STATE OF OREGON

WHITE PAPER

Department of Transportation
Transportation Planning Analysis Unit
Mill Creek Office Park
555 13th Street NE Suite 2
Salem, Oregon 97301-4178
(503) 986-4112 FAX (503) 986-4174

File Code:

Date: July 1, 2008

AUTHORS: Alex Bettinardi, P.E., Transportation Analyst
Doug Norval, P.E., Transportation Analyst Engineer
Peter Schuytema, P.E., Senior Transportation Analyst
Dorothy Upton, P.E., Senior Transportation Analyst

SUBJECT: Simulation Guidelines Project (February 2006 – May 2008)

Executive Summary

The purpose of this project was to advance ODOT's simulation procedures and guidelines for planning and project analysis. The findings from this project are to be used to update the Analysis Procedures Manual, Chapter 8. The project used Trafficware's SimTraffic/SYNCHRO software, however it was developed to be as independent of specific software as possible.

The project set out to determine the different calibration needs of study areas by area type; small urban, small-medium Metropolitan Planning Organization (MPO), Large MPO (METRO), and recreational areas. Two representative sites for each area type were selected by the team creating eight locations to analyze. For each location, field data was collected, the data was post processed and evaluated, and a series of calibration tests were run for determining the best calibration procedure by area type. The data collected and the calibration tests performed were designed after conducting a thorough literature review of the latest research and calibration methods. From that literature review, speed, headway, and driver reaction time were found to be the calibration parameters that best matched the project objectives.

After all of the data was collected and the calibration testing was complete the analysis showed that a consistent calibration procedure could not be found that was applicable to all locations or any subgroup of locations. The conclusion found by this work was similar to the findings from the literature review; calibration can be improved by collecting additional field data and incorporating it into one's model, however there is no ideal combination of calibration parameters and data across all models / projects, so engineering judgment needs to be applied during the calibration process to determine what/when/why/where/how adjustments to the model and/or additional data is needed to achieve an acceptable level of calibration.

This study did develop a series of guidelines, checklists, and thresholds to help aid analysts in the calibration process. Also, calibration requirements using SimTraffic's "vehicles exited" measure of effectiveness (MOE) was established from this study.

Introduction

The Transportation Analysis Planning Unit (TPAU) of the Oregon Department of Transportation (ODOT) is constantly working to improve the traffic analysis procedures that it recommends for use on Oregon highways. The Analysis Procedures Manual (APM) is a comprehensive guidebook of all the procedures that TPAU has developed. One of the sections of the APM that needed attention and refinement was the methodology suggested for micro-simulation analysis of projects (Chapter 8).

Prior to this work, the APM explained the inputs that went into a micro-simulation, but did not provide guidance on how to adjust these parameters and how to measure the performance of an analyst's micro-simulation, beyond citing the FWHA toolbox for additional information on calibration. To improve guidance on the proper way to apply micro-simulation analysis for the projects in the State of Oregon, TPAU formed this study to develop a set of procedures and criteria that would produce higher level of accuracy and precision from the micro-simulation analysis being conducted in Oregon. The study began February 2006 and ended May 2008.

To begin and frame this study a literature review was conducted, which totaled 23 manuals, reports, and articles. The list for the literature review came from searches from TPAU, as well as the State of Oregon Library which conducted searches to help form a comprehensive list of all existing research. The literature review contained calibration and validation processes for varying types of microsimulation applications, across a wide variety of software and methodologies. The review did not produce a clear set of steps, parameters, or measures of effectiveness to base this study on. However, the review did reinforce SimTraffic's calibration help documents which stated that the headway, turning speeds, and driver reaction time may need to be adjusted to achieve calibration.

The goal of this study was to create a set of simulation guidelines that would be independent of the software being used. However, the literature reviewed indicated that different process would have to be employed depending on the software being used. As part of this study, a list of Microsimulation software to be evaluated was developed; SimTraffic, VISSIM, PARAMICS, and CORSIM. However, after initial tests with calibrating these four software, it was determined that ODOT-TPAU did not have the time or budget to complete a study that would provide guidance on how to calibrate a Microsimulation across all software platforms. SimTraffic was used for this study for three reasons:

1. The ODOT-TPAU staff had been using SimTraffic for many years and was very familiar with the software. For this reason SimTraffic did not have the learning curve that some of the software had.
2. VISSIM and PARAMICS both offered dynamic assignment. After the literature review and testing the software it was determined that this feature was above and beyond the goals of this study. The literature reviewed suggested that it was harder to calibrate a model using dynamic assignment. In addition, most of the Microsimulation analysis performed by ODOT-TPAU, is at the corridor or small network level. Dynamic assignment is not as beneficial for smaller networks or corridors. The real benefit with dynamic assignment, and the primary use found in the literature, is for large congested networks; multiple parallel corridors, freeways with multiple access points being modeled

and a surrounding grid, or large downtown grids. Currently, ODOT does not deal with this scale of Microsimulation very often, although dynamic assignment will likely be more important for Oregon in the future and should be reevaluated then.

3. Both the literature review and the experience of ODOT-TPAU agreed that SimTraffic is closer to calibration "out-of-the-box" then the other software available at the time of this study and would therefore greatly simplify this work.

For these reasons only SimTraffic was used and, consequently, the guidelines are primarily for calibrating a network coded in SimTraffic*[*]. However, it is the hope of ODOT-TPAU that the guidelines for SimTraffic will be transferable to other software, or at least a good starting point in achieving calibration under other software.

After refining the study based on the literature review and what was available using SimTraffic, the following calibration parameters (independent variables, x) and measures of effectiveness (dependent variables, y) were used to develop the simulation guidelines:

Calibration Parameters

- X_1) Headway Factor
- X_2) 85^{th} Percentile Speed
- X_3) Driver Reaction Time

Measures of Effectiveness

- Y_1) Maximum Queue Length
- Y_2) Average Queue Length
- Y_3) 95^{th} Percentile Queue Length
- Y_4) Travel Time (Average Speed)
- Y_5) Vehicles Exited
- Y_6) Total Stops
- Y_7) Average Cycle Length

Site Selection

This study was scoped to test areas that represented projects from all areas of Oregon. The team wanted to ensure that results from this study could be used for all areas and projects across Oregon, not just the Willamette Valley (the majority of the population). The team picked eight locations that provided even coverage of the following project/area characteristics (note, more than eight locations was desired, however budget and time limitations only allowed for eight):

- Population – small urban areas, small MPO, medium MPO, large MPO (METRO)
- Project type – Expressway, Major arterial, pre-timed downtown grid
- Access Issues – Little to no access restriction to full access restriction
- Trip/Area Type – Commuter route, recreational route, urban area

[*] This study began using SimTraffic 6 (Build 612). During the course of this work SimTraffic 7 was released by Trafficware. The work was checked and completed using build 761.

All eight locations represented typical locations where a project would be needed or studied. Although, to keep control on the project, all study areas were kept fairly small, the largest area encompassing seven intersections. Larger networks would require data collection resources beyond the capability of this study and could have introduced extra noise to the calibration work making more difficult to draw conclusions. Larger networks may be addressed in future studies that will build off of the work conducted for this study.

Summary points for the eight locations chosen to meet the above criteria are provided here:

Albany, Oregon – US20, collections made July 25th & September 19th, 2006, and July 24th, 2007

- Small urban area
- Five-lane arterial
- Access-controlled
- Urban area
- Two signalized intersections
    - Spring Hill Drive
    - North Albany Road

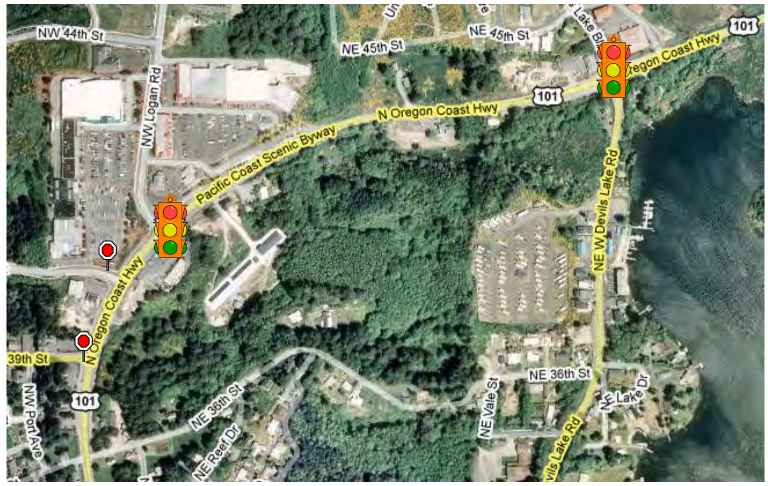
Notes:
Wide turns required turning speeds to be collected and input.

Bend, Oregon – US97, collections made July 19th, 2007 and November 11th, 2007

- Small MPO
- Five-lane-arterial – planned freeway
- Partial Access control
- Recreational area / Commuter route
- Two signalized intersections
    - Cooley Road
    - Robal Road
- Five unsignalized intersections/accesses
    - Clausen Drive
    - Lowe's Driveway
    - Chavre Way
    - Target Driveway
    - Nels Anderson Place

Notes:
This was a subsection of a much larger network for an active project for US97 in the northern end of Bend. There were many challenges (simulation run times, sparse field data for calibration, complexity incomparable to other locations) with using the full model area for the calibration testing.

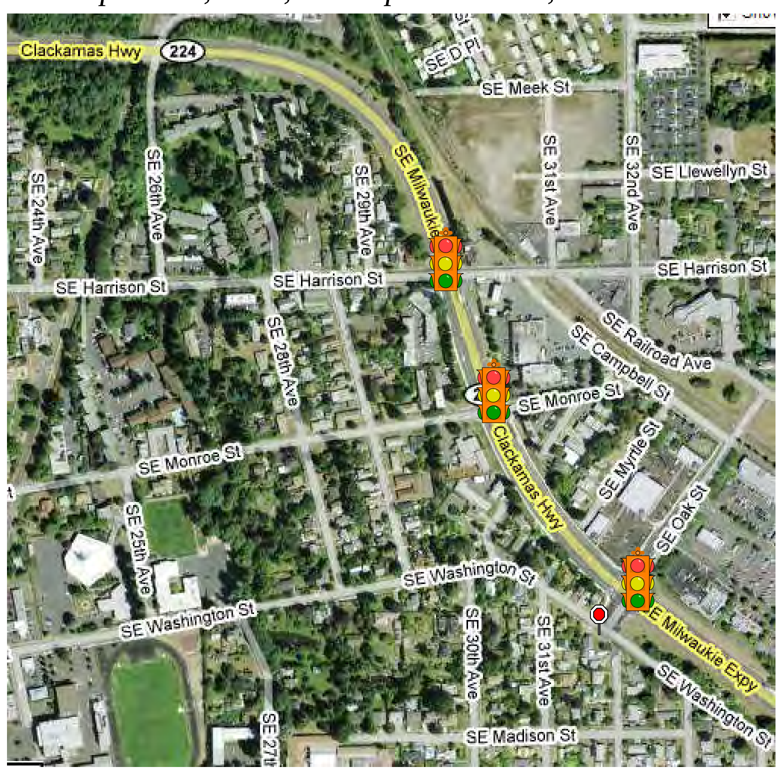
Lincoln City, Oregon – US101, collections made August 17th, September 11th & 28th, 2007



- Small urban area
- Two/Five-lane arterial
- Partial Access control
- Recreational area
- Two signalized intersections
  - West Devils Lake Road
  - Logan Road
- Two unsignalized intersections
  - 40th Street
  - 39th Street

Notes:
Wide turns required turning speeds to be collected and input. Also, short turn bays required special treatment with positioning distances.

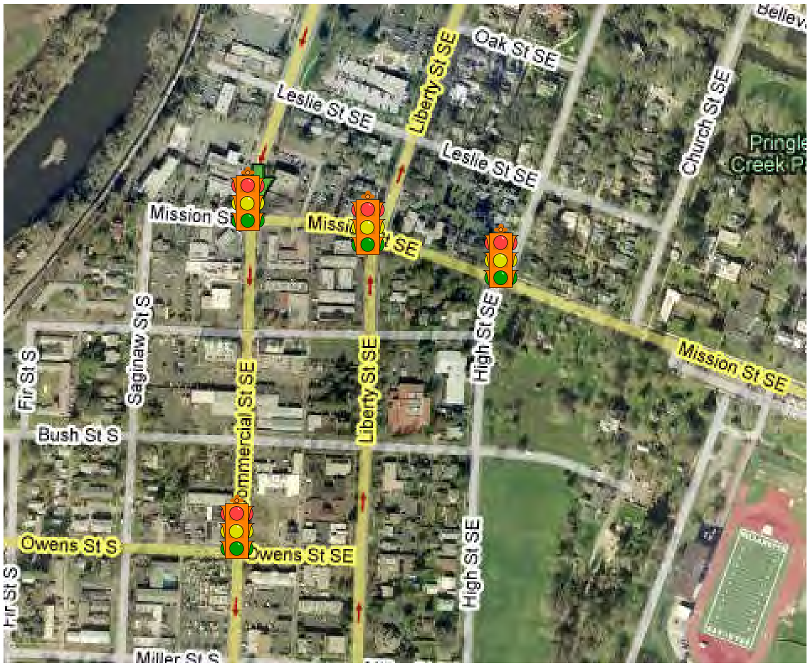
Milwaukie, Oregon – OR224, Clackamas Hwy, collections made October 12th, 2006, April 24th, 2007, and September 25th, 2007



- Large MPO (METRO)
- Five-lane expressway
- Access-controlled
- Urban commuter route
- Three signalized intersections
  - Harrison Street
  - Monroe Street
  - Oak Street
- One unsignalized intersection
  - Washington Street

Notes:
Due to the heavy turn moves during the peak hour, turn bay lengths were essential. With SimTraffic 7, taper lengths made a significant difference. Also, the <100 ft link length on Oak Street between OR224 and Washington Street caused problems with the turn moves on to Oak Street. Make note that signing near signals may cause improper behavior/operation at the signal. In this case, a yield sign on Oak Street in the simulation caused left turning vehicles to stop and check for clearance on their green, behavior not seen in the field. The yield sign was originally placed in the network when the network was in SYNCHRO 6 to remove improper long queues on Washington Street. However, in 7 the changes in behavior logic, removed the need for the yield sign and created a situation where its presence caused improper behavior at the intersection.

Salem, Oregon – Mission Street, collections made October 24th, 2006 and June 26th, 2007



- Medium MPO
- One-way grid
- Unrestricted access
- Urban commuter route
- Four signalized intersections
    - Commercial Street
    - Liberty Street
    - High Street
    - Owens Street

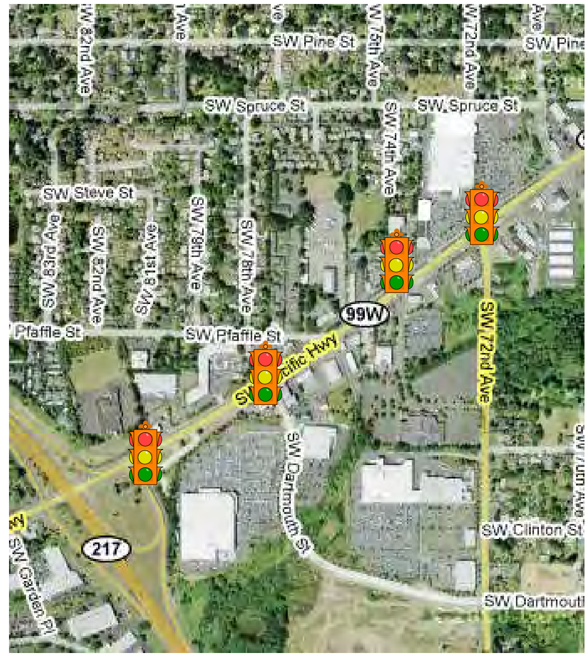
Notes:
Network performed better than witnessed in the field. Special attention needed to be given to the OD paths, specifically the heavy move from Mission Street to Commercial Street to Owens Street. This helped better model the congestion levels witnessed.

Salem, Oregon – OR22, collections made October 17th, 2006, May 15th and September 18th, 2007

- Medium MPO
- Expressway
- Access-controlled
- Urban commuter route
- Two signalized intersections
    - Airport Road
    - Hawthorne Avenue

Notes:
Driver expectance in this area did not follow the posted speed limit. The free flow speed was found to be less than the posted speed limit.

Tigard, Oregon – OR99W, collections made November 8th, 2006, May 8th and September 25th, 2007



- Large MPO (METRO)
- Five-lane arterial corridor
- Partial access control
- Urban Commuter Route
- Four signalized intersections
    - Northbound OR217 ramp terminal
    - Dartmouth Street
    - Theater access
    - 72nd Avenue

Notes:
~5% grade along OR99W at 72nd Street. Saturation flow measurements were taken for both the up and downgrade. After the ideal saturation flow was back-calculated both the up and downgrade rates were found to be very close (1683 and 1738 pcphpl, respectively). This finding helped to validate both the Highway Capacity Manual (HCM) saturation flow collection methodology and the field collection practices used for this study (the data was collected during the same time period by two different analysts).

Woodburn, Oregon – OR99E, collections made June 27th, 2006 and September 26th, 2006

- Small urban area
- Five-lane arterial
- Unrestricted access
- Urban area
- Two signalized intersections
    - Hardcastle Street
    - Lincoln Street

Notes:
This was the first collection site. The methodology was modified for the future sections based on this initial collection. Even with the change in methodology, there was not enough variation to warrant recollecting the data.

Data Collection

For this study, data that would be required for a project simulation analysis needed to be collected for each of the eight locations. Above and beyond that, new calibration data needed to be collected. Since the study would determine which calibration data would be important, many additional or to-be-determined unnecessary data also needed to be collected in order to get to the set of calibration data that ODOT would recommend or require be collected for simulation analysis. This created an extensive list of data to collect for each site, which is part of the reason why the test locations had to be limited to only a handful of intersections. The following is the list of data that was required at all eight locations:

- Roadway geometrics (Turn bay and taper lengths were critical)
- Classified vehicle counts
- Signal timing, phasing details, coordination
- Saturation flow measurements (following HCM methodology)
- Driver reaction time
- Queued vehicles (stopped counts)
- Average speed (using floating car measurements)
- Free-flow speed (85^{th} percentile measurements)
- Turning speeds (using either a probe vehicle in traffic or using LIDAR measurements)
- Lane utilization

The data currently required for projects (geometries, vehicle counts, Free-flow speed, turning speeds, signal timing) have tested processes for collection and was fairly straight forward. The additional detail required for the calibration work (saturation flow, driver reaction time, queued counts, average speed) was not as familiar to TPAU and posed the threat of being beyond the collection resources available for this study. Many of these measurements would typically be carried out by a team of two, one person to view traffic and one person to record the measurements. Another option would be to set up video cameras with time stamps and go back after the fact and record the data. To compound the problem, this study involved visiting and collecting data at each location a minimum of two times.

> *Part of the study was scoped to determine the importance of the calibration data being collected on the count day(s), or if there was a window of time around count days where the calibration data would be acceptable. It is anticipated that project schedules and resources will likely not allow for all of the vehicle counts and the calibration counts to occur on the same day.*

The large data collection requirements for this study created concern on the feasibility of this project with the resources available. Fortunately, a process using JAMAR counters was agreed on and developed during the planning stages of this project. The more complex JAMAR counters offer time stamp functionality, where the time and button number is marked each time a button pressed. For these time stamp units, JAMAR offers a methodology to collect saturation flow rates using the JAMAR units and their software, PETRA. TPAU did not have a current license of PETRA, in addition, TPAU saw the ability to collect additional data with JAMAR units during saturation flow collection. During the planning stages of the project TPAU wrote customized software in R that read and interpreted the text files reported from the JAMAR units.

Using the JAMAR counters allowed a single analyst to collect the following data for an approach all at the same time:

- Saturation flow rates
- Queued vehicle counts by cycle and period
- Driver reaction time
- Lost time for the approach
- Heavy truck counts (percentages for ideal saturation flow rate)
- Turning vehicle counts (percentages for ideal saturation flow rate)
- Vehicle counts
- Lane utilization
- Phasing detail, signal operation by cycle
- Arrival type (rating) for the approach

The JAMAR units made it possible to collect all the necessary data with the resources and budget available. TPAU has written up the process and instructions for how to use the JAMAR units as they were used for this study. On request, TPAU will provide instructions and the software.

Calibration Testing Methodology

Step 1

After all of the data was collected, the first step was to build a SimTraffic model for each location following all the recommendations currently in APM. These SimTraffic models served as the reference or base case for the calibration testing. To-date the "visual calibration" from the APM served as an acceptable level of calibration, the goal was to improve the calibration above and beyond this minimum level of prep-work. The "visual calibration" included measuring, adjusting, and fine tuning the following inputs:

- Reviewing and fixing all Error and Warning Reports
- Setting up the Seeding and Recording
- Setting the Random Seed Number to zero
- Vehicle composition (lengths and percentages)
- Turning speeds for irregular turns
- Adding full signal detail, including detection and detector spacing
- Proper geometry including turn bay and taper lengths
- Observed driver behavior
    - Lane (turning) alignment
    - Blocking intersections
    - Improperly using the shoulder or median
    - Positioning lengths
    - OD paths for major moves

After all of these were coded and a visual calibration or "laugh test" was performed to make sure that the microsimulation resembled the traffic conditions witnessed in the field. Each location required varying amounts visual calibration time and data collection, which was dependent on size and issues that were unique for every location.

Part of this study included developing a "Simulation Field Inventory Worksheet" for microsimulation work. This worksheet can be requested from TPAU. TPAU recommends that field collection / observation be made as close to the count date, or if there are multiple count dates, as close to the 30th highest hour as possible. The analyst being on site at the time of the count (or a representative day) will help ensure that the volume coded into the model from the counts (or adjusted counts as the case will probably be) can physically make it through the network. Most areas under analysis are near congestion. When counts are adjusted up, it is possible to create an input volume that is greater than the capacity at the intersection. For this reason, it is important to witness the driver behavior/movement during the count and noting where problems do and do not occur and to verify that volumes to be input into the model are within the capacity of all given intersections. If the analysts visits the site off of the count day, it is advised that a short (peak hour) count be performed at an important (or group of important) intersection or major approach. This will help ensure that any adjusted counts are in line with what's actually occurring on the day the site is visited.

Step 2

The next step was to review, clean, tabulate, and analyze all of the calibration data collected. For each location the measures of effectiveness (independent variables listed in the introduction) were tabulated by 15 minute collection periods for each location, and put into a text format identical to the text report created by SimTraffic. This process allowed for easy comparison between field conditions observed and model conditions. To help automate this comparison, custom software written in R was used to quickly compare, summarize, report, and plot the comparison of the thousands of measurements.

The use and application of the calibration parameters (dependent variables) was not as straight forward; each of the three variables had a different collection, tabulation, and application process applied. The three are described individually here:

Headway Factor (X_1) – TPAU had four questions about headway (saturation flow) to be answered by this study:

1. If the collection day had to be on the count day.
2. If the collection had to be within the peak hour or if it could shoulder on the surrounding hours, if 15 cycles (as required by the HCM) could not be collected within the peak.
3. In some small urban areas it may not be possible to get 15 cycles with eight or more queued vehicles (as required by the HCM) within the peak hour for a complete sat flow count. TPAU wanted to know how using cycles with less than eight queued vehicles affected the calibration. Headway factors were calculated using five or more queued vehicles, in addition to the eight or more required by HCM.

4. The HCM states to take an average of the 15 or more cycles to determine the field saturation rate. TPAU also calculated this measure using the median instead of the average, which removed the lower outliers, raising the saturation flow. TPAU wanted to know how using the headway factor calculated from the median sat flow would affect calibration.

These four "on-off" type questions created a factor of 16 extra tests to the existing four required tests for a full factorial design (discussed in *Step 3* of calibration testing), equaling a total of 64 possible calibration tests to be performed for each location. This type of testing was not possible for study.

In an attempt to still get at some of these questions and keep the study within reasonable bounds, it was decided to test the minimum headway factor and maximum headway factor found from these 16 cases, in addition to testing with and without the "standard or ideal" headway factor, which was collected on the count day, during the peak hour, using the average of 15 or more cycles which had eight or more queued vehicles. This created a total of 16 (8 x 2) calibration cases to test for each of the eight locations, which was a reasonable number to test.

85th Percentile Speed (X_2) – Unlike headway factor which was desired to be collected during the peak hour on the count day, the 85th percentile speed needed to be free flow, without congestion. That meant that it would not be collected during the peak and would ideally be the average over multiple days to ensure a true desired free flow speed. This measure did not have the same variability and questions regarding collection that headway factor posed. Therefore, the 85th Percentile Speed was simple on or off test during the calibration testing.

Driver Reaction Time (X_3) – Initially TPAU planned to collect and test driver reaction time for each location (meaning that driver reaction time would be area dependent). However, after the data was collected and tabulated, it was clear that there was far too much variability in driver reaction time from analyst and intersection even in the same area and collection period to recommend this as practice. However, what TPAU did find, was that over all locations, population sizes, and area types, the driver reaction time for Oregon drivers was fairly constant, and, more importantly, the collected times were greatly different than SimTraffic's defaults. In addition, SimTraffic recommended that their defaults be changed to match those collected. For these reasons, TPAU tested the benefit of using a statewide average not a location based driver reaction time.

SimTraffic Driver Reaction	State Wide Average Driver Reaction	State Wide Standard Dev. Driver Reaction
0.8	2.4	0.7
0.7	1.6	0.3
0.6	1.3	0.3
0.6	1.1	0.3
0.5	1.0	0.3
0.5	0.9	0.2
0.5	0.8	0.2
0.4	0.8	0.2
0.3	0.7	0.2
0.2	0.5	0.2

Step 3

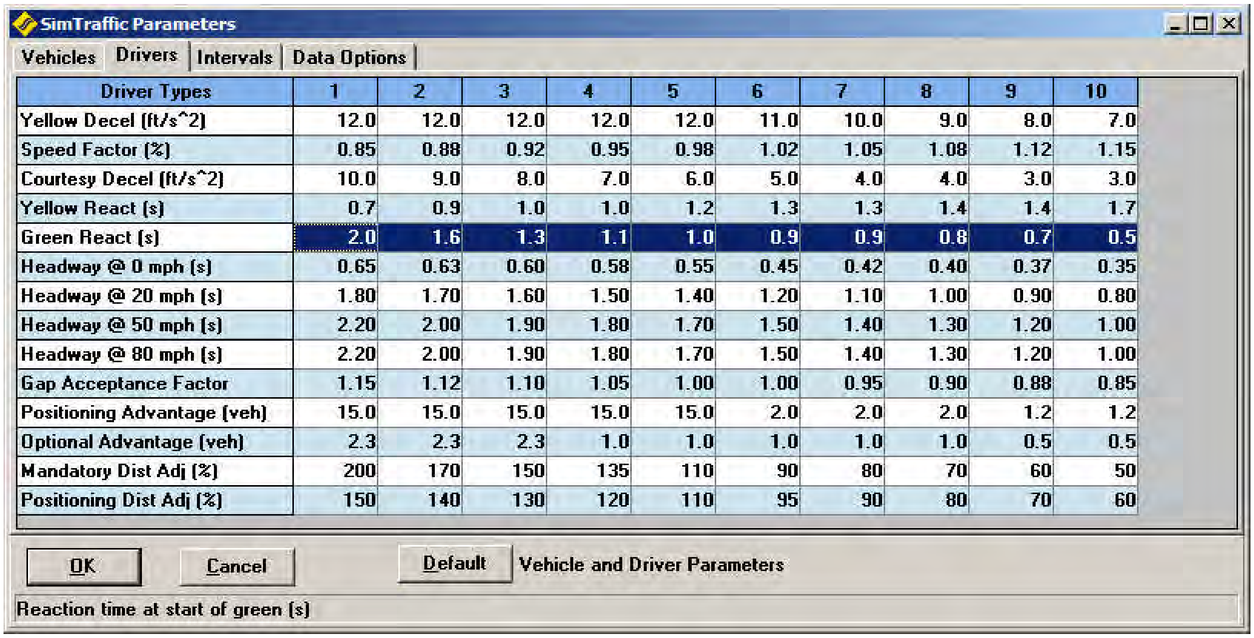
The final step was to run a series of 16 tests on each of the eight locations. The tests represented the 16 unique combinations of applying the three calibration parameters to the base or reference, "visually calibrated" network. As touched on in the discussion of the development of the calibration parameters (*Step 2*), there were four possible settings for headway factor and two possible settings (on or off) for 85th percentile speed and driver reaction time, creating a 4 x 2^2 factorial design (16 cases). The factorial design was applied as follows:

	X_1	X_2	X_3
Case 1 (Base)	-1	-1	-1
Case 2	1	-1	-1
Case 3	-1	1	-1
Case 4	1	1	-1
Case 5	-1	-1	1
Case 6	1	-1	1
Case 7	-1	1	1
Case 8	1	1	1
Case 9	Min	-1	-1
Case 10	Min	1	-1
Case 11	Min	-1	1
Case 12	Min	1	1
Case 13	Max	-1	-1
Case 14	Max	1	-1
Case 15	Max	-1	1
Case 16	Max	1	1

In this table (-1) and (1) represent off (default) and on (field measurement used), respectively. Each case has a unique set of inputs and by testing all cases the goal was to determine the importance of each input. For cases (models) where the headway factor was to be applied the headway factor was collected and calculated at the entrance approaches for the major arterial being studied, creating two points. The headway factors at the entrance points were assumed to carry through the study area. Example, if the major route was East-West, a WB headway factor would be collected at the intersection farthest east and a EB headway factor would be collected at the intersection farthest west. The EB and WB headway factor inputs would be adjusted to the calculated value at every intersection along the major route for any case that required that a measured headway factor be used (note, there were 16 different ways that the headway factors were calculated, three measurements were applied in the testing; the ideal (1), min, and max).

For cases where the 85th percentile speed was applied, the measured speed was used in place of the posted speed. The speed was measured along the major route and as close the study area as possible. For most of the locations, measurements could not be made within the study area since the close intersections did not allow for free flow speeds to be reached, or the area was not safe enough to setup for speed collection. If the measurement could not be made in the study area, it was made immediately outside the area on a representative section. The side streets were always assumed at the posted speed.

For cases where the driver reaction time was applied the default SimTraffic driver reaction time (Green React) by drive type was overwritten by the driver reaction times averaged over the 31 distributions collected from this study for ODOT (note the highest reaction time, 2.4 – driver type 1, was limited to 2.0 by the SimTraffic software).


SimTraffic Parameters

Vehicles | Drivers | Intervals | Data Options

Driver Types	1	2	3	4	5	6	7	8	9	10
Yellow Decel (ft/s^2)	12.0	12.0	12.0	12.0	12.0	11.0	10.0	9.0	8.0	7.0
Speed Factor (%)	0.85	0.88	0.92	0.95	0.98	1.02	1.05	1.08	1.12	1.15
Courtesy Decel (ft/s^2)	10.0	9.0	8.0	7.0	6.0	5.0	4.0	4.0	3.0	3.0
Yellow React (s)	0.7	0.9	1.0	1.0	1.2	1.3	1.3	1.4	1.4	1.7
Green React (s)	2.0	1.6	1.3	1.1	1.0	0.9	0.9	0.8	0.7	0.5
Headway @ 0 mph (s)	0.65	0.63	0.60	0.58	0.55	0.45	0.42	0.40	0.37	0.35
Headway @ 20 mph (s)	1.80	1.70	1.60	1.50	1.40	1.20	1.10	1.00	0.90	0.80
Headway @ 50 mph (s)	2.20	2.00	1.90	1.80	1.70	1.50	1.40	1.30	1.20	1.00
Headway @ 80 mph (s)	2.20	2.00	1.90	1.80	1.70	1.50	1.40	1.30	1.20	1.00
Gap Acceptance Factor	1.15	1.12	1.10	1.05	1.00	1.00	0.95	0.90	0.88	0.85
Positioning Advantage (veh)	15.0	15.0	15.0	15.0	15.0	2.0	2.0	2.0	1.2	1.2
Optional Advantage (veh)	2.3	2.3	2.3	1.0	1.0	1.0	1.0	1.0	0.5	0.5
Mandatory Dist Adj (%)	200	170	150	135	110	90	80	70	60	50
Positioning Dist Adj (%)	150	140	130	120	110	95	90	80	70	60

OK | Cancel | Default | Vehicle and Driver Parameters

Reaction time at start of green (s)

Prior to this study, the APM required that every microsimulation be run a minimum of five times so that a "true" average could be analyzed, helping to remove the effect of uncommon random effects in traffic arrival or patterns. At the beginning of this study TPAU recognized that, there can be significant variation in MOEs between each SimTraffic run. Because of this each of the 16 cases was run to the point where the MOEs produced a "static" average value. The following equation was used to determine how many runs were required for each case,

$$n = t^2\sigma^2 / (error)^2$$

For this study, a confidence interval of 90% was used. Each MOE's standard deviation and mean for the model area were measured and put into this equation to determine how many runs were required to achieve a model wide mean with 90% confidence. The study found that the number of runs required varied greatly depending on which measure was to be used. Queue lengths could require ~50 runs for some of the larger more congested networks (Bend, Milwaukie). For the smaller less congested networks (Woodburn, Albany), the standard 5 runs allowed all of the MOE's to be within a 90% confidence interval. After all the data had been tested an analyzed "Vehicles Exited" was the only MOE determined to be acceptable to recommend to be used to determine calibration. "Vehicles Exited" is a stable measure and never required more than five runs. It is for this reason that the APM will continue to recommend a minimum of five runs, however future work will have to fully investigate the sensitivity of queue lengths and how many runs is required to achieve a stable measure. Queue lengths are the primary reason to go to microsimulation so it is very important that they are reported correctly.

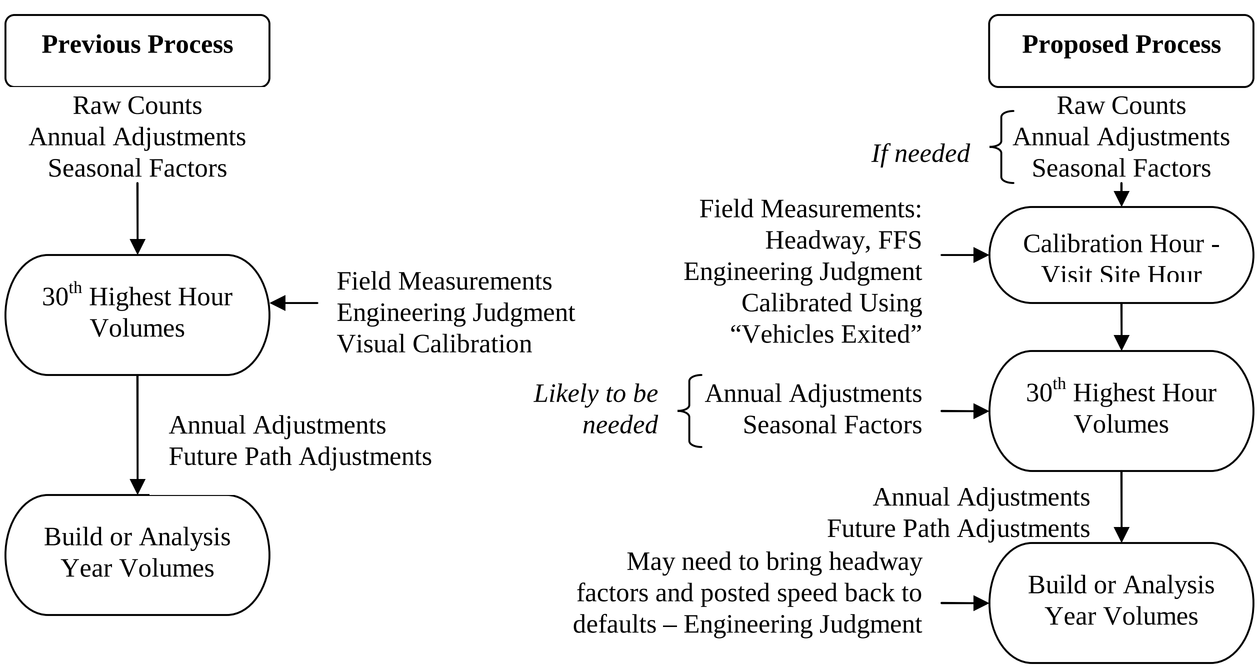
Conclusions / Proposed APM Changes

The work from this study improves on the current microsimulation process outlined by the APM. However, this work was not fully conclusive. From the comparison of the eight test sites, no clear combination of headway factors, 85th percentile speed, or driver reaction time, was found to improve the performance across all of the simulations and performance measures. This was also true for comparing test sites by their area characteristics; again no consistent combination of factors to achieve calibration could be found.

Consistent with the literature review conducted for this study, the APM should recommend that engineering judgment be used to correctly apply headway factors (HF) (saturation flow rates) and free-flow Speeds (85th percentile) to achieve an acceptable calibration. As indicated in Trafficware's Manual, free-flow speeds applies to turning speeds as well as corridor speeds.

That being the case, the ODOT driver reaction time should be used, as it was found that the SimTraffic default is grossly different than Oregon driver reaction times. In addition, Trafficware recommends that driver reaction times be adjusted as one of the three major parameters to achieve calibration. The ODOT driver reaction time will continue to be updated and refined as more data is collected and should be used for all projects on Oregon highways using microsimulation.

The major change to the prior APM methodology will be additional "network" or "build" that will need to be constructed for any given microsimulation project. Previously, the APM instructed users to build a "base year network" with 30th highest hour volumes. Then from that a future year (build year) network was saved with volumes grown from the 30th highest hour to the future year. The new process, which will include this calibration work, will be almost the same, except it will we require one more volume set for calibration. The following flowcharts illustrate the change:



The additional step in the "proposed process" would add an additional volume set during the base year development. To a point this work is already being done in the "previous process", however it is not formally documented and saved in a standard manner, as proposed. The proposed method, would adjust the raw counts up to the day that the area was visited by the analyst, who would then be able to verify the performance of the study area. From this "visit site hour", the visual calibration and "true" calibration, using "vehicles exited" could be completed. To reach "true" calibration, the analyst could use methods from the visual calibration or by incorporating field measured headway factors or 85th percentile speed, determined from the analyst's engineering judgment.

> *The study found that, although saturation flow rates can vary greatly from day-to-day (+200 vehicles per hour difference on different days), an area's saturation flow rates tend to average with a reasonable standard deviation over multiple days, under similar levels of congestion. If the saturation flow count can not be performed on the count / visit day, it is advised that the 15 or more cycles be collected during the peak hour over several days, to establish an average that the analyst has confidence in.*
>
> *Further, this study found that there was no benefit in deviating from the HCM's averaging the saturation flow counts (using median in place of average). In addition, the study findings agreed with the HCM when cycles with less than 8 queued vehicles are used. The findings showed that using cycles with less than 8 queued vehicles will decrease the saturation flow. Therefore, as the HCM states, only cycles with a queue of 8 or more vehicles may be used as part of the 15 or more cycles required to average the saturation flow rate for an approach.*
>
> *If an analyst feels that headway factors or 85th percentile speed would improve the calibration but the timeline, scale of project, and/or budget do not allow for field measurements, the following ranges (found from this study) can be applied in place of the SimTraffic/SYNCHRO defaults.*
>
>> *85th percentile speed – the average found was ~5 mph higher than the posted*
>>
>> *Headway factors – Using the APM defaults of 1900 veh/hour for METRO, Salem, and Eugene and 1750 for all other areas gives headway factors of 0.98 and 1.10, respectively.*

The "true" calibration is to be determined with the use of the MOE, "Vehicles Exited". The visual calibration can be considered as approaching calibration, or adding accuracy to the calibration. Using the quantitative measure of "Vehicles Exited" adds reassurance and precision to the calibration, quantitatively reaffirming that the number of vehicles seen passing through the system matches the number of vehicles passing through the model, and that there are no trouble spots where vehicles are incorrectly queuing or being blocked. "Vehicles Exited" represents the number of vehicles that make it through an intersection over a given period of time. This should equal the volume coded in the network for the "visit site hour". A tolerance of 1% for each intersection over the analysis period (hour) will be required to achieve calibration for the calibration volume set (not required for the 30th highest hour or build year network). The tolerances for each movement may vary depending on volume, but any movement over 100 vehicles/hour should be within 5% of the coded volume. Movements with less than 100 vehicles/hour should be checked to make sure that the vehicles exiting is reasonable. One of the extra benefits to using "Vehicles Exited" is that this is an automated report from SimTraffic and does not require any external software or manipulation of the data.

1001: Route 20 & Spring Hill Drive Performance by movement Entire Run

Movement	EBL	EBT	WBT	WBR	SBL	SBR	All
Total Stops	38	353	704	25	298	14	1432
Avg Speed (mph)	12	26	16	23	9	14	19
Vehicles Entered	39	1161	1198	476	317	19	3210
Vehicles Exited	39	1164	1200	476	320	19	3218
Hourly Exit Rate	39	1164	1200	476	320	19	3218
Input Volume	38	1169	1187	489	311	17	3211
% of Volume	103	100	101	97	103	112	100

Note that SimTraffic distributes vehicles to the network randomly based on coded volumes. Seeding offsets and random occurrence can cause the exiting volume to be greater than the coded volume. This is usually negligible, although if an analyst sees a consistent over assignment at a location further investigation may be necessary. The primary purpose for reviewing "Vehicles Exited" is to ensure that the number of vehicles exiting the intersection is not greatly less (~5%) than the coded volume. Less volume can indicate an improper blockage, bottle network, or miscoding in the network, and needs to be investigated and corrected.

After the "visit site hour" volume set calibration has been established, the second base year volume set, the 30th highest hour, would be grown to the 30th highest hour volumes and re-balanced (if the sight visit was at the 30th hour, this set could be skipped). This volume set would likely show worse conditions than the calibrated state.

The last volume set would have the volumes grown, adjusted, and balanced to the build year or analysis year (Design Hour Volume – DHV). For the Future No Build, any headways used for calibration would typically be brought back to statewide defaults (averages) if the headways in the field were greater than the defaults, ie., as congestion increases over time, headways will approach statewide averages, all else remaining the same. If the field measured headways were found to be less than the defaults, they could be kept for the future scenarios.

This simulation guidelines work took two years to complete and involved members from all over ODOT, not just within TPAU. Many important insights and notes were collected from different people involved through out the course of this work. These notes will be come apart of the APM Chapter 8, but are best communicated here as informative bullets:

- The analyst must visit the location on or within a reasonable time period of the count collection, as the analyst will need to fully understand what is actually occurring during the peak so that they can verify that the simulation is within reasonable bounds, before any calibration is attempted.

- Phasing detail improves calibration in SimTraffic 7 (S7), however in some cases, a short min gap on turn bays or side streets causes the phase to gap out while a queue still exists. This can be found by visual inspection and should be corrected by adjusting the minimum gap to two seconds or to the vehicle extension, if the vehicle extension is less than two seconds. Also note, for S7, lost time must be adjusted and detectors must be added. The detector spacing defaults provided by TPAU are for state facilities. Side streets should be treated as left turn bays, with call detectors if better information is not available. Without call detection, the side streets can be skipped for many cycles. Actual detector spacing and settings should be used in place of defaults if they are available.

- Correct turn bay length and geometry is crucial. If users are "lengthening" the turn bays by illegally using shoulder or the median, this should be coded in the model.

- The first step before calibration is to perform a visual "laugh test", addressing any problems through application of engineering judgment, further field collection, or discussing issues with the local jurisdictions responsible for the area or side streets (turning speeds, sat flow studies, speed studies, geometry or detection issues, validating/rechecking signal phasing, pedestrians, blocked intersections, OD patterns, lane (turn) alignment, taper & turn bay lengths, positioning lengths…)

- All simulation runs need to be checked to make sure that a run didn't blow up or freeze up. If the network freezes or performs much worse than expected (use judgment), either adjustments to the network must be made, or if it is just for random runs, remove (rerun) the poor runs.

- In running all of the tests it was found that five runs was enough to ensure that "Vehicles Exited" reached a reasonable average value. Other measures like stops, speed and queues see a lot more variability and would require more than five runs and extra work to ensure that convergences had been reached, but using "Vehicles Exited" allows five runs to continue to be the standard. Future work will have to be done to determine how many runs are needed to ensure that queue lengths have reached a stable average.

Literature Reviewed

A Report on the Use of Traffic Simulation Models in the San Diego Region. Institute of Transportation Engineers – California Border Section and Highway Capacity Task Force, March 2004.

Bayarri, M. J., J. O. Berger, G. Molina, N. M. Rouphail, and J. Sacks. *Assessing Uncertainties in Traffic Simulation: A Key Component in Model Calibration and Validation.* National Institute of Statistical Sciences, Research Triangle Park, NC, July, 2003.

Benekohal, R. F., and G. Abu-Lebdeh. Variability Analysis of Traffic Simulation Outputs: Practical Approach for TRAF-NETSIM. In *Transportation Research Record 1457*, TRB, National Research Council, Washington, D.C., 1994, pp. 198-206.

Chu, L., H. X. Liu, J.-S. Oh, and W. Recker. *A Calibration Procedure for Microscopic Traffic Simulation*. Cambridge Systematics, Irvine, CA, June 2002

Chu, L., and H. X. Liu. Caltrans TMS Master Plan – Calibration Methodology and Results for the I-405/I-5 Simulation Model. In *TRB 2006 Annual Meeting CD-ROM*, TRB, National Research Council, Washington, D.C., 2006.

Ciuffo, B., V. Punzo, and V. Torrieri. A Comparison Between Simulation-Based and Model-Based Calibrations of Traffic Flow Micro-Simulation Models. In *TRB 2008 Annual Meeting CD-ROM*, TRB, National Research Council, Washington, D.C., 2008.

Dowling, R., J. Holland, and A. Huang. *California Department of Transportation: Guidelines for Applying Traffic Microsimulation Modeling Software.* Dowling Associates, Oakland, CA, September 2002.

Dowling, R., A. Skabardonis, J. Halkias, G. McHale, and G. Zammit. Guidelines for Calibration of Microsimulation Models. In *Transportation Research Record 1876*, TRB, National Research Council, Washington, D.C., 2004, pp. 1-9.

Hellinga, B. R. *Requirements for the Calibration of Traffic Simulation Models.* Department of Civil Engineering, University of Waterloo, Ontario, Canada, 1998.

Hollander, Y. and R. Liu. *The Principles of Calibrating Traffic Microsimulation Models.* In *Transportation Vol 35 Number 3*, Kluwer Academic Publishers, Dordrecht, Boston, January, 2008, pp. 347-362.

Hourdakis, J., P. G. Michalopoulos, and J. Kottommannil. Practical Procedure for Calibrating Microscopic Traffic Simulation Models. In *Transportation Research Record 1852*, TRB, National Research Council, Washington, D.C., 2003, pp. 130-139.

Jayakrishnan, R., J.-S. Oh, and A.-E. K. Sahraoui. *Calibration and Path Dynamics Issues in Microscopic Simulation for Advanced Traffic Management and Information Systems.* Institute of Transportation Studies, University of California, Irvine, CA, December, 2000.

Jones, S. L., A. J. Sullivan, M. D. Anderson, and D. Malave. *Traffic Simulation Software Comparison Study.* University Transportation Center for Alabama, June, 2004.

Liu, H. X., L. Ding, J. X. Ban, A. Chen, and P. Chootinan. A Streamlined Network Calibration Procedure for California SR41 Corridor Traffic Simulation Study. In *TRB 2006 Annual Meeting CD-ROM*, TRB, National Research Council, Washington, D.C., 2006.

Long, G. A Unified Theory of Saturation Flow. In *TRB 2006 Annual Meeting CD-ROM*, TRB, National Research Council, Washington, D.C., 2006.

Park, B. and J. D. Schneeberger. Microscopic Simulation Model Calibration and Validation. In *Transportation Research Record 1856*, TRB, National Research Council, Washington, D.C., 2003, pp. 185-192.

Park, B., J. Won, and I. Yun. Application of Microscopic Simulation Model Calibration and Validation Procedure: Case Study of Coordinated Actuated Signal System. In *Transportation Research Record 1978*, TRB, National Research Council, Washington, D.C., 2006, pp. 113-122.

Pennsylvania Avenue Traffic Alternatives Analysis – Final Report. Parsons Transportation Group, P.C., October 2001.

Shaaban, K. S. and E. Radwan. A Calibration and Validation Procedure for Microscopic Simulation Model: A Case Study of SimTraffic for Arterial Streets. In *TRB 2005 Annual Meeting CD-ROM*, TRB, National Research Council, Washington, D.C., 2005.

Tian, Z. Z., T. Urbanik II, R. Engelbrecht, and K. Balke. Variations in Capacity and Delay Estimates from Microscopic Traffic Simulation Models. In *Transportation Research Record 1802*, TRB, National Research Council, Washington, D.C., 2002.

Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software. FHWA, U.S Department of Transportation, July 2004.

Traffic Analysis Toolbox Volume IV: Guidelines for Applying CORSIM Microsimulation Modeling Software. FHWA, U.S Department of Transportation, January 2007.

Wan, B., N. Rouphail, and J. Sacks. *Failure Detection and Diagnosis in Micro-Simulation Traffic Models.* National Institute of Statistical Sciences, Research Triangle Park, NC, July, 2005.