ABSTRACT

TOPL is a suite of software tools for specifying freeway operational improvement strategies, such as ramp metering, demand and incident management, and for quickly estimating the benefits of such improvements. TOPL is based on the macroscopic cell transmission model. The paper summarizes the theory of the cell transmission model and describes the procedure to carry out a TOPL application. The procedure is illustrated for the 26-mile long I-210W freeway in California, whose model is calibrated using loop detector measurements of volume and speed. The measurements show that congestion originates in a bottleneck and moves upstream, as predicted by the theory. The simulations show that appropriate ramp metering can dramatically reduce total congestion delay and mainline travel time.

1 INTRODUCTION

In November, 2006 California voters approved a $20 billion bond measure to improve transportation. Subsequently, the California Department of Transportation (Caltrans) launched an ambitious ‘corridor management program’ to design and implement operational improvements—emphasizing ramp metering, incident management, traveler information, and demand management (including using tolls)—that would reduce congestion in 2025 by 40 percent [1]. This paper describes TOPL (Tools for Operations Planning), a suite of software tools for (1) specifying such operational improvements and (2) quickly estimating the benefits such improvements are likely to provide.

TOPL is based on the macroscopic Cell Transmission Model (CTM). Traditionally, transportation planning investigations favor use of microscopic models, and indeed Caltrans has let contracts for microsimulation models. However, data collection and model calibration efforts are significant for microscopic models, thereby slowing these efforts [2]. In contrast, the CTM model is based on aggregate variables such as volume or flow and density, which for California freeways, are routinely measured and archived [3]. Consequently, TOPL models are very quickly specified, calibrated and run to generate useful results.

This paper illustrates the procedure used in TOPL to calibrate the macroscopic Cell Transmission model. Section 2 reviews the
Cell Transmission model used for freeway corridor simulation. It also highlights main results advocating the use of ramp metering to alleviate excessive demand [4]. Section 3 explains the procedure adopted in TOPL for freeway network calibration, imputation, analysis and simulation. Finally, section 4 illustrates the application of the methodology to simulate and study the 26-mile long I-210W freeway in the Los Angeles area.

2 CELL TRANSMISSION MODEL

The cell transmission model used to simulate freeway traffic in TOPL [5]. This section summarises the model and the traffic behavior predicted from the analysis of this model [4]. Figure 1 shows the freeway divided into $N$ sections or cells, each with a maximum of one on- and one off-ramp. Vehicles move from right to left. There are two boundary conditions. Free flow prevails downstream of section 0, and vehicles from upstream of the freeway enter an “on-ramp” with specified inflow $r_N$. The flow accepted by section $(N-1)$ is $f_N(k)$ vehicles in period $k$. The cumulative difference forms a queue of size $n_N(k)$.

![Figure 1. FREEWAY WITH N SECTIONS.](image)

Table 1 lists the model variables and parameters. The length of all sections is normalized to 1 by absorbing differences in length in the speeds $v_i, w_i$. Since the lengths have been fixed by the segmentation scheme described before, the time period of the simulation is chosen to effect $0 < v_i, w_i < 1$, so that the vehicles do not cross an entire cell in one period, making the scheme unstable. The off-ramp flow is assumed to be a portion $\beta_i(k)$ of the total flow leaving the section:

$$s_i(k) = \beta_i(k)(s_i(k) + f_i(k)), \quad \text{or} \quad s_i(k) = \left[\beta_i(k)/(1 - \beta_i(k))\right]f_i(k).$$

Assume $\beta_N(k) = 0 \forall k$. With $\bar{\beta}_i(k) = 1 - \beta_i(k)$, the model is, for $k \geq 0$,

$$n_i(k + 1) = n_i(k) - f_i(k)/\bar{\beta}_i(k) + f_{i+1}(k) + r_i(k), \quad 0 \leq i \leq N - 1, \quad (1)$$

$$f_i(k) = \min\{\bar{\beta}_i(k)v_i n_i(k), w_{i-1}[n_{i-1} - n_{i-1}(k)], F_i\}, \quad 1 \leq i \leq N, \quad (2)$$

$$f_N(0) = \min\{\bar{\beta}_0(k)v_0 n_0(k), F_0\}, \quad (3)$$

$$n_N(k + 1) = n_N(k) - f_N(k) + r_N(k). \quad (4)$$

Flow conservation in section $i \leq N - 1$ is expressed by

$$n_i(k + 1) = n_i(k) - f_i(k) + f_{i+1}(k) + r_i(k) - s_i(k), \quad (5)$$

which is equivalent to (1), using $s_i(k) = \beta_i(k)/\bar{\beta}_i(k)f_i(k)$; and in section $N$ by (4). The flow $f_i(k)$ from section $i$ to $i-1$ is governed by the ‘fundamental diagram’ (2) with this interpretation: $\bar{\beta}_i v_i n_i(k)$ is the number of vehicles that can move from section $i$ to $i-1$ in period $k$: $w_{i-1}[n_{i-1} - n_{i-1}(k)]$ is the number that $i-1$ can accept; and $F_i$ is the capacity or maximum possible flow from section $i$ to $i-1$. Equation (3) indicates there is no congestion downstream of section 0. It is tacitly assumed that the flows $s_i(k)$ are not constrained by off-ramp capacity.

The parameters in Table 1 correspond to the fundamental diagram of Figure 2. The state of the system is the $N$-dimensional vector $n(k) = (n_0(k), \ldots, n_{N-1}(k))$. Assume stationary demands $r_i(k) \equiv r_i$. Then the following propositions hold [4].

**Proposition 1.** A feasible demand $r$ has a unique uncongested equilibrium $n^u(r)$:

$$n^u_i(r) = (\bar{\beta}_i v_i)\left[f_i(r), \quad 0 \leq i \leq N - 1. \quad (6)\right]$$

In addition to the uncongested equilibrium (6), there is an infinite number—in fact, a continuum—of congested equilibria. Let $E = E(r)$ be the set of equilibria. Then, Each on-ramp demand vector...
Proposition 2 indicates that congestion starts at a bottleneck section and spreads upstream. Also, the flows in all sections are the same at every equilibrium in the set (11), even though at (say) the equilibrium \( n^* \) no section is congested and vehicles move at free flow speed, whereas at \( n^{N-1} \) every section is congested and vehicles move at lower speed. Thus the presence of congestion is not an indication of excess demand.

**Proposition 3.** (i) \( \hat{r}_N < r_N \) is the largest upstream flow for which the demand \( \hat{r} = (r_0, \cdots, r_{N-1}, \hat{r}_N) \) is feasible. The corresponding equilibrium flow \( \hat{\phi} \) is

\[
\hat{\phi}_N = \hat{r}_N, \quad \hat{\phi}_i = \hat{\beta}_i (\hat{\phi}_{i+1} + r_i), \quad 0 \leq i \leq N - 1.
\]

(ii) With demand \( r \), under the no-metering strategy the system converges to the (unique) most congested equilibrium \( n^{con} \in E(\hat{r}) \) corresponding to demand \( \hat{r} \). The queue \( n_N(k) \) at the upstream ramp grows at the rate of \( (r_N - \hat{r}_N) \) vehicles per period.

(iii) \( \hat{r}_0 < r_0 \) is the largest flow for which the demand \( \hat{r} = (\hat{r}_0, r_1, \cdots, r_N) \) is feasible. The corresponding equilibrium flow \( \hat{\phi} \) is

\[
\hat{\phi}_N = r_N, \quad \hat{\phi}_i = \hat{\beta}_i (\hat{\phi}_{i+1} + r_i), \quad 1 \leq i \leq N - 1, \quad \hat{\phi}_0 = \hat{\beta}_0 (\hat{\phi}_1 + \hat{r}_0).
\]

Under the ramp metering strategy that reduces the on-ramp flow in section 0 from \( r_0 \) to \( \hat{r}_0 \), the system converges to some equilibrium in \( E(\hat{r}) \). The queue at the on-ramp in section 0 grows at the rate of \( (r_0 - \hat{r}_0) \) vehicles per period.

(iv) Flows under the ramp-metering strategy are larger throughout the freeway:

\[
\hat{\phi}_i < \phi_i, \quad 1 \leq i \leq N \quad \text{and} \quad \hat{\phi}_0 = \hat{\phi}_0 = F_0.
\]

Suppose \( \hat{\beta}_i > 0 \) for some \( i \geq 1 \), so that there is non-zero off-ramp flow in at least one section. Then the total discharge under the ramp-metering strategy is strictly larger than under the no-metering strategy. Moreover,

\[
\mu = \frac{r_N - \hat{r}_N}{r_0 - \hat{r}_0} = (\hat{\beta}_1 \cdots \hat{\beta}_N)^{-1} > 1.
\]

**Proposition 4.** There is a ramp metering strategy that achieves the metering gain. With a small sacrifice in capacity, the strategy achieves the uncongested equilibrium.

### 3 TOPL PROCEDURE

The TOPL procedure for calibration, imputation and scenario simulation require the following steps.
3.1 Network specification

The freeway network must be defined in the form of Figure 1, i.e., the freeway must be divided into cells, each with (at most) one on- and one off-ramp. Each cell should be homogeneous in terms of number of lanes and grade, so that it is sensible to represent the behavior of traffic in the cell by a single aggregate fundamental diagram. In order to facilitate calibration, it is desired that each cell should have a vehicle detector station that measures volume and speed.

In TOPL, network specification begins with a GIS map from which the freeway geometry (number of lanes and position of ramps) is extracted. An algorithm takes the specified geometry and the location of detector stations, obtained from PeMS [3] and produces a cell division of the freeway. The resulting cell structure is manually inspected and changed if needed.

3.2 Fundamental diagram Calibration

The calibration of the fundamental diagram parameters shown in Figure 2 is performed on data provided by PeMS. Each cell of the network contains a vehicle detector station (vds) that measures volume and speed. A health history report is generated to determine the days when the detector was reported to function properly, and data is obtained for days when the detector registered congestion, i.e., the speed across the vds became less than 40 mph for at least 5 minutes during the day. The congestion is required to observe capacity and estimate congestion speed \( w \). Each cell is then calibrated on the flow vs density scatter plot (Figure 3).

First, the free-flow speed \( v \) is estimated by performing a least-squares fit on the flow vs density data at the time instants where the speed was reported to be above 55 mph by the vds. Then the maximum observed flow is recorded and it is projected horizontally to the free-flow line previously determined by \( v \). This projected point constitutes the Capacity for this cell of the freeway and the corresponding density is labeled as the critical density where the freeway section hits its capacity and starts getting congested.

Finally, the congestion parameter \( w \) and the jam density are estimated. The main assumption in the calibration process is that the freeway sections rarely reach their ideal states and hit their actual capacity due to random external effects such as driver behavior, weather conditions, road work, incidents, etc. Thus, the calibration aims at the upper end of the distribution in terms of capacity and \( w \). This ensures that the simulations are performed with the best estimates of the actual capacities and congestion wave speeds, and hence, we can simulate the effects of different external factors which reduce the external capacities. For this reason, an approximate Quantile Regression [6] method is implemented rather than a least-squares fit on the congested flow vs density data. In this procedure, the flows corresponding to densities higher than the critical density are partitioned into bins so that each bin corresponds to one density value and contains 10 flows. Among these 10 flows, the maximum non-outlier is taken to represent the bin. This maximum non-outlier is determined as follows:

\[
Bin = \{f_1, f_2, ..., f_{10}\} \\
Bin Flow = \max_{f_i} (f_i | f_i \in Bin, \ f_i < Q_3 + 1.5 \times IQR)
\]

where \( f_1 \) through \( f_{10} \) and \( f_i \) are the flow values inside one such bin, \( Q_3 \) is the 75th percentile of the data points in the bin and IQR is defined as the difference between the 25th percentile and the 75th percentile of the data.

After all the Bin Flows are determined, a least-squares fit is performed on these data points. The regression line is constrained to pivot at the projected maximum flow point to complete the triangular Fundamental Diagram of the CTM. The final results of the calibration for the example vds is depicted below in (Figure 3). For those cells which do not function properly, nominal values are used. Also, in case a vds station does not record congestion, nominal values are used for Capacity and congestion wave speeds, while the calibrated value is used for free-flow speed.

3.3 Ramp flow imputation

On-ramps and off-ramps also contain vehicle detector stations, and the flow data obtained from them forms an essential input for the simulation of the freeway. However, these data are often
found to be missing or incorrect. TOPL has an elaborate procedure to impute missing and incorrect ramp data. The CTM model is utilized to determine the flow profile recorded in the vehicle detector station. The procedure has been extended from the repetitive control technique described in [7, 8]. It is assumed that the density and ramp flow profile is 24 hour periodic, and the on-ramp and off-ramp flows are represented as a convolution of a kernel on a constant periodic ramp parameter vector.

\[ r(k) = K_r(k)^T c_r, \quad s(k) = K_s(k)^T s_r, \quad (14) \]

where \( K_r(k) \) and \( K_s(k) \) represent a periodic, time dependent kernel vector. Some typical kernel functions include a unit-impulse or a gaussian window centered at time \( k \). The imputation procedure assumes and initial estimate for the ramp parameter vectors \( \hat{c}_r \) and \( \hat{c}_s \). These estimates are then adapted so that the model calculated densities match with the density profile recorded in the vehicle detector station. The input parameters for the imputation procedure are indicated in Figure 4. The procedure for the imputation is outlined below.

\[
\hat{f}_{up}(k) = \min(v_{up}n_{up}(k) - s_{up}(k), w(\hat{n}_I - \hat{n}(k)), F_{dn})
\]

\[
\hat{f}_{dn}(k) = \min(v_Rn_{dn}(k) - s_{dn}(k), w(\hat{n}_o - n_d(k)), F)
\]

\[
\hat{n}(k + 1) = n + (\hat{n}(k) + f_{up}(k) - f_{da}(k))
+ K_r(k)^T \hat{c}_r(k) - K_s(k)^T \hat{c}_s(k) - a(n(k) - \hat{n}(k))
\]

If ( Mode is FF and \( \hat{n}(k + 1) < 0 \) or
\[
\text{min}(w_{dn}(\hat{n}_d - n_{dn}(k)), F) \geq f_{dn}(k)) \) or
\[
\text{(Mode is C and min}(w_{dn}(\hat{n}_I - n_{dn}(k)), F) \geq f_{dn}(k)) \)

\[
\hat{n}(k + 1) = \frac{\hat{n}(k + 1)}{1 + K_r(k)F K_s(k)^T}
\]

\[
\hat{c}_r(k + 1) = \frac{\hat{c}_r(k) + F K_r(k) \hat{n}(k + 1)}{F}
\]

\[
\hat{c}_s(k + 1) = \frac{\hat{c}_s(k) - K_s(k)^T \hat{c}_r(k)}{F}
\]

The system is in Freeflow (FF) mode if \( \min(w_{dn}(\hat{n}_d - n_{dn}(k)), F) \geq v_R(k) - K_s(k)^T \hat{c}_s(k) \). Otherwise, it is in Congestion (C).

The parameter \( a \) is chosen so that the error equation is asymptotically stable. The adaptation procedure is carried out through the entire density profile multiple times, so as to reduce the “error” \( \sum |n_t(k) - \hat{n}_t(k)| \). This procedure is repeated until the error is acceptable. A similar procedure is adopted when only one of the ramp flows need to be imputed. Figure 5 shows an example of imputation performed for one of the cells in the I210W freeway. Finally, the ramp flows are now used to generate split ratios,

\[
\hat{f}_{dn}(k) = \frac{\hat{f}_{dn}(k)}{1 + K_r(k)F K_s(k)^T}
\]

\[
\hat{c}_r(k + 1) = \frac{\hat{c}_r(k) + F K_r(k) \hat{n}(k + 1)}{F}
\]

\[
\hat{c}_s(k + 1) = \frac{\hat{c}_s(k) - K_s(k)^T \hat{c}_r(k)}{F}
\]

which are usually time-varying.

### 3.4 Base case and Scenarios

At the end network specification, calibration and imputation, the CTM model is fully specified. The model is run for a ‘base case’, which simply means a particular day or several days for which good data are available. The model output is compared with actual field data in terms of (1) location of bottlenecks and speed.
contour plots, (2) hourly delays, (3) travel time, and other performance measures.

Several scenarios are specified. A scenario is created by specifying changes in the fundamental diagram and in the on-ramp demand. These include (1) increasing demand by (say) five percent relative to the base case; (2) modeling an incident in a particular cell by reducing the capacity of the cell by one or two lanes for a certain time period; (3) a demand management scheme that reduces on-ramp flows at some locations by a certain amount.

TOPL provides several ramp metering control laws, including ALINEA [9] and those described in [10]. Running each specified scenario with and without ramp metering in place provides an estimate of the benefits of ramp metering.

4 TOPL APPLICATION: I-210

This section is devoted to a TOPL case study of I-210W, a 26-mile long freeway in Pasadena, Southern California, shown in the map of Figure 6. The map shows the location of three major bottlenecks. The bottlenecks are visible in the speed contour plots for two days shown in Figure 7. Observe in both plots that congestion starts at a bottleneck and moves upstream, forming the characteristic low speed triangles, as predicted by Proposition 2. Figure 8 is a screen shot of a display window of the TOPL simulation of I-210W for the base case, after steps 1-3 of the TOPL procedure are executed. TOPL simulation require the boundaries to be in free-flow, so that flow boundary conditions may be specified for the simulations. Hence, the initial section of the freeway (which was in congestion), was disregarded in the simulation. Comparisons with empirical data (Figures 8 and 7) indicate that the estimated CTM model conforms reasonably well with measurement. For example, the contour plots clearly show the three major bottlenecks. Similarly, the delay in the morning peak is much larger than the afternoon peak, which is to be expected, because the morning commute direction is West, toward downtown Los Angeles. Also Figure 9 shows that the simulated performance measurements - Vehicle Miles Travelled (VMT), Vehicle Hours Travelled (VHT) and Delay (Hrs) compare well with the actual data. These comparisons lend confidence to the results of the scenario analysis considered next. Figure 10 shows the sim-
that this 5% increase in demand causes a 38.1% increase in travel time. Nevertheless, ramp metering can still keep the mainline free flowing as indicated in Figure 14, while limiting the increase in total travel time to 7.8%.

5 CONCLUSION

The macroscopic CTM model is much easier than microscopic models to calibrate and use to specify strategies to improve freeway operations and to evaluate their potential benefits. The calibrated CTM model for I-210W generates behavior that agrees closely with empirical measurement, including location of bottlenecks, propagation of congestion upstream from bottlenecks, hourly delay and travel time. The validity of the model is further confirmed by comparing its performance under a simulated accident with empirical measurements. This lends confidence to the model’s prediction of major reductions in delay by appropriate ramp metering.

TOPL can be used to provide benefit assessment for various ramp metering strategies. Critical sections can be identified for metering, and different algorithms can be tested for evaluating their efficacy. Scenarios like capacity (lane) expansions can also be
simulated, and their benefits assessed. Overall, TOPL can be effectively used as a critical tool for freeway operations planning.

ACKNOWLEDGMENT

This work is supported by the California Department of Transportation through the California PATH Program. The contents of this paper reflect the views of the author and not necessarily the official views or policy of the California Department of Transportation.

REFERENCES


