

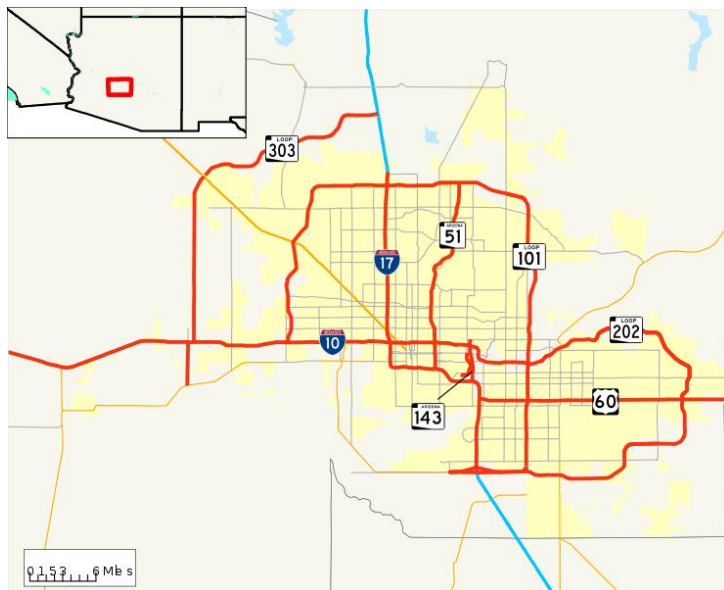
Analysis, Modeling, and Simulation (AMS) Testbed Development and Evaluation to Support Dynamic Mobility Applications (DMA) and Active Transportation and Demand Management (ATDM) Programs

Calibration Report for Phoenix Testbed

www.its.dot.gov/index.htm

Final Report — October 2016

FHWA-JPO-16-379



U.S. Department of Transportation

Produced by
Booz Allen Hamilton
U.S. Department of Transportation
Intelligent Transportation System (ITS) Joint Program Office (JPO)

[Cover Image by Arizona State University]

Notice

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The U.S. Government is not endorsing any manufacturers, products, or services cited herein and any trade name that may appear in the work has been included only because it is essential to the contents of the work.

Technical Report Documentation Page

1. Report No. FHWA-JPO-16-379	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Analysis, Modeling, and Simulation (AMS) Testbed Development and Evaluation to Support Dynamic Mobility Applications (DMA) and Active Transportation and Demand Management (ATDM) Programs — Calibration Report for Phoenix Testbed		5. Report Date October 2016	
		6. Performing Organization Code	
7. Author(s) Balaji Yelchuru, Xuesong Zhou, Pengfei Li, Pitu Mirchandani, Ismail Zohdy, Raj Kamalanathsharma		8. Performing Organization Report No.	
9. Performing Organization Name And Address Booz Allen Hamilton 20 M St SE, Suite 1000 Washington, DC 20003		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. DTFH61-12-D-00041	
12. Sponsoring Agency Name and Address U.S. Department of Transportation Intelligent Transportation Systems—Joint Program Office (ITS JPO) 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes FHWA Government Task Manager: James Colyar			
16. Abstract <p>The primary objective of this project is to develop multiple simulation Testbeds/transportation models to evaluate the impacts of DMA connected vehicle applications and the active and dynamic transportation management (ATDM) strategies. The outputs (modeling results) from this project will help USDOT prioritize their investment decisions for DMA and ATDM programs.</p> <p>The primary purpose of this report is to document the calibration process for the Phoenix Testbed. The report discusses the field observed data used to calibrate the simulation models, calibration methodology, calibrated data and also the margins of error for the efforts for each baseline operational condition.</p>			
17. Key Words ATDM, DMA, AMS, Selection criteria, Testbeds, active transportation demand managements, modeling, dynamic mobility application, Phoenix		18. Distribution Statement	
19. Security Class if. (of this report)	20. Security Class if. (of this page)	21. No. of Pages 38	22. Price

Acknowledgements

The Booz Allen Hamilton team thanks the U.S.DOT and project team members for their valuable input.

Name	Organization
Alex Skabardonis	Kittelson & Associates
Brandon Nevers	Kittelson & Associates
Chung Tran	Federal Highway Administration (FHWA)
David Roden	AECOM
James Colyar	Federal Highway Administration (FHWA)
Jim Sturrock	Federal Highway Administration (FHWA)
John Halkias	Federal Highway Administration (FHWA)
Karl Wunderlich	Noblis
Khaled Abdelghany	Southern Methodist University (SMU)
Matthew Juckes	Transport Simulation Systems (TSS)
Meenakshy Vasudevan	Noblis
Peiwei Wang	Noblis
Richard Dowling	Kittelson & Associates
Roemer Alfelor	Federal Highway Administration (FHWA)
Sampson Asare	Noblis
Thomas Bauer	Traffic Technology Solutions (TTS)

Table of Contents

Chapter 1.	Introduction	1
Chapter 2.	Testbed Description	3
2.1	Testbed Overview	3
2.2	Cluster Analysis Results	4
Chapter 3.	Model Calibration Methodology	7
3.1	An Overview of the Calibration Methodology.....	7
3.2	Network and Data Preparation.....	8
3.3	OD-Matrix Generation.....	8
3.4	Model Validation and Calibration	11
Chapter 4.	Calibration Results	14
4.1	Calibration Metrics	14
4.1.1	Calibration process	14
4.2	Scenario I (July 17, 2014)	14
4.2.1	Comparison of Initial Simulation Outputs and Field Observations.....	15
4.2.2	Comparison of Calibrated Simulation Outputs and Field Observations.....	17
4.3	Scenario 2 (May 21, 2014).....	20
4.3.1	Comparison of Initial Simulation Outputs and Field Observations.....	20
4.3.2	Comparison of Calibrated Simulation Outputs and Field Observations.....	23
4.4	Scenario 3 (June 29, 2014).....	26
4.4.1	Comparison of Initial Simulation Outputs and Field Observations.....	26
4.4.2	Comparison of Calibrated Simulation Outputs and Field Observations.....	29
4.5	Scenario 4 (November 22, 2013).....	32
4.5.1	Comparison of Initial Simulation Outputs and Field Observations.....	32
4.5.2	Comparison of Calibrated Simulation Outputs and Field Observations.....	35
4.6	Calibration Results Summary	38

List of Tables

Table 2-1: Summary of Four Clusters in PM Peak Period in Phoenix Area	5
Table 4-1: Link Count Calibration Statistics	38
Table 4-2: R-squared Calibration Statistics	38

List of Figures

Figure 2-1: Google Maps Glimpse of the Greater Phoenix Area [Source: Google Maps] ..	3
Figure 2-2: Process of building HD-DTA network [Source: ASU]	4
Figure 2-3: Cluster Analysis Results and Representative Days Selection [Source: ASU] .	6
Figure 3-1: HD-DTA Network Calibration Process Diagram [Source: ASU]	8
Figure 3-2: Locations of counting sensors within the boundary of Tempe HD-DTA network (Screenshot in NexTA) [Source: ASU]	9
Figure 3-3: Illustration of OD-cut from Phoenix regional DTA network to Tempe HD-DTA network [Source: ASU]	10
Figure 3-4: Initial O-D matrix for Tempe HD-DTA network (16:45 PM~17:00 PM) [Source: ASU]	10
Figure 3-5: Centroid Adjustment for Each Zone [Source: ASU]	11
Figure 4-1: Scenario 1 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]	15
Figure 4-2: Scenario 1 - Comparison of Initial Link Counts and Field Observations [Source: ASU]	16
Figure 4-3: Scenario 1 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]	17
Figure 4-4: Scenario 1 - Regression Analysis of Link Counts after Calibration [Source: ASU]	18
Figure 4-5: Scenario 1 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]	19
Figure 4-6: Scenario 1 - Calibrated Trip Production and Attraction Matrices of Tempe [Source: ASU]	20
Figure 4-7: Scenario 2 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]	21
Figure 4-8: Scenario 2 - Comparison of Initial Link Counts and Field Observations [Source: ASU]	22
Figure 4-9: Scenario 2 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]	23
Figure 4-10: Scenario 2 - Regression Analysis of Link Counts after Calibration [Source: ASU]	24
Figure 4-11: Scenario 2 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]	25
Figure 4-12: Scenario 2 - Calibrated Trip Production and Attraction Matrices of Tempe [Source: ASU]	26
Figure 4-13: Scenario 3 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]	27
Figure 4-14: Scenario 3 - Comparison of Initial Link Counts and Field Observations [Source: ASU]	28
Figure 4-15: Scenario 3 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]	29
Figure 4-16: Scenario 3 - Regression Analysis of Link Counts after Calibration [Source: ASU]	30

Figure 4-17: Scenario 3 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]	31
Figure 4-18: Scenario 3 - Calibrated Trip Production and Attraction Matrices of Tempe [Source: ASU]	32
Figure 4-19: Scenario 4 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]	33
Figure 4-20: Scenario 4 - Comparison of Initial Link Counts and Field Observations [Source: ASU]	34
Figure 4-21: Scenario 4 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]	35
Figure 4-22: Scenario 4 - Regression Analysis of Link Counts after Calibration [Source: ASU]	36
Figure 4-23: Scenario 4 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]	37
Figure 4-24: Scenario 4 - Calibrated Trip Production and Attraction Matrices of Tempe [Source: ASU]	38

Chapter 1. Introduction

The United States Department of Transportation (USDOT) initiated the Active Transportation and Demand Management (ATDM) and the Dynamic Mobility Applications (DMA) programs to achieve transformative mobility, safety, and environmental benefits through enhanced, performance-driven operational practices in surface transportation systems management. In order to explore a potential transformation in the transportation system's performance, both programs require an Analysis, Modeling, and Simulation (AMS) capability. Capable, reliable AMS Testbeds provide valuable mechanisms to address this shared need by providing a laboratory to refine and integrate research concepts in virtual computer-based simulation environments prior to field deployments.

The foundational work conducted for the DMA and ATDM programs revealed a number of technical risks associated with developing an AMS Testbed which can facilitate detailed evaluation of the DMA and ATDM concepts. Therefore, instead of selecting a single Testbed, it is desirable to identify a portfolio of AMS Testbeds and mitigate the risks posed by a single Testbed approach by conducting the analysis using more than an "optimal" number of Testbeds. At the conclusion of the AMS Testbed selection process, six (6) AMS Testbeds were selected to form a diversified portfolio to achieve rigorous DMA bundle and ATDM strategy evaluation: San Mateo (US 101), Pasadena, ICM Dallas, Phoenix, Chicago, and San Diego Testbeds.

In a preceding set of deliverables, the analysis plans developed for the selected AMS Testbeds are presented. These analysis plans describe the baseline operation scenarios to be considered for each Testbed. These baseline scenarios were obtained based on a cluster analysis that is conducted to determine common operational conditions for each Testbed. A primary objective of this task is to calibrate the traffic network simulation models that are used to simulate the traffic conditions of these Testbeds to ensure that the models are capable of replicating the observed traffic patterns in the network.

The primary purpose of this report is to document the model calibration effort for the Phoenix testbed to represent the different baseline scenarios. Given the available data in this area, the efforts primarily focus on the calibration of DTA network to have valid origin-destination demands to synchronize the simulated link counts and actual link counts collected by Arizona DOT. Through add-on tools developed by the team, the origin-destination demands and all possible paths generated in DTA (Dynamic Traffic Assignment) network will be automatically parsed into the microscopic simulation engine, VISSIM, to have a consistent network representation in both layers of simulation engines.

This Testbed will be used to test several DMA/ATDM strategies considering a proactive network management approach that adopts simulation-based prediction capabilities. The Testbed is developed using a newly developed multi-resolution simulation platform developed at Arizona State University.

This report is organized into four chapters as follows:

- Chapter 1 – Introduction: This chapter presents the report overview and objectives.
- Chapter 2 – Testbed Description: This chapter presents the regional characteristics of the Testbed, the proposed operational conditions, the results of the cluster analysis and the selection of the baseline scenarios.

- Chapter 3 – Model Calibration Methodology: This chapter presents the methodology used to calibrate the High-Definition Dynamic Traffic Assignment (HD-DTA) model against the operational conditions of the baseline scenarios selected for the Testbed. The methodology describes the process used to adjust the different model parameters.
- Chapter 4 – Calibration Results: This chapter summarizes the model calibration results. It provides a comparison between the operational conditions observed for one scenario and the corresponding model results.

Chapter 2. Testbed Description

2.1 Testbed Overview

The proposed Phoenix testbed is located in the Maricopa County which is home to more than 1.5 million households and 4.2 million inhabitants according to the 2010 Census. The region covers an area of 9,200 square miles and is characterized by a low density development pattern with population density just about 253 people per square mile. The region has one city with more than 1 million people (Phoenix) and eight cities/towns with more than 100,000 people each. Figure 2-1 shows a Google Maps glimpse of the Maricopa County region.

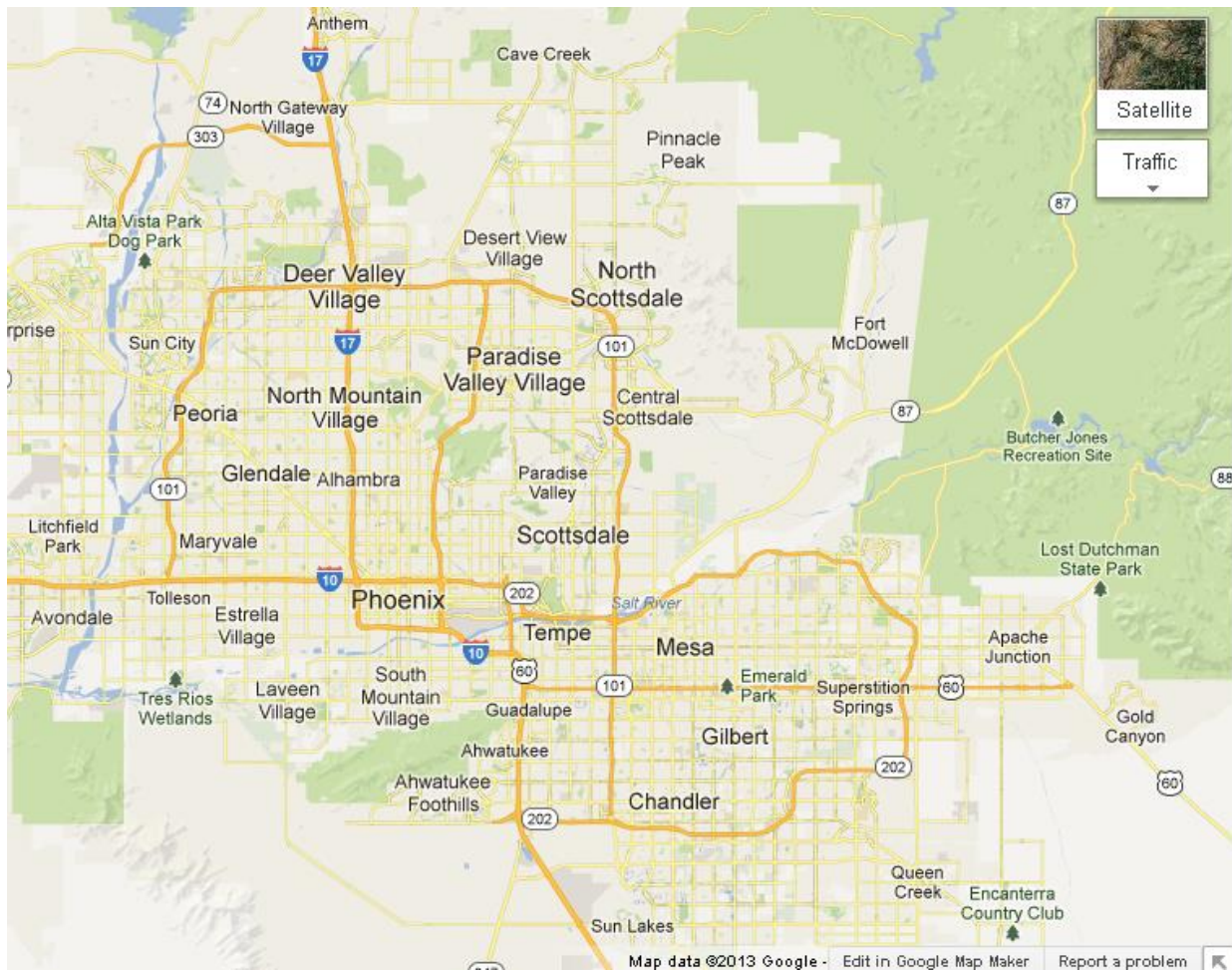


Figure 2-1: Google Maps Glimpse of the Greater Phoenix Area [Source: Google Maps]

Given that the Greater Phoenix testbed is too large to be simulated as a whole, the project team decided to focus on one of its member cities, the City of Tempe. The Tempe HD-DTA network can represent 90% of road conditions and geometries, covering the freeway-type roads like I-10, 143, 101, 202 and US 60 as well as most major arterials including 59 signalized intersections and 19 ramps within Tempe. The total length of roads within the network is over 150 miles. For this project, the team adopted a “bottom-up” method to build the network in which the microscopic simulation model was built first and then migrated to higher level DTA model for the Tempe region as shown in Figure 2-2. As a result, the Tempe HD-DTA network was built containing more details than typical DTA networks. One major difference is that the HD-DTA network contains more details within intersections so modeling advanced traffic signal control in DTALite is feasible. In contrast, intersections are represented with points in typical DTA networks and so it is difficult to simulate high-fidelity traffic signal control mechanism.

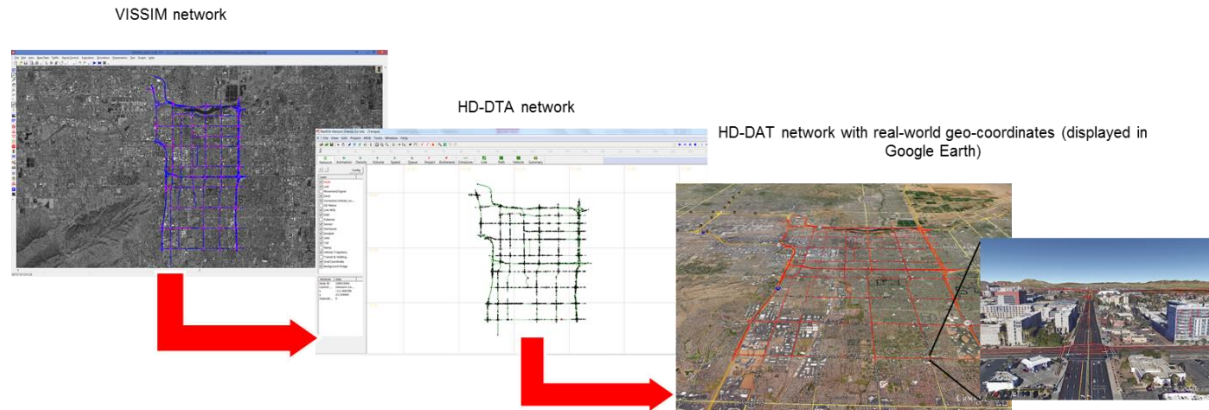


Figure 2-2: Process of building HD-DTA network [Source: ASU]

2.2 Cluster Analysis Results

A cluster analysis was performed to determine the main operational conditions of the Phoenix Testbed. The detailed approach and results of the cluster analysis are presented in the Analysis Plan document for the Testbed. Based on the cluster analysis conducted for the Phoenix Testbed, four main clusters are determined to represent various evening-peak (15:00 to 19:00) scenarios composed of different volumes, incidents and travel speeds. Each cluster represents a bin of multiple days in the analysis year and one representative day was selected for each cluster that is closest to the cluster centroid.

Table 2-1 provides a description of the main four clusters obtained based on this analysis including the representative day, the average hourly volumes, travel-speeds and incident severity values. Comparing the values of these variables against the average values for all data records, we can generally obtain some meaningful description of these four clusters. The clusters are named based on the representative values of traffic demand, travel speeds and incident severity. The traffic demand is represented by the average hourly volume in the network. The travel-speed is represented by the average speed of vehicles on the freeways in miles per hour and incident severity is represented by the product of number of incidents and the number of lane closures resulted from it. For example, Cluster 1 consist of higher traffic volumes, higher vehicle speeds and low number of incidents. The location of incidents is of extreme importance in modeling and is computed using the data patterns (loop-detector data) from freeways. Clusters 1 through 3 represent dry weather conditions, while Cluster 4 is associated with wet pavement (or rain at 0.01 in/hour).

Table 2-1: Summary of Four Clusters in PM Peak Period in Phoenix Area

		Cluster 1	Cluster 2	Cluster 3	Cluster 4
Representative Day		7/17/2014	5/21/2014	6/29/2014	11/22/2013
Cluster Name		High traffic, high speed, low incidents, dry	High traffic, high speed, high incidents, dry	low traffic, high speed, low incidents, dry	high traffic, low speed, medium incidents, wet
Volume (veh/hr)	15:00	9550	9630	6229	8958
	16:00	9291	9993	6324	8226
	17:00	8762	9226	6227	8222
	18:00	7820	8153	6018	7349
	19:00	6494	6906	5223	5784
Average Volume (veh/hr)		8383	8782	6004	7708
Average Speed (mph)	15:00	62	66	64	34
	16:00	64	64	63	27
	17:00	66	66	65	26
	18:00	66	65	67	43
	19:00	67	66	68	62
Average Speed (mph)		65	65.4	65.4	38.4
Number of Incidents X Lanes closed	15:00	2	6	1	5
	16:00	3	10	0	3
	17:00	4	5	1	3
	18:00	0	0	1	7
	19:00	0	1	0	5
Average Incident Severity (lanes closed)		1.8	4.4	0.6	4.6

The project team chose four representative days according to the Euclidean distances of samples away from the centroid values with each clusters. As discussed, only the PM attributes were used to calculate the Euclidean distances. Figure 2-3 shows the Euclidean distances of all samples in each clusters and four representative days with the minimum distance (without considering the weather attributes) are: July-17-2014 (cluster 1), May-21-2014 (cluster 2), June-29-2014 (cluster 3) and November-22-2013 (cluster 4). If the weather is considered, the weather-favored sample (11/22/2013) belongs to cluster 4 (the red bar in the cluster 4) and therefore the final selected representative days and the descriptions of four clusters are:

- July-17-2014: (Cluster 1: high traffic, high speed, low incidents, dry)
- May-21-2014: (Cluster 2: high traffic, high speed, high incidents, dry)
- June-29-2014 (Cluster 3: low traffic, high speed, low incidents, dry)
- November-22-2013 (Cluster 4: high traffic, low speed, medium incidents, wet)

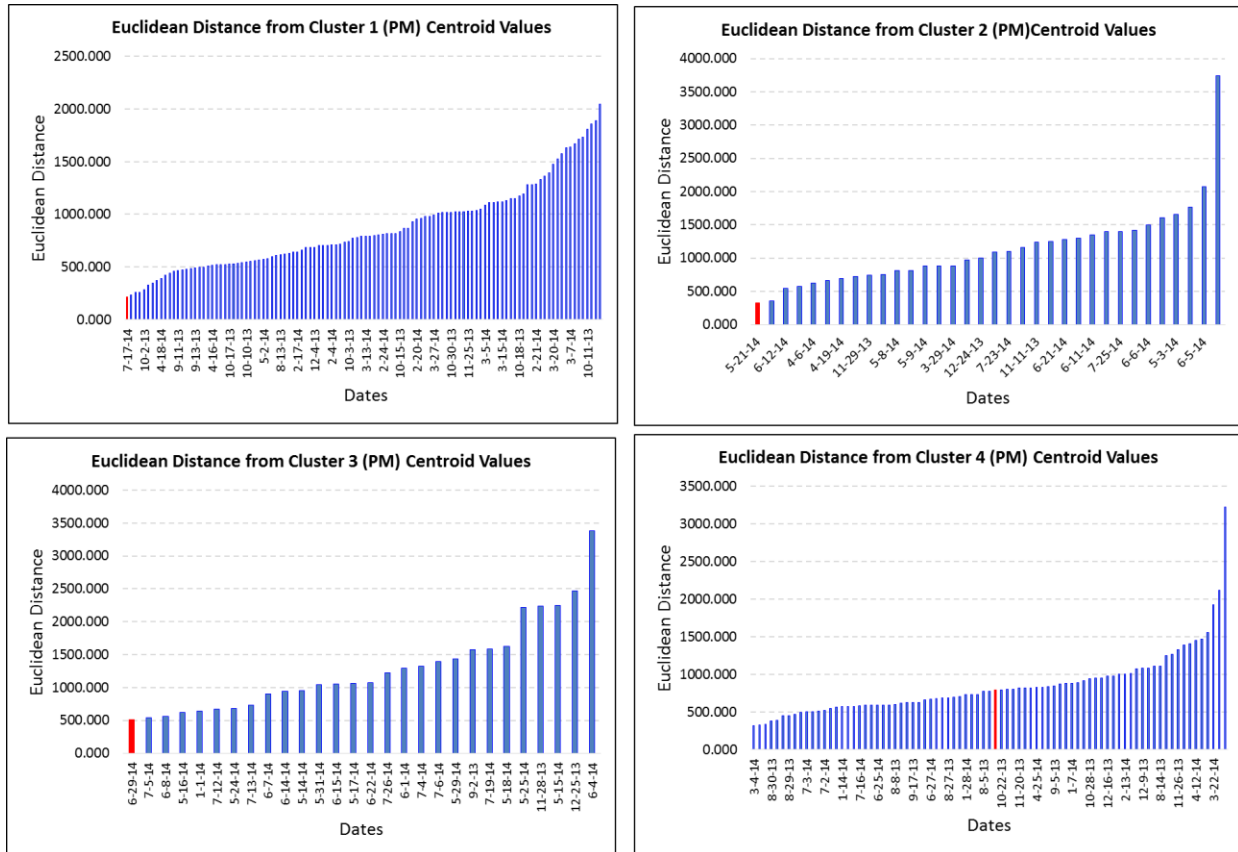


Figure 2-3: Cluster Analysis Results and Representative Days Selection [Source: ASU]

Chapter 3. Model Calibration Methodology

This chapter describes the calibration methodology against the selected baseline scenarios identified through the cluster analysis and illustrates how the different model parameters are adjusted. These adjustments attempt to replicate the observed traffic pattern and associated congestion phenomena.

3.1 An Overview of the Calibration Methodology

The objective of model calibration is to ensure that the baseline simulation results are consistent with the observed data. Calibration is often performed in conjunction with model validation. There are three components in the simulation network that can affect the simulation outputs; these are link capacity, travel demand and available paths between origins and destinations.

Road capacity is typically hard to measure and it often needs lots of local experiences. Toward that goal, we compared the link capacities in the Tempe HD-DTA network with the existing DTA network of the Greater Phoenix area and considered the following link capacities are appropriate for the HD-DTA network:

- Basic saturation rate for general links: 1,900 vehicles per hour per lane;
- Saturation rate for reduced-capacity links, such as ramps or turning links at intersections: 1,500 vehicles per hour per lane;

The available Origin-Destination (OD) matrix in Phoenix is for the regional DTA network and the scopes of most Traffic Analysis Zones (TAZs) are too wide for the Tempe HD-DTA network. To address this issue, the team proposed a new approach, referred to as the “O-D cut”. The Tempe HD-DTA network is first divided into multiple smaller zones and the trips between those smaller zones are derived according to vehicle trajectories generated from the Maricopa Association of Governments’ (MAG) DTA model. According to the vehicle trajectories, the first zone where a vehicle enters the boundary of Tempe is considered as its origin and the last zone where the vehicle leaves Tempe is considered as its destination. Then the time-dependent O-D matrix can be derived via scanning all vehicle trajectories.

Once the initial O-D matrix is determined, the next step is to calibrate the simulated output using the observed data in the field. The available field data are the counts collected at over 20 locations along freeways in Tempe. Turning movement count at intersections were collected on different days during morning and evening peak hours by Maricopa Association of Governments were used as the arterial data. After comparing the initial simulated output with the observations in the field, it is noticed that substantial effort is required to calibrate and subsequently build a valid model. As such, extensive efforts were carried out to adjust the centroids in each zone, initial O-D matrix and initial paths. After several rounds of adjustments, the calibrated HD-DTA network can generate results approximate to the field observation. For more details, please refer to Chapter 4.

In summary, the calibration process is shown as in Figure 3-1.

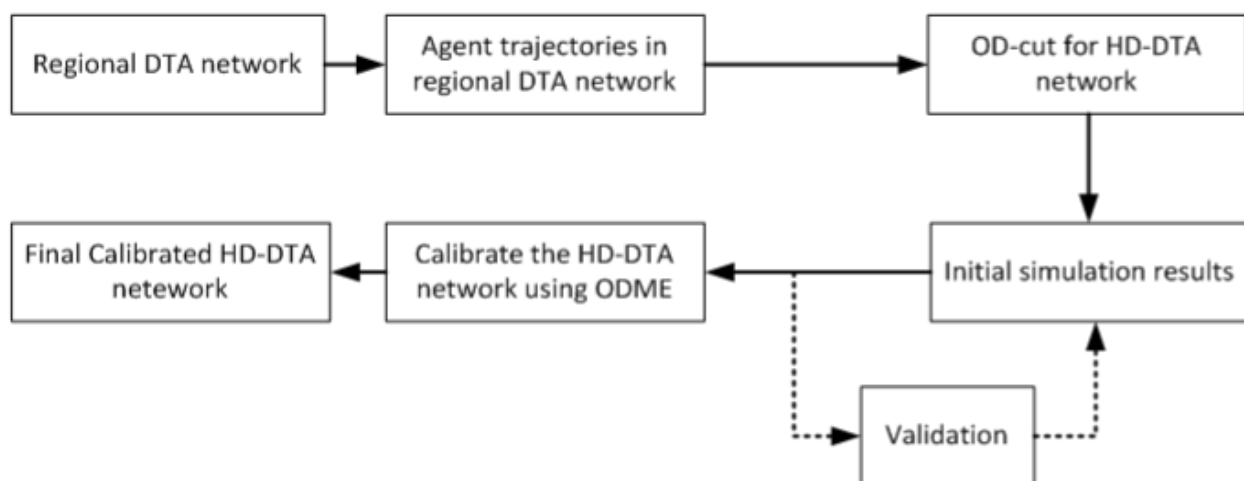


Figure 3-1: HD-DTA Network Calibration Process Diagram [Source: ASU]

3.2 Network and Data Preparation

Traffic volume data: The available volume data sources for this task is the continuous link traffic counts along the freeway and major arterials systems in the Greater Phoenix area. Totally there are over 60 counting stations located with the boundary of the Tempe HD-DTA network. Figure 3-2 shows the locations of counting stations (the green blocks). Each counting stations can provide continuously archived counts with various resolutions and the project team adopted link counts with 15-min intervals. As discussed in the analysis plan document, the evaluation period for the Phoenix testbed is from 15:00 to 19:00.

Regarding the incidents dataset, Arizona DOT has provided a list of all incidents on freeways in 2013 and 2014. We also expanded the arterial counts from one or two hours to four hours at each locations on arterials. It is known that incidents will lower the road capacity, increase delay, and make vehicles to take detours. In simulation, the locations and durations of incidents within the simulated period are identified and the corresponding link capacities are reduced for certain periods, respectively. The reduced link capacities were identified from the MAG data-set and then derived empirically for individual lane-closure incidents.

Phoenix Regional DTA Network: Maricopa Association of Governments (MAG) provided the project team with a well-calibrated DTA model for the entire Phoenix area. The regional DTA model includes rich information, such as link capacities, link speeds and travel demands.

3.3 OD-Matrix Generation

Concept of OD-cut: It is noticed that the O-D demand retrieved from the regional DTA model cannot be directly applied to the Tempe HD-DTA network because the number of TAZs falling into the boundary of Tempe is too small to be reasonable origins. To address this issue, the project team used a new approach named “O-D cut” to generate an O-D matrix with high fidelity for the Tempe HD-DTA network. As shown in Figure 3-3, we first further divided the whole HD-DTA network into smaller zones (just like the TAZ zones in typical DTA network). In order to determine the departure and arrival vehicles within each zone, we retrieved all the vehicle trajectories from the simulation outputs of the regional DTA network. Each vehicle was identified whether it passed the Tempe or not. If so, the first zone where the subject agent entered the boundary of Tempe was considered its origin within Tempe and the last zone where the subject agent left Tempe was considered its destination within Tempe. After all the generated trajectories were scanned and we got a time-dependent O-D matrix for all the new smaller zones in

Tempe. Figure 3-3 illustrates the trajectory of three example vehicles which resulted in three inter-zonal trips between zones. Figure 3-4 shows one of twenty O-D matrices from 16:45 to 17:00.

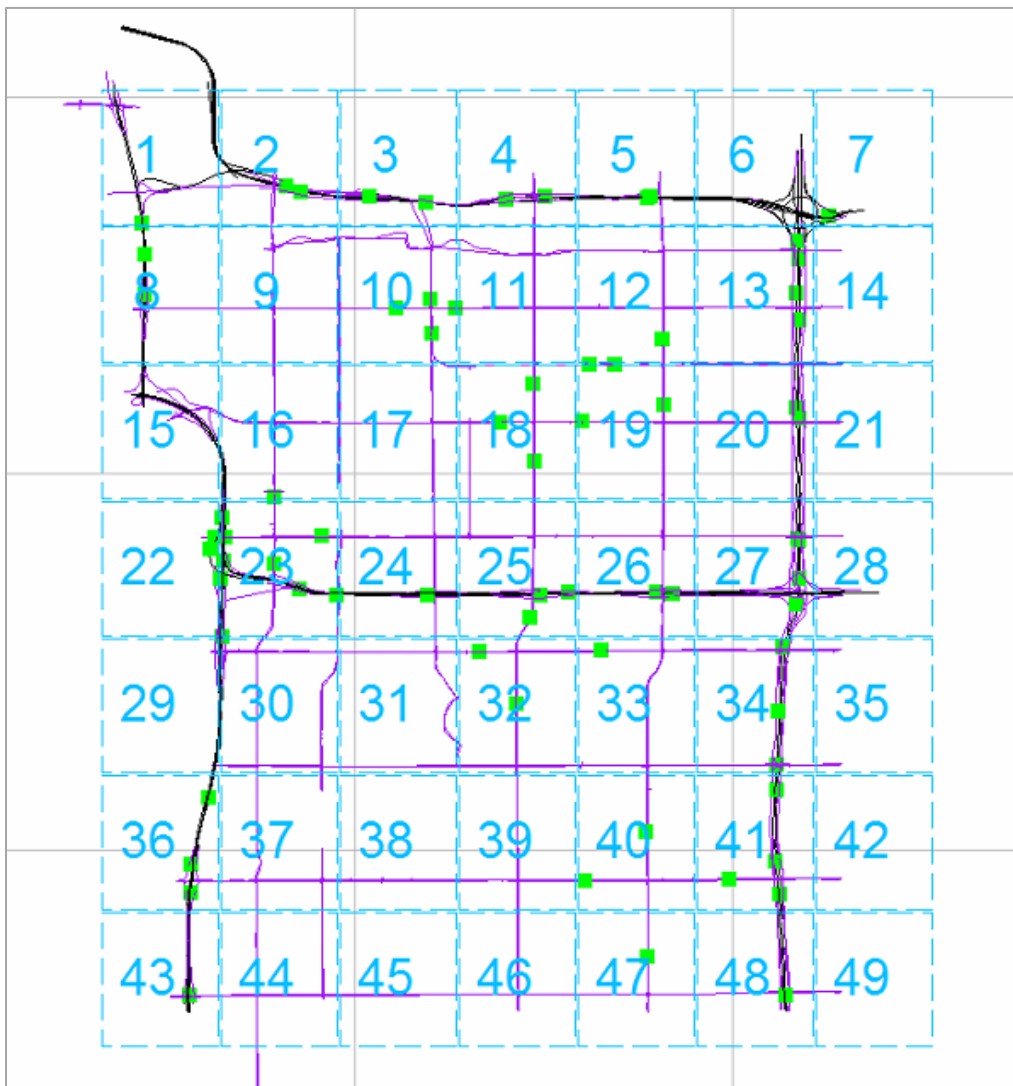


Figure 3-2: Locations of counting sensors within the boundary of Tempe HD-DTA network (Screenshot in NexTA) [Source: ASU]

3.4 Model Validation and Calibration

Initial simulation of Tempe HD-DTA network revealed that the simulated link counts on freeways in Tempe HD-DTA network is substantially different from the link counts collected by ADOT. The difference between simulated link counts and observed link counts was more than 100% in many locations and there existed certain unrealistic bottlenecks during the simulation. These issues cannot be corrected simply by adjusting the time-dependent O-D matrix using Origin-Destination Matrix Estimation (ODME). After closely checking the results, the following issues were observed:

1. Selection of centroids: unlike the traditional DTA networks composed of unidirectional links, the Tempe HD-DTA network is only composed directional links. As a result, the zone centroids must not be placed on those “dead-end” links. Otherwise, many vehicles will be piled up at certain locations. To address this issue, the project team members adjusted the centroids of each zone based on their engineering judgement. Figure 3-5 demonstrates one centroid adjustment at one zone.
2. Link capacities in the HD-DTA network needs to be further adjusted: although the link capacities in the regional DTA network provides good references for model calibration. In practice, the project team had to adjust the link capacities empirically because the topologies of road networks in Tempe are quite different between the regional DTA network and HD-DTA network.

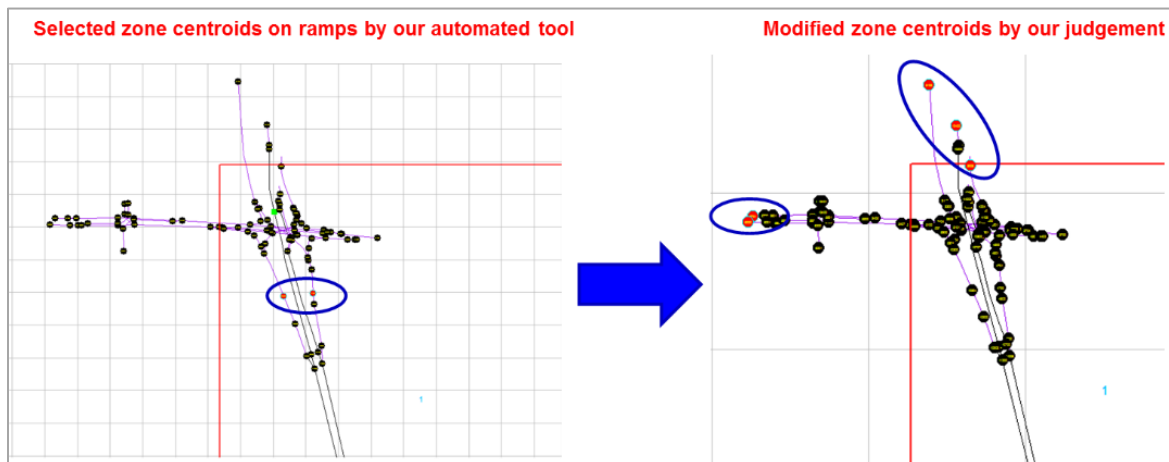


Figure 3-5: Centroid Adjustment for Each Zone [Source: ASU]

3. Inconsistency between the regional DTA network and observed link counts: Even though the regional DTA network was well calibrated, it turned out that the link counts out of regional DTA network was still quite different from the observed value after excluding the impact of different years.
4. The magnitude of link numbers in the HD-DTA network is insufficient for the automated ODME tool to match the observed link counts (e.g., the ODME does not have too many alternative paths while attempting re-assign traffic), especially when the initial simulation results are rather different from the field observations. Therefore, model validation is necessary before the calibration.

To address the above issues, the project team first scanned and adjusted all the initially assigned origin and destination centroids in each zone to ensure the centroid points with each zone are located where there is full access to all other zones. Such efforts considerably reduce the unreasonable bottlenecks. Through preliminary simulation and careful review of regional DTA model of Phoenix area, we empirically increase or decrease some link capacities to remove or relocate bottlenecks to reflect the local facts.

After the first steps of adjustments, the ODME tool was used to reduce the gap between the simulated link counts in DTALite and observed link counts. In essence, the ODME tool is to seek the best O-D demands between TAZ zones to generate certain indicators close to the observed ones, in this context, the link counts on freeways and arterials are used. In essence, DTALite is a dynamic network loading (DNL) model based on Newell's kinematic wave theory to capture congestion phenomena and shock wave propagation in the process of solving dynamic traffic assignment problems. Given observed link flows and target O-D demands, the proposed time-dependent path flow estimation model can be represented with a nonlinear programming problem using the path flows $r(w, \tau, p)$, $\forall w, \tau, p$ and least path travel times $\pi = \{\pi(w, \tau), \forall w, \tau\}$, as the decision variables. If $c = \{c(w, \tau, p), \forall w, \tau, p\}$, $q = \{q(l, t), \forall l, t\}$ and $k = \{k(l, t), \forall l, t\}$. Mathematically, the problem can be formulated as:

$$\begin{aligned} \text{Min } Z = & \beta_d \sum_w [\sum_{\tau \in H_d} \sum_p r(w, \tau, p) - \bar{d}(w)]^2 + \sum_{l \in S} \sum_{t \in H_o} \left\{ \beta_q [q(l, t) - \bar{q}(l, t)]^2 + \right. \\ & \left. \beta_k [k(l, t) - \bar{k}(l, t)]^2 \right\} \text{Min } z(r) = \beta_d f(r) + \beta_q h^q(r) + \beta_k h^k(r) + \beta_T h^T(r) \end{aligned} \quad (3-1)$$

Subject to

$$(c, q, k) = \text{DNL}(r), \quad (3-2)$$

$$g(r, \pi) = \sum_w \sum_{\tau} \sum_p \{r(w, \tau, p)[c(w, \tau, p) - \pi(w, \tau)]\} = 0, \quad (3-3)$$

$$c(w, \tau, p) - \pi(w, \tau) \geq 0, \forall w, \tau, p, \quad (3-4)$$

$$\pi(w, \tau) \geq 0, \forall p \in P(w, \tau), \forall w, \tau \quad (3-5)$$

$$r(w, \tau, p) \geq 0, \forall w, \tau, p. \quad (3-6)$$

where:

DNL: Dynamic Network Loading Function

A: set of links

W: set of OD pairs

P: set of paths

S: set of links with sensors, $S \subseteq A$

H_d : set of discretized departure time intervals

H_o : set of discretized observation time intervals

t : index of simulation time intervals, $t = 0, \dots, T$.

τ : index of departure time intervals, $\tau \in H_d$

w : index of OD pairs, $w \in W$

p : index of paths for each OD pair, $p \in P$

l : index of links, $l \in A$

$\bar{q}(l, t)$: observed number of vehicles passing through an upstream detector on link l during observation interval t

$\bar{k}(l, t)$: observed density on link l during observation interval t

$\bar{d}(w)$: target demand, which is the total traffic demand for OD pair w over a planning horizon

$r(w, \tau, p)$: estimated path flow on path p of OD pair w and departure time interval τ

$c(w, \tau, p)$: estimated path travel time on path p of OD pair w and departure time interval τ

$\pi(w, \tau)$: estimated least path travel time of OD pair w and departure time interval τ

$P(w, \tau)$: a set of paths as set of paths for OD pair w at departure time interval τ

$q(l, t)$: estimated number of vehicles passing through an upstream detector on link l during observation interval t

$k(l, t)$: estimated density on link l during observation interval t

$d(w, \tau)$: estimated demand of OD pair w and departure time interval τ
 β : weighting factors

The solution to this problem in ODME is a heuristic method based on Lagrangian relaxation and sub-gradient method. Through iterations, the low bound provided through the Lagrangian relaxation and sub-gradient method; the upper bound provided from the DTALite will be closer and closer. The iterations are stopped after 80 runs or when the target performance in terms of link counts is reached as well as additional calibration metrics such as trip production and attraction and R-squared statistic, defined in next chapter, are also satisfied.

Chapter 4. Calibration Results

This chapter demonstrates the calibration process for the selected clusters, or scenarios which are represented by data from unique days as mentioned in Table 2-1.

4.1 Calibration Metrics

Since the calibration focuses on link counts, the objective of this task is to calibrate the travel demand between zones, or O-D matrix. Three methods were applied to the results evaluation: regression validation (R^2 approach), link counts comparison at counting locations and trip production and attraction at each zone. The regression validation allows us to have an overall comparison between simulated results and field observations. Location-specific comparison offer some information of zone-to-zone trips. The zone-based trip attractions and productions allows for comparison between the simulation outputs and the real world traffic patterns. The main approach for model calibration, ODME has been explained in details in the report of cluster analysis.

4.1.1 Calibration process

The goal of the calibration task is to provide a realistic simulation platform to evaluate the potential benefits of ATDM and DMA applications. Given the magnitude of the HD -DTA network and huge difference between initial simulation outputs and observed link counts, our realistic goal is to (a) improve the value of R^2 to at least 50%, (b) optimize the gap between simulated and observed link counts ranges between -50% and 50% at most locations and (c) generate verifiable zone-based trip productions and attractions that match with the local experiences. These three criteria are demonstrated for all the four scenarios before and after the calibration process.

Since the initial O-D demands are critical for the performance of ODME, we decided to use the final calibrated O-D demands as the initial O-D demands for the next scenario. As a result, the initial results of Scenario 1 is very biased since we had to start with the estimated O-D demands from the Phoenix regional DTA model which turned out considerably different from the Tempe HD-DTA mode. In contrast, the initial results of Scenarios 2, 3 and 4 are much better because they used the calibrated O-D demands of other scenarios as the starting points.

As far as the calibration results goes, the next subsections are categorized into three types of calibration metrics for the four scenarios before and after the calibration process. They are:

1. Regression Analysis Results
2. Comparison of Link Counts and
3. Trip Production and Attraction Zones.

4.2 Scenario I (July 17, 2014)

For scenario 1, the initial simulation output and the calibrated output are compared to the field observations in the following subsections.

4.2.1 Comparison of Initial Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-1 shows the regression analysis between initial simulation results and field observations. Apparently, the initial travel demand retrieved from the regional DTA model cannot be used to evaluate ATDM/DMA applications as the R² is only about 25%.

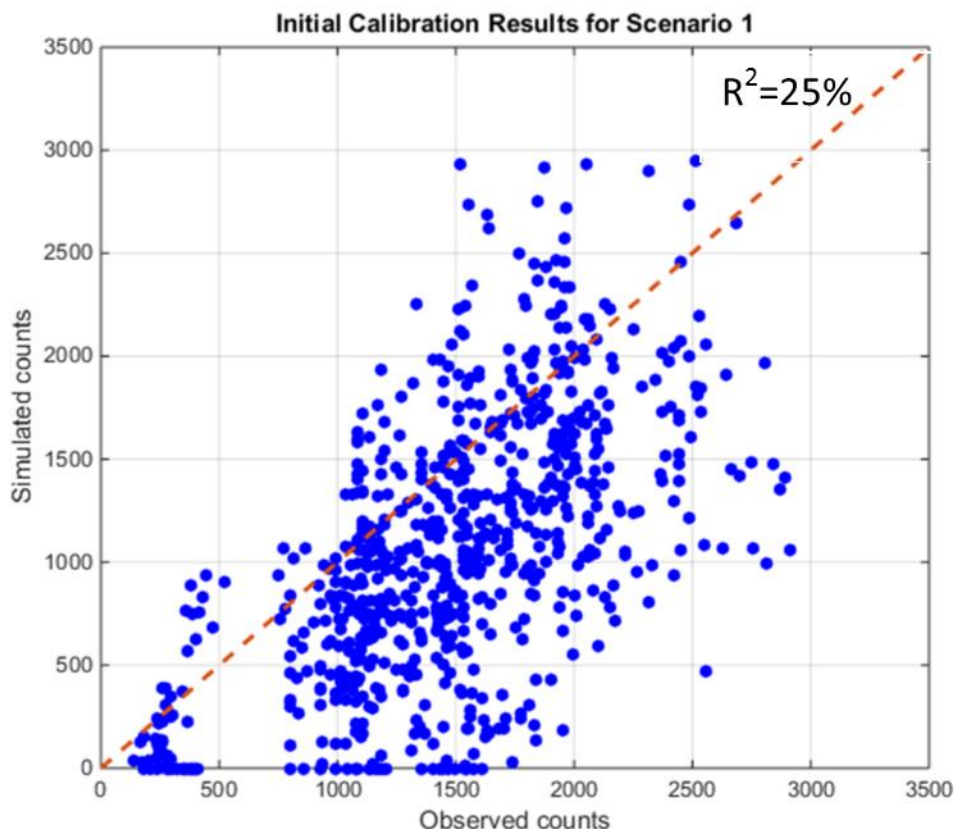


Figure 4-1: Scenario 1 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]

Counts comparison at different locations:

Figure 4-2 shows the location-specific counts comparison between initial simulation results and observed values. Figure 4-2 shows the relative difference between initial simulated results and the field observation as demonstrated by $(\% = \frac{\text{simulated} - \text{observed}}{\text{observed}})$. As shown, the green bar is where the simulated results are close to the field observation and it was seen that as underestimation of initially simulated link counts are demonstrated at most locations.

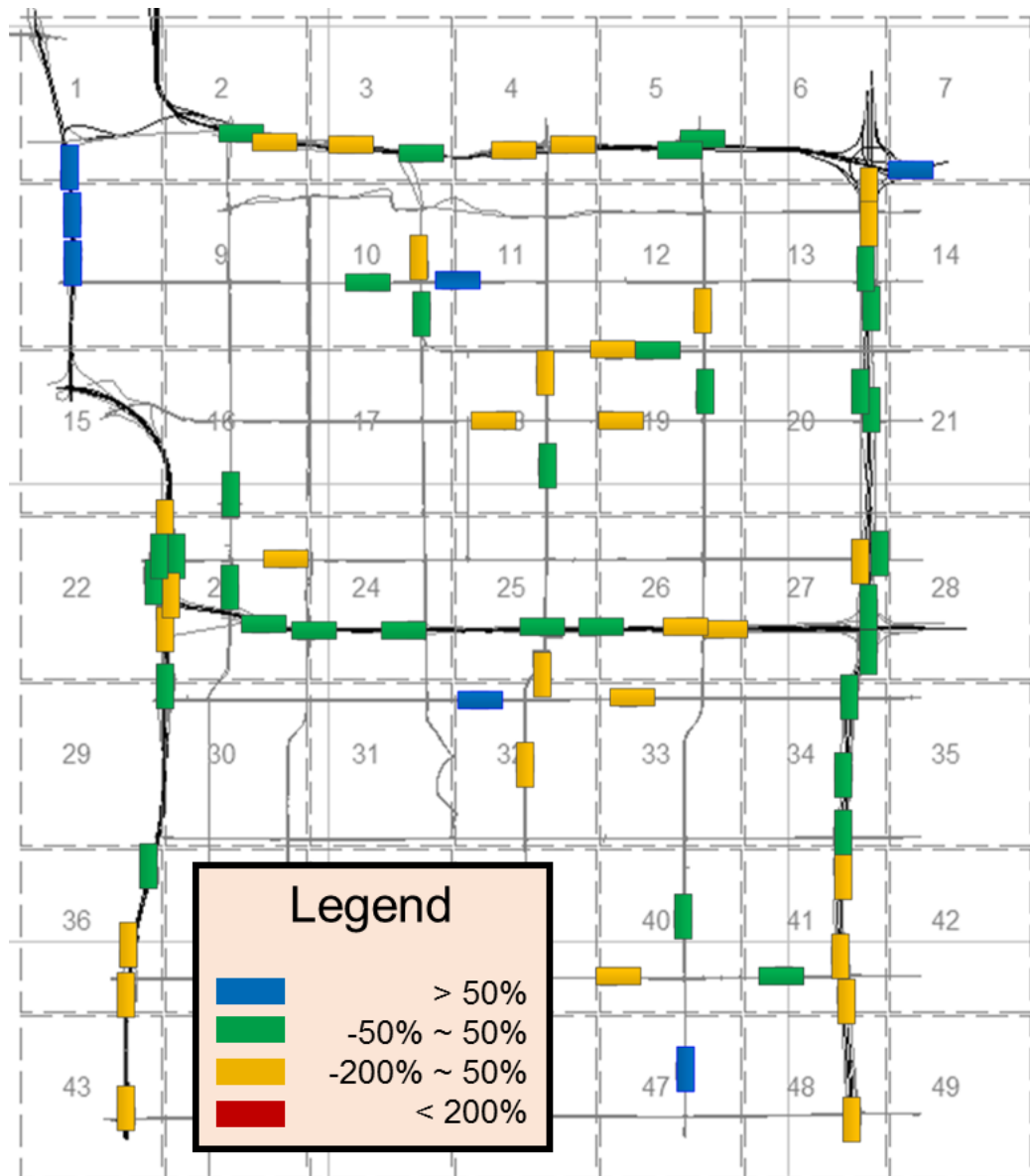


Figure 4-2: Scenario 1 - Comparison of Initial Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Zone-based trip production analysis offers an intuitive representation of traffic pattern in simulation and can be used to compare the simulation results with local experiences. These are shown in Figure 4-3.

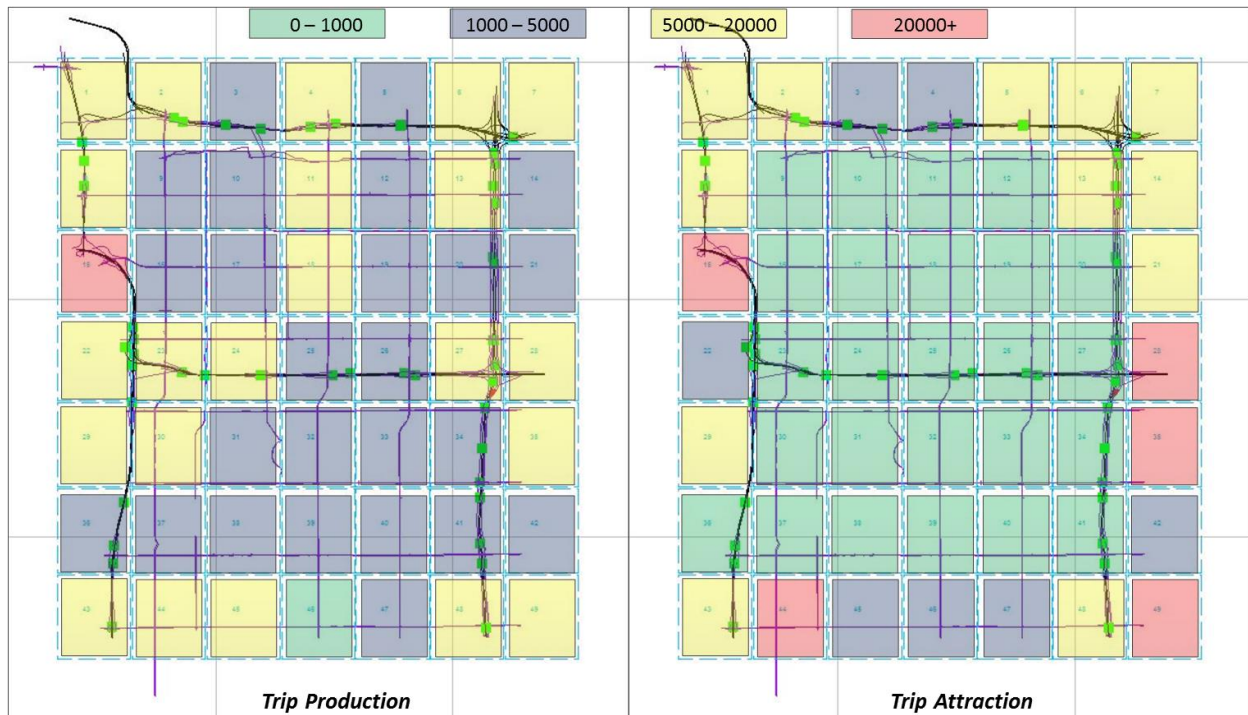


Figure 4-3: Scenario 1 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]

From Figure 4-3, the generated trip on arterials (the inner zones) are unrealistic and the generated trips along freeways are neither balanced nor realistic. For instance, most arterials in the HD-DTA network are major roads and the link volumes during PM peak hours are higher than 5,000 vehicles. Many vehicles leave the City of Tempe during the PM peak hours and therefore the zones containing freeways should have much higher trip attractions than the current pattern.

4.2.2 Comparison of Calibrated Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-4 shows the regression analysis between calibrated simulation results and field observations. Apparently, the calibrated travel demand retrieved from the regional DTA model is much more valid to evaluate ATDM/DMA applications as the R^2 is increased to 56%.

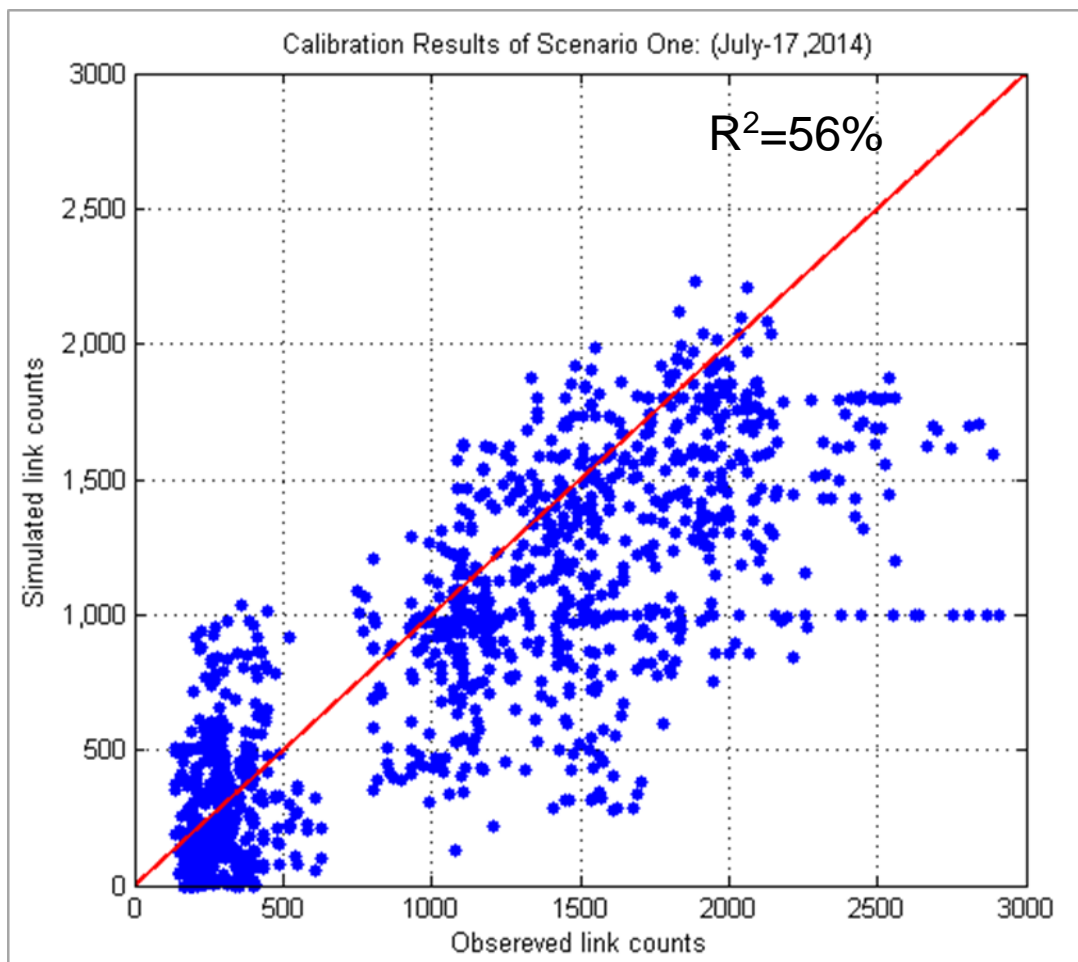


Figure 4-4: Scenario 1 - Regression Analysis of Link Counts after Calibration [Source: ASU]

Comparison of Link Counts:

Figure 4-5 shows the location-specific counts comparison between calibrated simulation results and observed values. It shows the relative difference ($\% = \frac{(\text{simulated} - \text{observed})}{\text{observed}}$) between calibrated simulated results and the field observation and at most locations the difference between simulation results and field observation is much smaller than the initial simulation results.

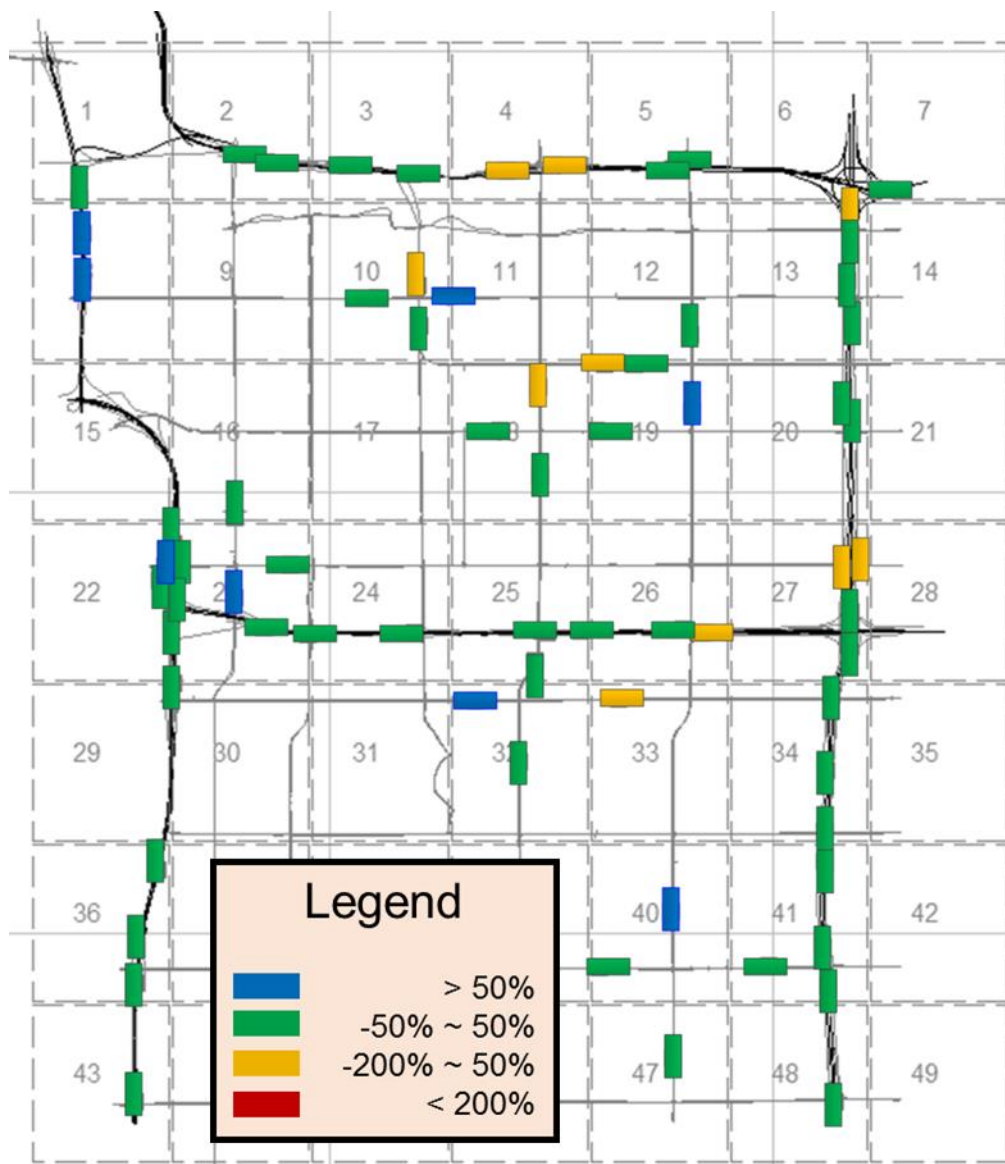


Figure 4-5: Scenario 1 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Figure 4-6 shows the trip production and attraction in the calibrated simulation model. It is consistent with the local experiences. Specifically, more trips are generated within the scope of Arizona State University because a huge amount of commuters leave campus for their homes. The 101 freeway is very congested as well. On the other hand, the inner zones attract a small amount of trips whereas the boundary zones (freeway segments) have much higher volumes than the inner zones. It also make senses since most travelers take freeways to leave Tempe and the last zones where they leave Tempe are considered their destinations. As a result, it is reasonable that the boundary zones have high trip attractions.

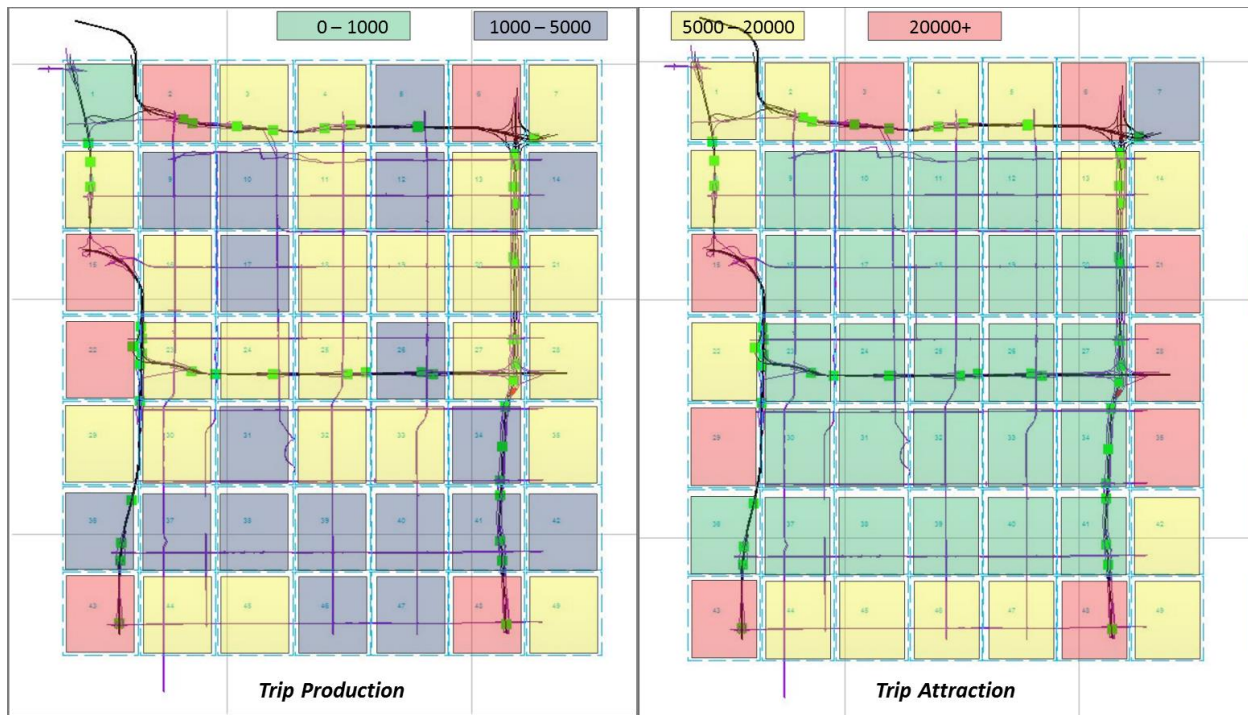


Figure 4-6: Scenario 1 - Calibrated Trip Production and Attraction Matrices of Tempe
[Source: ASU]

4.3 Scenario 2 (May 21, 2014)

For scenario 2, the initial simulation output and the calibrated output are compared to the field observations in the following subsections.

4.3.1 Comparison of Initial Simulation Outputs and Field Observations

Figure 4-7 shows the regression analysis between initial simulation results and field observations. Apparently, the initial travel demand retrieved from the regional DTA model cannot be used to evaluate ATDM/DMA applications as the R^2 is about 41%.

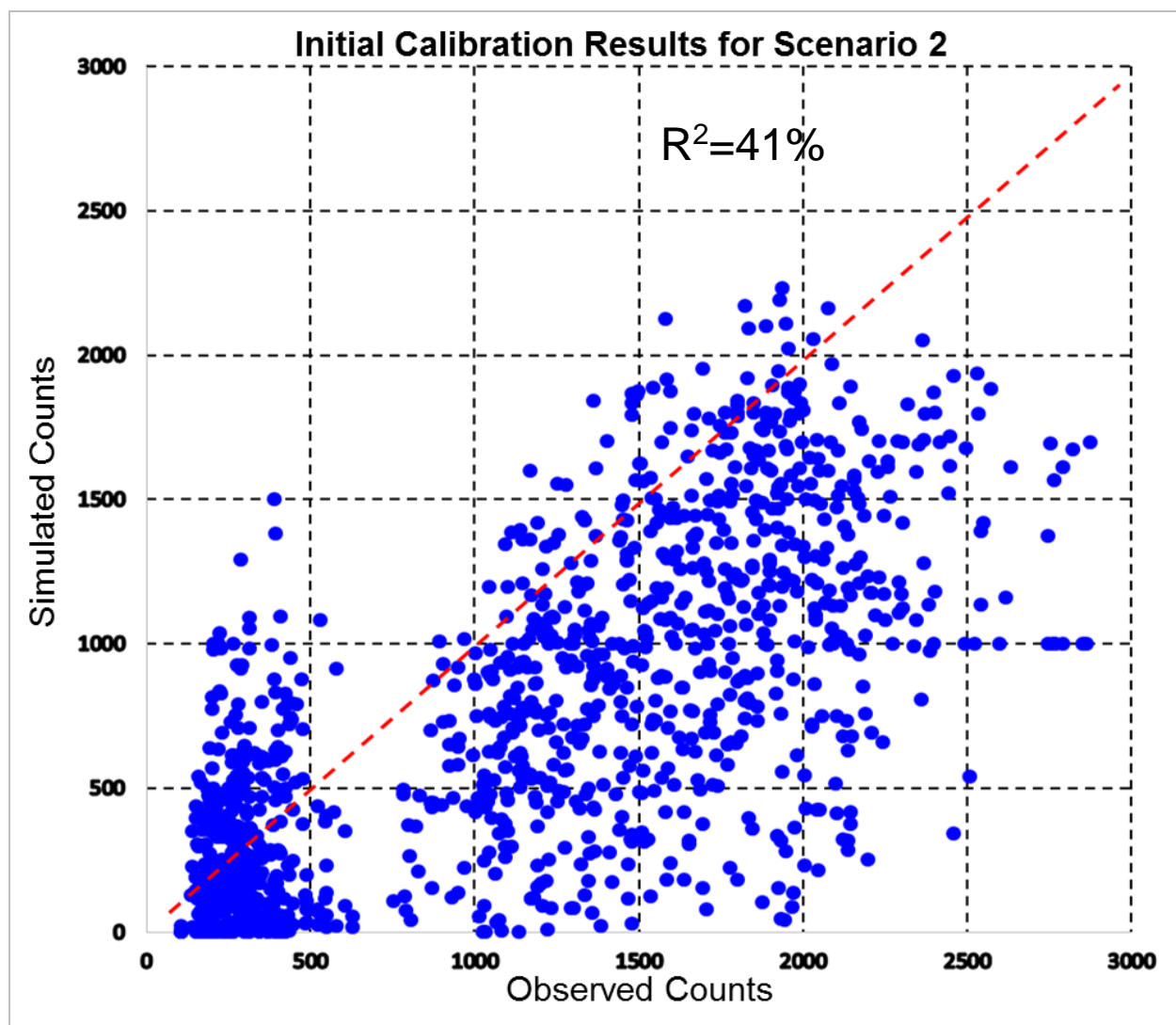


Figure 4-7: Scenario 2 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]

Comparison of Link Counts:

Figure 4-8 shows the location-specific counts comparison between initial simulation results and observed values. It shows the relative difference between initial simulated results and the field observation as given by $(\% = \frac{\text{simulated} - \text{observed}}{\text{observed}})$. The green bar is where the simulated results are close to the field observation. As shown, there is underestimation of initially simulated link counts at most locations.

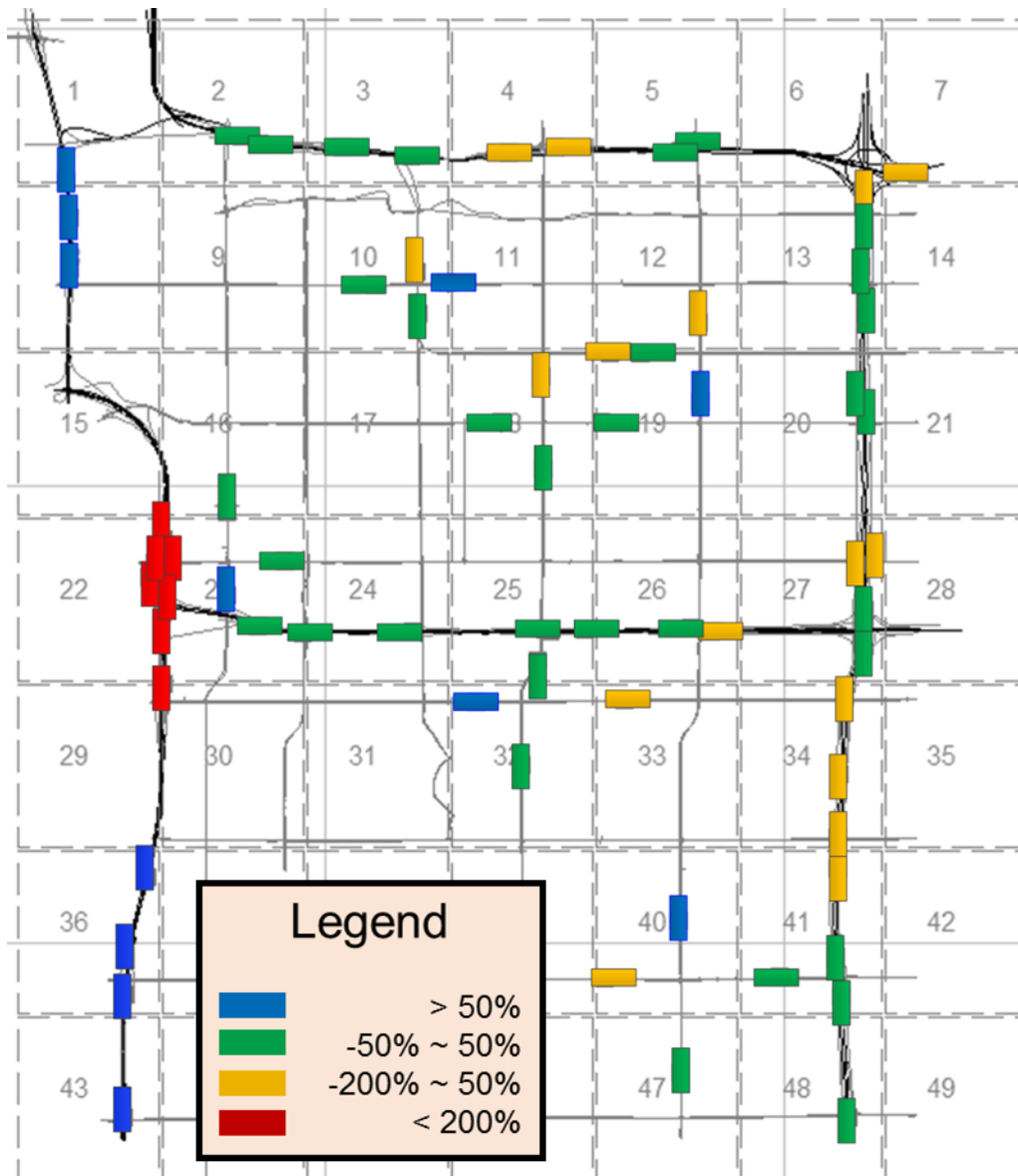


Figure 4-8: Scenario 2 - Comparison of Initial Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Zone-based trip production analysis offers an intuitive representation of traffic pattern in simulation and can be used to compare the simulation results with local experiences.

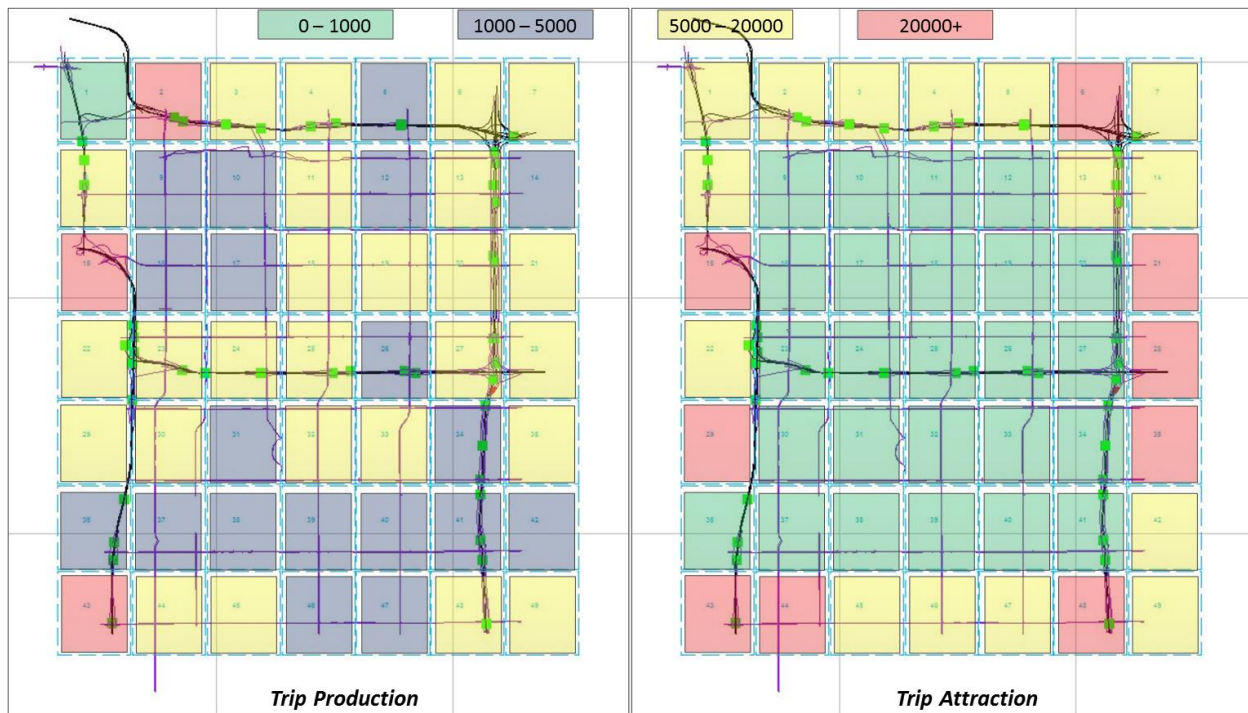


Figure 4-9: Scenario 2 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]

As shown in Figure 4-9, the generated trip on arterials (the inner zones) are unrealistically and the generated trips along freeways are neither balanced nor realistic. On the other hands, most freeway segments attract insufficient trips which is inconsistent with a fact that most commuters will leave Tempe during the PM peak.

4.3.2 Comparison of Calibrated Simulation Outputs and Field Observations

Regression validation:

Figure 4-10 shows the regression analysis between calibrated simulation results and field observations. Apparently, the calibrated travel demand retrieved from the regional DTA model is much more valid to evaluate ATDM/DMA applications as the R^2 is increased to 55%.

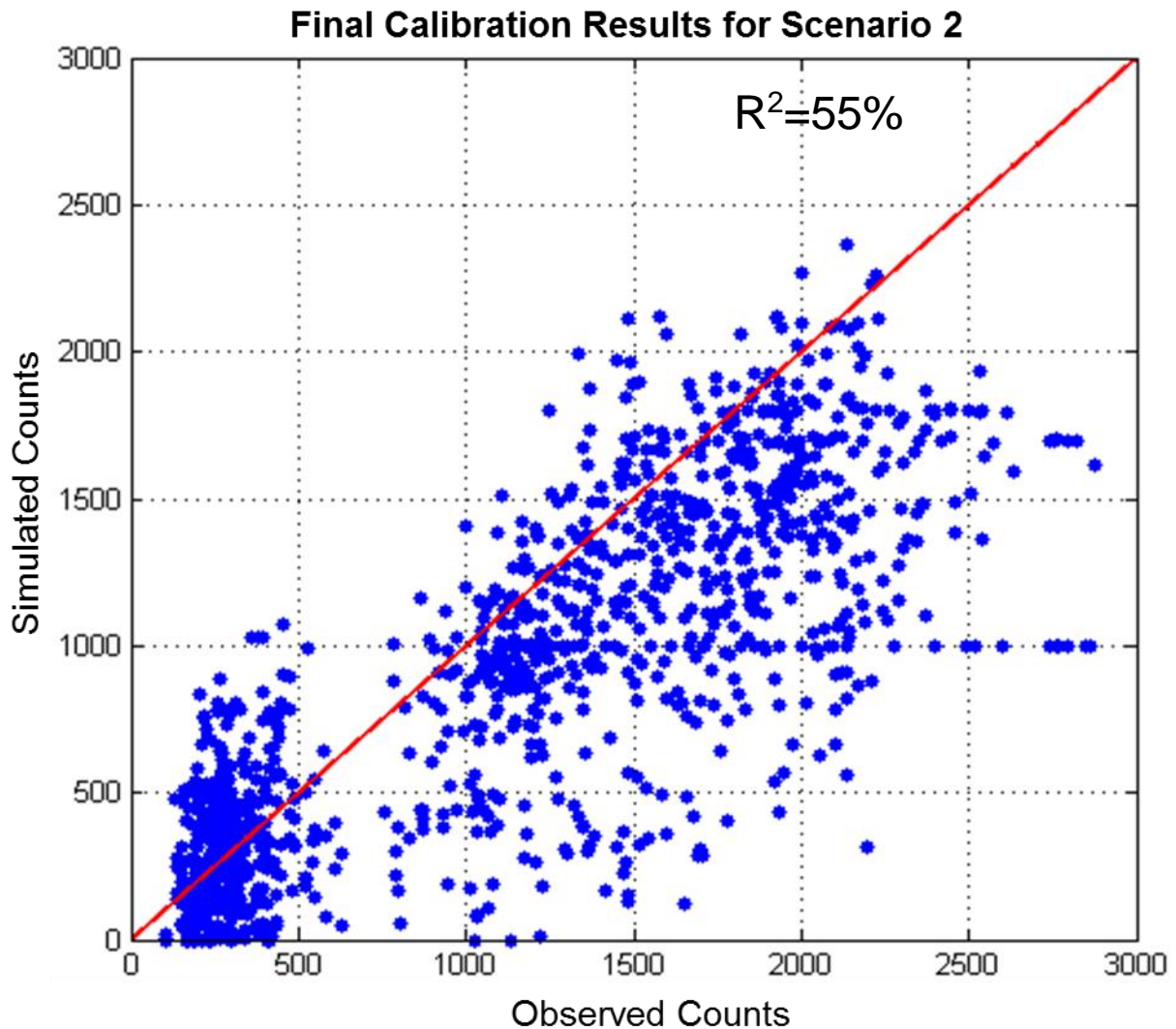


Figure 4-10: Scenario 2 - Regression Analysis of Link Counts after Calibration [Source: ASU]

Comparison of Link Counts:

Figure 4-11 shows the location-specific counts comparison between calibrated simulation results and observed values. It shows the relative difference ($\% = \frac{(simulated - observed)}{observed}$) between calibrated simulated results and the field observation and at most locations the difference between simulation results and field observation is much smaller than the initial simulation results.

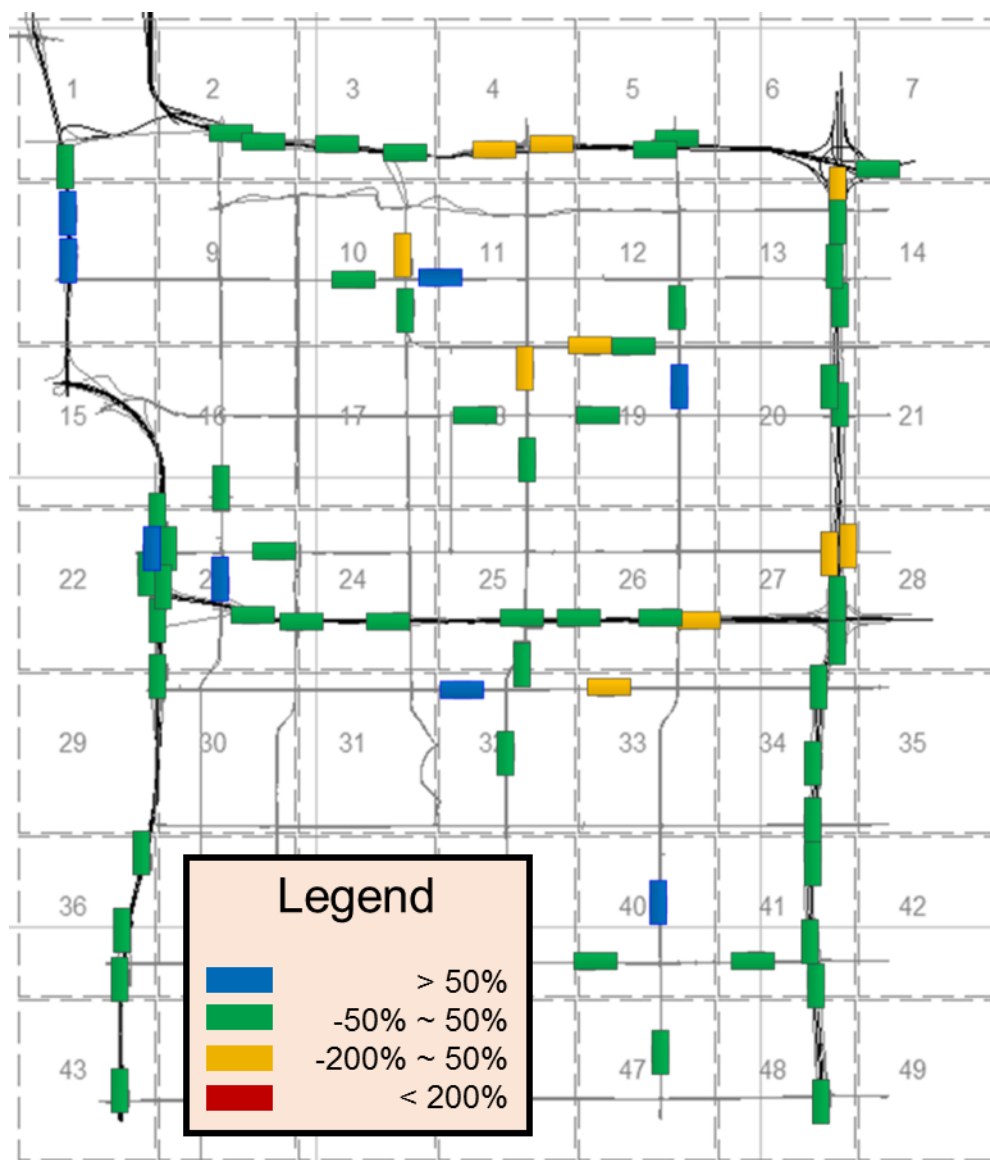


Figure 4-11: Scenario 2 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Figure 4-12 shows the trip production and attraction in the calibrated simulation model. It is consistent with the local experiences. Specifically, more trips are generated within the scope of Arizona State University because a huge amount of commuters leave campus for their homes. The 101 is congested as well. On the other hand, the inner zones attract a small amount of trips whereas the boundary zones (freeway segments) have much higher volumes than the inner zones. It also make senses since most travelers take freeways to leave Tempe and the last zones where they leave Tempe are considered their destinations. As a result, it is reasonable that the boundary zones have high trip attractions.

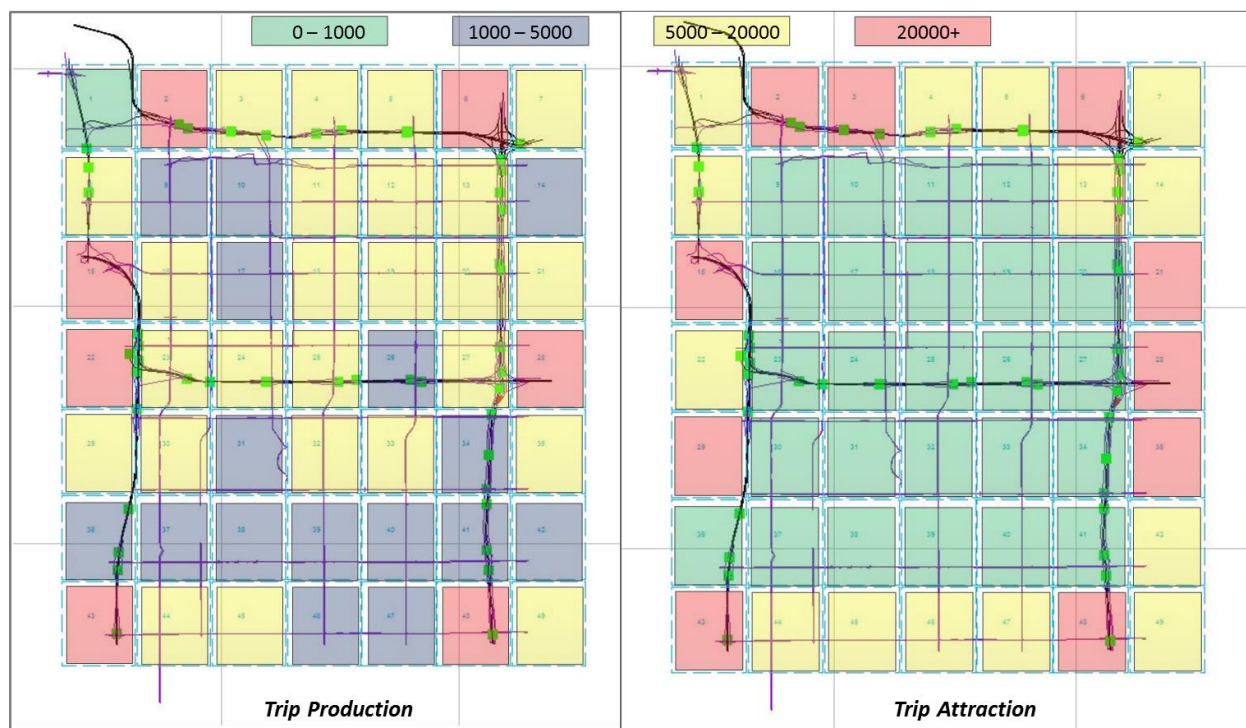


Figure 4-12: Scenario 2 - Calibrated Trip Production and Attraction Matrices of Tempe [Source: ASU]

4.4 Scenario 3 (June 29, 2014)

For scenario 3, the initial simulation output and the calibrated output are compared to the field observations in the following subsections.

4.4.1 Comparison of Initial Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-13 shows the regression analysis between initial simulation results and field observations. Apparently, the initial travel demand retrieved from the regional DTA model cannot be used to evaluate ATDM/DMA applications as the R^2 is about 61%.

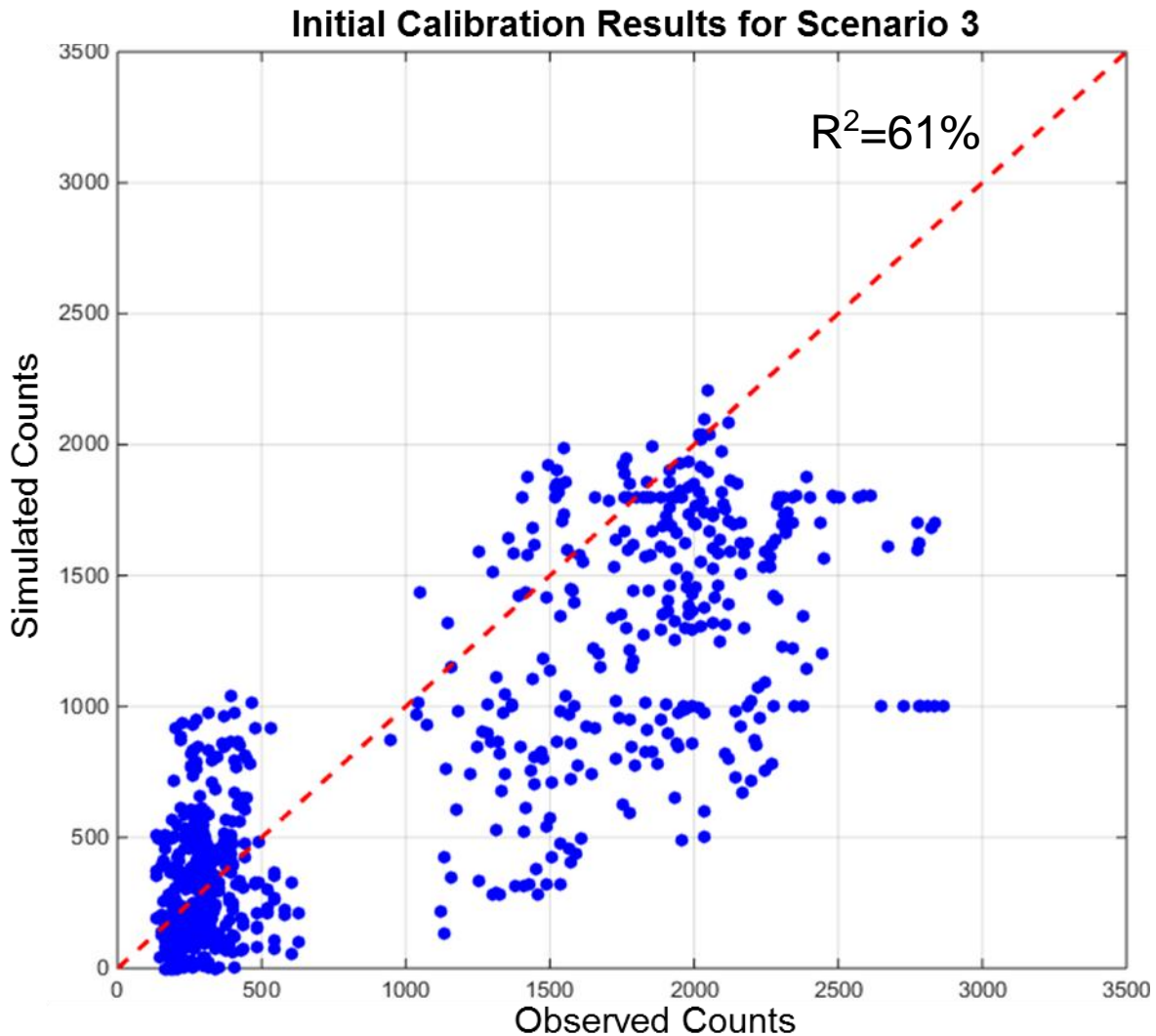


Figure 4-13: Scenario 3 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]

Comparison of Link Counts:

Figure 4-14 shows the location-specific counts comparison between initial simulation results and observed values. It shows the relative difference between initial simulated results and the field observation given as $(\% = \frac{(simulated - observed)}{observed})$. The green bar is where the simulated results are close to the field observation.

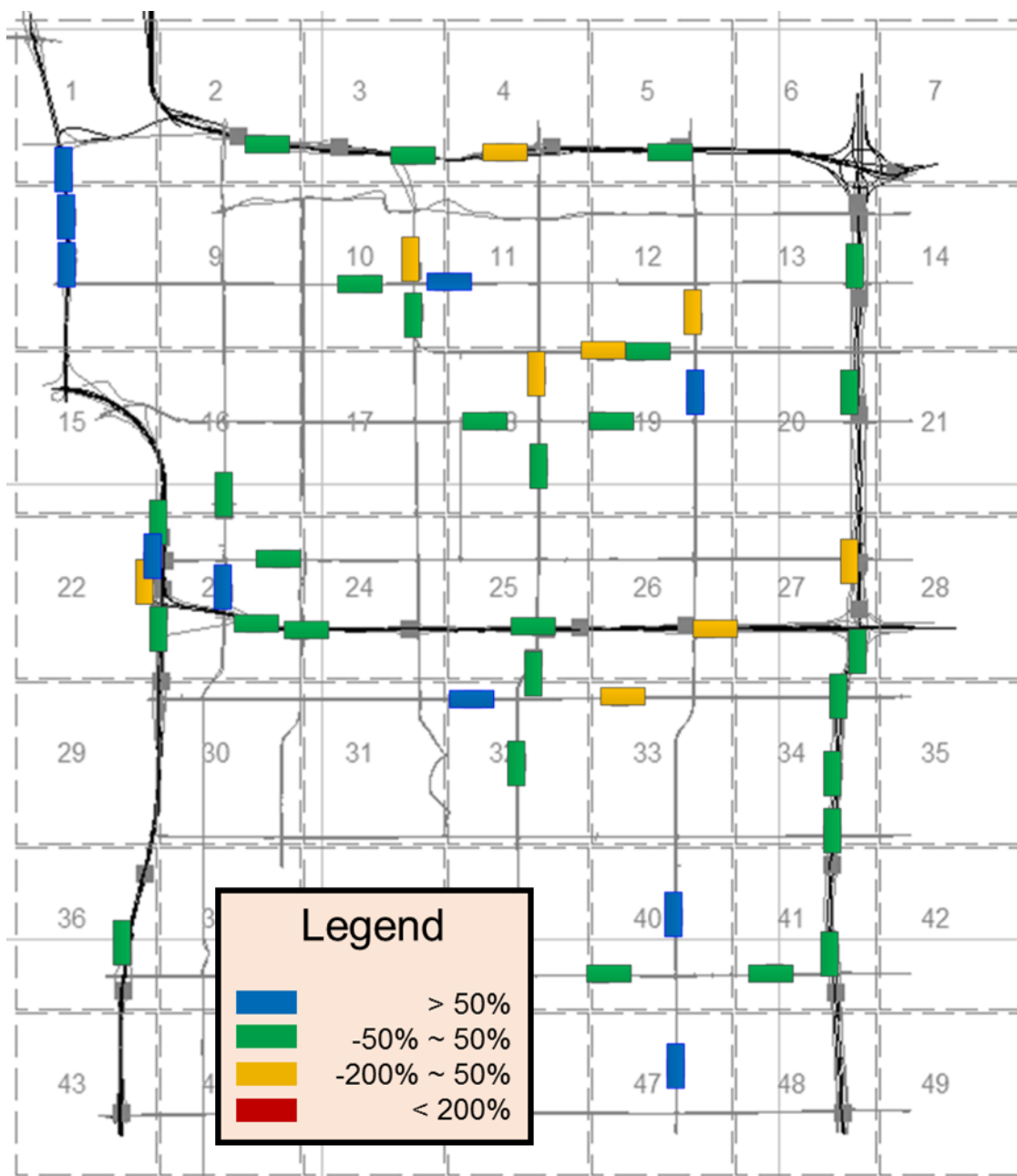


Figure 4-14: Scenario 3 - Comparison of Initial Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Zone-based trip production analysis offers an intuitive representation of traffic pattern in simulation and can be used to compare the simulation results with local experiences.

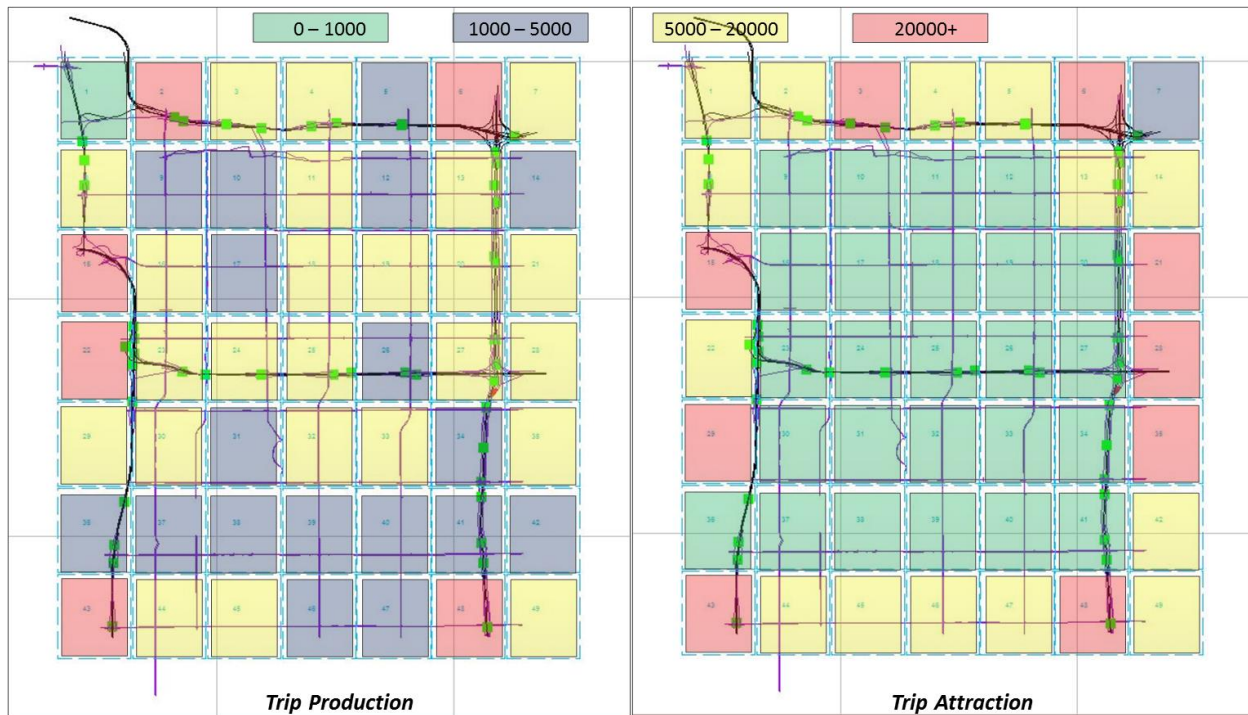


Figure 4-15: Scenario 3 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]

From Figure 4-15, the generated trip on arterials (the inner zones) are unrealistically and the generated trips along freeways are neither balanced nor realistic. On the other hands, most freeway segments attract too few trips which is inconsistent with a fact that most commuters will leave Tempe during the PM peak.

4.4.2 Comparison of Calibrated Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-16 shows the regression analysis between calibrated simulation results and field observations. Apparently, the calibrated travel demand retrieved from the regional DTA model is much more valid to evaluate ATDM/DMA applications as the R^2 is increased to 68%.

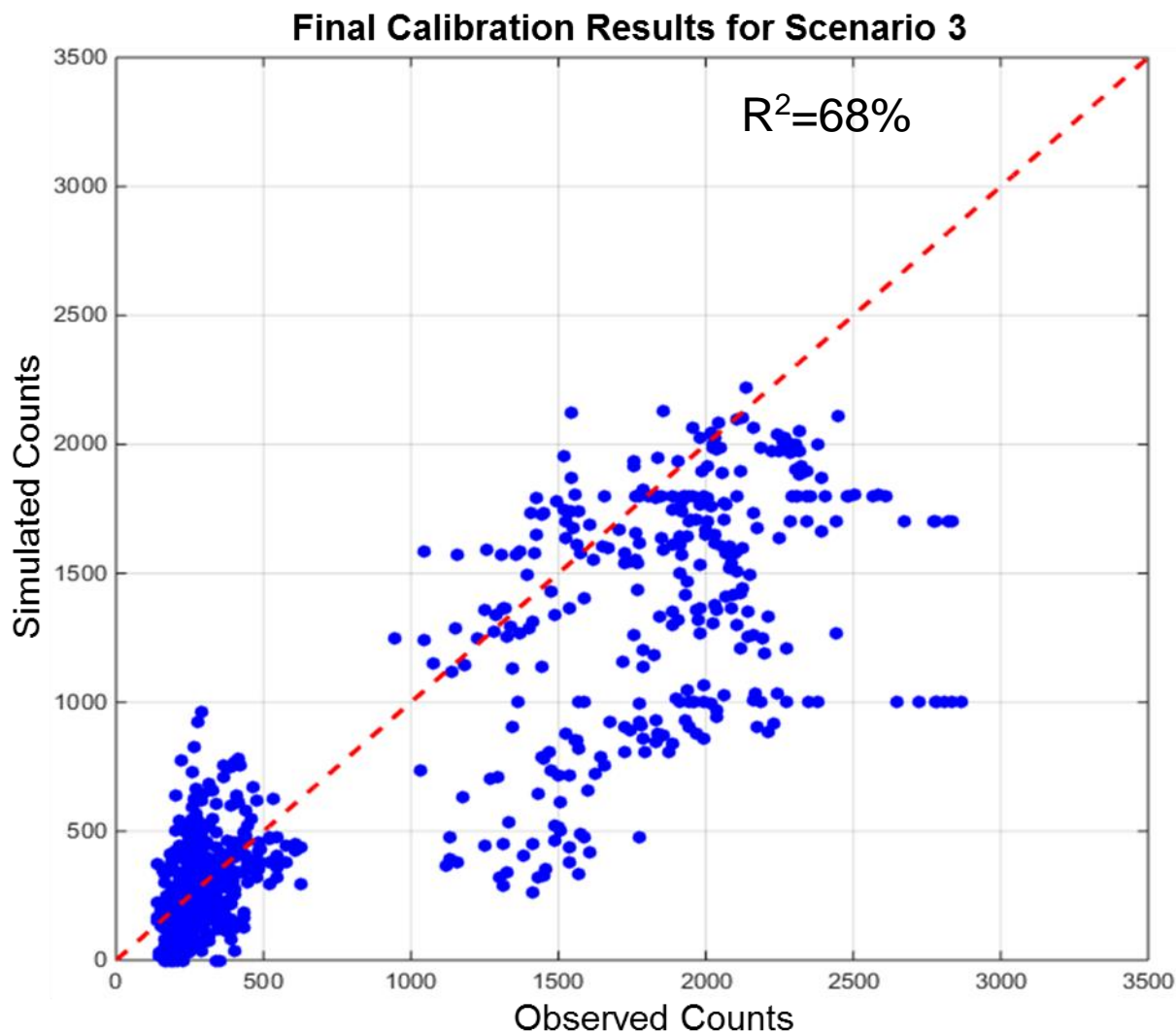


Figure 4-16: Scenario 3 - Regression Analysis of Link Counts After Calibration [Source: ASU]

Counts comparison at different locations: Figure 4-16 shows the location-specific counts comparison between calibrated simulation results and observed values. Figure 4-17 shows the relative difference ($\% = \frac{(simulated - observed)}{observed}$) between calibrated simulated results and the field observation and at

most locations the difference between simulation results and field observation is much smaller than the initial simulation results.

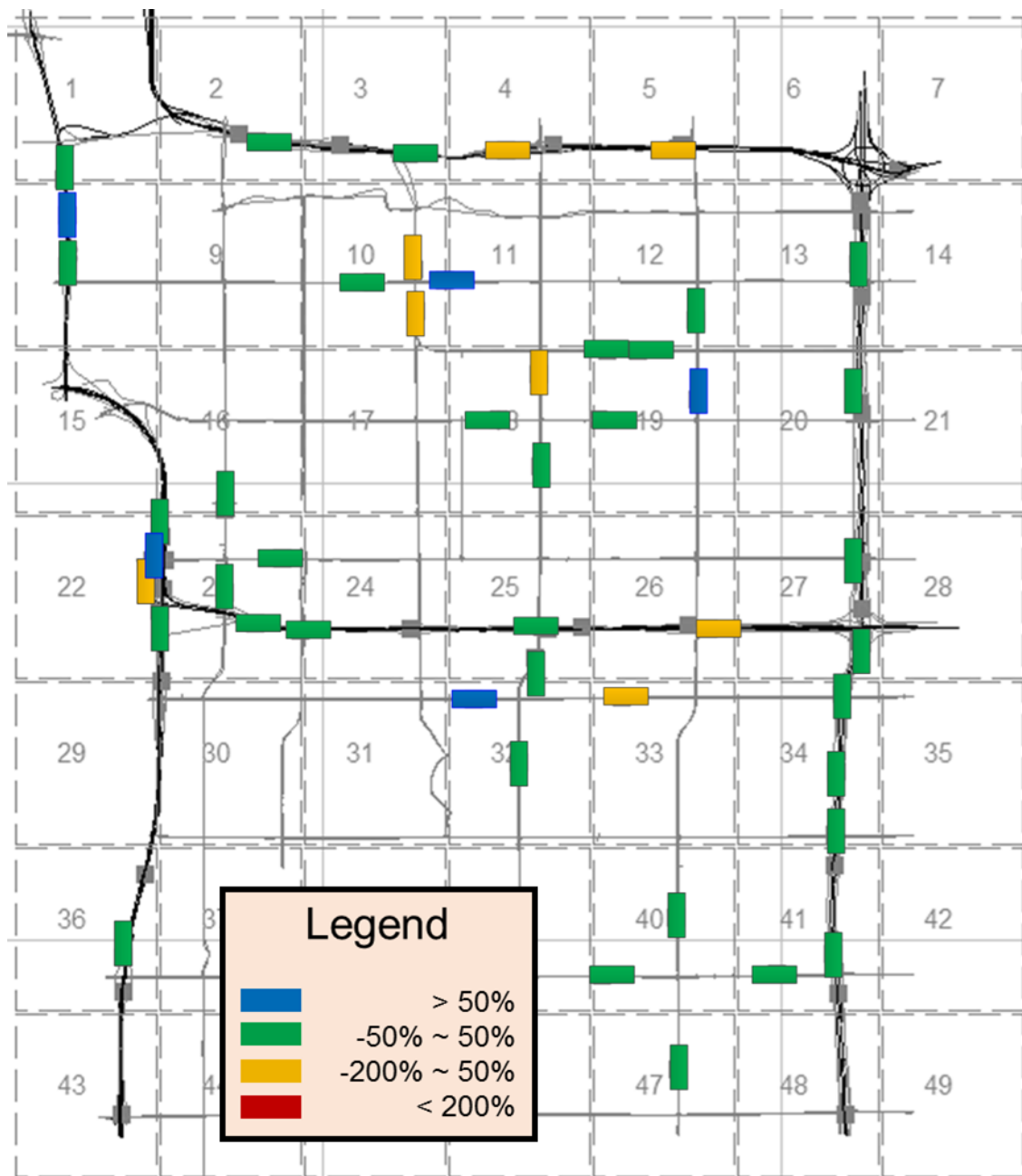


Figure 4-17: Scenario 3 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Figure 4-18 shows the trip production and attraction in the calibrated simulation model. It is consistent with the local experiences. Specifically, more trips are generated within the scope of Arizona State University because a huge amount of commuters leave campus for their homes. The 101 is also very congested as well. On the other hand, the inner zones attract a small amount of trips whereas the boundary zones (freeway segments) have much higher volumes than the inner zones. It also make senses since most

travelers take freeways to leave Tempe and the last zones where they leave Tempe are considered their destinations. As a result, it is reasonable that the boundary zones have high trip attractions.

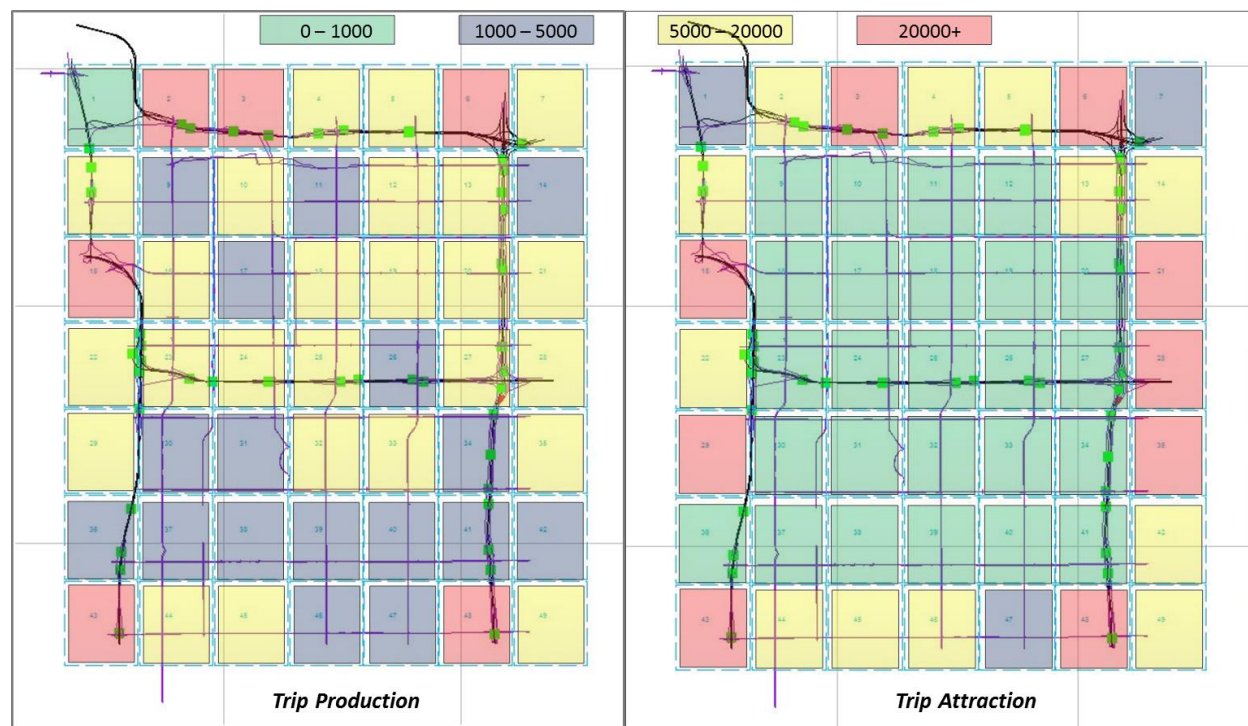


Figure 4-18: Scenario 3 - Calibrated Trip Production and Attraction Matrices of Tempe
[Source: ASU]

4.5 Scenario 4 (November 22, 2013)

For scenario 4, the initial simulation output and the calibrated output are compared to the field observations in the following subsections.

4.5.1 Comparison of Initial Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-19 shows the regression analysis between initial simulation results and field observations. Apparently, the initial travel demand retrieved from the regional DTA model cannot be used to evaluate ATDM/DMA applications as the R^2 is about 65%.

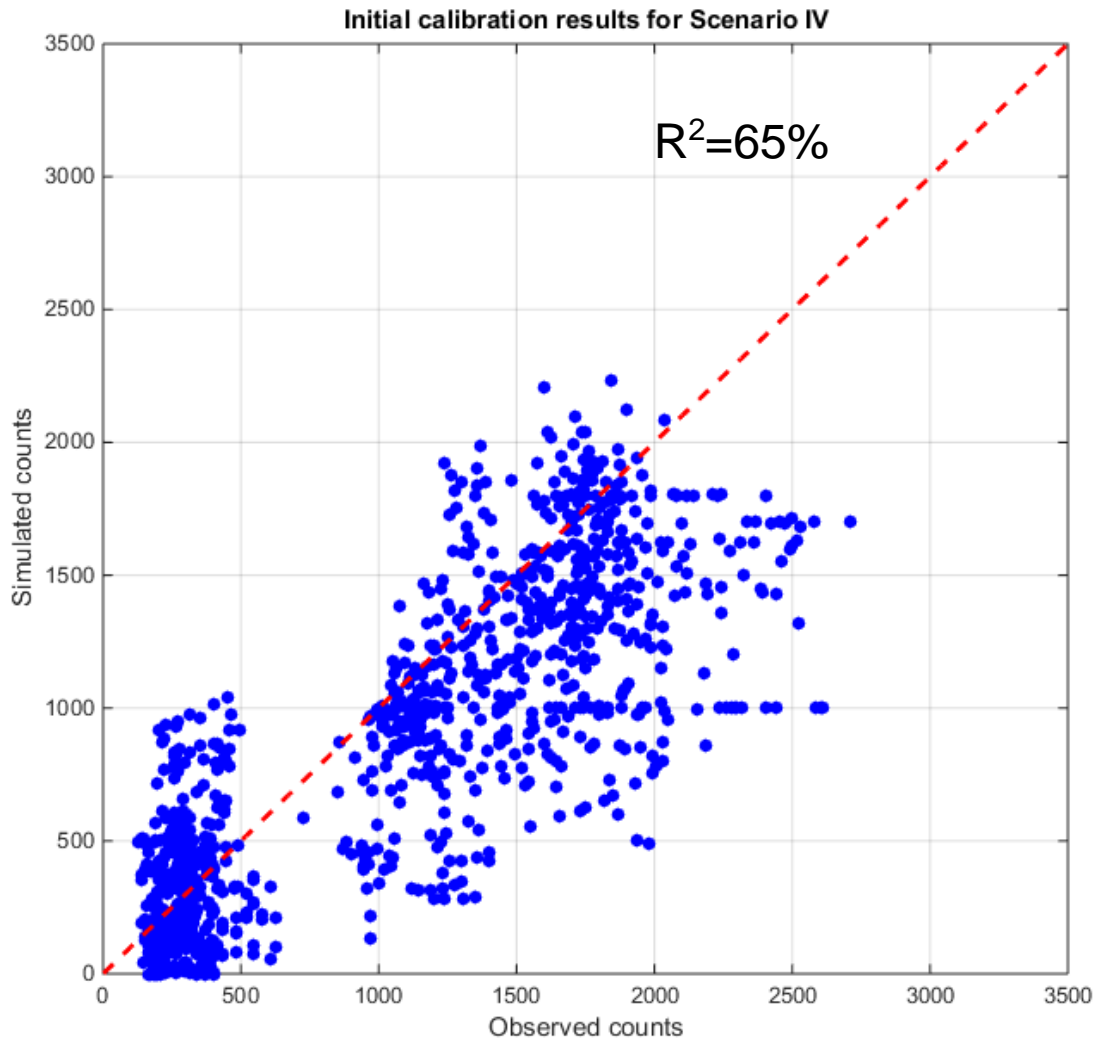


Figure 4-19: Scenario 4 - Regression Analysis of the Initial Simulated Link Counts [Source: ASU]

Comparison of Link Counts:

Figure 4-20 shows the location-specific counts comparison between initial simulation results and observed values. It the relative difference ($\% = \frac{(simulated - observed)}{observed}$) between initial simulated results and the field observation. The green bar is where the simulated results are close to the field observation.

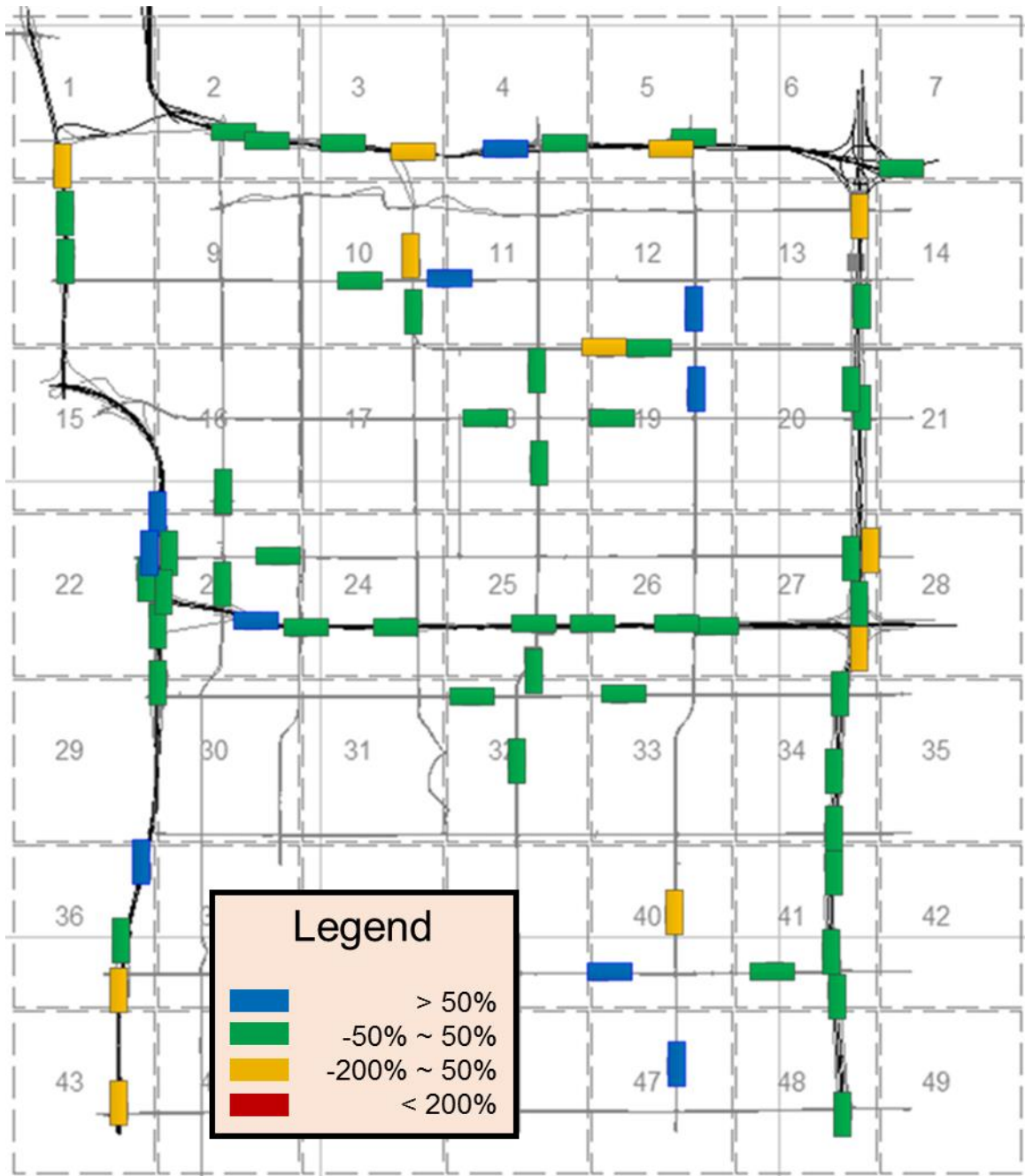


Figure 4-20: Scenario 4 - Comparison of Initial Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Zone-based trip production analysis offers an intuitive representation of traffic pattern in simulation and can be used to compare the simulation results with local experiences. This is shown in Figure 4-21.

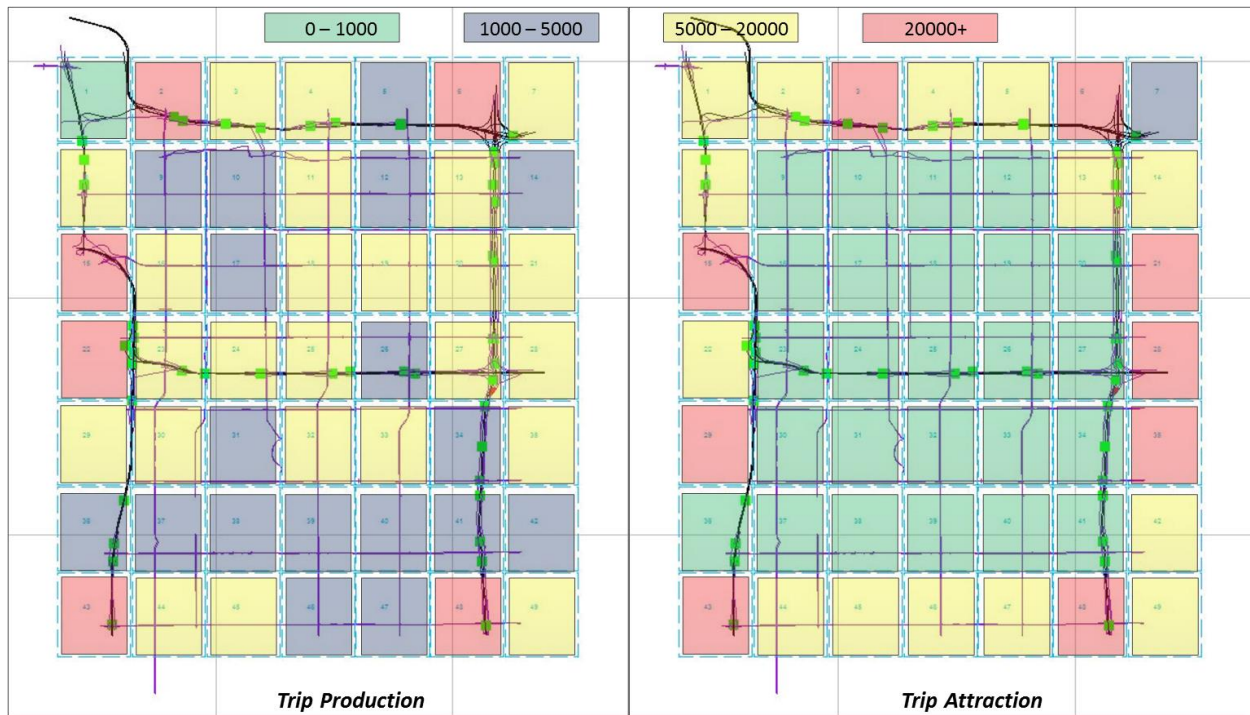


Figure 4-21: Scenario 4 - Initial Trip Production and Attraction Matrices of Tempe [Source: ASU]

4.5.2 Comparison of Calibrated Simulation Outputs and Field Observations

Regression Analysis:

Figure 4-22 shows the regression analysis between calibrated simulation results and field observations. Apparently, the calibrated travel demand retrieved from the regional DTA model is much more valid to evaluate ATDM/DMA applications as the R^2 is increased to 69%.

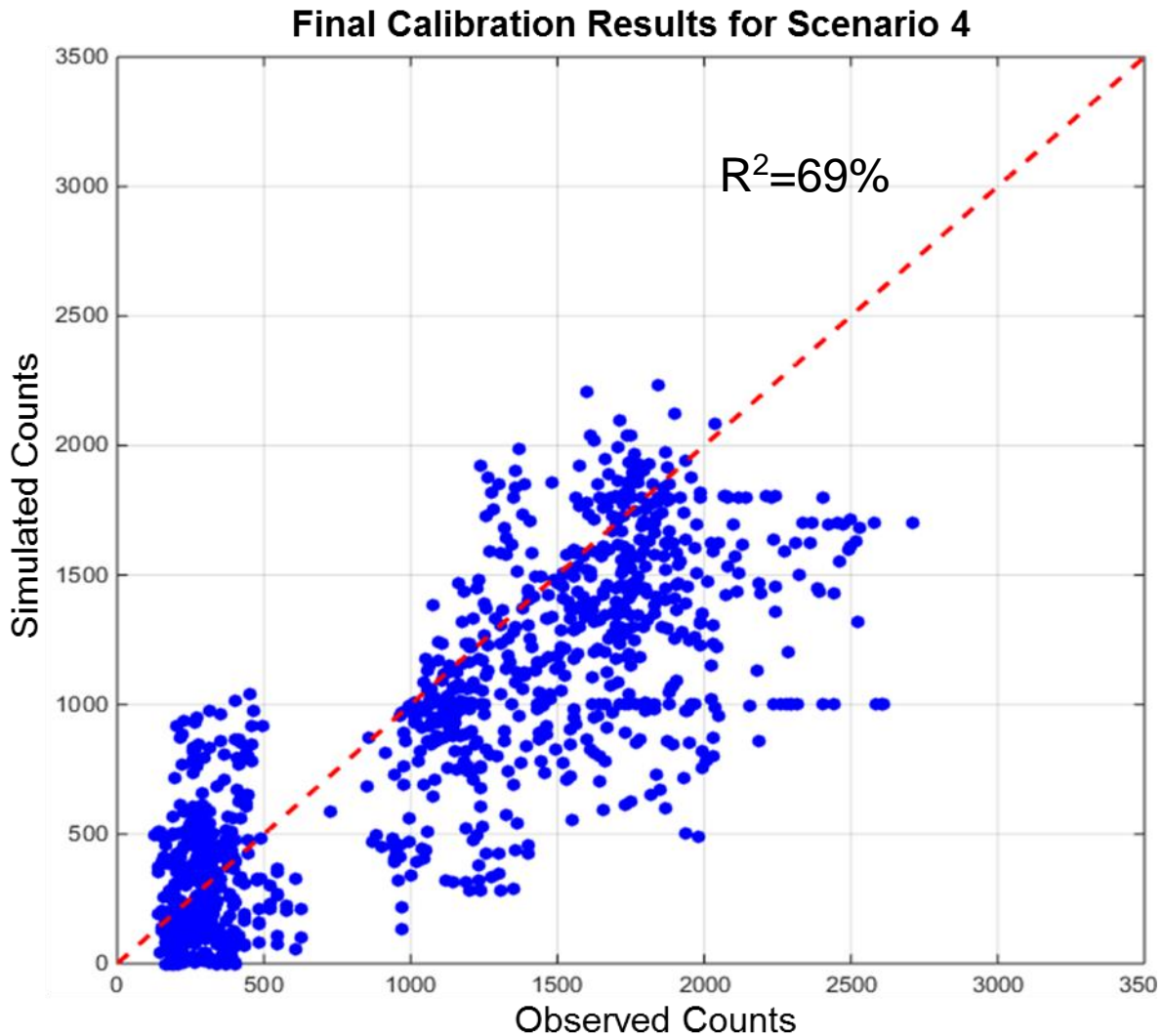


Figure 4-22: Scenario 4 - Regression Analysis of Link Counts after Calibration [Source: ASU]

Comparison of Link Counts:

Figure 4-23 shows the location-specific counts comparison between calibrated simulation results and observed values. Figure 4-23 shows the relative difference ($\% = \frac{(simulated - observed)}{observed}$) between calibrated simulated results and the field observation and at most locations the difference between simulation results and field observation is smaller than the initial simulation results.

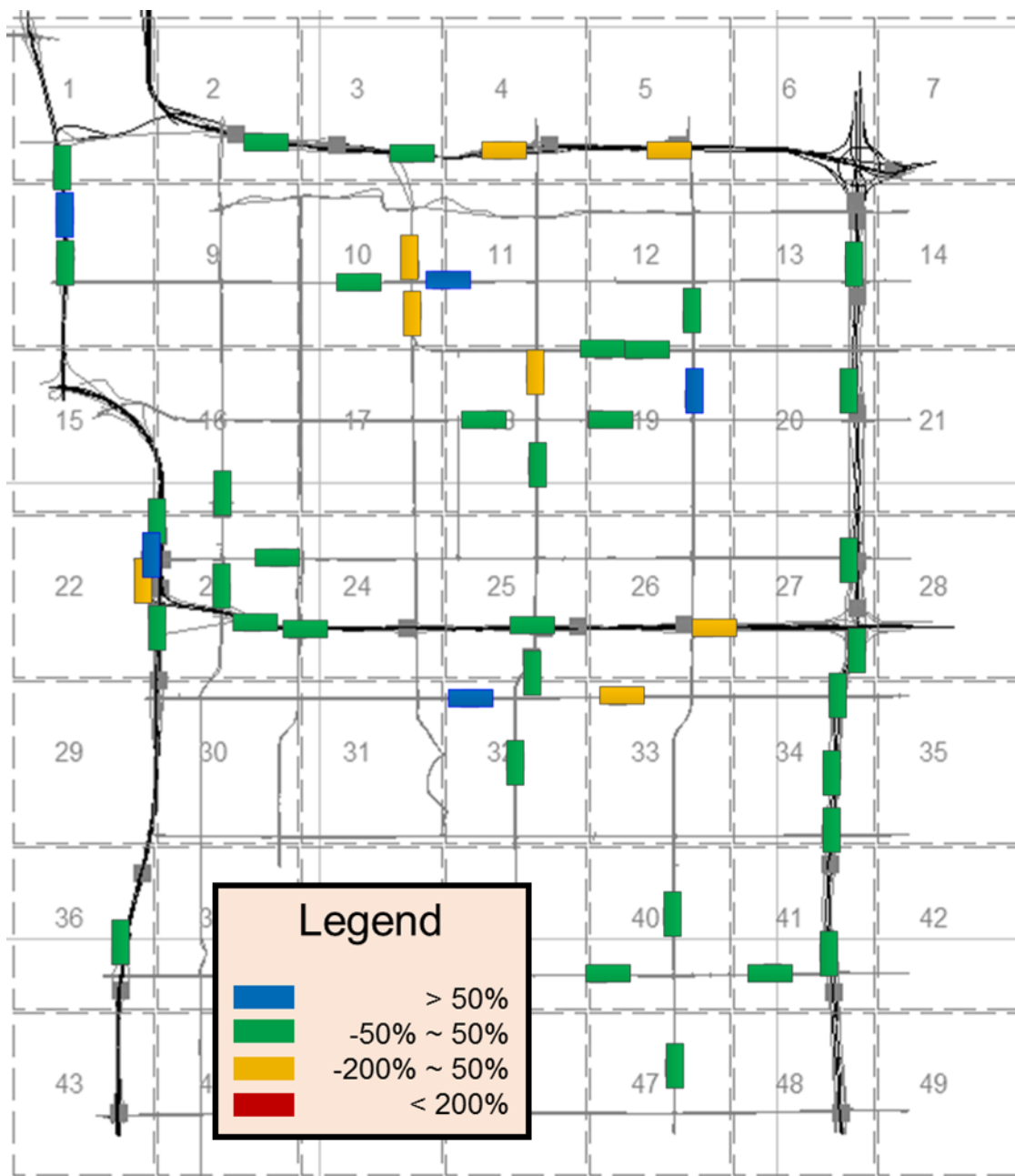


Figure 4-23: Scenario 4 - Comparison of Calibrated Link Counts and Field Observations [Source: ASU]

Trip Production and Attraction Results:

Figure 4-24 shows the trip production and attraction in the calibrated simulation model. It is consistent with the local experiences.

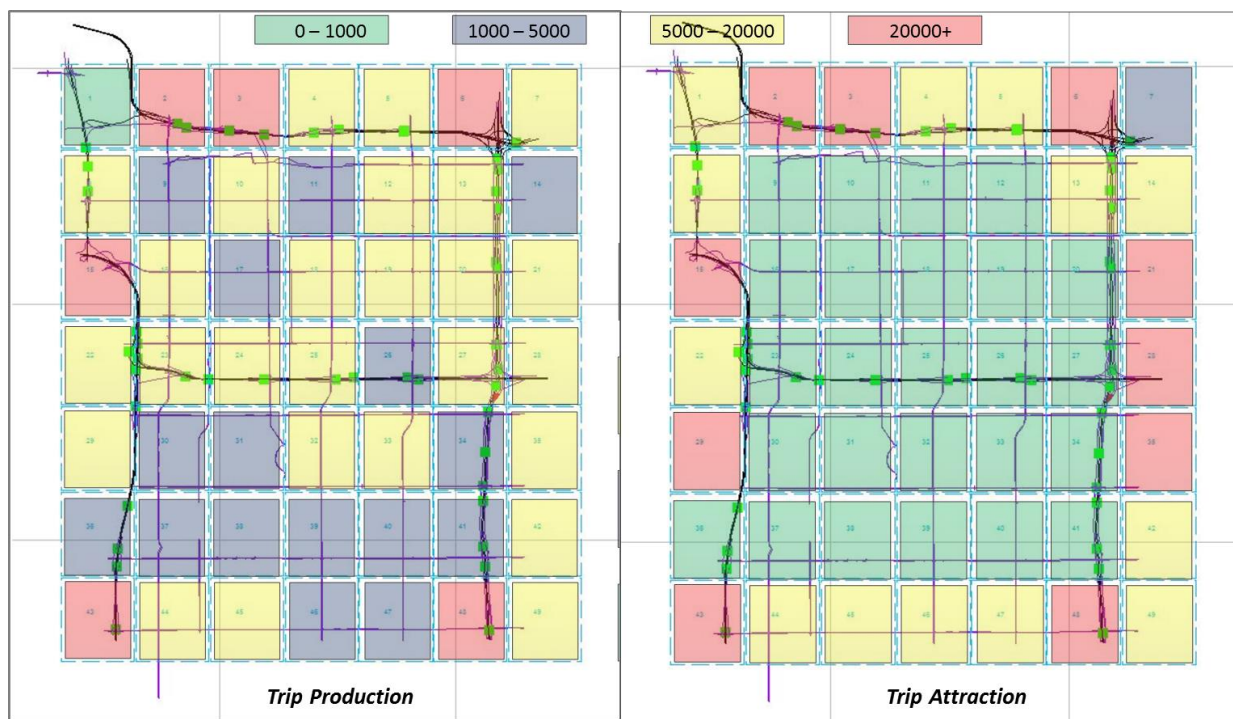


Figure 4-24: Scenario 4 - Calibrated Trip Production and Attraction Matrices of Tempe
 [Source: ASU]

4.6 Calibration Results Summary

The four scenarios were calibrated using a special tool, referred to as OD Matrix Estimation (ODME), to minimize the differences between observed and simulated counts. From the final calibration results, all four scenario are more valid than their initial condition. The simulated counts are close to the observed values and the trip attractions and generations are also consistent with our local experiences. Table 4-1 and Table 4-2 shows the summary calibration statistics in terms of link counts and R-squared values.

Table 4-1: Link Count Calibration Statistics

	Before Calibration			After Calibration		
	Under Estimation	Within Target Measure	Over Estimation	Under Estimation	Within Target Measure	Over Estimation
Scenario 1	10.0%	47.1%	42.9%	11.4%	74.3%	14.3%
Scenario 2	17.1%	48.6%	34.3%	11.4%	74.3%	14.3%
Scenario 3	21.3%	59.6%	19.1%	10.6%	72.3%	17.0%
Scenario 4	21.3%	55.3%	23.4%	10.6%	72.3%	17.0%

Table 4-2: R-squared Calibration Statistics

	Before Calibration	After Calibration
Scenario 1	25%	56%
Scenario 2	41%	55%
Scenario 3	61%	68%
Scenario 4	65%	69%

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free "Help Line" 866-367-7487
www.its.dot.gov

FHWA-JPO-16-379



U.S. Department of Transportation