MEMO

TO: Michelle O'Neill, PE

FROM: Lauren Warren, PE, PTOE; Trevor J. Kirsch, MS, EIT; Matt Hill, PE, PTOE

SUBJECT: **Calibration and Validation Memo**

DATE: **October 30, 2019**

VALIDATION CRITERIA

A validation methodology is necessary to ensure that the microsimulation model is as representative of real-world traffic conditions as possible. This is achieved through a rigorous calibration process to ensure adequate model reliability and the validity of calculated measures of effectiveness (MOEs). Best practice for microsimulation modeling is to have at least two separate validation criteria to ensure the existing condition microsimulation model is representative of the provided data. For this analysis, three metrics were utilized for validation:

1. Traffic volumes on critical segments
2. Traffic volumes within the network
3. Queue patterns

Traffic volumes on critical segments: The first measure of validity is how closely the microsimulation traffic volumes match the real-world traffic volumes temporally on critical segments within the modeled study area. A simple percentage difference between the model and real-world characteristics is not an accurate temporal comparison of the wide range of mainline segment or turning movement volumes possible in the model. Thus, a universal measure to temporally compare the microsimulation data with the real-world data is the GEH statistic. The GEH formula is displayed below:

$$GEH = \sqrt{\frac{2(m-c)^2}{m+c}}$$

where *m* (vehicles/hour) is the traffic volume on the desired segment from the microsimulation model and *c* (vehicles/hour) is the traffic volume on the desired segment from the real-world data. The intent of the GEH analysis is to ensure that microsimulation volumes are temporally reflective of real-world conditions. Based on best practices, generally acceptable criteria for GEH statistics are shown in Table 1.

Table 1. GEH Statistic Criteria

Facility Type	Criteria
Mainline Segments	GEH<3.0
On Ramp/Off Ramp Segments	GEH<3.0
Network Entry and Exit Segments	GEH<3.0
Other Local Segments	GEH<5.0 for at least 85% of applicable segments

Traffic volumes within the network: The second measure of validity is to compare the entire traffic volume within the study area of the microsimulation model with the traffic volume in the real-world over the entire analysis period. The intent of this comparison is to ensure that the satisfied traffic demand under real-world conditions is accurately reflected in the microsimulation model. Based on best practices, the traffic volume within the microsimulation model during the entire analysis period should be within +/-5% of the real-world traffic volume.

Queue patterns: The third measure of validity is how closely the microsimulation model queue patterns match the real-world queue patterns. The queue patterns of interest for the study area were established from MDOT feedback, field review, and video observation.

ASSUMPTIONS

To begin, default driver behavior parameters (Wiedemann 99) were assumed for all interstate segments and entry ramp segments within the modeled influence area. Adjustments were made as necessary during the calibration process to more appropriately match real-world conditions. Lane change distances were also adjusted to ensure that congestion was formed as expected based on the previously mentioned observations.

Likewise, the default driver behavior parameters (Wiedemann 74) were assumed for all surface street segments and interstate exit ramp segments within the modeled influence area. Adjustments were made as necessary during the calibration process to more appropriately match real-world conditions. Lane change distances were also adjusted to ensure that congestion was formed as expected based on the previously mentioned observations and feedback provided by MDOT.

INITIAL FINDINGS

After the initial simulation was completed, the desired GEH criteria and traffic volume requirement were not met. The queue patterns were also not generated as expected. Based on the results, the simulation vehicles were not aggressive enough at merge, diverge, and weave segments throughout the study area. To correct this, the default driver behaviors were adjusted for the merge, diverge, and weave segments to more accurately represent the aggressiveness that is present in real-world conditions.

CALIBRATION

To correct the under aggressiveness of the simulation vehicles on merge, diverge, and weave segments, the lane change driver behavior was adjusted to increase the willingness of simulation vehicles to complete their desired lane changes more aggressively. The lane change characteristics control the cooperative attributes of the simulation vehicles during lane change interactions, such as minimum headway and allowable deceleration rate, among others. The lane change characteristics that were changed from their default values are listed in Table 2:

Table 2. Adjusted Lane Change Characteristics

Parameter	Definition	Default	Adjusted
-1 ft/s^2 per Distance	Controls the acceptable distance needed to decelerate to facilitate a lane change	200 ft	100 ft
Minimum Headway	Minimum distance between two vehicles that must be available to complete a lane change	1.64 ft	1.00 ft
Safety Distance Reduction Factor	Reduction factor that controls the safety distance	0.60	0.35
Maximum Deceleration for Cooperative Braking	The maximum acceptable rate of deceleration to allow a vehicle to change lanes	-9.84 ft/s^2	-25.00 ft/s^2
Cooperative Lane Change Maximum Speed Difference	The maximum speed difference at which a vehicle will not change lanes to facilitate another vehicles lane change	6.71 mph	10.00 mph

Following these adjustments to the driver behavior at merge, diverge, and weave segments, the resultant MOEs passed most of the validation criteria. Based on these results, the microsimulation model was considered to be calibrated appropriately.

MODEL CONFIDENCE

Because VISSIM is a dynamic traffic microsimulation software, each simulation is controlled by a random seed number. This random seed number is correlated to various distributions within the microsimulation model. As such, each simulation run uses a different random seed number, therefore changing the interactions between simulation vehicles and generating different MOEs. Just as real-world traffic conditions are not identical every day, each simulation run is different than the previous based on this random seed number. Because of this, the confidence level in the microsimulation models must be calculated to ensure that significant differences are not present in varying simulation runs that would skew the reporting of MOEs. The confidence level is a statistical test that quantifies how reliable a specific metric is based on a range of values. In short, the confidence level defines how accurate the models are based on the measured variability in a parameter of interest.

For this analysis, the confidence level was established using various travel times throughout the study area. The travel time metric was selected to ensure that the experienced congestion in each simulation run was within a statistically reasonable threshold and no outliers interfered with the reported MOEs.

To determine the amount of simulation runs required to meet the confidence threshold, the travel time along I-94 westbound (WB), US-131 northbound (NB), and I-94 WB to US-131 NB were captured. The travel time MOEs were averaged over a period of ten simulation runs. This quantity of simulation runs was initially selected based on best practices. Each of the three travel times were analyzed per hour, meaning that each travel time has four results given the four-hour analysis period. Table 3 contains the confidence interval results at an 85% confidence level:

Table 3. Travel Time Confidence Intervals

Route	Time Period	Travel Time (s)	Standard Deviation (s)	Simulation Runs Required
I-94 WB	3:00 PM-4:00 PM	415	1	0
	4:00 PM-5:00 PM	417	1	0
	5:00 PM-6:00 PM	424	6	0
	6:00 PM-7:00 PM	418	10	1
US-131 NB	3:00 PM-4:00 PM	217	1	0
	4:00 PM-5:00 PM	216	1	0
	5:00 PM-6:00 PM	217	0	0
	6:00 PM-7:00 PM	213	1	0
I-94 WB to US-131 NB	3:00 PM-4:00 PM	434	6	0
	4:00 PM-5:00 PM	449	23	3
	5:00 PM-6:00 PM	541	38	8
	6:00 PM-7:00 PM	492	28	5

As depicted in Table 3, the maximum number of simulation runs required to maintain an 85% confidence level in the microsimulation model is eight. As previously stated, ten simulation runs were conducted to establish this confidence threshold. Because the amount of simulation runs utilized is greater than those required, the microsimulation model is considered accurate at an 85% confidence level. Although a higher confidence interval could be utilized, this would require significantly more simulation runs, which would increase the level of effort for post-processing results and have a marginal impact on resultant MOEs. Therefore, an 85% confidence level was considered acceptable for this analysis.

Note that although the microsimulation model is acceptable at an 85% confidence level, most of the variability is in the area of interest (the interchange ramp between I-94 WB to US-131 NB). This variability is especially prevalent during the 5:00 PM to 6:00 PM period, which contains most of the PM peak for this analysis (4:45 PM to 5:45 PM). Also MDOT feedback indicated the congestion in this area of interest is frequent but volatile, as the typical queue length in this area can range from localized slowing to extreme backups which persist along the mainline. This variability in congestion is captured in the microsimulation model as this location and this peak hour time period has the greatest standard deviation and requires the most simulation runs for acceptability, as shown in Table 3.

COMPARISON

After determining that the microsimulation model was accurate at an 85% confidence level, the resultant MOEs from the ten simulation runs were compared with the GEH, traffic volume, and queue pattern validation criteria. The GEH criteria were compared on a per hour interval for the four-hour analysis period, as well as a total for the entire analysis period. Table 4 contains the results from the mainline segment GEH validation:

Table 4. Mainline Segment GEH Validation

Route	To	From	GEH				
			3:00 PM-4:00 PM	4:00 PM-5:00 PM	5:00 PM-6:00 PM	6:00 PM-7:00 PM	Total
I-94 WB	Lovers Ln	Westnedge Ave Off Ramp	0.3	0.1	0.2	0.5	0.1
	Westnedge Ave Off Ramp	Westnedge Ave On Ramp	0.7	0.3	0.3	1.1	0.0
	Westnedge Ave On Ramp	Oakland Dr Off Ramp	1.1	0.8	0.1	1.0	0.7
	Oakland Dr Off Ramp	Oakland Dr On Ramp	1.0	0.5	0.1	1.6	0.1
	Oakland Dr On Ramp	US-131 NB Off Ramp	2.1	2.0	1.9	2.2	2.3
	US-131 NB Off Ramp	US-131 NB On Ramp	0.7	0.8	0.1	2.9	0.3
	US-131 NB On Ramp	US-131 SB Off Ramp	1.5	1.7	0.9	2.7	1.1
	US-131 SB Off Ramp	US-131 SB On Ramp	1.2	1.5	0.3	2.5	0.6
	US-131 SB On Ramp	9th St Off Ramp	1.1	1.4	0.2	2.4	0.5
	9th St Off Ramp	9th St On Ramp	0.8	1.0	0.4	1.9	0.1
	9th St On Ramp	6th St	0.8	1.0	0.5	2.1	0.1
US-131 NB	Milham Ave	I-94 EB Off Ramp	0.1	0.0	0.2	0.0	0.2
	I-94 EB Off Ramp	I-94 EB On Ramp	0.2	0.3	0.5	0.2	0.2
	I-94 EB On Ramp	I-94 WB Off Ramp	1.0	0.6	1.2	0.4	1.7
	I-94 WB Off Ramp	I-94 WB On Ramp	0.0	0.3	0.3	0.1	0.1
	I-94 WB On Ramp	Stadium Dr Off Ramp	1.1	0.9	2.0	2.8	0.9
	Stadium Dr Off Ramp	Stadium Dr On Ramp	0.6	0.2	1.4	3.1	0.3
	Stadium Dr On Ramp	KI Ave	0.7	0.2	1.3	2.7	0.2

Similarly, the resultant MOEs were compared to the GEH validation criteria for on ramp and off ramp segments within the influence area. The results of this comparison are in Table 5:

Table 5. On Ramp and Off Ramp Segment GEH Validation

Route	Segment	GEH				
		3:00 PM-4:00 PM	4:00 PM-5:00 PM	5:00 PM-6:00 PM	6:00 PM-7:00 PM	Total
I-94 WB	Westnedge Ave Off Ramp	1.2	0.9	0.7	0.1	1.4
	Westnedge Ave On Ramp	0.5	1.0	0.9	0.3	1.4
	Oakland Dr Off Ramp	0.8	1.3	1.0	0.5	1.4
	Oakland Dr On Ramp	0.4	0.4	0.5	0.1	0.7
	US-131 NB Off Ramp	0.4	0.4	1.4	3.9	0.5
	US-131 NB On Ramp	0.6	0.4	0.6	0.9	0.5
	US-131 SB Off Ramp	0.4	0.0	0.7	1.9	0.2
	US-131 SB On Ramp	0.1	0.1	0.1	0.0	0.2
	9th St Off Ramp	0.6	0.8	0.2	2.5	0.5
	9th St On Ramp	0.1	0.1	0.1	0.0	0.2
US-131 NB	I-94 EB Off Ramp	0.1	0.4	0.0	0.2	0.4
	I-94 EB On Ramp	0.0	0.1	0.0	0.1	0.0
	I-94 WB Off Ramp	See "US-131 NB On Ramp" above				
	I-94 WB On Ramp	See "US-131 NB Off Ramp" above				
	Stadium Dr Off Ramp	1.6	1.5	2.1	1.4	2.1
	Stadium Dr On Ramp	0.1	0.1	0.1	0.1	0.2

Likewise, the resultant MOEs were compared to the GEH validation criteria for the network entry and exit segments within the influence area. Table 6 contains the network entry segment comparison, while Table 7 contains the network exit segment comparison.

Table 6. Network Entry Segment GEH Validation

Segment	GEH				
	3:00 PM-4:00 PM	4:00 PM-5:00 PM	5:00 PM-6:00 PM	6:00 PM-7:00 PM	Total
I-94 WB	0.3	0.1	0.2	0.5	0.1
Westnedge Ave NB	0.0	0.0	0.0	0.0	0.0
Westnedge Ave SB	0.0	0.0	0.0	0.0	0.0
I-94 EB to Westnedge Ave	0.2	0.3	0.3	0.2	0.5
Oakland Dr NB	1.1	1.1	1.2	1.0	2.2
Oakland Dr SB	0.7	0.8	0.8	0.7	1.5
I-94 EB to Oakland Dr	0.2	0.2	0.2	0.2	0.4
US-131 SB to I-94 WB	0.1	0.1	0.1	0.0	0.2
9^{th} St to I-94 WB	0.1	0.1	0.1	0.0	0.2
US-131 NB	0.1	0.0	0.2	0.0	0.2
I-94 EB to US-131 NB	0.0	0.1	0.0	0.1	0.0
Stadium Dr to US-131 NB	0.1	0.1	0.1	0.1	0.2

Table 7. Network Exit Segment GEH Validation

Segment	GEH				
	3:00 PM-4:00 PM	4:00 PM-5:00 PM	5:00 PM-6:00 PM	6:00 PM-7:00 PM	Total
Westnedge Ave NB	0.2	0.2	0.2	0.6	0.4
Westnedge Ave SB	0.6	0.1	0.1	0.8	0.7
Westnedge Ave to I-94 EB	0.2	0.0	0.0	0.8	0.5
Oakland Dr NB	0.1	0.4	0.3	0.7	0.2
Oakland Dr SB	0.0	0.5	0.1	0.7	0.0
Oakland Dr to I-94 EB	0.6	0.4	0.1	0.4	0.2
I-94 WB to US-131 SB	0.4	0.0	0.7	1.9	0.2
9th St Off Ramp	0.6	0.8	0.2	2.5	0.5
I-94 WB	0.8	1.0	0.5	2.1	0.1
I-94 EB Off Ramp	0.1	0.4	0.0	0.2	0.4
Stadium Dr Off Ramp	1.1	0.9	1.5	1.9	0.9
US-131 NB	0.8	0.1	1.5	2.7	0.0

Furthermore, the resultant MOEs from the other local segments were compared to the applicable GEH validation criteria. The results of this comparison are displayed in Table 8:

Table 8. Other Local Segment GEH Validation

Segment	GEH				
	3:00 PM-4:00 PM	4:00 PM-5:00 PM	5:00 PM-6:00 PM	6:00 PM-7:00 PM	Total
Westnedge Ave NB to I-94 WB	1.2	1.1	1.0	1.0	2.2
Westnedge Ave NB to I-94 EB	0.5	0.5	0.9	0.9	1.4
Westnedge Ave SB to I-94 EB	0.6	0.2	0.5	1.4	0.7
Westnedge Ave SB to I-94 WB	0.4	1.1	0.8	1.3	1.7
Oakland Dr NB to I-94 WB	1.2	1.1	1.5	1.0	2.4
Oakland Dr NB	0.5	0.6	0.1	0.0	0.6
Oakland Dr NB to I-94 EB	2.1	2.2	1.9	1.3	3.8
Oakland Dr SB to I-94 EB	1.0	0.6	0.3	0.0	1.0
Oakland Dr SB	0.4	0.3	0.3	0.9	0.9
Oakland Dr SB to I-94 WB	0.0	0.1	0.1	0.1	0.0

Additionally, the entire traffic volume from the microsimulation model was compared to the real-world traffic volume within the study area over the entire analysis period to determine if the model satisfied the traffic volume validation criteria. The results of this comparison are in Table 9:

Table 9. Traffic Volume Validation

Real World Traffic Volume			**Microsimulation Traffic Volume**	
Total	Lower Bound (5%)	Upper Bound (5%)	Total	Percent Difference
273,387	259,718	287,056	271,451	1%

Lastly, the queue patterns of the ten simulation runs were analyzed to determine if the microsimulation model was accurately representing the congestion as determined by MDOT feedback, field review, and video observation. The results of this qualitative analysis are discussed in the next section.

SUMMARY

After the rigorous calibration of the microsimulation model and establishing confidence in the results, the calculated MOEs were compared to the relevant validation criteria contained in Table 1 and listed in the first section of this memo. Table 4 contains the validation results of the mainline segments within the influence area. Based on these results, most of the mainline segments pass the GEH statistic threshold, except for the US-131 NB mainline between the Stadium Dr off ramp and the Stadium Dr on ramp. This was the only mainline segment to not pass the validation criteria, with a GEH of 3.1 during the final hour (6:00 PM to 7:00 PM) of the analysis period. This is likely due to the volatility of the congestion experienced at the upstream I-94 WB to US-131 NB interchange. Because this location is immediately upstream of this mainline segment, the desired traffic demand is highly sensitive to the time at which this congestion dissipates and vehicles are able to successfully merge onto US-131 NB. Despite this, 94% of the mainline segments meet the validation criteria.

Similarly, Table 5 contains the validation results of the on ramp segments and off ramp segments within the influence area. Based on these results, one of the ramps does not meet the required GEH statistic. The interchange ramp between I-94 WB and US-131 NB has a GEH statistic equal to 3.9 during the 6:00 PM to 7:00 PM period. Although this does not meet the GEH threshold, this area experiences frequent congestion that is volatile in nature, as determined by MDOT feedback. This is also the period with the most volatility, as noted during the confidence interval calculation. Based on the computation of the GEH, it seems that the microsimulation model is temporally shifting the congestion later in the analysis period in comparison to real-world operations, meaning that the congestion in the model is occurring later than in the real-world. Because of this variability, it is difficult to maintain a consistent GEH statistic which passes the validation criteria at this location because the traffic counts vary slightly between the model and the real-world due to this temporal shift. Despite this, the GEH criteria is met for all remaining on ramp and off ramp segments during all other time periods. Based on the validation results, 94% of all ramp segments pass the validation criteria.

Table 6 and Table 7 depict the validation results for the network entry segments and exit segments within the study area, respectively. Based on these results, all the network entry and exit segments pass the validation criteria with GEH statistics less than 3.0 under all time periods considered.

Furthermore, Table 8 contains the validation results for the other local segments within the influence area. Based on these results, all the applicable local segments pass the validation criteria with GEH statistics less than 5.0 under all time periods considered.

As mentioned previously, the total traffic volume in the microsimulation model must be within 5% of the real-world traffic volume within the influence area over the entire analysis period. Table 9 outlines the results of this comparison. Ultimately, the microsimulation model passes this validation criteria. The traffic volume in the microsimulation model is within 1% of the real-world traffic volume, which indicates that the model should be accurately representing the existing conditions.

Lastly, the queue patterns of the model were analyzed to determine if the congestion in the microsimulation model was representative of the congestion documented through MDOT feedback, field review, and video observation. As mentioned previously, most of the congestion within the study area is generated from the interchange ramp between I-94 WB and US-

131 NB. This area of the microsimulation was observed during the entire analysis period, and the resultant congestion shown in the ten simulation runs was determined to be representative of the documented congestion. The extent of the typical queue in the microsimulation models mirrored the queue length observed in the video observations, while the volatility of the queue was recognized in the various simulation runs due to the random seed number.

In conclusion, the results of each validation that was performed on the microsimulation model are summarized below:

- Mainline Segments – All the mainline segments meet the appropriate validation criteria over all the time periods considered.
- On Ramp/Off Ramp Segments – Most of the on ramp/off ramp segments meet the validation criteria. The only ramp segment to not meet the validation criteria was the system interchange ramp between I-94 WB and US-131 NB during the last hour of the analysis period. As previously discussed, this is likely due to the congestion volatility that is present under existing conditions.
- Network Entry and Exit Segments – All the network entry and exit segments meet the appropriate validation criteria over all the time periods considered.
- Other Local Segments – All the local segments within the microsimulation model meet the appropriate validation criteria over all the time periods considered.
- Traffic Volume - The microsimulation traffic volume is within the acceptable tolerance range of the real-world traffic volume for the entire analysis period
- Queue Patterns – The queue patterns present in the existing condition models are representative of current, real-world congestion, based on MDOT feedback