Chapter 9

Model Calibration

John Hourdakisis
Center for Transportation Studies, U of Mn
Hourd001@tc.umn.edu
Why Calibrate?

• Computers Cannot Magically Replicate Reality!
  • Simulation Models Are Designed to be General
  • Driver Behavior and Road Characteristics Depend on
    ▪ Location i.e. Minnesota vs California
    ▪ Vehicle Characteristics (Horsepower, Size, etc.)
    ▪ Weather Conditions (Dry, Wet, Ice, etc.)
    ▪ Day or Night
  • Microscopic Simulators Can Adapt and Replicate Almost Any Condition if the Model Parameters Are Properly Adjusted
  • What is Realistic and What is Not?
Example of a Disaster

- Half of an Ordinary Double Cloverleaf
- Freeway Speed 65 mph, Ramps 45 mph
- Turning Warnings More Than 2000 Feet Away
- Lot’s of Traffic: 6500 veh/hour on Mainline
- Three Vehicle Types: Car, Truck, and Semi-Trailer
Simulation Results From 5 Scenarios

- Same Demands, Speed Limits, Turnings, Most of Model Parameters.
- Changed Car and Semi-Trailer:
  - Acceleration
  - Normal Deceleration
  - Maximum Deceleration (Emergency Stop)

<table>
<thead>
<tr>
<th></th>
<th>Average Speed</th>
<th>Average Flow</th>
<th>Total Travel Time</th>
<th>Av. Delay</th>
<th>LOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>+2f/s²</td>
<td>45.5 mph</td>
<td>6831 vph</td>
<td>257 hours</td>
<td>9 sec/veh</td>
<td>D</td>
</tr>
<tr>
<td>+1f/s²</td>
<td>29.1 mph</td>
<td>6200 vph</td>
<td>484 hours</td>
<td>41 sec/veh</td>
<td>F</td>
</tr>
<tr>
<td>BASE</td>
<td>46.7 mph</td>
<td>6396 vph</td>
<td>234 hours</td>
<td>9 sec/veh</td>
<td>D</td>
</tr>
<tr>
<td>-1f/s²</td>
<td>35.3 mph</td>
<td>6152 vph</td>
<td>393 hours</td>
<td>30 sec/veh</td>
<td>E</td>
</tr>
<tr>
<td>-2f/s²</td>
<td>27.4 mph</td>
<td>5500 vph</td>
<td>536 hours</td>
<td>57 sec/veh</td>
<td>F</td>
</tr>
</tbody>
</table>
Calibration Procedure

**Historical system input data**

- **Actual system**
  - **System output data**
  - **Compare**
    - **Valid?**
      - **No**
        - **Calibrate**
  - **Simulation Model**
  - **Model output data**

- **Historical system input data**
REALITY
Current Traffic Measurements, etc.

• Need to KNOW the Network to be Replicated
• Depending on the Scope of the Project, Two or More of the Following Are Needed:
  • Mainline Volumes (Every X Feet)
  • Mainline Speeds (Every X Feet)
  • Travel Times (Link or Between O/D Pairs)
  • Bottleneck Capacity (Measured, Not Theoretical)
  • Entrance Ramp Queues
  • Intersection Queues and Queue Discharge Rates
Modeling Parameters

- **Global Parameters**
  - Vehicle Parameters
    - Size and Power
    - Driver Behavior
  - Car-Following Model Parameters
    - Reaction Time
    - Speed Distribution Among Lanes (Overtaking Manuevers)

- **Local Section Parameters**
  - Curvature and Grade
  - Speed Limit or Free Flow Speed
  - Lane Changing Distances
  - Headway/Hesitation Factor
Model Parameter Issues

• Driver/Vehicle Characteristics
  • Literature Does Not Provide Adequate Information
  • Not Common to All Simulators
  • No Info on Effects of Weather/Pavement/Ambient Conditions

• Car-Following Parameters
  • Depend on the Simulator’s Car-Following Model Employed.
  • Most Simulators Do Not Adequately Describe the Modeling Process (if at all).
    • User Has No Sense of Model Parameter Effects on Results.
Calibration Issues

- Very Important For Model Accuracy and Robustness
- Accuracy Depends on Measurement Granularity
  - Averages Over Several Days is a Bad Choice
  - Might Need Additional Information to be Collected in Turbulent Sections (Bottlenecks, Weaving Areas, etc.)

- Simulation Objective Affects Calibration
  - When Adaptive Control Strategies Are Simulated, Stricter Validation is Needed
  - Modeling of an Isolated Interchange in Rural Minnesota Will be Restrictive
Calibration Issues

• VERY TIME CONSUMMING PROCESS
  • Currently Simulators Do Not Provide a Methodology or Tools to Assist in Calibration
  • Often Users End Up in Endless Trial-and-Error Cycles
  • Sporadic Attempts Made in Literature to Streamline the Process But:
    • Focused on a Particular Simulator
    • Too Complex or too Naive to be Effectively Used in Practice
    • No Widely Accepted Methods/Standards Currently Available
Before Calibration!

- Check Geometry For Correctness
  - Disjoined Sections
  - Stuck Vehicles (Sizes of Accel/Decel Lanes)
  - Verify Location of Detectors
- Check Input For Accuracy
  - Entrance Volume Comparison (Perfect Match)
  - Exit Volume Comparison (Match Sum Over All Hours)
  - Volume Totals on Mainline Stations Should Match
Practical Calibration Methodology (Employed on Twin Cities Freeways)

- Need Simultaneous Boundary and Mainline Station Measurements.
- Technique:

  upstream

  mainline detector stations

- Objective is to Match the Simulated and Actual Mainline Traffic Measurements
- Traffic Measurements Used: Volume and Speed
  - Occupancy Affected by Detector Sensitivity (Unknown)
- Perform Calibration in Stages:
  - First 2 Stages Based on Volume and Speed in That Order
  - Further Improvements in Optional 3rd Stage:
    - Depending on Objective i.e. For Ramp Control -> Queue Length
Goodness-of-Fit Test Measures

- Recommended Goodness-of-Fit Measures:

  1. RMS Percent Error = \[
  \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \frac{x_i - y_i}{y_i} \right)^2}
  \]
    (Measures Overall % Error)

  2. Correlation Coefficient = \[
  CORL = \frac{1}{n-1} \sum_{i=1}^{n} \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}
  \]
    (Measures Linear Association)

Where:
- \(x_i\) is the Simulated Traffic Measurement at Time \(i\)
- \(x'\) is the Mean of the Simulated Traffic Measurements
- \(y_i\) is the Actual Traffic Measurement at Time \(i\)
- \(y'\) is the Mean of the Actual Traffic Measurements
- \(\sigma_x\) is the Standard Deviation of the Simulated Traffic Measurement
- \(\sigma_y\) is the Standard Deviation of the Actual Traffic Measurements
- \(n\) is the Number of Traffic Measurement Observations
3. Theil’s U (Considers the Disproportionate Weight of Large Errors)

- **3 Components of Theil’s U**
  - **Us**
    (Measure of Variance Proportion, Close to 1 Satisfactory)
  - **Uc**
    (Measure of Covariance Proportion or Unsystematic Error, Close to 0 Satisfactory)
  - **Um**
    (Measure of Bias Proportion or Systematic Error, Close to 0 Satisfactory)

Where $\sigma_y$ and $\sigma_x$ are the Standard Deviations of the Actual and Simulated Series

$r$ is the Correlation Coefficient Between the Two Series

$$U_s = \frac{n(\sigma_y - \sigma_x)^2}{\sum_{i=1}^{n}(y_i - x_i)^2}$$

$$U_c = \frac{2(1-r)n\sigma_y\sigma_x}{\sum_{i=1}^{n}(y_i - x_i)^2}$$

$$U_m = \frac{n(\bar{y} - \bar{x})^2}{\sum_{i=1}^{n}(y_i - x_i)^2}$$
Examples

Figure 1(a): Illustration of unsatisfactory Um

Figure 1(b): Illustration of unsatisfactory Uc

Figure 1(c): Illustration of unsatisfactory Uc

Figure 1(d): Illustration of unsatisfactory Us
Stage 1: Volume-Based Calibration

- Objective is to Match Simulated and Actual Mainline Station Volumes
- Simulation Model Calibrated Beginning Upstream and Proceeding Downstream
- Global Parameters Are Calibrated First:
  - Usually Accomplished in First Few Stations
  - Trial & Error Iterative Process For Each Parameter
  - RMSP, r and U are the Metrics Used in Each Iterations
- Local Parameters Calibrated at All Stations
  - Um, Uc and Us are the Metrics Used at This Point.
Stage 2: Speed-based Calibration

- Objective is to Match Simulated and Actual Mainline Station Speeds and Bottleneck Locations
- Actual Speeds Derived From Volume, Occupancy, and Effective Vehicle Length (For Single Loop Detectors)
- Speed Contour Graphs Used For Comparing (Visually) Simulated and Actual Speeds
  1. If Speed Contours Exhibit Significant Discrepancy, Revise Global Parameters From Stage 1
  2. Beginning Upstream and Proceeding Downstream, Calibrate Local Parameters Until Mainline Speeds and All Bottleneck Locations in the 2 Contour Graphs Match
Stage 3: Application Dependent

- For Adaptive Ramp control:
  - Compare Queue Lengths
  - Local Simulation Parameters Affecting Detector Output
    Need Further Calibration in This Stage

- Simulated Entrance Ramp Queues Should Match Actual Ones
- Queue Measurements From 30 sec Detector Counts
- Queue = Metered Demand - Actual Demand (Upstream)

- Over-Calibration to be Avoided to Ensure Generality
  - Repeat Simulation With Different Random Seeds
  - Simulate Additional Days
Example of Calibration

- AIMSUN Microsimulator
- Mn/DOT Ramp Metering Evaluation.
- 2 Test Sites (Only One Presented)
  - TH 169 Northbound in Minneapolis, MN
    - 12 Miles Long: From I-494 to 63rd Avenue N
    - 24 Entrance Ramps, 25 Exit Ramps
    - 30 Detector Stations.
    - 5-Minute Volume and Occupancy
    - March 21st to 23rd, 2000
    - 14:00 to 20:00 hrs
Test site 1: TH 169NB
Stage 1 Results (Volume)

- 500 Simulator Iterations Required, 2 Months (!)
- Irregularities in Input Data Observed Due to Sensor Misplacement

<table>
<thead>
<tr>
<th>Goodness-of-Fit Measure</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>0.960</td>
</tr>
<tr>
<td>RMSP</td>
<td>7.39%</td>
</tr>
<tr>
<td>U</td>
<td>0.002</td>
</tr>
<tr>
<td>Um</td>
<td>0.088</td>
</tr>
<tr>
<td>Us</td>
<td>0.031</td>
</tr>
<tr>
<td>Uc</td>
<td>0.881</td>
</tr>
</tbody>
</table>
Actual Speed Contour
Stage 2 Results
(Speed)

Simulated Speed
After Stage 1

Simulated Speed
After Stage 2
(About 200 Iterations)

Actual Speed
Stage 3 Results (Queues)

- Simulated and Actual Queues Did Not Match Before This Stage
- 3\textsuperscript{rd} Stage Required About 100 Iterations

Example Ramp

Simulation vs Real Queues
Validation Accuracy (Volume)

<table>
<thead>
<tr>
<th></th>
<th>RMS%</th>
<th>r</th>
<th>U</th>
<th>Um</th>
<th>Us</th>
<th>Uc</th>
</tr>
</thead>
<tbody>
<tr>
<td>March 21st (validation)</td>
<td>10.620</td>
<td>0.980</td>
<td>0.004</td>
<td>0.309</td>
<td>0.011</td>
<td>0.681</td>
</tr>
<tr>
<td>March 22nd (validation)</td>
<td>6.420</td>
<td>0.970</td>
<td>0.001</td>
<td>0.124</td>
<td>0.054</td>
<td>0.823</td>
</tr>
<tr>
<td>March 23rd (Calibration)</td>
<td>7.390</td>
<td>0.960</td>
<td>0.002</td>
<td>0.088</td>
<td>0.031</td>
<td>0.881</td>
</tr>
</tbody>
</table>

Over All Stations.
## Results (Calibrated Parameters)

**AIMSUN Microscopic Simulator**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Initial</th>
<th>After stage 1</th>
<th>After stage 2</th>
<th>After stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max. desired speed (kmph)</td>
<td>105.000</td>
<td>110.000</td>
<td>110.000</td>
<td>110.000</td>
</tr>
<tr>
<td>Max. acc. rate (m/s²)</td>
<td>4.500</td>
<td>3.000</td>
<td>3.000</td>
<td>3.000</td>
</tr>
<tr>
<td>Normal dec. rate (m/s²)</td>
<td>-4.500</td>
<td>-5.000</td>
<td>-5.000</td>
<td>-5.000</td>
</tr>
<tr>
<td>Max. dec. rate (m/s²)</td>
<td>-5.000</td>
<td>-5.500</td>
<td>-5.500</td>
<td>-5.500</td>
</tr>
<tr>
<td>Reaction time (sec)</td>
<td>0.700</td>
<td>0.590</td>
<td>0.610</td>
<td>0.610</td>
</tr>
<tr>
<td>Percent overtake</td>
<td>0.950</td>
<td>0.950</td>
<td>0.940</td>
<td>0.940</td>
</tr>
<tr>
<td>Percent recover</td>
<td>1.000</td>
<td>1.000</td>
<td>0.990</td>
<td>0.990</td>
</tr>
<tr>
<td>Max. speed difference (kmph)</td>
<td>40.000</td>
<td>40.000</td>
<td>60.000</td>
<td>60.000</td>
</tr>
<tr>
<td>Max. speed difference on-ramp (kmph)</td>
<td>50.000</td>
<td>50.000</td>
<td>70.000</td>
<td>70.000</td>
</tr>
<tr>
<td>Av. section speed (regular section, kmph)</td>
<td>110.000</td>
<td>100.000</td>
<td>105.000</td>
<td>105.000</td>
</tr>
<tr>
<td>Av. section speed (weaving section, kmph)</td>
<td>90.000</td>
<td>75.000</td>
<td>70.000</td>
<td>70.000</td>
</tr>
<tr>
<td>Av. section speed (ramp section, kmph)</td>
<td>60.000</td>
<td>60.000</td>
<td>55.000</td>
<td>55.000</td>
</tr>
</tbody>
</table>
Conclusion

• Garbage In >>>> Garbage Out

• Simulation Useless/Dangerous Without Calibration

Questions?