Jordan Journal of Civil Engineering, Volume **6**, No. **2**, **2012**

Methodological Notes on the Regional Level Validation of a Microscopic Traffic Simulation Model

Hana Naghawi

Assistant Professor, Civil Engineering Department, Jordan University, Amman, Jordan

ABSTRACT

Traffic simulation models have been increasingly used to evaluate and compare alternative complex real-world traffic problems. Simulation is safer, less expensive and faster than field testing. The past few years have witnessed substantial development of transportation network modeling tools and stronger emphasis on addressing the need to model large-scale networks more accurately and efficiently. While these simulation models can be helpful to transportation engineers, the models must be well calibrated and validated before they can provide credible results. However, simulation models have been often conducted under default parameters. This is mainly due to either the difficulties in field data collection or the lack of knowledge of the appropriate procedure to calibrate and validate traffic simulation models.

This paper presents the results of a recent effort to microscopically simulate the regional evacuation plan for New Orleans Metropolitan Area during the hurricane Katrina. The model involved over 300,000 vehicles moving within a road network that covered several thousand square miles over a 48 hour period. Output statistics were generated on a second-by-second basis for each traveler in the system. Model validation was based upon a comparison of the TRANSIMS generated traffic volumes to the corresponding traffic volumes actually observed during the 2005 hurricane Katrina evacuation. The validation process included the percent error estimation and the regression analysis between the simulated and observed traffic volume data. This study was unique in that it is among the first to develop validation criteria for a regional model based on actual traffic data collected during a live regional mass evacuation.

Analysis was performed utilizing percent errors estimation based on direct comparisons of the hourly volumes at each counting station. Also, an alternative validation approach was carried out using regression analysis between the cumulative observed and simulated volumes for the same stations by analyzing the fit for the regression line $y = a + bx + \varepsilon$. The error percentage and the fit were found to be reasonable with an error percentage less than 25 percent and an R-squared value of over 0.80. This indicated that the TRANSIMS simulation model was a realistic representation of the evacuation operations observed during the hurricane Katrina.

KEYWORDS: Validation, Regional validation, Microscopic simulation, Regional simulation, TRANSIMS.

INTRODUCTION

Traffic simulation models are increasingly used to evaluate and compare alternative complex real-world traffic problems. Early studies seeking to apply to traffic simulation models were limited in their geographical scales and time durations (Theodoulou and Wolshon, 2004; Kwon and Pitt, 2005; Jha et al., 2004). Recent simulation models such as TRANSIMS,

Accepted for Publication on 23/11/2011.

DynaSmart, VISSIM, DynusT and CORSIM, have the ability to model second-by-second movements of hundreds of thousands of individual vehicles, moving over vast geographic areas, for periods as long as several days. It is well known that even the most detailed input into the most detailed simulation models has the potential to yield unrealistic or even useless results. Thus, a critical requirement in the development of any simulation model is the validation of the output results. Validation helps to insure, or at least demonstrate, the level of accuracy, so that the output results of the model are reasonably close to those of the essence of the actual system that is being modeled. A validated model also gives a base point from which it is possible to make changes and assess modifications to the system. In such an arrangement, the theory is that once a model is able to reproduce the desired essence reality in the base case, then any different outcomes that result from modifications to the system can logically be assumed to be a consequence of the changes.

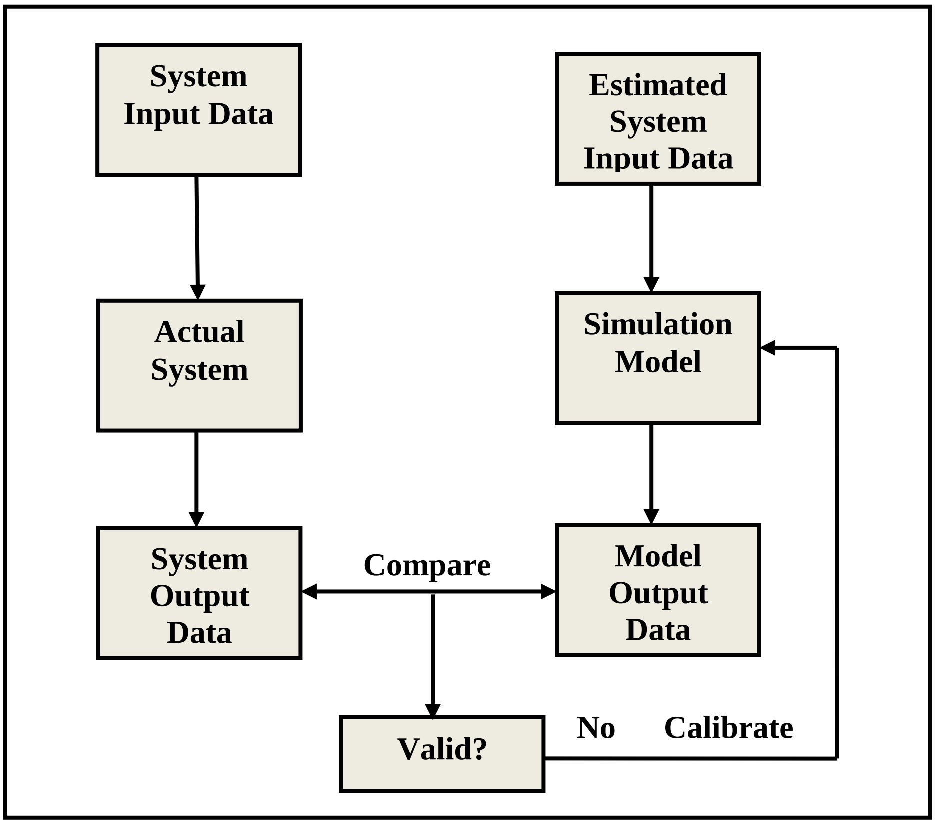

Figure 1: Logical Diagram for Model Validation
Source: (Law and Kelton, 1991)

To date, the review of the state-of-the-art traffic simulation models reveals that many studies addressing the calibration and validation process of microscopic and macroscopic simulation models were limited in their geographical scales and time durations. Kunde (2002) used the speed–density relationship and capacity to calibrate DynaMIT-P. Kim and Rilett (2003) used the simplex algorithm to optimize the degree-of-fit for their models in CORSIM and TRANSIMS. Ma and Abdulhai (2002) and Lee et al.

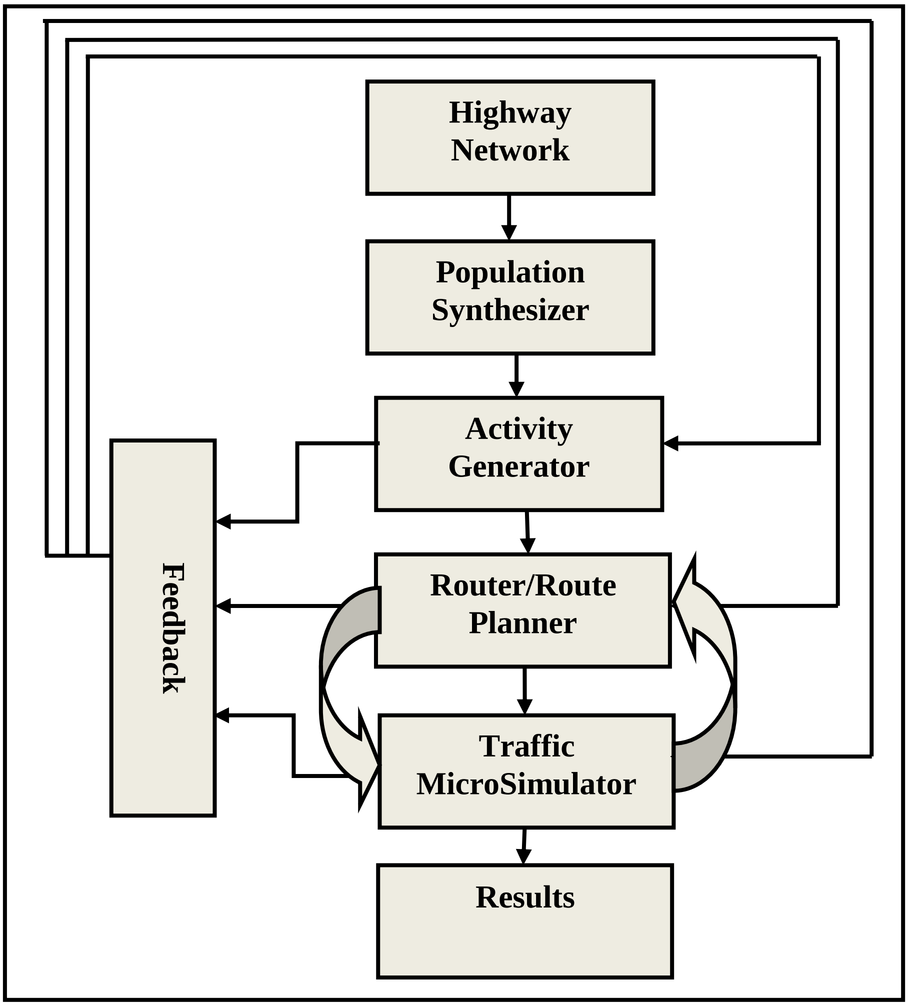

Figure 2: Coding methodology
Source: (TRANSIMS Open Source)

(2001) used PARAMICS traffic simulation model to determine values for the calibration parameters of mean target headway and mean reaction time. Recently, Chiu et al. (2008) conducted a regional scale traffic simulation model using DynusT for the Houston-Galveston area during hurricane Rita, to evaluate various evacuation scenarios, but due to the lack of available data the simulation output could not be calibrated to Rita evacuation. Unfortunately, the calibration and validation of a microscopic regional traffic simulation model output are extremely challenging due to the large geographical area, the detailed network in addition to the long simulation duration. Model calibration and validation form naturally a statistical process in which the uncertainty due to data and model errors should be accounted for. Nearly any statistical test would reject the results of models at this level, even those that were reasonably accurate. Therefore, the choice of the statistical test used to compare the observed and simulated values is a critical task. The conceptual framework for the validation methodology is described in Figure 1. The key question in Figure 1 is: "Is the model valid/do model results realistically represent reality"? The statistical techniques provide a quantified answer to this question. According to Rouphail and Sacks (2003), the probability that the difference between the observed output and the simulated output is less than a specified tolerable difference within a given level of significance can be written as:

$$P\ \{\ |\text{observed-simulated output}\ | \leq d\}\ > \alpha$$

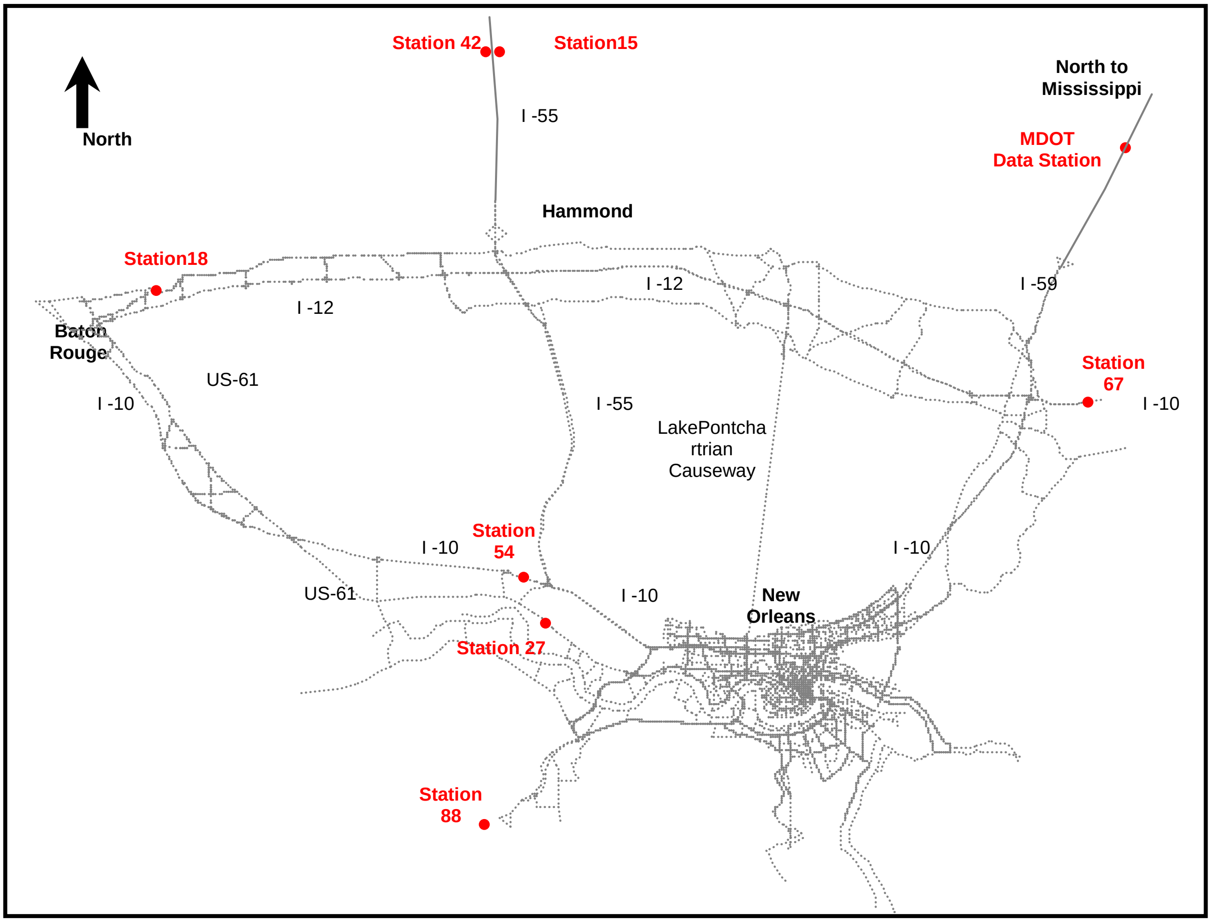

Figure 3: Location of LA DOTD count stations used for volume comparison

where:

d: tolerable difference threshold indicating how close the model is to reality;

α: significance level that tells how the results are obtained from the simulation model.

The key methodological steps for building a valid and credible simulation model are (Law and Kelton, 1991):

- Verification: which is concerned with building the model correctly to ensure the model performance,
- Validation: which is concerned with the accuracy of the model and
- Credibility of the model: which is concerned with the acceptance of the model by the user.

Balci (1998) defines a successful simulation model to be *"the one that produces a sufficiently credible solution that is accepted by decision makers"*. This involves the assessment of the simulation model quality throughout the verification and validation of the simulation models.

In this paper, system validation was based upon a comparison of the TRANSIMS generated traffic volumes to the corresponding traffic volumes actually observed during the 2005 hurricane Katrina evacuation in New Orleans Metropolitan Area. The validation process incorporated a number of steps leading up to quantitative comparisons of the data sets to evaluate the results. The following sections summarize the various data sources and methods used in the validation process as well as the results gained from them.

Table 1. LA DOTD data station observed evacuation volume

	Time (hours after midnight 8/27/05)	Westbound			Eastbound	Northbound	Southbound	Total
		Station 54 I-10 Westbound	Station 27 US-61 Westbound	Station 18 US-190 Westbound	MDOT Station I-59 Northbound	Station15 I-55 Northbound	Station 88 US-90 Southbound	
		433	146	82	306	224	140	**1,331**
	1	323	102	69	202	155	115	**966**
	2	217	81	58	151	98	117	**722**
	3	235	57	41	149	93	90	**665**
	4	206	130	55	171	79	78	**719**
	5	350	127	109	208	136	174	**1,104**
	6	502	183	283	230	230	208	**1,636**
	7	693	225	428	338	384	230	**2,298**
	8	950	234	567	559	499	356	**3,165**
	9	1,317	326	726	793	651	519	**4,332**
	10	1,838	374	784	1,062	853	602	**5,513**
	11	1,816	571	819	1,143	1,158	749	**6,256**
	12	1,743	881	716	1,059	1,196	982	**6,577**
	13	1,704	1,342	731	1,271	1,498	1,201	**7,747**
	14	1,630	1,686	663	1,418	1,616	1,652	**8,665**
	15	1,064	1,785	761	1,112	2,121	1,792	**8,635**
Contraflow Period	16	1,446	1,675	893	1,168	2,148	2,095	**9,425**
	17	2,412	1,743	940	1,526	2,001	2,313	**10,935**
	18	2,174	1,670	970	1,694	2,395	1,994	**10,897**
	19	1,815	1,565	1,022	1,200	2,451	1,771	**9,824**
	20	1,939	1,279	929	612	2,537	2,119	**9,415**
	21	1,901	583	819	532	2,215	1,272	**7,322**
	22	1,805	544	923	438	1,474	991	**6,175**
	23	1,795	513	670	434	1,032	778	**5,222**
	24	1,761	496	342	222	922	750	**4,493**
	25	1,797	511	381	282	1,160	655	**4,786**
	26	1,778	413	292	197	956	615	**4,251**
	27	1,968	567	191	272	1,028	904	**4,930**
	28	2,349	927	255	485	1,565	1,199	**6,780**
	29	2,134	1,344	566	700	1,844	2,006	**8,594**
	30	2,525	1,731	730	1,138	2,599	1,692	**10,415**
	31	2,637	1,881	1,009	1,409	3,280	1,925	**12,141**
	32	2,505	1,804	1,108	1,571	4,017	2,309	**13,314**
	33	2,493	1,760	1,259	1,943	4,407	2,325	**14,187**
	34	2,554	1,695	1,400	1,887	4,660	2,037	**14,233**
	35	2,442	1,660	1,352	2,134	4,742	2,052	**14,382**
	36	2,574	1,708	1,343	2,212	4,833	2,112	**14,782**
	37	2,504	1,696	1,372	2,043	4,710	2,057	**14,382**
	38	2,353	1,684	1,313	1,789	4,893	2,035	**14,067**
	39	2,477	1,680	1,404	1,609	4,695	2,114	**13,979**
	40	2,210	1,733	1,300	2,303	4,600	2,146	**14,292**
	41	1,432	1,540	1,373	3,009	3,951	2,139	**13,444**
	42	573	816	1,228	2,097	2,766	2,177	**9,657**
	43	275	52	673	1,901	1,888	2,214	**7,003**
	44	163	22	305	682	531	2,162	**3,865**
	45	119	17	217	64	255	1,876	**2,548**
	46	81	7	130	28	183	1,336	**1,765**
	47	54	6	75	8	118	272	**533**

Table 2. TRANSIMS simulated evacuation traffic volume

	Time (hours after midnight 8/27/05)	Westbound			Eastbound	Northbound	Southbound	Total
		Station 54 I-10 Westbound	Station 27 US-61 Westbound	Station 18 US-190 Westbound	MDOT Station I-59 Northbound	Station15 I-55 Northbound	Station 88 US-90 Southbound	
		256	14		83	24	99	**476**
	1	463	12		163	355	210	**1,203**
	2	332	9		106	286	162	**895**
	3	258	11		104	189	140	**702**
	4	252	8		101	192	146	**699**
	5	378	8		128	187	166	**867**
	6	532	15		220	326	241	**1,334**
	7	681	29		320	491	407	**1,928**
	8	1,085	25		379	609	476	**2,574**
	9	1,583	35		559	906	710	**3,793**
	10	2,010	53		705	1,229	944	**4,941**
	11	2,395	58		918	1,495	1,083	**5,949**
	12	2,658	78		940	1,749	1,276	**6,701**
	13	2,895	87		1,085	1,890	1,369	**7,326**
	14	3,056	102		1,229	2,016	1,496	**7,899**
	15	3,127	144		1,335	2,410	1,573	**8,589**
Contraflow Period	16	2,448	775		1,351	2,716	1,609	**8,899**
	17	3,706	108		1,431	2,577	1,875	**9,697**
	18	3,646	726	3	1,645	2,682	1,993	**10,695**
	19	3,591	1,178	11	1,472	3,419	2,112	**11,783**
	20	2,905	1,103	55	1,390	2,961	1,988	**10,402**
	21	2,816	719	15	1,219	2,636	1,661	**9,066**
	22	2,534	358	1	980	2,010	1,360	**7,243**
	23	2,322	245	1	871	1,727	1,091	**6,257**
	24	2,090	86		731	1,542	942	**5,391**
	25	1,949	79	5	664	1,256	943	**4,896**
	26	2,005	61	1	652	1,347	866	**4,932**
	27	1,955	53	3	680	1,241	850	**4,782**
	28	2,401	54	2	735	1,348	1,107	**5,647**
	29	3,123	65	6	1,032	1,943	1,510	**7,679**
	30	3,552	100	5	1,299	2,441	1,676	**9,073**
	31	3,863	206	7	1,575	2,601	1,902	**10,154**
	32	4,050	1,233	4	1,824	2,856	2,186	**12,153**
	33	4,057	2,039	10	2,044	4,442	2,420	**15,012**
	34	3,085	2,271	1	2,062	4,227	2,475	**14,121**
	35	2,927	2,289		2,121	3,942	2,503	**13,782**
	36	2,921	2,181		2,088	3,923	2,363	**13,476**
	37	2,937	2,251		2,083	3,972	2,330	**13,573**
	38	2,942	2,260		2,041	3,975	2,435	**13,653**
	39	2,810	2,232		2,007	3,779	2,431	**13,259**
	40	867	2,978		2,208	4,045	2,597	**12,695**
	41	2,664	2,575		2,143	3,743	2,402	**13,527**
	42	3,014	2,030	4	1,772	3,534	2,224	**12,578**
	43	2,908	731	5	1,342	2,952	2,121	**10,059**
	44	2,275	200	19	911	2,080	1,993	**7,478**
	45	1,417	49	2	539	1,122	1,964	**5,093**
	46	925	18		341	737	1,075	**3,096**
	47	491	7		169	486	226	**1,379**
	Total	**107,157**	**31,948**	**160**	**51,797**	**98,616**	**67,728**	**357,406**

Table 3. Comparison of volumes – temporal aggregation

Time Interval	General Travel Direction	Location	Observed Traffic Volume (vph)	Simulated Traffic Volume (vph)	Volume Difference (vph)	Error %	Error % by Time Increment
0-15 Prior to Contraflow	West	I-10 Westbound	15,021	21,961	6,940	**46.20**	**-7.38**
		US 61 Westbound	8,250	688	-7,562	**-91.66**	
		US 190 Westbound	6,892	N/A	-6,892	**-100.00**	
	East	I-59 Northbound	10,172	8,375	-1,797	**-17.67**	
	North	I-55 Northbound	10,991	14,354	3,363	**-30.60**	
	South	US 90 Southbound	9,005	10,948	1,493	**16.58**	
16-39 During Contraflow	West	I-10 Westbound	52,138	70,635	18,497	**35.48**	**-1.38**
		US 61 Westbound	31,129	22,672	-8,457	**27.17**	
		US 190 Westbound	21,483	130	-21,353	**-99.39**	
	East	I-59 Northbound	27,497	33,997	6,500	**23.64**	
	North	I-55 Northbound	66,564	65,563	-1,001	**-1.50**	
	South	US 90 Southbound	40,120	42,628	2,508	**6.25**	
39-47 After Contraflow	West	I-10 Westbound	4,907	14,561	9,654	**196.74**	**24.10**
		US 61 Westbound	4,193	8,588	4,395	**104.82**	
		US 190 Westbound	5,301	30	-5,271	**-99.43**	
	East	I-59 Northbound	10,092	9,425	-667	**-6.61**	
	North	I-55 Northbound	14,292	18,699	4,407	**30.84**	
	South	US 90 Southbound	14,322	14,602	280	**-1.96**	
Total			**352,369**	**357,406**	**5,037**	**1.43**	**1.43**

Table 4. Comparison of volumes – temporal and spatial aggregation

Time Interval	General Travel Direction	Observed Traffic Volume (vph)	Simulated Traffic Volume (vph)	Volume Difference (vph)	Error %	Error % by Time Increment
0-15 Prior to C/F	West	30,163	22,649	-7,514	**-24.91**	**-7.38**
	East	10,172	8,375	-1,797	**-17.67**	
	North	10,991	14,354	3,363	**-30.60**	
	South	9,005	10,948	1,493	**16.58**	
16-39 During C/F	West	104,750	93,437	-11,313	**-10.80**	**-1.38**
	East	27,497	33,997	6,500	**23.64**	
	North	66,564	65,563	-1,001	**-1.50**	
	South	40,120	42,628	2,508	**6.25**	
39-47 After C/F	West	14,401	23,179	8,778	**60.95**	**24.10**
	East	10,092	9,425	-667	**-6.61**	
	North	14,292	18,699	4,407	**30.84**	
	South	14,322	14,602	280	**-1.96**	
Total		**352,369**	**357,406**	**5,037**	**1.43**	**1.43**

METHODOLOGY

Model Development

The simulation involved over 300,000 vehicles moving within a road network that covered several thousand square miles and temporally during a 48 hour period during which output statistics were generated on a second-by-second basis. Figure 2 shows a schematic diagram summarizing the general flow of the coding methodology that translated the assumed Katrina evacuation characteristics into TRANSIMS model. The first step in the process required the creation of the model *Highway Network* of the region including its key characteristics (speed, number of lanes, control,… etc.). This network also served as an input to spatially distribute the synthetic population. The second step of the development process involved the creation of a representative population of people and households in the study area using the TRANSIMS *Population Synthesizer* module. The synthetic population was based on the 2000 UScensus aggregated data and the disaggregated data from Public Use Microdata Samples (PUMS). Land use data were also used to locate households relative to the transportation networks. The synthetic population and the household activity survey files were used to feed the TRANSIMS *Activity Generator* module. The *Activity Generator* assigned travel activity patterns to individual household members and distributed these activities to locations and modes. The synthetic activity served as input to the TRANSIMS *Router/Route Planner* module to generate travel plans for evacuation trips. Finally, all of the movements and their interactions within the network were generated by the TRANSIMS *Microsimulator* module using the travel plans generated by the *Router*. Feedback is applied to the equilibration process iterating between *Route Planner* and *Traffic Microsimulator*. Through feedback module, some routes may be found infeasible. These activities are then passed back to the *Activity Generator* to determine appropriate alternatives. Some trip plans cannot be followed in the *Traffic Microsimulator* because of time-dependent road closures and other causes. In this case, individuals with those plans are passed back to the *Route Planner* for new routing suggestions. Finally, TRANSIMS can create aggregate results comparable to traditional analysis tools. The *Microsimulation* can lead to highly detailed snapshot data; for example, the exact location of every traveler at any given time.

Table 5. Comparison of volumes – cumulative directional aggregation

Time Interval (hr)	General Travel Direction	Observed Traffic Volume (vph)	Simulated Traffic Volume (vph)	Volume Difference (vph)	Error %	Error % by Time Increment
0-47	**West**	149,314	139,265	-10,049	**-6.73**	**1.43**
	East	47,761	51,797	4,036	**8.45**	
	North	91,847	98,616	6,769	**7.37**	
	South	63,447	67,728	4,281	**6.75**	
Total		**352,369**	**357,406**	**5,037**	**1.43**	

Data Sources

The validation process of the TRANSIMS New Orleans hurricane evacuation model was based on traffic volume data. While it has been suggested that other validation measures of effectiveness (MOE), like vehicle speeds, headways, occupancies,… etc., could have also been used to evaluate the model results, none of these parameters were available at the time of the evacuation. As a result, the basic goal of the validation was to have the modeled traffic patterns reproduce traffic patterns similar to those observed during the Katrina evacuation of 2005.

The traffic volume data used for this study were collected by the Louisiana Department of Transportation and Development (LA DOTD) Office of Planning and Programming as part of their statewide traffic data collection program. The objective of this program is to continuously record traffic volumes to monitor long-term traffic trends on a statewide level. The data are used primarily for aggregate-level planning and trend analyses. However, they can also be extracted more frequently and compiled for the assessment of traffic conditions associated with particular events, such as, in this case, the evacuation for hurricane Katrina.

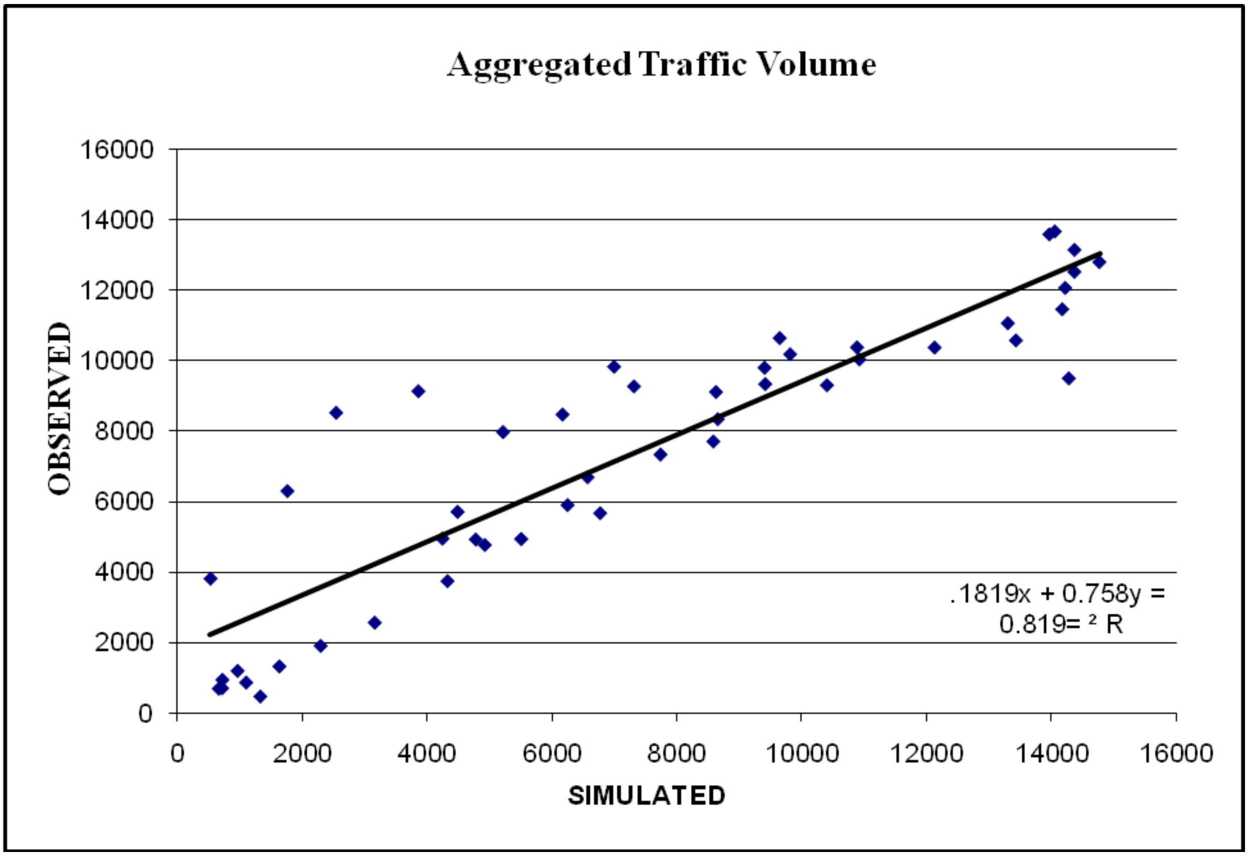

Figure 4: Regression comparison of aggregated traffic volumes

As part of the LA DOTD monitoring program, traffic volumes are collected on a routine basis using a network of 82 permanent count stations located on various roads across the state. These automated recorders are arranged to provide a representative sample of traffic on all road classifications (freeway, arterial, collector,… etc.) across the urbanized and non-urbanized regions of the state. During August 2005, 67 of the 82 LA DOTD data recorders were in operation, of these, 16 of the stations were located on Interstate (I) Freeways, 22 were on US Highways and the remaining 29 were on Louisiana State Highway (LA) system roads.

For this study, data from a total of eight stations located on the major outbound evacuation routes from the New Orleans Metropolitan Areas were used for comparison. The approximate locations of these stations are illustrated in Figure 3. These stations were selected because they were the stations that monitored output routes in the New Orleans area while limiting the potential inclusion of local (i.e., non-evacuation specific) traffic. Several of them were also located near, or in the case of Station 42 – directly on, the contra flow segments.

The LA DOTD data used for the validation encompassed a 48 hour period from 12:00 am Saturday, August 27th through 12:00 am Monday, August 29th, reflecting the Katrina evacuation process.

During this period, the hourly traffic volumes fluctuated at various times. However, the cumulative volume trend, aggregated for all stations, resulted in the characteristic Double-S cumulative distribution curve. The observed traffic volumes are shown in tabular form in Table 1.

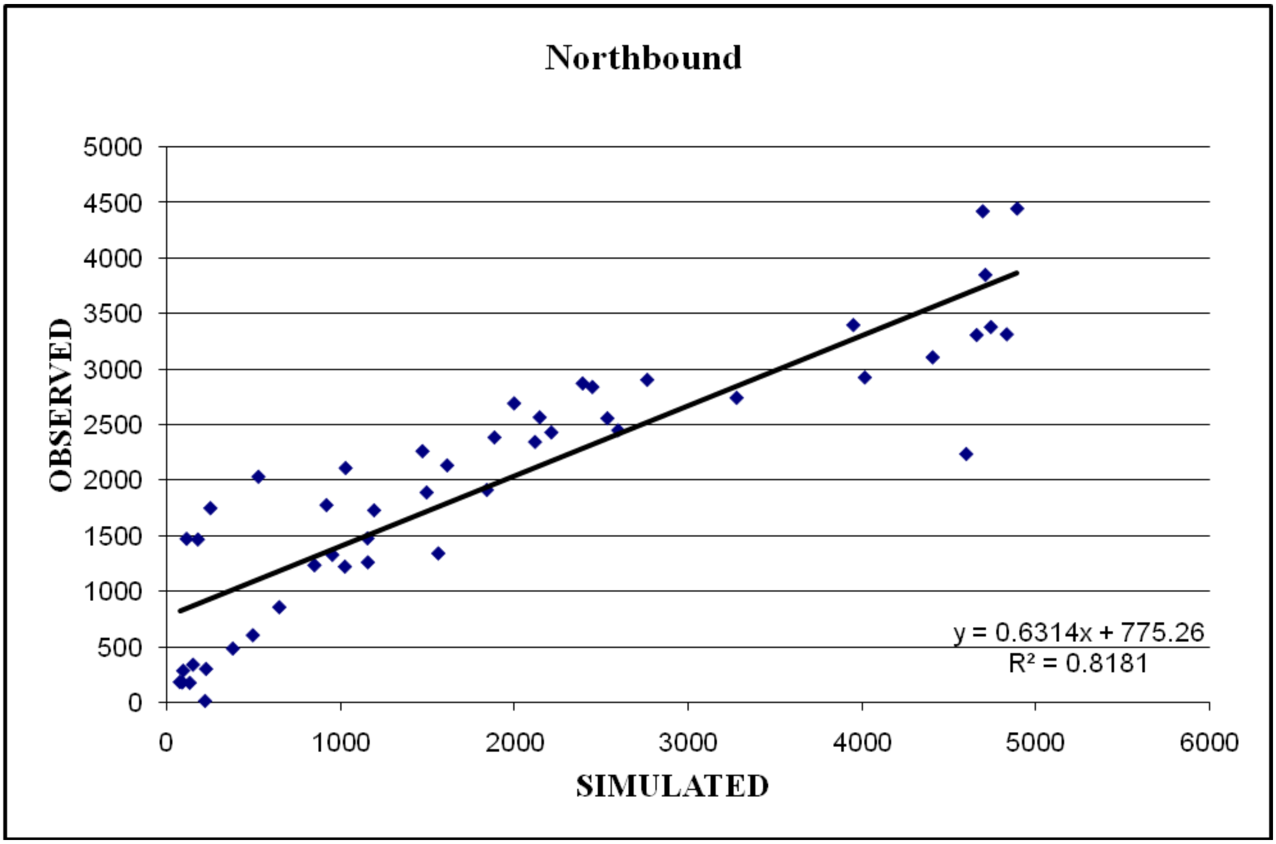

Figure 5: Regression comparison of the northbound observed traffic volume *versus* simulated traffic volumes

Validation Procedure

The goal of the calibration and validation process was to ensure that the TRANSIMS generated traffic volumes were similar to those observed on the field during the hurricane Katrina evacuation. The validation procedure used in this project followed a multi-step iterative process. The initial part of this process typically involved the execution of the TRANSIMS Router about ten times. After this routine, the next part of the process involved a series of combined Router-Microsimulator runs to reach convergence. Typically, about ten of these combined trials were required because it was necessary to route the trips assessing the outcomes of this routing arrangement to determine if the traffic distribution was reflective of a realistic condition. After this lengthy process of model execution and adjustment, a final set of model volumes was produced. These data are shown in Table 2. Similar to Table 1, the volume data were generated on an hourly basis at each station during the 48 hour evacuation period. Conveniently, this arrangement permitted a one-to-one hourly comparison of traffic volume at each station permitting validation to be undertaken on both spatial and temporal bases.

VALIDATION RESULTS

The process of calibration and validation of regional simulation models is crucial to ensure that the model is realistically representative of actual traffic. The results of the validation process are presented in the following sections based on the series of analyses conducted. The first results were based on direct comparisons of the hourly volumes at each station. Based on these results, further analyses were conducted using the same data set, after grouping the data into various sets based on selected time periods or travel directions. The results are presented based on both direct computations of percentage error as well as regression analysis.

First Approach: Statistical Validation Based on Percentage Error

Temporally Aggregated Analyses

Comparison of the observed and simulated traffic volumes at each of the station locations was carried out using volumes aggregated by the time periods. In this analysis, the time periods were aggregated based on the three segments of the evacuation process. The first included the initial 15 hours of the evacuation prior to the implementation of contra flow. This spanned the period from midnight on Saturday the 27^{th} (hour zero) to 4:00 pm on Saturday afternoon (hour 15). The second period included 24 hours of contra flow operation between 4:00 pm on Saturday afternoon (hour 16) to 4:00 pm the following Sunday afternoon (hour 39). The third period included the last 8 hours of the evacuation from 4:00 pm on Sunday afternoon until midnight after the termination of contra flow and as evacuation volume ebbed to a trickle as travel conditions deteriorated.

The results of the period aggregation analyses are shown in Table 3. In the table, it can be seen that, similar to the disaggregate analysis, the error percentages at the individual data stations are quite substantial; ranging from nearly 200 percent along westbound I-10 to less than two percent at some of the north and southbound station locations. As also evidenced in the table, a persistent discrepancy occurred in the assignment of traffic along the parallel routes of I-10 and US-61 that carried westbound traffic out of New Orleans. The issue was related to a condition in which the TRASNIMS Router tended to overutilize I-10 and underutilize US-61. This condition was particularly noteworthy during the periods without contra flow.

In an attempt to correct this problem, the link free-flow speeds on these routes were modified to achieve an assignment closer to the observed values. Unfortunately, this had only a minor effect on the numbers. Adjustments were also made to the functional classification preference setting but, once again, this also had a very limited effect on the assignment disparities between the two routes.

As would be expected, Table 3 also shows that when volume comparisons were performed on more aggregate bases, the level of error was reduced. In the table, this is most notable in the right-most column values where the error percentage decreased to less than ten percent when volumes were combined over the first two phases of the evacuation. The error percentage shows an error of just over 24 percent during the last eight hours of the evacuation after contra flow operations were terminated. It can be argued that the majority of this error is based largely on the discrepancies observed at the westbound data stations.

The side-by-side comparison of Table 3 also shows that TRANSIMS overpredicted the total number of trips out of the New Orleans area by 5,037 trip or 1.43 percent. This number is potentially significant for several reasons. First, it suggests the tremendous predictive accuracy that has resulted from the TRANSIMS *Population Synthesizer*. Results within two percent of the observed values are well below the initial expectations of the system. This discrepancy is also well below those of prior modeling systems that have been applied for evacuation simulation purposes. In recent studies (Brockfeld et al., 2004; Chiu et el., 2008), it has been suggested that error percentages between modeled and actual/observed volumes in the range of 15 to 25 percent are acceptable for this type of modeling.

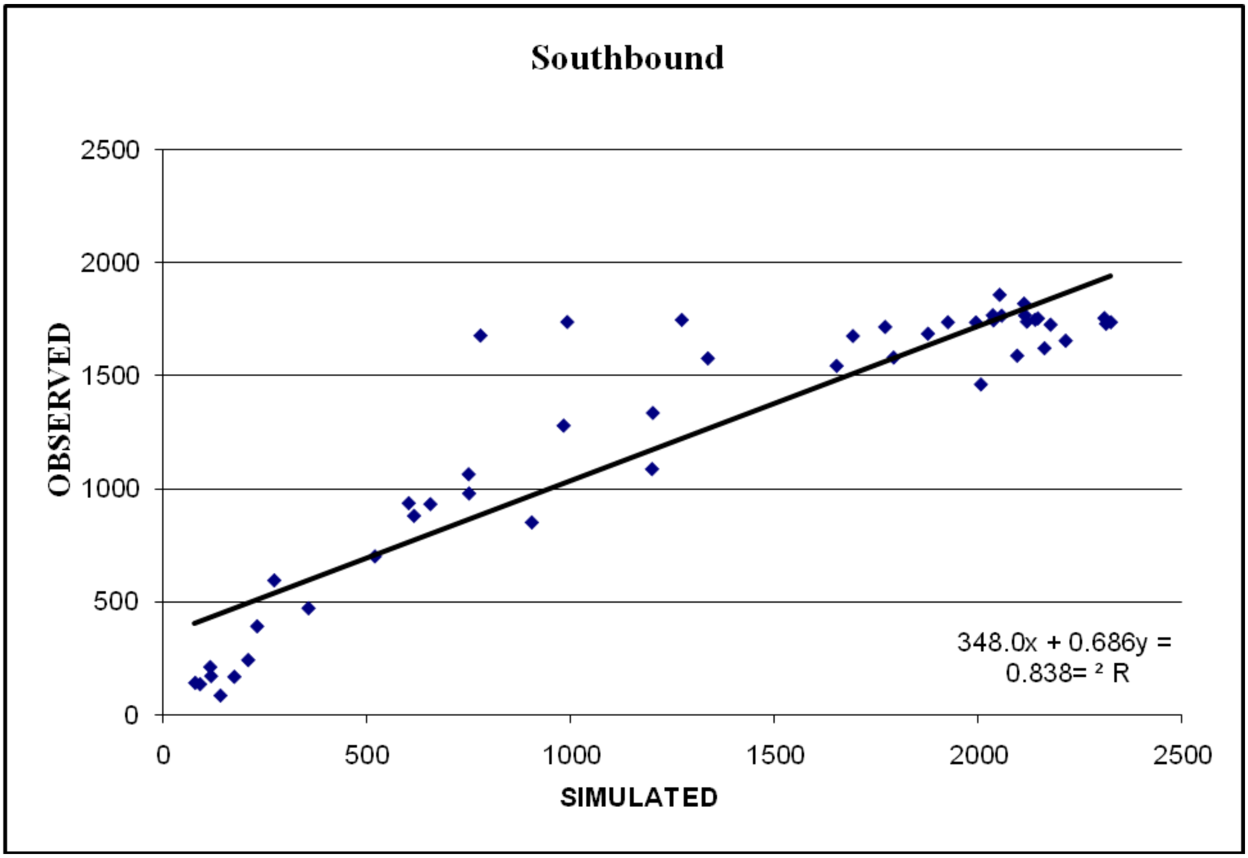

Figure 6: Regression comparison of the southbound observed traffic volume *versus* simulated traffic volumes

Further, this modest error value becomes more interesting when it is recognized that there were other factors that likely complicated any potential straight-forward evacuation. First, it is well recognized that a percentage of the New Orleans population did not evacuate for the storm. Although no one precisely knows what this number is, it has been widely suggested to have been as high as 100,000 people or 7.7 percent of the regional population. Given typical evacuation vehicle occupancy rates of 2.2 to 2.5 persons per vehicle, this could reflect approximately 40,000 vehicles. Traffic observed at the LA DOTD count stations is also recognized to include locally-generated non-evacuation traffic. Since there is no present method to distinguish these specific vehicle groups from the overall counts, there is no way to consider or determine the effect of these vehicles on the validation process.

Spatially Aggregated Analyses

After reviewing the results of the prior assessments, the data were further aggregated to evaluate the conditions more specifically associated with the directional distribution of evacuation traffic. To accomplish this task, the preceding temporal aggregations were further grouped by the general direction of travel of the evacuees. As shown in Table 4, the aggregation of the eastbound, northbound and southbound evacuees was effectively moot because there was only a single station in each of these directions. However, the westbound aggregation, which combined three stations, resulted in significantly lowered errors. This decrease in error percentage is quite logical and should be expected based on error percentage calculation, since some of the stations overpredicted the demand and others underpredicted it leading to an overall "cancellation effect" between the two extremes. Another expected phenomenon illustrated in the table was the consistent cumulative error percentages for the sums across the pre-, during and post-contra flow time periods. Since the cumulative numbers did not change between the two tables, it should follow that they would be consistent.

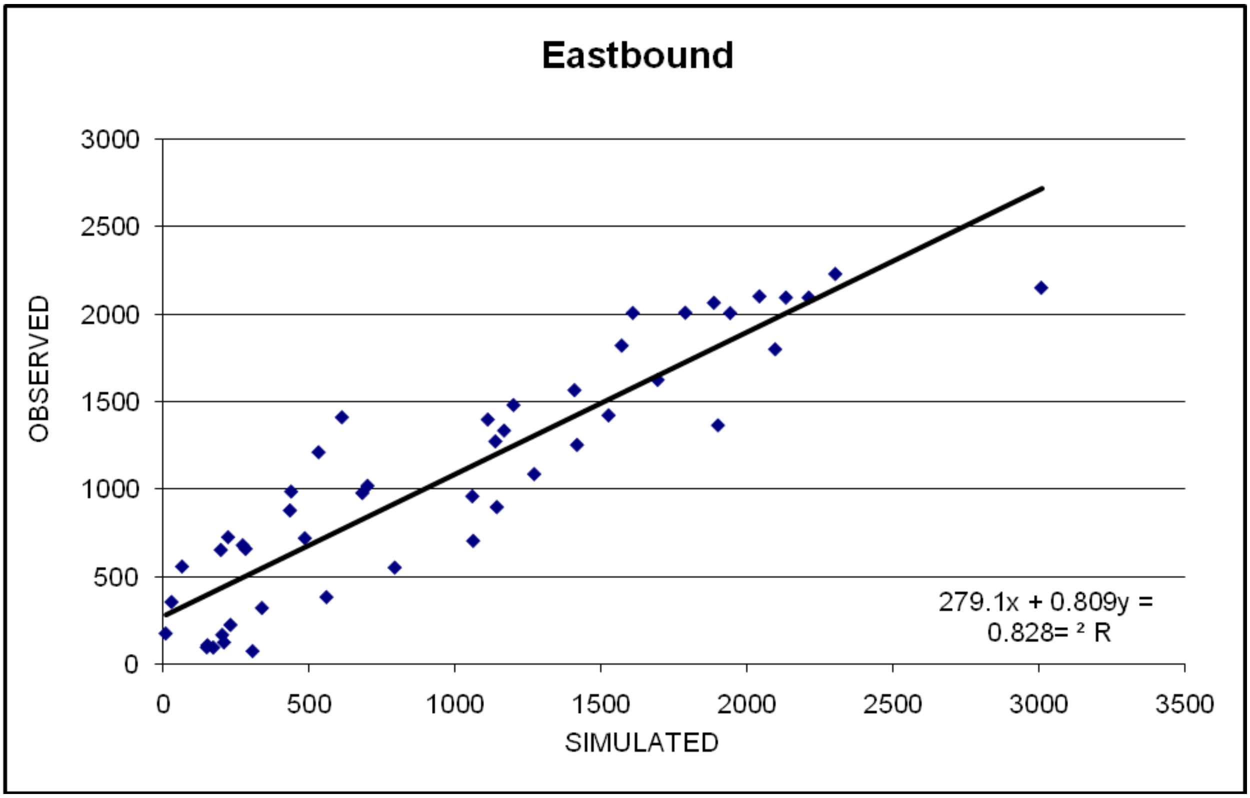

Figure 7: Regression comparison of the eastbound observed traffic volume *versus* simulated traffic volumes

The results of the Table 4 aggregations suggest that while the TRANSIMS Router module seemingly experienced difficulty in replicating the route choices of the evacuees during the process, it was able to demonstrate an improvement in terms of assigning the trips toward a particular direction given the various routes. This concept was tested as further examined in the final set of aggregation groupings.

In the final set of aggregation groupings, each of the directional volume sets was summed over the entire 48 hour evacuation period. The results of this analysis are shown in Table 5. The error percentages here show that at this level of temporal aggregation, all of the directional assignment error percentages were less than ten percent. This was quite encouraging from the stand point of the objectives of the validation. However, because of the enormous number of variables involved in the 2005 New Orleans evacuation, it is also difficult to completely attribute positive quantitative results purely to the quality of the TRANSIMS system or to accurate data sets, assumptions and/or analytical methodologies.

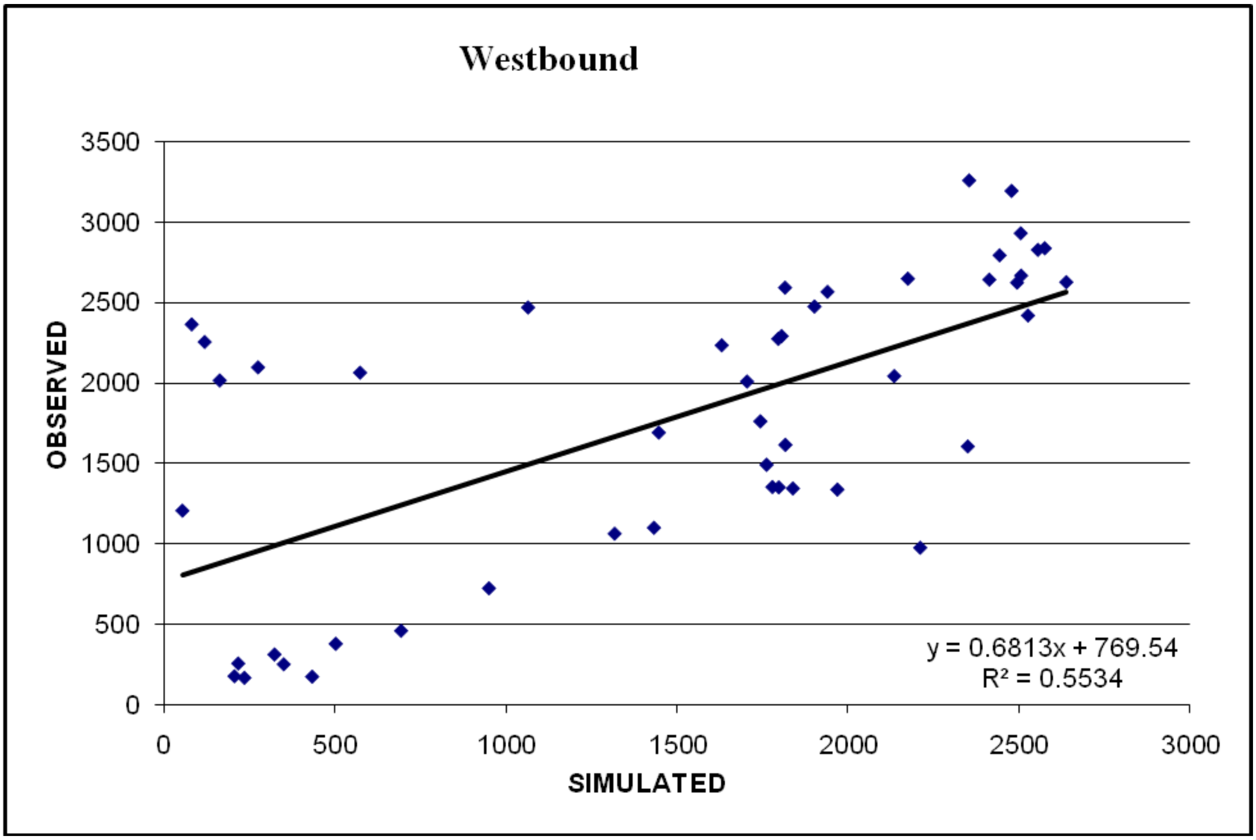

Figure 8: Regression comparison of the westbound observed traffic volume *versus* simulated traffic volumes

Second Approach: Statistical Validation Based on Regression Analysis

A different family of statistical tests for the validation of traffic simulation models is the regression analysis. According to Barcelo and Casas (2003), this method has been used to statistically compare the output from the simulation and the observed data for microscopic traffic simulation models in a situation in which only aggregated values are available (Flow counts at different stations aggregated to the hour). In this case, the observed traffic volumes are considered as the original data and the simulated traffic volumes are considered as a prediction of the observed data.

Regression analysis was conducted for the total traffic and then for each evacuation direction separately. The performance of the model is evaluated based on the R-squared (R^2) value of the regression line $y = a + bx + \varepsilon$. R^2 indicates how closely traffic volumes prediction matched the observed data. Its value lies between 0 and 1.The higher the R-squared value, the better the performance of the model. Although higher values are desirable, smaller values are considered reasonable considering the large scale of our model. A regression analysis was performed on the cumulative volumes as well as the hourly volumes between observed and simulated data.

Figures 4 through 8 show the regression comparisons between cumulative volumes observed in the field and cumulative simulated volumes over 48 hours for the total aggregate traffic volumes in the study area, northbound (station 15 located on I-55), southbound (station 88 located on US-90), eastbound (station 67 located on eastbound I-10) and westbound (station 54 on I-10 in LaPlace, station 27 on US-61 in LaPlace and station 18 on US 190), respectively.

The regression analyses at these locations were found to have acceptable fits with R-squared values greater than 0.80. This indicated that the TRANSIMS simulation model was a valid model that was able to realistically replicate the traffic patterns observed during the evacuation of hurricane Katrina, except for the westbound traffic volume,in which it is obvious that TRANSIMS underestimated traffic volumes due to the fact that the DOTD counts included the through traffic from other neighboring states.

SUMMARY AND CONCLUSION

This paper proposed a methodological pattern to validate regional level microscopic traffic simulation models. This study is one of the first to utilize actual field observed evacuation data. Model validation was based upon a comparison of the TRANSIMS generated traffic volumes to the corresponding traffic volumes actually observed during the 2005 hurricane Katrina evacuation. Analysis was performed utilizing percent errors estimation based on direct comparisons of the hourly volumes at each station. Further analyses were conducted using the same data sets by grouping the data into various sets based on selected time periods or travel directions. Also, an alternative validation approach was carried out using regression analysis between the cumulative observed and simulated volumes for six stations by analyzing the fit for the regression line $y=a + bx + \varepsilon$. The error percentage and the fit were found to be reasonable with an error percentage less than 25 percent and an R-squared value of over 0.80. This indicated that the TRANSIMS simulation model was a realistic representation of the evacuation operations observed during hurricane Katrina.

ACKNOWLEDGEMENTS

The author of this paper gratefully acknowledges the support provided for this research by the United States Department of Transportation (USDOT) under project no. DTFH61-06-R-00042 *"Application of TRANSIMS for the Multimodal Microscale Simulation of the New Orleans Emergency Evacuation Plan."*

REFERENCES

Balci, O. 1998. Verification, Validation and Testing, in: Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice, Ed. by J. Banks, John Wiley.

Barcelo, J. and Casas, J. 2003. Dynamic Network Simulation with AIMSUM, Proceedings of the International Symposium on Traffic Simulation, Yokohama, Kluwer.

Brockfeld, E., Kuhne, R.D. and Wanger, P. 2004. Calibration and Validation of Microscopic Traffic Models, Delft University of Technology, Netherlands, Transportation Research Record No. 1876.

Chiu, Y., Zheng, H., Villalobos, J.A., Peacock, W. and Henk, R. 2008. Evaluating Regional Contra-Flow and Phased Evacuation Strategies for Texas Using a Large-Scale Dynamic Traffic Simulation and Assignment Approach. *Journal of Homeland Security and Emergency Management*, 5 (1), Article 34.

Jha, M., Moore, K. and Pashaie, B. 2004. Emergency Evacuation Planning with Microscopic Traffic Simulation, TRB 83rd Annual Meeting CD-ROM.

Kim, K.O. and Rilett, L.R. 2003. Simplex-based

Calibration of Traffic Microsimulation Models with Intelligent Transportation Systems Data, *Transportation Research Record: Journal of Transportation Research Board*, No. 1855, 80-89.

Kunde, K.K. 2002. Calibration of Mesoscopic Traffic Simulation Models for Dynamic Traffic Assignment, Master Thesis, *Massachusetts Institute of Technology*.

Kwon, E. and Pitt, S. 2005. Evaluation of Emergency Evacuation Strategy for Downtown Event Traffic Using a Dynamic Network Model, TRB 84th Annual Meeting CD-ROM.

Law, A.M. and Kelton, W.D. 1991. Simulation Modeling and Analysis, McGraw-Hill.

Lee, D. H., Yang, X. and Chandrasekar, P. 2001. Parameter Calibration for PARAMICS Using Genetic Algorithm, 80th Annual Meeting Preprint CD-ROM, Transportation Research Board 01-2399.

Ma, T. and Abdulhai, B. 2002. Genetic Algorithm-based Optimization Approach and Generic Tool for Calibrating Traffic Microscopic Simulation Parameters, *Transportation Research Record: Journal of Transportation Research Board*, No.1800: 6-15.

Rouphail, Nagui M. 2003. NC State University and Jerry Sacks, National Institute of Statistical Science, North Carolina, USA, Thoughts on Traffic Models Calibration and Validation, Paper Presented at the Workshop on Traffic Modeling, Sitges, Spain, June 2003.

Theodoulou, G. and Wolshon, B. 2004. Modeling and Analyses of Freeway Contra-flow, TRB 83rd Annual Meeting CD-ROM.

TRANSIMS Open Source, http://www.transims-opensource.net/Accessed March, 2011.