2014 American Control Conference (ACC)
June 4-6, 2014. Portland, Oregon, USA

Macroscopic Freeway Model Calibration with Partially Observed Data, a Case Study

Gunes Dervisoglu [1] and Alexander Kurzhanskiy [2] and Gabriel Gomes [3] and Roberto Horowitz [4]

Abstract— **In this paper, we present a case study on the macroscopic model calibration of the I-680 freeway in Northern California. This calibration effort posed two major challenges: 1) Only about half of the detectors on the mainline were functioning, and 2) no detection was available on the ramps. The only available ramp flow information were the Census counts on about 15 days back in 2009. The model calibrated based on sparsely available data from 2009 was subsequently used for projecting the state of the freeway in 2013. This is due to the fact that data quality significantly deteriorates after 2009 and almost no data is available after 2012. We demonstrate the algorithms developed to automatically calibrate a couple of macroscopic traffic flow models derived from the Cell Transmission Model. The algorithms estimate the unobserved input flows and replace faulty or missing mainline measurements. The unknown input estimation is achieved by an adaptive imputation scheme that uses the partial knowledge of the ramp flows as lower and upper bounds for the estimates. Whereas the fault diagnosis algorithm identifies bad measurements on the mainline and replaces them with profiles that agree with the model dynamics. Due to the inherent unobservability of the system in the absence of both mainline and ramp flow data, this heuristic substitution method is employed rather than a more rigorous state estimation scheme. The resulting calibrated model is able to reproduce the observed congestion patterns on the mainline for the 2009 data while keeping the estimated boundary flow inputs within the expected bounds to the extent where they don't violate the model dynamics. The model also performs well under projected conditions for 2013.**

I. Introduction

Macroscopic models of traffic flow, such as the Cell Transmission Model [5] adopted in this study, provide the basis for cost and time efficient simulation tools to evaluate traffic operation and planning strategies like ramp metering, infrastructure enhancement, demand management, etc. These models replicate the traffic behavior by expressing it in terms of aggregate quantities like flow, density and speed as functions of time and space. The main advantages of such models over their microscopic counterparts, which model every individual vehicle separately, is their ease of calibration due to their less stringent data collection requirements and their simulation speed due to their relative simplicity.

The model building exercise follows several well-defined steps [13]: 1) Model construction: the freeway to be modeled is represented as sequential segments (cells) where on-ramps and off-ramps are sources and sinks that feed flow into and out of the mainline. The sensors on the freeway are associated with the cells they are situated on. 2) Fundamental diagram calibration: The flow and density data in each sensor is treated with a regression technique [7] to estimate the segment-specific empirical flow-density relation, which is deemed the fundamental diagram of traffic flow, a concept first introduced by Greenshields' seminal work [10]. 3) Mainline sensor fault detection: The faulty mainline sensors are automatically flagged by the fault detection algorithm [8] and the appropriate fault handling needs to be carried through for each sensor before proceeding with the next step. 4) Unknown ramp flow imputation: When the measurements on the ramps are not available, they need to be estimated using an adaptive learning-based imputation scheme that learns the ramp flows by iteratively running the calibrated model and comparing to the mainline measurements [12]. 5) Simulation: The last step is to run an open loop simulation for a whole day and compare the results to the measurements in order to visually and quantitatively assess the quality of the calibrated model.

The data required for model building is gathered by inductive loop detectors embedded in the pavement on most California freeways. The Performance Measurement Systems (PeMS) [2] collects, processes and archives these data and makes them available for professional and academic use. In an ideal situation where all ramp flows are measured and mainline measurements are all available and fault-free, steps 3 and 4 in the procedure delineated above would not be required. However, this is practically never the case and the rest of this paper will demonstrate the performance of the algorithms developed for handling data imperfections for the model building process applied to a 34 mile portion of Interstate 680 in Northern California.

The paper is organized as follows: Section II provides

[1]Gunes Dervisoglu is a researcher with UC Berkeley PATH gunesder@berkeley.edu
[2]Alexander Kurzhanskiy is a researcher with UC Berkeley PATH akurzhan@eecs.berkeley.edu
[2]Gabriel Gomes is a researcher with UC Berkeley PATH gomes@path.berkeley.edu
[2]Roberto Horowitz is with Faculty of Mechanical Engineering UC Berkeley horowitz@me.berkeley.edu

an overview of the data that is used in this case study and shows the model specification step on the northbound Interstate 680. Section II shows the fundamental diagram fitting. Section IV briefly describes the fault detection and handling schemes. Section V explains the constrained imputation algorithm that comes up with the unknown demands into the model. Section VI demonstrates how the model is reconstructed after the artificial mainline measurements are injected where measurements are missing, and finally, section VII presents the results of the open loop simulation of the calibrated model in juxtaposition with the measurements from 2009 and projections for 2013.

II. MODEL SPECIFICATION AND DATA AVAILABILITY

The modeled section of the northbound Interstate 680 stretches from the Interstate 580 connection in the South up to the entrance of the Carquinez Bridge in the North. The model specification is done using the network editor tool developed as part of the TOPL (Tools for Operational Planning) project [1]. The specification of the model is done directly on the map by placing nodes and links where appropriate and making the corresponding connections to build a unidirectional freeway with onramps and offramps that serve as sources and sinks to the freeway, respectively. Between the dates of January 2009 and July 2013, there are a reported total of 84 mainline detectors and 0 ramp detectors in the PeMS repository in the specified stretch. However, only 65 are reported to be active during the year of 2009. Figure 1 shows the overall observation percentage for all the detectors in the range over the analyzed 3.5 year period. This figure also shows that the overall percentage of data collection is significantly higher for 2009. However, it should be noted that these values of around 75% are skewed due to the fact that the active number of detectors were reported as 65 during this year. So the given figure of 75% is only among the 65 reported and drops to 58% when all 84 physically existing detectors are considered. In fact, on a good day in 2009, about 15 detectors on average are flagged as not-working by PeMS and another 7 are flagged by our fault detection algorithm. Hence about 41 detectors out of a total of 84 are unavailable to us on a good day, and this is accompanied by no detection on the ramps.

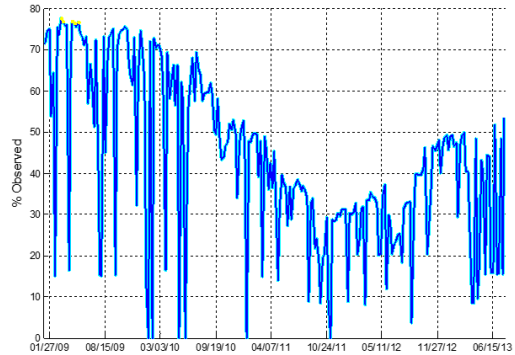

Fig. 1: Overall detector health over the 3.5 year period

In this study, we substitute the missing ramp flow information by the scarcely and sparsely available Census counts[1], which were obtained on several days in 2009 for the ramps on I-680. Since these flows are not necessarily on the same day for all the ramps, they were used to calculate upper and lower bounds for the learning algorithm, rather than definite ramp flow knowledge. Figure 2 shows an example ramp with individual flow contours for each day and the calculated upper and lower bounds. Note that these counts were reported every hour within the day they were taken on.

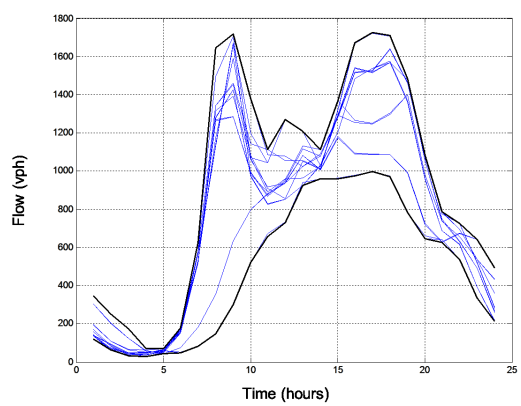

Fig. 2: Census ramp flow counts with calculated upper and lower bounds, example ramp: offramp to Treat Blvd

The Cell Transmission model designates the density of each cell to be the state of the system and the flow between cells are determined by the send-and-receive relationship dictated by the fundamental diagram parameters, the calibration of which are shown in the next section. To be specific, the adopted model is the so-called Link-Node Cell Transmission Model (LN-CTM),

[1]Vehicle counts are taken by the state Department of Transportation on all state highways every three years. Counts are also taken for ramps and local cross-streets.

which places nodes at the flow exchange interfaces between cells. These nodes have nxm split ratio matrices associated with them, where n is the number of input links to the node and m is the number of output links. These split ratio matrices specify what portion of flow in an input link go to which output link. Hence, the entries of these matrices are between 0 and 1 and the rows add up to 1. The terms link and cell are used interchangeably throughout the rest of this paper unless otherwise specified. A schematic representation of the LN-CTM is given in Figure 3.

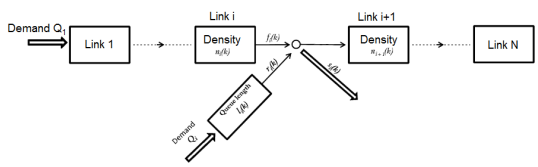

Fig. 3: The Link-Node Cell Transmission Model

The model equations are given in Eqn. (1), where n_i is the density in link i and f_i is the flow leaving link i. F_i, v_i and w_i are capacity, free flow speed and congestion wave propagation speed of each link and together they constitute the fundamental diagram of the link. The metrics are normalized per lane and per mile. Hence the densities are in vehicles/mile and the flows are in vehicles/hour/lane and the geometry of links are captured in the normalized fundamental diagram parameters, which are the topic of the next section.

$$\begin{aligned} n_i(k+1) &= n_i(k) + f_{i-1}(k) - f_i(k) \\ f_i(k) &= min(Demand_i(k), Supply_{i+1}(k)) \\ Demand_i(k) &= min(F_i, v_i n_i(k)) \\ Supply_{i+1}(k) &= min(F_{i+1}, w_{i+1}(n^J_{i+1} - n_{i+1}(k))) \end{aligned} \tag{1}$$

III. FUNDAMENTAL DIAGRAM CALIBRATION

Upon completion of the previous step, the freeway is represented as successive links and some of these links have detectors associated with them. The fundamental diagram relation between flow and density is found using a regression scheme described in [7]. A resulting example fundamental diagram is shown in Figure 4. Here, vf stands for free flow velocity and corresponds to the slope of the left hand side line, whereas w is the congestion wave propagation speed and corresponds to the slope of the right hand side line. Rhojam corresponds to n^J and specifies the jam density, i.e. the density value where the congestion line intersects the horizontal axis. And finally, qmax stands for the capacity. The units are as specified in the graph.

It is almost guaranteed that there will be more links than there are sensors for the specified freeway. Therefore, the fundamental diagrams that are calibrated for each sensor need to be assigned to nearby links so that every link has a fundamental diagram defined for it. A downstream assignment scheme is adopted for this. Figure 5 demonstrates this assignment method schematically. In this scheme, each individual sensor's fundamental diagram is assigned to the link the detector is in, and then it is assigned to the links downstream until another sensor is encountered. The justification for this assignment scheme lies in the fact that during free flow we do not expect any significant difference in the normalized density between the measurement location and inside of the cell whereas when the cell is congested, congestion spills back in the upstream direction, which is better captured by the detector upstream of the cell.

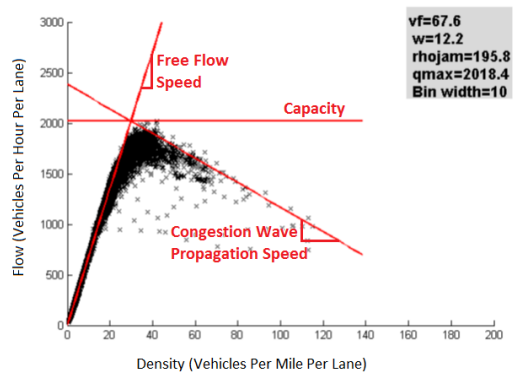

Fig. 4: An Example Fitted Fundamental Diagram for Detector ID 400450

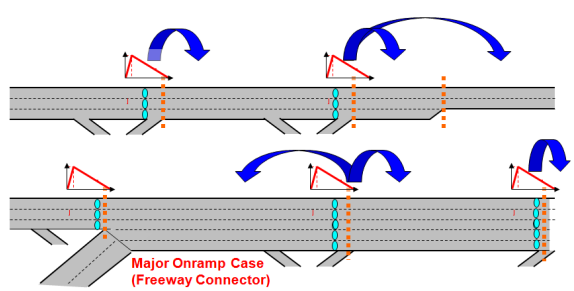

Fig. 5: Downstream assignment of sensor fundamental diagrams

IV. MAINLINE SENSOR FAULT DETECTION AND HANDLING

Once the unidirectional freeway model is specified as in the previous section, the first step is to ensure data health. PeMS readily implements a statistical method

[4] to flag aberrant data in the loops and specifies an observation percentage for each detector station, as shown in figures 2 and 3. However, this scheme cannot detect more structural problems that are statistically consistent, such as deficient lanes, a systematic bias in measurements due to incorrect hardware settings or the position of the sensor with respect to a nearby ramp.

One of the most common problems is the mismatch between sensor-reported number of lanes and the physical geometry of the freeway. There are two cases: 1) the detector station has more loops than existing lanes: This is usually the case when excessive sensors are included to the detector station from neighboring ramps and/or from the opposite freeway direction. 2) The detector station has less loops than existing lanes: this is often the case when there are auxiliary lanes which lack detector coverage. The detection of such lane mismatches are trivial, because the network geometry is specified on the satellite map and sensors are associated with links. If the lanes of the sensor and the lanes of the link it is associated with do not match, a warning is given and the necessary modification needs to be applied to the data. For lane deficient detection stations, either an average of the existing measurements or a vector of zeroes are appended as the additional lanes measurements. It has been empirically observed that for auxiliary lanes, the zero flow approach yields more accurate results rather than appending the average. For lane abundant detectors, the data needs to be inspected and the lane that corresponds to the outlying flow profile needs to be eliminated from the data. This was done by visual inspection of the data for this study but it can be automated for future practices if it proves to be a common occurrence.

The automatic fault detection methodology is a model-based scheme described in detail in [8] and [6]. It learns the missing ramp flows and runs the model with the learnt demands to establish residuals between measurements and simulation. A decision logic evaluates these residuals and determines the detectors that are inconsistent with respect to the model. The block diagram of the algorithm is given in figure 8. This scheme is based on a simplified CTM called the ACTM which has been introduced to model unidirectional freeways [9]. This model makes the assumption that onramp flows enter the mainline without any restriction from the downstream link, i.e. the downstream supply constraint of equation 1 is allowed to be violated for onramps. This simplification allows for the decoupling of cells in terms of unknown ramp flow imputation and hence, the fault signatures are also isolated to the location of the faulty detector. This decoupling is not possible with the LN-CTM as will be shown in the next section. Here, we note that this fault detection algorithm has been developed for the cases where ramp data is either completely available or entirely lacking. In this study, the unknown ramp flow imputation algorithm has been extended to incorporate the partial information about the ramp flows. However, this extension does not readily apply to the fault detection scheme. Therefore we apply the algorithm unchanged, as devised in [8] and [6].

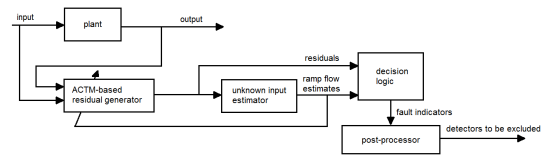

Fig. 6: Block Diagram of the Fault Diagnosis Scheme

The fault handling is achieved by the simple exclusion of the bad detectors. Obviously, this has ramifications as far as the model accuracy is concerned. Irrespective of how many links are created on the freeway during model specification, the model effectively has as many cells (collections of links) as there are detectors. Therefore, the simplification of the model due to omission of detectors eventually needs to be reversed. This is not a straight-forward task and is described in section 6. For the days that have been analyzed, the fault detection algorithm yielded 6 faulty mainline detectors on average in addition to the 15 reported by PeMS. So in the end, there remain 44 healthy mainline detectors out of 84 total to work with.

V. UNKNOWN RAMP FLOW IMPUTATION USING CONSTRAINTS

Once the fundamental diagram parameters are defined and the faulty detectors are removed, the next step is to estimate the unknown ramp flows. The so-called imputation algorithm that iteratively learns the ramp flows is described in [15] and [14] for the LN-CTM and ACTM models respectively. The algorithms are based on adaptive repetitive learning [11] and adaptive iterative learning paradigms [16], [3], respectively. Here, we provide a brief overview of the imputation algorithm for the LN-CTM based imputation. To present the algorithm we need to introduce the following definitions:

- Capacity Adjusted Free Flow Speed:
  $\bar{v}(k) = min(v, \frac{F}{n(k)})$
- Capacity Adjusted Congestion Wave Speed:
  $\bar{w}(k) = min(w, \frac{F}{n^J - n(k)})$
- Total Effective Demand:
  $c(k) = n(k)\bar{v}(k)(1 - \beta(k)) + d(k)$

, where definitions of individual variables and parameters are given in Table I. With these added definitions,

the demand and supply functions in Equation (1) can be written as

$$D_i(k) = min(n_i(k)v_i, F_i) = n_i(k)\bar{v}_i(k)$$
$$S_i(k) = \bar{w}_i(k)(n_i^J - n_i(k)) \quad (2)$$

And the model update equations are given by:

- Density Update:

$$n_i(k+1) = n_i(k) + f_i^{in}(k) - f_i^{out}(k) \quad (3)$$

- Flow Update:

$$f_i^{in}(k) = min(\bar{w}_i(k)(n_i^J - n_i(k)), c_{i-1}(k)) \quad (4)$$

$$f_i^{out}(k) = n_i(k)\bar{v}_i(k)min(1, \frac{\bar{w}_{i+1}(k)(n_{i+1}^J - n_{i+1}(k))}{c_i(k)}) \quad (5)$$

$$c_i(k) = n_i(k)\bar{v}_i(k)(1 - \beta_i(k)) + d_i(k) \quad (6)$$

$$d_i(k+1) = d_i(k) + Q_i(k) - r_i(k) \quad (7)$$

$$s_i(k) = \beta_i(k)f_i^{out}(k) \quad (8)$$

$$r_i(k) = \frac{1}{c_i(k)d_i(k)}min(c_i(k), \bar{w}_{i+1}(k)(n_(i+1)^J - n_{i+1}(k))) \quad (9)$$

Symbol	Name	Unit
F_i	flow capacity of Link i	veh/period
v_i	free flow speed of Link i	section/period
w_i	congestion wave speed of Link i	section/period
n_i^J	jam density of Link i	veh/section
k	simulation time step (period)	dimensionless
$\beta_i(k)$	split ratio at node i	dimensionless
$f_i(k)$	flow out of Link i	veh/period
$n_i(k)$	number of vehicles (vehicle density) in Link i	veh/period
$s_i(k), r_i(k)$	off-ramp, on-ramp flow in node i	veh/period
$d_i(k)$	on-ramp i demand	veh/period
$l_i(k)$	queue length on on-ramp i	veh/period
$c_i(k)$	effective demand into Link i	veh/period
C_i	flow capacity for on-ramp i	veh/period
L_i	queue capacity for on-ramp i	veh/period
$Q_i(k)$	input flow for on-ramp i	veh/period

TABLE I: LN-CTM variables and parameters

With the above formulation, the imputation algorithm solves a two-step problem:

1) Estimate the effective demand vectors $c_i(k)$ for each link on the freeway using an adaptive learning scheme to match the density measurements on the mainline
2) Using the estimated effective demands, evaluate incoming and outgoing flows to and from each link. Using these flows and the measured flows (usually measured at the location depicted in Figure 7, solve a linear program to determine estimated on-ramp flows r and off-ramp split ratios β.

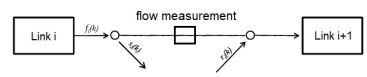

Fig. 7: Measurement Location with respect to the LN-CTM

As shown in section II, Figure 2, only a few days worth of hourly counts are available for the ramp flows. These are translated into upper and lower bounds of flow which need to be further transformed into split ratio and demand profiles so that they can be imposed in the learning step of the imputation algorithm, because it is based on learning the demands rather than the flows that the demands induce on the freeway. During free flow conditions, this mapping is easy since all of the onramp demand is available to enter the freeway. Hence, the bounds on demands and split ratios are readily defined by:

$$d_i^{upper}(k) = r_i^{upper}(k)$$
$$d_i^{lower}(k) = r_i^{lower}(k)$$
$$\beta_i^{upper}(k) = \frac{s_i^{upper}(k)}{n_i(k)\bar{v}_i(k)}$$
$$\beta_i^{lower}(k) = \frac{s_i^{lower}(k)}{n_i(k)\bar{v}_i(k)} \quad (10)$$

During congestion however, the following optimization problem needs to be solved to find the demands d and split ratios β that correspond to the given upper and lower bounds of flow through the ramps.

$$min(|w_{i+1}(n_{i+1}^J - n_{i+1})d_i - (n_i v_i(1-\beta_i) + d_i)r| + |w_{i+1}(n_{i+1}^J - n_{i+1})n_i v_i \beta_i - (n_i v_i(1-\beta_i) + d_i)s|) \quad (11)$$

,where the $|.|$ operator is the l1-norm, rendering the problem an LP. In the above formulation, note that $w_{i+1}(n_{i+1}^J - n_{i+1})$ is the downstream supply at a given cell i+1 and $(n_i v_i(1-\beta_i) + d_i)$ is the corresponding upstream demand. Hence, the result of the above problem yields d and β values that minimize the difference between model-reproduced and given ramp flows, weighted by the corresponding demands and supplies at the appropriate node.

Once the corresponding upper and lower bounds for demands and split ratios are established, they are fed to

the first step of the imputation algorithm that learns the total effective demands entering each cell in the model. The lower and upper bounds are used to establish upper and lower bounds on the total effective demands and these are imposed at the end of each iteration on top of the learned value by snapping it to the nearest bound if the value lies outside of it.

In the second step of the imputation algorithm, the actual flow values are imposed again instead of the derived demand and split ratio bounds. Looking at figure 7, it can be seen that this problem has enough degrees of freedom to satisfy the constraints when both the onramp and the offramp exist at the node. However, when this is not the case, the constraints on a single ramp node may cause infeasibility. In such cases, we allow the ramp flow profile to violate the bounds at those time instances of infeasibility in order to keep accordance with the model dynamics. Figure 8 demonstrates an example of this. Note that the ramp flow profile (blue solid) that is estimated stays within the given bounds (red dashed) most of the time, except for a few periods when it violates the bounds.

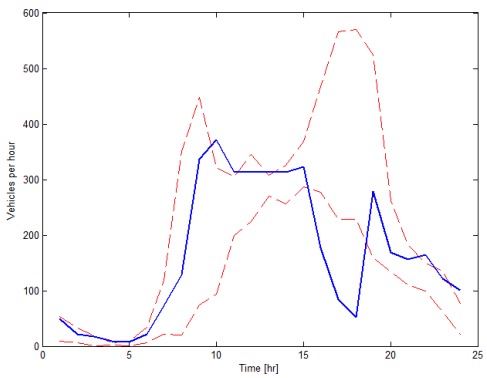

Fig. 8: Example imputed offramp flow with upper and lower bounds

Obviously, imposing these constraints increases model errors on the mainline but it dramatically increases the accuracy of the estimated ramp flows, significantly more so than the deterioration in mainline observation matching performance. Hence, it can be argued that the constrained ramp flow imputation provides an overall enhancement in the model accuracy, but not without conceding the fact that this argument is based on the visual assessment of the plots presented in the results section rather than a quantifiable comparison.

VI. MODEL RECONSTRUCTION

In section IV, it was mentioned faulty detectors are discarded from analysis. This usually leads to a topological downgrading of the model accuracy since, in effect, the number of available sensors determine the number of cells in the model. Even when eventually the cells can be broken down into smaller links, links that belong to a particular cell end up having identical parameters.

From the perspective of the imputation algorithm, the removal of cells causes the creation of the so-called mega-cells that may contain multiple onramps and off-ramps. Figure 9 shows an example of such a mega-cell being created due to the removal of a detector at a certain point. Besides the fact that this violating the model assumptions, it is also impossible for the imputation algorithm to discern between two onramp flows, when its not even possible in some cases to find a unique solution to the problem with only one onramp and one off-ramp. In the end result, however, we want a model that topologically agrees with the actual freeway geometry so the estimated ramp flows that enter and leave the mega-cells that have multiple ramps need to be separated into individual ramp flows.

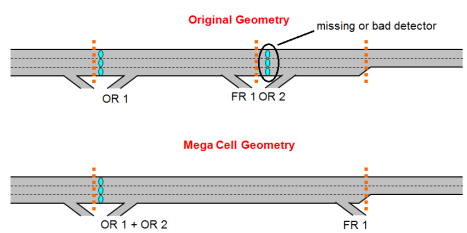

Fig. 9: Mega-cell creation example

This is a much more delicate problem than can be handled with an arbitrary distribution of the flow values to the available ramps because excessive flows at a particular ramp can disrupt the model dynamics dramatically. Therefore, an iterative mega-cell splitting algorithm has been developed that makes use of the imputation algorithm at each iteration to ensure the new individual ramp flows agree with the model dynamics. This algorithm is presented more precisely in [6].

The algorithm aims at finding artificial measurements in the location of the missing or removed detector by solving the linear program in Equation (12). The variables in the equation are depicted in figure 11, where the red star shows the location of the missing sensor, i.e. cells 1 and 2 constitute the mega-cell in this particular picture. In the cost function, the pseudo-densities are aimed to track realistic density profiles provided by the spatially closest sensor. The result of this LP is density and flow profiles at the bad detector location on the freeway that matches the flow at the boundaries of the mega-cell. In practice, however, the exact matching of these flows are virtually impossible. Therefore, that constraint is relaxed by including the difference between

flows in the cost function of the optimization problem. This way, the flows at the cell boundaries are kept close to the simulated flows before the splitting of the mega-cell and this ensures that the model performance does not deteriorate outside of the mega-cell boundaries.

$$\begin{array}{ll} \underset{r_1,s_1,r_2,s_2,n_1^P,n_2^P,f_m id}{\text{minimize}} & |n_1^P - n_1| + |n_2^P - n_3| \\ \text{subject to} & \text{model dynamics} \end{array} \tag{12}$$

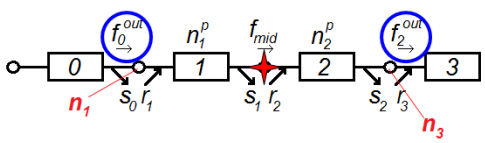

Fig. 10: Variables in the Mega-cell Splitting LP formulation

Once these artificial mainline measurements are found, the imputation algorithm is run with this new additional knowledge of the mainline to estimate the individual ramp flows. It must be noted that even when all sensors are healthy, mega-cells can still be created due to bad sensor placement. Therefore, this algorithm becomes an important extension to the imputation algorithm of section V.

VII. PROJECTION AND SIMULATION RESULTS

The challenge of the Interstate 680 study was that the final simulation model had to reflect the state of traffic in 2013, for which no consistent measurements existed. There were two main reasons for that: 1) significant construction was performed in the last 3 years, and the majority of detectors were broken or needed re-wiring; and 2) new lanes appeared at certain stretches of freeway, and as a result of lane marking changes some loops ended up between two lanes, which adversely affected their counts. At the same time, there were several scattered sources of different measurements, collected on I-680 over the past two years. These included:

1) Census flow counts for some on- and off-ramps collected in 2012. The ramps covered by 2012 Census measurements were a subset of those having 2009 Census counts.
2) Several drivers were continuously driving in circles over the studied segment of I-680 on one workday afternoon in March and one workday afternoon in April 2013. This provided hints about the afternoon dynamics of speed on I-680.
3) At two places with overpasses 14-hour (from 6 am to 8 pm) videos were taken on two different days. From these videos, vehicle flows could be inferred.
4) Finally, some of the loop detectors that were healthy in 2009 remained intact in 2013.

Items 3 and 4 allowed us to come up with the growth factor that could be applied to 2009 demand to obtain the demand of 2013.

To obtain the simulation model of I-680 for 2013 projection, we followed the steps below:

1) Generate the desired 5-minute speed profiles for all the freeway links using individual traces obtained by drivers as well as the video footage from two overpasses. The result is the 2013 Projection speed contour in Figure 11.
2) Apply 2009-2013 demand growth factor to the density data on freeway links. We cannot blindly increase the demand flows by this growth factor and run the simulation, because split ratios have changed as well. The result is the 2013 Projection density contour in Figure 11.
3) Introduce new flow profiles to ramps with existing 2012 Census counts. This results in the relaxation of min/max bounds for those ramps allowing wider ranges of possible flows.
4) Compute flow profiles for all the freeway links: multiply densities by speeds and impose CTM restrictions (1) fundamental diagrams remain unchanged after calibrating them with 2009 data. The result is the 2013 Projection flow contour in figure 11.
5) Consider all the mainline detectors healthy, each detector receiving flow and density profiles from the link it is snapped to.
6) Run imputation algorithm for the projected 2013 mainline data using relaxed ramp flow constraints and demand flow and split ratio profiles. This concludes the creation of 2013 simulation model.

Simulation results for I-680 North are shown in the right part of Figure 11. These are heat maps showing the corresponding state of the freeway at each cell (columns) at each time step of the day (rows). In the plots, traffic flows from left to right and the time of day increases along the vertical axis. In Figure 12 hourly Vehicle Miles Traveled (VMT) and Vehicle Hours Traveled (VHT) computed from the 2013 projected data are compared with those computed by the simulation. VMT is the measure of demand, and VHT is the measure of delay.

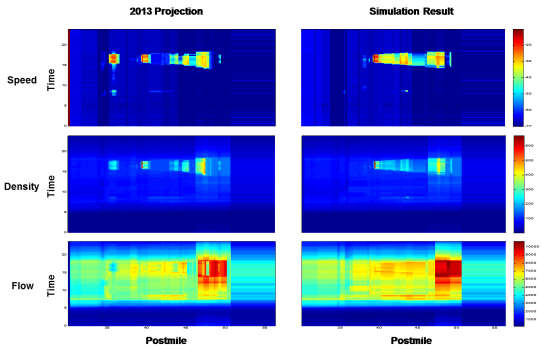

Fig. 11: 2013 I-680 North mainline data projection (left) and the corresponding simulation result (right).

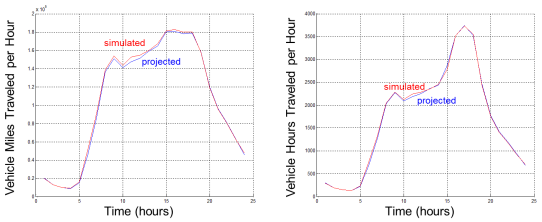

Fig. 12: Comparison of hourly Vehicle Miles (left) and Vehicle Hours (right) Traveled computed from the projected 2013 data (blue) and the simulation (red).

VIII. Conclusion

Given a traffic model, in our case LN-CTM, the on-ramp demand specified by the corresponding flows and the off-ramp demand specified by split ratios, the forward problem in classical sense is to find how the system state (density in the freeway cells) evolves in time. In this paper we solved the inverse problem: given the traffic state measured on the freeway, find on-ramp flows and off-ramp split ratios. This problem was motivated by the absence of ramp flow measurements on many freeways in California. It was solved in application to the Interstate 680 in Northern California, as part of the Corridor System Management Project conducted by Caltrans. The challenges of this project were that (1) freeway mainline measurements were sparse, creating gaps with one or more on- and off-ramps between measurement points; (2) even the officially healthy detectors sometimes malfunctioned; (3) the resulting ramp flows were supposed to lie within prescribed bounds; and (4) the available measurement data were collected in 2009, whereas we had to model 2013 traffic conditions. The first challenge affected the calibration of fundamental diagrams and their assignment to cells. The second challenge prompted the need for the fault detection mechanism. The configuration of mega-cells depends on location of non-faulty detectors. To address the third challenge, the original ramp flow imputation algorithm was modified to incorporate the restricting flow bounds. Finally, the initial 2009 freeway model was adjusted to reflect 2013 traffic conditions that could not be directly measured due to poor detection, but were reconstructed through diverse sources, such as probe vehicles and video.

IX. Acknowledgment

This work is supported by the California Department of Transportation (Caltrans) under the Connected Corridors program and by the National Science Foundation (NSF) through grant CDI- 0941326.

References

[1] *http://gateway.path.berkeley.edu/topl/*, 2013.
[2] *http://pems.ca.dot.gov*, 2013.
[3] S.K. Arimoto and F. Miyazaki. Bettering operation of robots by learning. *Journal of Robotic Systems*, 1984.
[4] C. Chen, J. Kwon, J. Rice, A. Skabardonis, and P. Varaiya. Detecting errors and imputing missing data for single loop surveillance systems. *Transportation Research Record*, 1855, 2003.
[5] C. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research. Part B*, 28(04), 1994.
[6] G. Dervisoglu. *Automatic Calibration of Freeway Models with Model Based Sensor Fault Detection*. PhD thesis, University of California, Berkeley, CA., 2012.
[7] G. Dervisoglu, G. Gomes, J. Kwon, R. Horowitz, and P. Varaiya. Automatic calibration of the fundamental diagram and empirical observations on capacity. In *TRB 88th Annual Meeting*, 2009.
[8] G. Dervisoglu and R. Horowitz. Model based fault detection of freeway traffic sensors. In *Proceedings of the Dynamic Systems and Control Conference*, 2011.
[9] G. Gomes and R. Horowitz. Optimal freeway ramp metering using the asymmetric cell transmission model. *Transportation Research, Part C*, 14(4):244–262, 2006.
[10] B. Greenshields. A study of traffic capacity. In *Proceedings of the 14th Annual meeting of the Highway Research Board*, 1934.
[11] W. Messner, R. Horowitz, W-W. Kao, and M. Boals. A new adaptive learning rule. *Journal of Robotic Systems*, 1984.
[12] A. Muralidharan. *Tools for Modeling and Control of Freeway Networks*. PhD thesis, University of California, Berkeley, CA., 2012.
[13] A. Muralidharan, G. Dervisoglu, and R. Horowitz. Freeway traffic flow simulation using the link node cell transmission model. In *Proceedings of the American Control Conference*, 2009.
[14] A. Muralidharan and R. Horowitz. Imputation of ramp flow data for freeway traffic simulation. *Transportation Research Record*, 2099, 2009.
[15] A. Muralidharan and R. Horowitz. Analysis of adaptive learning algorithm for freeway ramp flow imputation. In *Proceedings of the Dynamic Systems and Control Conference*, 2011.
[16] J. Xu and Y. Tan. In *Linear and Nonlinear Iterative Learning Control*. Springer-Verlag, 2003.