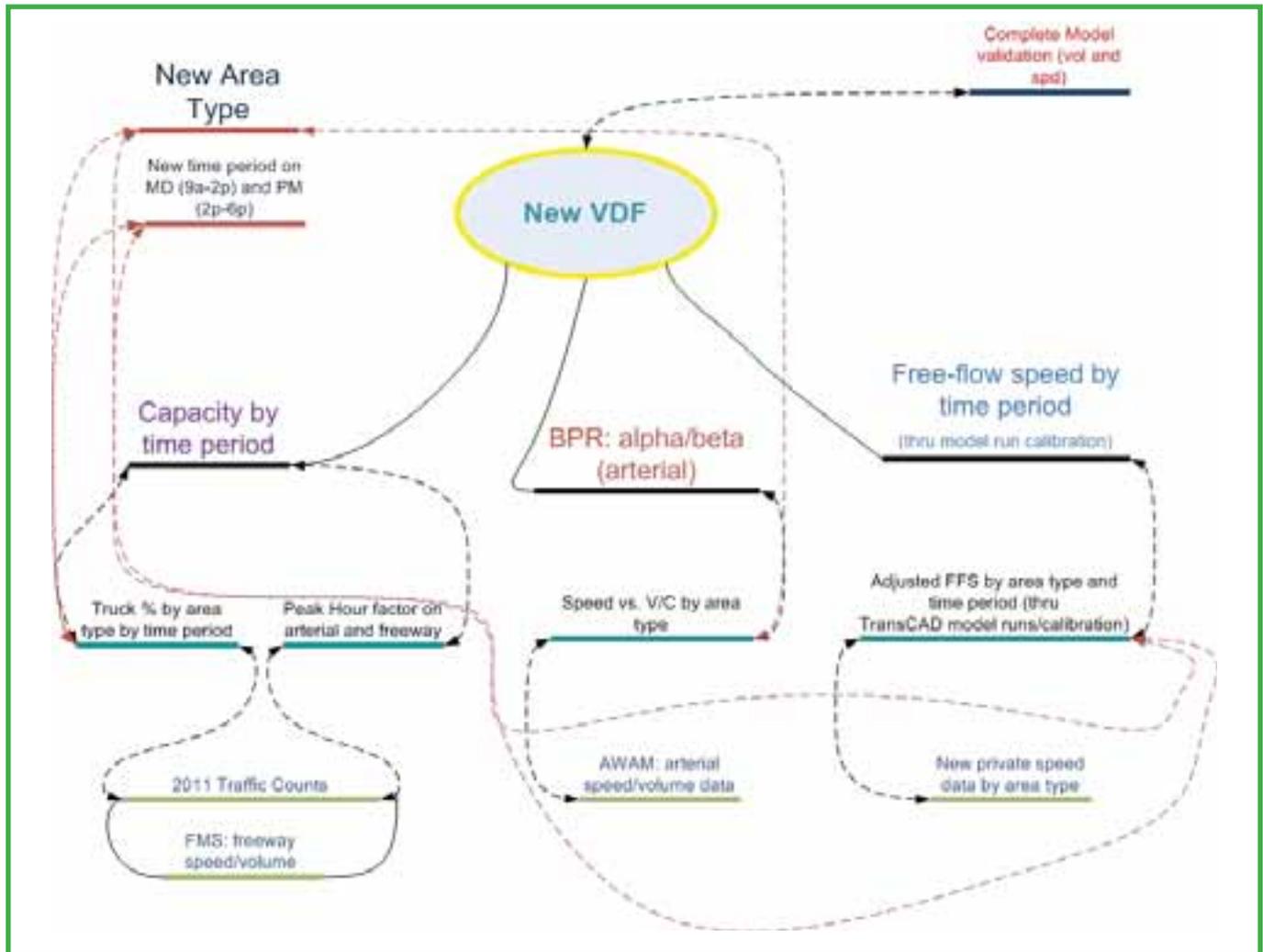


Model Application of New Area Type

A series of travel demand model calibrations and updates were conducted after the introduction of the new area type methodology. The following is a summarization.

Once the new area type definition was applied to the travel demand model network, a VDF calibration was performed. The VDF essentially determines link travel speed which affects the trip assignment results in the model. In the calibration, the network links' capacities, free-flow speeds and VDF parameters are studied by area type and functional class. Specifically, traffic data are re-grouped and analyzed according to the new area type definition. The links' capacities by area type and functional class are recalculated from traffic counts, vehicle classifications and peak hour factors, as illustrated in the left part of *Figure 5*.

Figure 5: *VDF Calibration Flow Chart*



In the next step, VDF parameters from the BPR (Bureau of Public Roads) were calibrated by area type. Since the new area type definition reflects the economic intensity, the regrouped travel delay data allows modelers to adjust the BPR parameters further in order to best-fit the VDF curves to the field data. This parameter calibration process determines how travel delay (speed) responds to a specific volume/capacity ratio at the link level. The *Figures 6, 7 and 8* identify BPR-Old Area Type (prior to calibration) versus BPR-New Area Type (after calibration) curves when compared to the field data. It is clear that the BPR curves based on the new area type methodology perform better than the BPR curves created with the old methodology.

Since the area type is relevant to roadway links' capacities, free-flow speed and VDF, it particularly affects the supply part of the model in terms of providing specific travel capacity in certain areas of the network. Consequently, the model outputs such as link traffic flow and travel speed (delay) are highly sensitive to the area type. This method ensures that the link level area type reflects the economic intensity with high fidelity. Furthermore area type values can be more easily updated in various model scenarios and horizon years. The travel demand model results are better validated, especially in travel speed and traffic flow estimation.

Figure 6: *VDF BPR Parameters Calibration—Area Type 1*

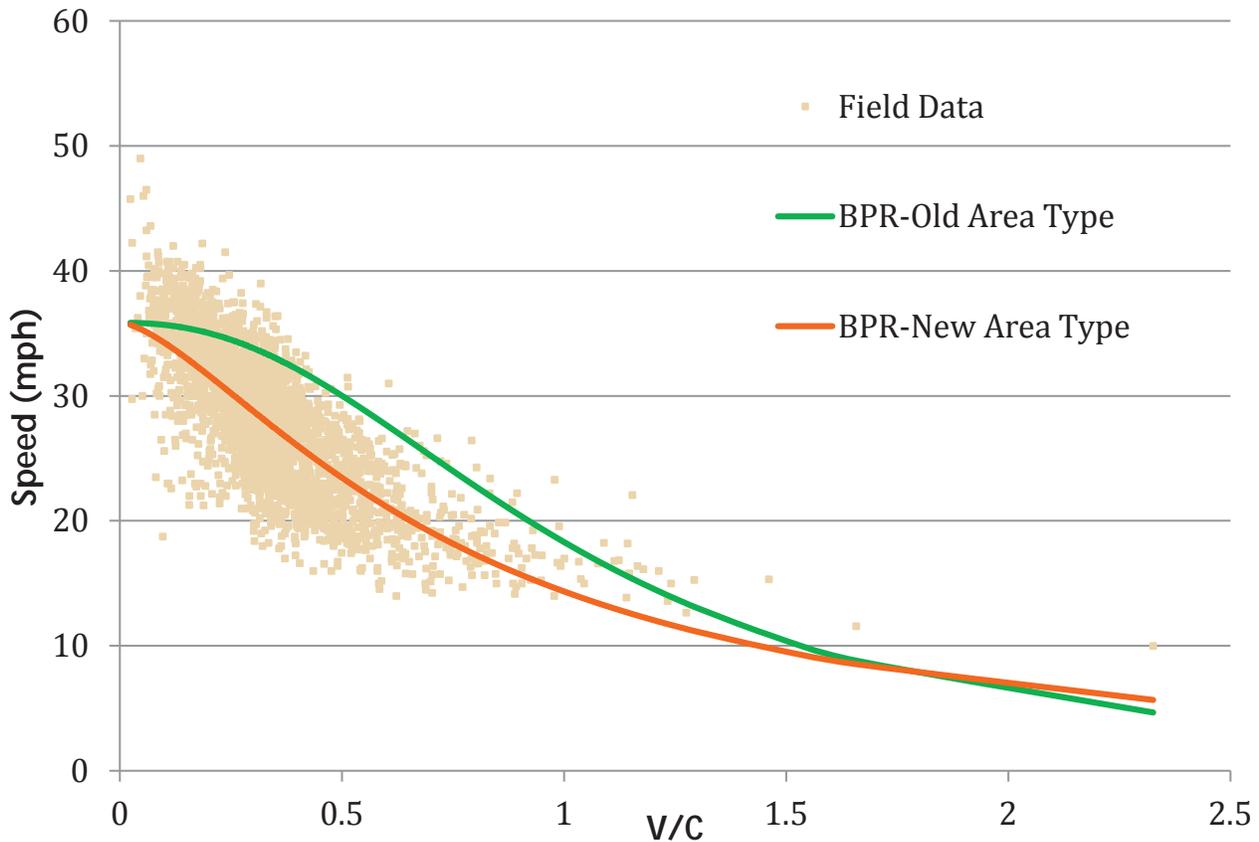


Figure 7: VDF BPR Parameters Calibration—Area Type 2

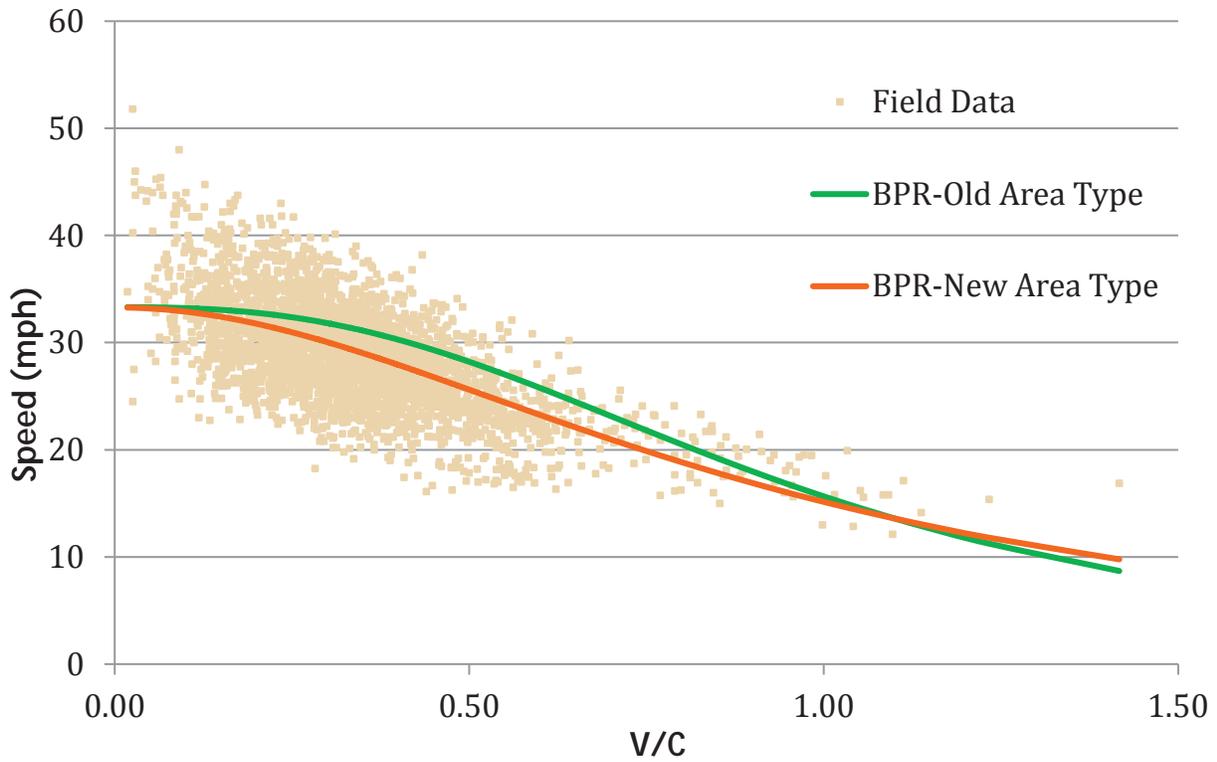
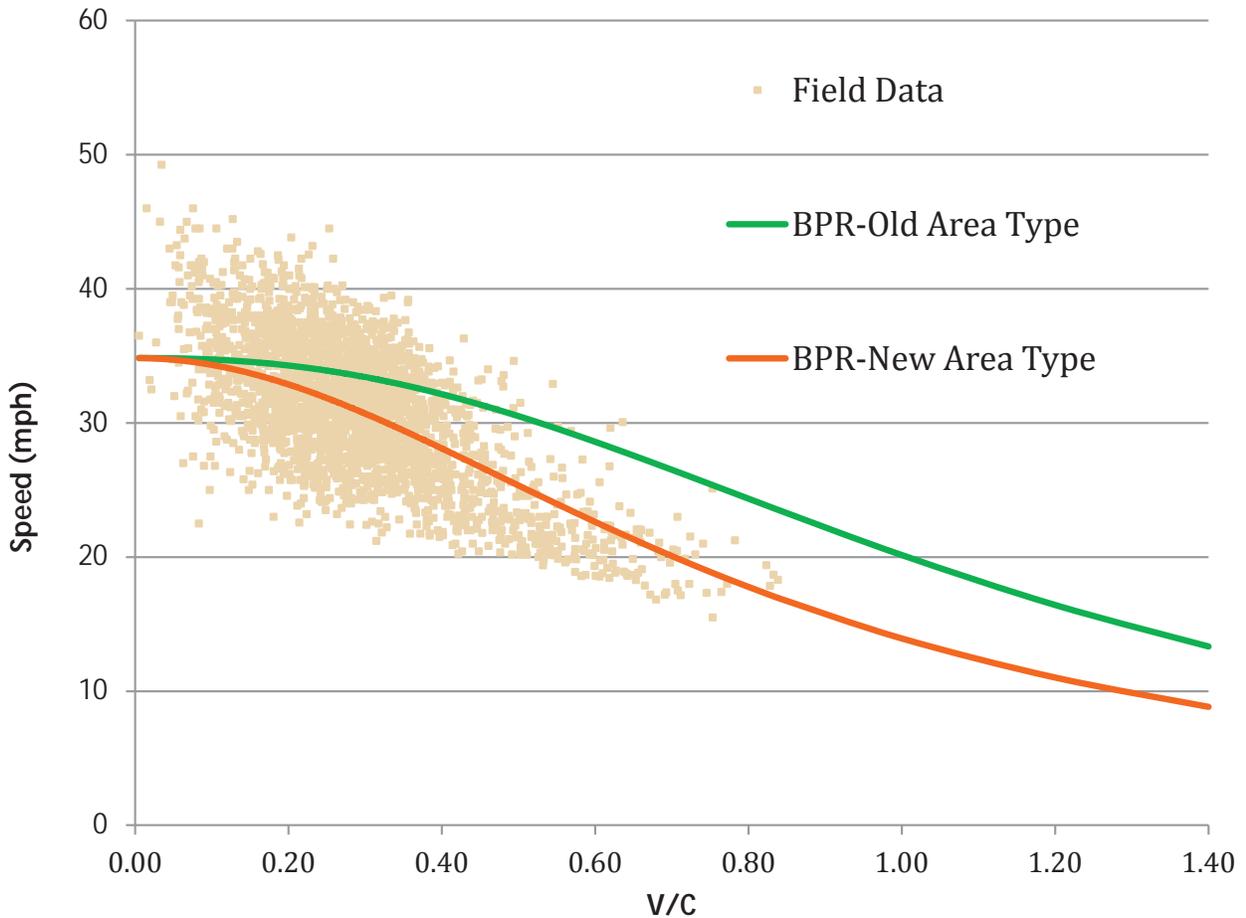


Figure 8: VDF BPR Parameters Calibration—Area Type 3, 4 and 5



Conclusion

GIS procedures and technology can effectively be used to estimate economic intensity and area type for the purposes of travel demand modeling. The GIS based approach allows users to achieve higher productivity due to automation. More importantly the new area type methodology presented in this paper can provide for better validations in terms of speeds, flows and delays.

Abbreviations and Terms:

ASU: Arizona State University

CBD: Central Business District

C/D Road: Collector / Distributor Road

BPR: Bureau of Public Roads. The agency is responsible for a widely used formula to determine the volume-delay function for roads depending on their functional class.

GIS: Geographic Information Systems

HOV: High Occupancy Vehicle Lane: also known as carpool lane

MAG: Maricopa Association of Governments: the metropolitan planning agency for the greater Phoenix Area

AF: Adjustment Factor

RAZ: Regional Analysis Zone

V/C: Volume / Capacity Ratio

VDF: Volume-Delay Function

TAZ: Traffic Analysis Zone

References

“Local Polynomials for Data Trending and Interpolation in the Presence of Barriers”, Gribov and Krivoruchko (2010)

“Highway Capacity Manual 2000”, Transportation Research Board of the National Academies of Science (2000)