Identifying Chronic Splashover Errors at Freeway Loop Detectors

Ho Lee, PhD Candidate
Graduate Research Associate
Department of Civil, Environmental, and Geodetic Engineering,
The Ohio State University,
Columbus, OH 43210
E-mail: lee.2406@osu.edu

Benjamin Coifman, PhD
Associate Professor
The Ohio State University
Joint appointment with the Department of Civil, Environmental, and Geodetic Engineering, and the Department of Electrical and Computer Engineering
Hitchcock Hall 470
2070 Neil Ave, Columbus, OH 43210
Phone: (614) 292-4282
E-mail: Coifman.1@osu.edu

Abstract

Loop detectors are the most commonly used vehicle detectors for freeway management. There has been considerable research to screen the quality of loop detector data using macroscopic measurements (e.g., five min aggregate data) and more recently using microscopic data (i.e., individual vehicle actuations). However, some significant detector errors have not received much attention due to the difficulty of identifying their occurrence. This paper examines one such error, splashover: the erroneous detection in one lane of a vehicle from an adjacent lane. If not caught, the splashover events will cause the flow and occupancy measurements to be too high. We examine the nature of splashover and develop an algorithm to find detectors exhibiting chronic splashover problems.

At the crux of this algorithm, an erroneous pulse arising from splashover in one lane should usually be bounded by the valid pulse from the vehicle in its lane of travel. The algorithm calculates the rate of suspected splashover: the percentage of pulses in the subject lane that are bounded by a pulse in an adjacent lane. However, a difficulty arises because any given splashover event in the data stream is usually indistinguishable from the non-splashover event of two vehicles passing the detector station at the same time yielding valid concurrent actuations. To control for non-splashover events we calculate a dynamic threshold rate of false positives as a function of the observed traffic conditions. Recognizing the fact that chronic splashover is non-transient, for the reasons explained in the paper, the algorithm is intended to be applied during free flow periods since these conditions are opportune.

Concurrent loop detector data and video ground truth are used for the evaluation from a total of 21 directional data sets, covering 94 adjacent lane pairs from 68 loop detectors at 15 loop detector stations, both with and without actual splashover. The algorithm exhibited good performance. Finally, with only minor modifications, the algorithm should also be applicable to other detector technologies, e.g., side-fire microwave radar sensors.

1. Introduction

Loop detectors are the most commonly used vehicle detector for freeway management. They are effectively metal detectors embedded in the pavement. A typical loop detector station will have either one or two loop detectors per lane, i.e., single or dual loop detectors, respectively. Conventional single loop detectors can measure flow (the number of vehicles per unit time) and occupancy (the percent time the detector is occupied). From these measurements one can estimate average speed, but not measure it. Dual loop detectors overcome the speed estimation problem; they can measure individual vehicle speeds from the difference in arrival times at the two loops and the known loop separation, and then calculate the average speed.

Loop detector data are used for applications such as ramp metering (e.g., Papageorgiou et al., 1997; Hourdakis and Michalopoulos, 2002), incident detection (e.g., Payne and Tignor, 1978; Payne and Thompson, 1997; Williams and Guin, 2007), travel time prediction (e.g., Kwon et al., 2000; Coifman and Krishnamurthy, 2007), and vehicle classification (e.g., FHWA, 2001; Coifman and Kim, 2009). The performance of such applications greatly depends on the accuracy of the detector data, but data collected from loop detectors are prone to various errors caused by hardware and software problems. Detector errors degrade the quality of detector data, and the impact of these errors will propagate to subsequent measurements such as flow, occupancy, and speed from the loop detectors. In the end, data incorporating detector errors could affect the traffic control decisions and traveler information based on the detector's data.

In an effort to improve detector data quality, this paper examines splashover: the erroneous detection in one lane of a vehicle from an adjacent lane. We examine the nature of splashover and develop an algorithm to automatically find detectors exhibiting chronic splashover problems.[1] A loop detector consists of a physical loop of wire embedded in the pavement connected to a sensor located in a

[1] Note that for this paper we assume that adjacent lane cross-talk is functionally the same as splashover in terms of the reported data, though the underlying electronic fault and subsequent resolution may be different. The algorithm only seeks to catch the presence of the over-counting error, it does not differentiate between possible sources of the error, e.g., misaligned detection zone, an over-sensitive detector, or cross-talk.

nearby cabinet. The given loop detector installation in the pavement will impact how responsive the detector is to passing vehicles. The sensor detects the presence or absence of vehicles over the loop and typically allows a user to manually select the sensitivity level of operation to accommodate for a wide range of responsiveness from the physical loop. A loop detector will yield poor performance if the sensitivity is not set correctly for the responsiveness. In conventional practice, however, it is difficult to know the physical loop's responsiveness, which makes selecting the appropriate sensitivity level difficult. According to the Traffic Detector Handbook (Klein et al., 2006) splashover usually occurs when the sensitivity level of a loop detector is set too high or a loop is too close to an adjacent lane, though there are many other factors that could cause splashover as well. While the handbook offers some advice for fixing this problem, it does not offer characteristics of detector data with splashover or a means of detecting the presence or absence of splashover in the detector data.

The present work only seeks to detect the presence of splashover, the algorithm developed herein does not attempt to diagnose the origin of the splashover problem. When splashover is found, the underlying hardware or software error needs to be fixed. Often the necessary fix could be as simple as adjusting the loop sensor card's sensitivity setting; however, if an immediate fix is not feasible, one must recognize that the data contain errors.

Consider the hypothetical example of a splashover event on a two-lane freeway in Fig. 1. The time-space diagram shows a vehicle passing in lane 2. A schematic of the roadway is shown coincident with the distance axis. For this vehicle the detection zone in the lane of travel, DZ_2, is larger than the detection zone in the splashover lane, DZ_1^X (we use superscript "X" to denote the fact that this detection zone is erroneous and, ideally, should not respond to a lane 2 vehicle). The resulting pulses from both lanes are shown below the time-space diagram, coincident with the time axis. The loop detector in lane 1 erroneously responded to this lane 2 vehicle for an on-time denoted OnT_1 while the valid actuation in lane 2 had an on-time of OnT_2. In general, the resulting double count can lead to inaccurate measurements, e.g., higher flow and occupancy. A loop detector's sensitivity is inversely proportional to the square of the distance to the vehicle (Klein et al., 2006). So typically DZ_1^X for the splashover event should be smaller

than DZ_2 for the valid vehicle actuation. Thus, a pulse arising from splashover in one lane should usually be bounded by the valid pulse from the vehicle in its lane of travel, e.g., the pulses at the bottom of Fig. 1 and call these *suspected splashover events*. However, as will be discussed shortly, a difficulty arises because any given splashover event in the data stream is usually indistinguishable from the non-splashover event of two vehicles passing the detector station at the same time, yielding *valid concurrent actuations*.

"Place Fig.1 about here"

While there have been numerous efforts to automatically identify detector errors, splashover has not received much attention due to the difficulty of differentiating the error from valid concurrent actuations. Reviewing the literature to place our work in context, the general body of detector error tests have been implemented both at the macroscopic and microscopic levels. Macroscopic tests embody the formalization of heuristics to check average measurements from a given sample period against statistical tolerance, while microscopic tests examine the individual vehicle actuations (i.e., the detector "pulses"), when the detector turns "on" and "off" for each vehicle that passes over a loop detector.

The macroscopic tests are more common because conventional traffic monitoring practice discards the microscopic data within the controller after aggregation to macroscopic metrics, e.g., flow, occupancy and average speed. Jacobson et al. (1990) introduced a macroscopic test for setting limits for acceptable values of flow for any given occupancy on the basis of plausible ratios between flow and occupancy within specific occupancy ranges. The examples showed the algorithm was effective for detecting some intermittent failures by the loop detectors, e.g., short pulses (attributed to "hanging–off" by the authors). Cleghorn et al. (1991) presented several screening methods using macroscopic measurements. They claimed to have obtained a tighter upper boundary from feasible flow-occupancy pairs than Jacobson et al. to screen data from a single loop detector. Cleghorn et al. presented additional screening for dual loop detectors that compares the measurements between the upstream and downstream loop detectors as well as expanding the two dimensional flow-occupancy feasible region to a calibrated

three-dimensional speed-flow-occupancy "acceptable region". There has been a renewed interest in macroscopic tests arising from several recent deployments of archived data systems that collect the real time data for performance monitoring and analysis. Two examples of such work are Turochy and Smith (2000) and Chen et al. (2003). Turochy and Smith (2000) developed an integrated data-screening procedure including a critical threshold value of measurements such as occupancy and flow, and tests utilizing the relationships between speed, flow, and occupancy. Among those tests included in the procedure, a maximum hourly flow threshold test (e.g., 3,100 vehicles/lane/hr) was used to catch detector errors causing unreasonably high flow. Chen et al. (2003) developed a macroscopic error detection test using the time series of flow and occupancy measurements. Statistics computed over a whole day at each detector are used to differentiate between a "bad" or "good" detector with respect to various loop detector malfunctions, e.g., sticking or hanging. The algorithm did not find chattering or pulse breakup (i.e., a single pulse expected from a vehicle is separated into two or more pulses if the detector "drops out" in the middle), but they suggested that additional constraint such as checking for consistently high flow should be useful to detect loop detectors with these errors.

At the microscopic level, Chen and May (1987) may have been the first to use individual vehicle actuations to assess detector data quality. They examined the ratio of a detector's average on-time to the average on-time of all detectors at the detector station. This on-time ratio test provided an indication of detector status, e.g., the sensitivity setting. Chen and May found pulse breakups and they surmised that the breakups might be caused by the loop detector's sensitivity setting being too low. In addition, they found unexpected detector actuations in their data reportedly due to lane change maneuvers over the loop detectors, splashover, and phantom actuations that did not appear to be due to any vehicles. Coifman (1999) presented a microscopic method utilizing the redundancy of a pair of loops to assess the performance of dual loop detectors and to identify detector errors; namely, that during free flow conditions the on-times for a given vehicle from each detector in the pair of loops should be virtually identical regardless of vehicle length. The method detected a chronic hanging-on problem and cross-talk problems. Coifman and Dhoorjaty (2004) developed eight detector validation tests using microscopic data

to identify various errors both at single and dual loop detectors. That work specifically classified errors into seven groups: either the rising or the falling edge being premature or delayed, pulse breakup, missed vehicle, and wrong detection. Lee and Coifman (2012a) presented the daily median on-time test as a measure of performance for single loop detectors. The test calculates median on-time over a day. Assuming most vehicles are free flowing passenger cars, this daily median on-time should fall within a small range. The test indirectly detects an inappropriate loop detector sensitivity setting. Cheevarunothai et al. (2007) developed an algorithm to improve the quality of dual-loop truck data so as to identify and correct detector problems such as pulse breakups, cross-talk, and the difference of sensitivity between the pair of loops in a dual loop detector. Lang and Coifman (2006) employed microscopic data in a different way. They used the temporal relationships between different loop detectors at a given station to correct for discrepancies between the assumed map from freeway lane to input port on the controller. Their method compared each detector at the station to all of the other detectors, it automatically identified whether a given pair of loop detectors were in a dual loop, and if so, identified which one was upstream. Then, whether dual loop or single loop detectors, Lang and Coifman used the time series speed to group lanes into common directions based on the times congestion occurs and dissipates at the detectors. As splashover errors often arise from a detector that is too sensitive, pulse breakup errors often arise from a detector that is not sensitive enough. Lee and Coifman (2012b) developed a test to identify pulse breakup errors that accounts for the traffic conditions.

Even when splashover is explicitly discussed in the published tests, it does not receive sufficient attention due to the difficulty of differentiating the error from valid concurrent vehicle passages. Any given splashover event in the data stream is usually indistinguishable from the non-splashover event of valid concurrent actuations from two vehicles passing the detector station at the same time. Our approach overcomes this problem. The algorithm calculates the percentage of pulses in the subject lane that are suspected splashover events, i.e., the *Rate of Suspected Splashover* (RSS). To control for non-splashover events in RSS we calculate a dynamic *Threshold Rate of False Positives* (TRFP) as a function of the observed traffic conditions. If RSS exceeds TRFP the lane pair is suspected of exhibiting splashover.

While the focus of this paper is on loop detectors, with only minor modification the algorithm should also be applicable to other detector technologies, e.g., side-fire microwave radar.

The remainder of this paper is organized as follows. Section 2 presents the procedure of collecting the ground truth data and the characteristics of splashover revealed from the ground truth data. Next Section 3 presents the splashover detection algorithm. Then Section 4 evaluates the algorithm with field data from 21 directional detector stations: 4 with splashover and 17 without. Finally Section 5 presents the conclusions.

2. The nature of splashover events

This section reviews the data used in this study, including: collection, ground truth data reduction, and the emergent trends that led to the splashover detection algorithm.

2.1 Ground truth data

Microscopic loop detector data with concurrent video ground truth are used in this study to develop and validate the algorithm to identify loop detectors with splashover. After time synchronization between the loop detector and video data we step through all of the loop detector actuations individually, loading the video corresponding to a given loop detector actuation, and manually classify each detector actuation. This comparison is enabled by a purpose built software graphical user interface (GUI) tool to semi-automate the process. Inspired by VideoSync (Caltrans, 2007), the GUI can step through the detector data in a given lane and display both the time series detector data for a few seconds before and after the actuation along with the frame corresponding to the actuation time. From the direct comparison between concurrent detector and video data, a user can mark actuations as detector errors such as splashover or other events such as a vehicle changing lanes. As the user enters the classification, the GUI jumps to the next detector actuation in that lane. The process was repeated for each visible lane during the entire time period of video data. While this research examined all 65 loop detector stations in Columbus, Ohio, this paper only presents the results for stations that had concurrent ground truth data. The video data

predominately came from the CCTV cameras recorded on a VCR in the Traffic Management Center, though in some cases a suitable view allowed for video recording in the field with a camcorder.

2.2 Types of splashover

Fig. 2 shows three examples of actual splashover in lane 1 at a single loop detector station (station 104 EB). Fig. 2a shows the pulses from all three lanes as the detectors respond to vehicles (throughout this example lane 1 is at the bottom and lane 3 is at the top). According to the concurrent video frames, Figs. 2b-d, all three of the pulses in lane 1 were erroneous. First, in Fig. 2b, a vehicle passing in lane 2 is also recorded in lane 1 in the absence of any vehicles in lane 1. We term this event a *"unique splashover"* because it only involved one vehicle. In contrast, Figs. 2c-d both show examples where the on-time of a short vehicle in lane 1 is extended due to splashover from lane 2 (in both cases the on-times are roughly twice as long as they should be for the given lane 1 vehicle). We term events like these as *"combined splashover"* because an otherwise unique splashover merged with a valid actuation. This distinction between the two types of splashover appears to be novel. When splashover is discussed in the literature it is usually limited to what we are calling unique splashover and we are not aware of any references to combined splashover. The difference between the two forms depends simply on whether or not a concurrent vehicle happens to be in the target lane. Since the time gaps between vehicles are typically much larger than the on-times during free flow, unique splashover is more common in free flow traffic. We suspect combined splashover becomes more prevalent during congestion, though we did not look for such events.

"Place Fig.2 about here"

From 10 min of video at this station we observed 164 vehicles in lane 1 and 347 in lane 2. Reviewing the loop detector data for lane 1, all of the lane 1 vehicles were detected, as well as 92% of the lane 2 vehicles[2]. Of these actual splashover events, 273 are unique splashover and 45 are combined splashover. In terms of macroscopic measurement errors, flow in lane 1 would be over-counted by the

[2] We deliberately chose station 104 eastbound for this example because it had the highest rate of splashover among the 94 lane pairs studied and it happened to offer a concise example illustrating both unique splashover and combined splashover.

number of unique splashover events (66% in this case). While the combined splashover events do not impact flow, both types of splashover will lead to an erroneous increase in occupancy because the detector reports that it is on for a longer cumulative time (i.e., $\sum OnT_1$) than it is actually occupied by vehicles.

2.3 The relationship between splashover and valid vehicle detection

Fig. 3a plots the on-time for the erroneous splashover events seen in lane 1 versus the given vehicle's valid on-time seen in lane 2, the vehicle's lane of travel (i.e., OnT_1 versus OnT_2 in the notation from Fig. 1). The dashed line shows the set of points where the two on-times are equal. As expected, unique splashover errors tend to fall below the reference line while most of the combined splashover errors fall above the reference line. Figs. 3b-c tabulate the cumulative distribution function (CDF) of the difference of on-times between the two lanes for unique splashover and combined splashover, respectively. A negative value indicates that the splashover on-time is shorter than that of the valid pulse in lane 2. About 90% of the unique splashover pulses in lane 1 have shorter on-times than the corresponding valid pulses in lane 2, and about 90% of the combined splashover pulses in lane 1 have longer on-times than the corresponding valid pulses in lane 2.

"Place Fig.3 about here"

The pulse from each detector actuation is defined by a rise transition time (RTT) and fall transition time (FTT). Comparing these across the two lanes, the difference in rising transition times (DRTT) and the difference in falling transition times (DFTT) between a splashover pulse in lane 1 and the valid pulse in the vehicle's lane of travel, lane 2, are used to quantify the characteristics of splashover. Fig. 4a illustrates the various relationships between DFTT and DRTT. The plane is divided into four regions (numbered I to IV, counterclockwise from top right) with boundaries at zero seconds on both axes. Hypothetical examples of the relationship are shown between the splashover pulse in lane 1 and the valid pulse in lane 2 for each region. If the lane 1 pulse begins before the lane 2 pulse, the result will fall to the left of the vertical division line (region II or III), otherwise it will fall to the right of the division line

(region I or IV). One would expect most unique splashover events from lane 2 to lane 1 to be like Fig. 1 and fall in region IV, while the combined splashover events should fall in regions I and III. Using the data from Fig. 3, Fig. 4b shows the relation of DFTT and DRTT between each splashover actuation in lane 1 and the corresponding valid detection in lane 2 at station 104 EB. Within each region the brackets tally the total number of observations of [unique splashover, combined splashover] seen. As expected, 86% of unique splashover events (236 out of 273) fall in region IV. The concurrent video data reveals that the unique splashover observed in region II occurred when the lane 2 vehicle traveled much closer to the boundary with lane 1 than the vehicles observed in region IV. Since a loop detector's sensitivity is in inversely proportional to the square of the distance to the vehicle (Klein et al., 2006), typically the detection zone for the splashover actuation (DZ_1^X in Fig. 1) should become larger and the detection zone for valid vehicle actuation (DZ_2 in Fig. 1) should become smaller as a lane 2 vehicle gets closer to the lane boundary. Meanwhile, 82% of the combined splashover events are observed in regions I and III (37 out of 45). Not surprisingly, the concurrent video data reveal that for the combined splashover events in region I a vehicle in lane 2 causing splashover is followed by a vehicle in lane 1 (e.g., Fig. 2c), while in region III a vehicle in lane 1 is followed by a vehicle in lane 2 causing splashover (e.g., Fig. 2d). Combined splashover events in region II and region IV usually occurred when two vehicles passed over the loop detectors at roughly the same time. In this case, about 76% of all splashover events are bounded by the valid pulses from the vehicle's actual lane of travel (e.g., Fig. 1), consistent with the fact that unique splashover is more frequent than combined splashover in free flow traffic. This distribution across the four regions is typical of the ground truth data from other stations with splashover during free flow conditions.

"Place Fig.4 about here"

3. Development of the splashover detection algorithm

The splashover detection algorithm is designed to identify a loop detector with chronic splashover errors. As seen in Fig. 4b, a splashover pulse is usually bounded by the valid pulse from the actual lane of travel.

This feature of splashover is used to identify pairs of pulses in adjacent lanes that are suspected of arising from splashover (termed as suspected splashover event) in Section 3.1. However, a suspected splashover event does not always correspond to actual splashover. To find a loop detector exhibiting chronic splashover errors we need to control for the non-splashover events that were labeled suspected splashover. To this end we estimate the expected number of false positive suspected splashover events (or simply "false positives") as a function of the observed traffic conditions in Section 3.2.

3.1 Number of suspected splashover events

A pair of loop detectors in adjacent lanes is selected for identifying a loop detector with splashover. One loop is arbitrarily taken as the "source lane", and for the test is assumed to provide an accurate record of vehicle actuations in that lane. The other loop is taken as the "target lane," and is evaluated as to whether or not it may be recording splashover pulses in response to vehicles in the source lane. The algorithm checks each pulse in the source lane to see if it spans a pulse in the target lane (i.e., falling in region IV in Fig. 4a). Any time the check is true it is considered to be a suspected splashover event and the number of suspected splashover events (nSS) is incremented by one. For example, Fig. 5a shows a hypothetical case where the pulse in the source lane, s1, spans a pulse in the target lane, t1. Thus, t1 would be a suspected splashover event and nSS would be incremented by one as a result of the comparison between s1 and t1. The algorithm repeats the process over all pulses in the source lane for the data set (in this study the duration of the ground truth video sequence). The roles of the lanes are exchanged and the process is repeated, then it is repeated for every other pair of adjacent lanes at the station. So like Lang and Coifman (2006), the present work uses the temporal relationships between different detectors to identify errors.

"Place Fig.5 about here"

The first two rows of Table 1 summarize the actual splashover number (ASn) from ground truth data and nSS from the algorithm for each pair of adjacent lanes at station 104 EB. The three loop detectors at the station have a total of four adjacent lane pairs. First consider the case with lane 2 as the

source and lane 1 the target: ASn is 318 and the algorithm reported nSS is 240 in lane 1 from vehicles in lane 2. In this case, all 240 of nSS turn out to be actual splashover events. However, any given suspected splashover event could be a false positive arising from non-splashover events in adjacent lanes. For example, in this data set the algorithm found a positive nSS for all of the remaining lane pairs, but ASn show that these lanes have no actual splashover events. According to the concurrent video, all suspected splashover events between lane 2 and lane 3 arise from valid concurrent actuations, but suspected splashover events with source lane 1 and target lane 2 arise both from valid concurrent actuations in the adjacent lanes and from lane 2 vehicles splashing over to lane 1 and falling into region II in Fig. 4a (which then appears to be region IV when lane 1 is used as the source lane). The situation is further complicated by the fact that it is not uncommon for a vehicle changing lanes over a detector station to actuate the loop detectors in both lanes, i.e., a detector error of a slightly different nature (see, e.g., Coifman, 2006a). As such, occasional splashover or lane change maneuver errors are within normal tolerance of conventional loop detector stations and such intermittent errors are below the resolution of the present work.

"Place Table 1 about here"

3.2 Expected number of false positives

The task of detecting splashover is complicated by the fact that one cannot differentiate between an actual splashover event and non-splashover event in the suspected splashovers using only the detector data. Instead, we estimate how many non-splashover events are expected in nSS. If vehicle arrivals are independent in the two lanes and there were no splashover events or lane changing maneuver errors, nSS is simply the total number non-splashover events, i.e., valid concurrent actuation of two vehicles passing the detector station at the same time by chance. The frequency of valid concurrent actuation events between the two lanes depends on the demand in each of the lanes. For the given traffic conditions, the number of valid concurrent actuation events is a random variable with some expected value. When all arrivals in the source lane are shifted by a small time offset of a few seconds relative to the target lane's

clock, the demand in the two lanes does not change. In most cases the time-shifted vehicle arrivals in the source lane remain independent of the un-shifted vehicle arrivals in the target lane. The number of valid concurrent actuation events after the time-shift is also a random variable. Since the traffic conditions did not change and arrivals are assumed to be independent in the two lanes, it has the same expected value as the un-shifted result. The assumption is used to estimate the number of non-splashover events in nSS, yielding the Estimated number of False Positives (EnFP). Provided the time offset is sufficiently large, the shift will eliminate the possibility that the given target pulse arose from the same vehicle that caused the given source pulse. Thus, when comparing "concurrent" pulses, the time shifted pulse train will be free of actual splashover events.

Formalizing the process to derive EnFP, like nSS before, when a pulse in the time-shifted source lane pulse train spans a pulse in the target lane, EnFP for this lane pair is incremented by one. For example, Fig. 5b revisits the data from Fig. 5a, only now it shows the arrival s1 in the source lane shifted by ε.[3] In this case, target lane pulse t3 begins within the time-shifted window from the source lane pulse, but it is not spanned by the time-shifted pulse, since t3 ends after the time-shifted s1 and it would not contribute to EnFP.

In the absence of splashover events, if the only contributing factor were valid concurrent actuations, both nSS and EnFP that are random variables with the same expected value will not always cancel each other because of randomness. Of course concurrent vehicle actuations are not completely random and the arrivals in adjacent lanes likely exhibit some small dependency (e.g., a driver may momentarily slow down for safety as they overtake a vehicle in an adjacent lane). Thus, in the absence of splashover errors, nSS is likely to be slightly larger than EnFP due to vehicle interactions, e.g., as evidenced in the third and fourth rows of Table 1. Table 2 enumerates the various factors that contribute to nSS and EnFP.

"Place Table 2 about here"

[3] This paper arbitrarily sets ε to five seconds. Although not shown, we repeated the analysis for ε set to each integer value from 1 to 10 seconds and found similar results.

To control for the non-splashover events included in nSS, we use a dynamic Threshold number of False Positives (TnFP) that effectively incorporates a demand responsive adjustment factor to ensure its expected value is slightly larger than that of nSS when there are no actual splashover events, Instead of requiring the entire pulse in the target lane to fall within the time-shifted source lane pulse, a target lane pulse will be counted in TnFP if the target pulse's rising edge is bounded by the time-shifted source lane pulse (i.e., regions I and IV in Fig. 4a). Thus, TnFP can never be smaller than EnFP and will generally be larger than EnFP in an attempt to compensate for the small variability between nSS and EnFP in the given sample, as well as driver interaction and lane changing maneuvers that potentially bias EnFP. Like EnFP, however, TnFP varies by as a function of demand in the two lanes.[4] With this definition, pulse t3 in Fig. 5b would contribute to TnFP. Returning to the on-going example at station 104 the results for TnFP are tabulated in the final two rows of Table 1. TnFP far outnumbers nSS in the three lane pairs that had zero ASn, as evidenced by negative values in the final row. However, TnFP does not exceed the impact of the non-zero ASn on nSS from source lane 2 to target lane 1. In general, after subtracting TnFP from nSS we interpret a positive difference in any lane pair as being indicative of a loop detector with chronic splashover errors. The last column of Table 2 shows the factors contribute to TnFP, compared to nSS and EnFP.

3.3 The adjusted rate of suspected splashover

To illustrate the fundamental ideas, thus far the narrative has talked about the numbers of events (ASn, nSS, and TnFP). The corresponding rate of these events is more practical to work with. We define the Actual Splashover Rate (ASR), Rate of Suspected Splashover events (RSS), and the dynamic Threshold Rate of False Positives (TRFP) relative to the total pulses in the source lane, N, via Equation (1)

[4] We have also contemplated linear comparisons between EnFP and nSS and have found that these comparisons require a large calibration data set- larger than our evaluation data set. At present our best performance with the linear models is comparable to TnFP and for brevity are not presented herein. See Lee (forthcoming) for greater details. In principle the linear models offer greater control and provide more insight, so they remain a topic for future research.

$$ASR = \frac{ASn}{N};\ RSS = \frac{nSS}{N};\ TRFP = \frac{TnFP}{N} \tag{1}$$

This process of reducing RSS by TRFP to account for non-splashover events is formalized in Equation 2, which is termed Adjusted Rate of Suspected Splashover (ARSS). The primary information from the test is the sign of ARSS, indicating whether RSS exceeds TRFP, i.e., a positive ARSS during free flow traffic is an indicator of chronic splashover. The magnitude of ARSS provides a coarse measure of confidence, but not with sufficient precision for diagnostics since the rate of two vehicles passing in adjacent lanes simultaneously depends on demand in both lanes, and thus, the magnitude of ARSS is not in itself a fair measure of comparison between lanes (especially if N inadvertently includes many non-vehicle pulses due to a detector error in the source lane).

$$\mathrm{ARSS} = \max\left(\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\left(P_{ij} - Q_{ij}\right)}{N},\quad 0 \right) \tag{2}$$

where,

$$P_{ij} = \begin{cases} 1, \text{ if } RTT_i^S \le RTT_j^T \le FTT_i^S \ \&\ RTT_i^S \le FTT_j^T \le FTT_i^S \\ 0, \text{ otherwise} \end{cases}$$

$$Q_{ij} = \begin{cases} 1, \text{ if } RTT_i^S + \varepsilon \le RTT_j^T \le FTT_i^S + \varepsilon \\ 0, \text{ otherwise} \end{cases}$$

and

ARSS= adjusted ratio of suspected splashover between source lane and target lane,
N = total number of pulses in the source lane,
M = total number of pulses in the target lane,
i = i-th pulse in the source lane (i=1, 2, ⋯, n),
j = j-th pulse in the target lane (j=1, 2, ⋯, m),
P_{ij} = suspected splashover of pulse j in the target lane matched to pulse i in the source lane (i.e., contribution to RSS),
Q_{ij} = non-splashover of pulse j in the target lane matched to pulse i shifted by the constant delay in the source lane (i.e., contribution to TRFP),
ε = constant delay for shifting a pulse in the source lane, set to five seconds in the presented results,

RTT_i^S = i-th pulse rising transition time in source lane,
FTT_i^S = i-th pulse falling transition time in source lane,
RTT_j^T = j-th pulse rising transition time in target lane,
FTT_j^T = j-th pulse falling transition time in target lane.

Of course the ARSS calculation assumes that a given target and source lane have roughly the same sensitivity level. This sensitivity depends on the hardware installation and the settings of the loop sensor cards. If the sensitivity in the target lane is significantly greater than the source lane, it is possible to see an inversion, e.g., most unique splashover events in target lane 1 from source lane 2 would fall in region II instead of region IV of Fig. 4a. In which case the algorithm will still detect the splashover event in the lane pair, but attribute it to the wrong lane.[5]

3.4 Optimal application periods

As speeds decrease due to congestion, TRFP will increase simply because the source lane vehicles reside over the loop detectors for a longer on-time. Since chronic splashover usually arises due to a hardware fault and is non-transient, the present work is meant to be applied during (predominantly) free flow conditions. There are many ways to ensure predominantly free flow traffic conditions. As noted previously, in this study the period was determined by the duration of the ground truth video sequence. In practice, one can use the macroscopic measures from the detectors to ascertain when conditions are free flowing and then apply the algorithm over several hours within one or more of these periods each day. One could also use time of day, e.g., when evaluating the Columbus detectors without ground truth (not shown), we have found that applying the tests daily during the mid-day period (9:00-15:00) yields good performance.

The free flow restriction will not increase the possibility of a false detection. Though it is possible that some errors may become more pronounced during congestion, e.g., if drivers tend to shift their lane

[5] Fortunately, most such cases of sensitivity errors can be detected via the median on-time test (Lee and Coifman, 2012) if one wants to ensure that a detected splashover is not attributed to the wrong lane in the pair. In extreme situations when the majority of the pulses at a detector actually arise from splashover, at such an over-sensitive detector it may become difficult for the median on-time test alone to distinguish between low sensitivity and the dominant, erroneous short pulses from splashover. While we have yet to observe such a situation, in these cases the impacts should be readily evident in other metrics (e.g., conventional single loop detector speed estimation) or tests (e.g., Jacobson et al., 1990; or the combined results of our splashover test and median on-time during uncongested low flow periods).

position over the detector in response to vehicles in other lanes. Such a situation can easily arise on a curve, drivers may tend to travel close to the inside lane line if there are no vehicles there, but be more conservative when there is a vehicle in the inside lane. So the algorithm may miss some of the lane pairs exhibiting splashover if the events occur predominantly during congestion. If coverage during congestion proves desirable, one might be able to use a rate derived from EnFP and examine the distributions of RSS and ERFP sampled over many days.

4. Evaluation of the algorithm

A total of 21 directional station ground truth data sets were generated in free flow conditions. Four of these directional station sets exhibited some degree of actual splashover events, as enumerated in Table 3. The remaining 17 data sets did not have any observed splashover events. The station, direction, date, time and duration of all 21 samples are summarized in the first few columns of Table 4. The stations were selected strictly on the basis of whether they were safely viewable; the dates and times were chosen arbitrarily to fit our availability, though any congested periods were excluded from this study. Reviewing the stations with actual splashover events in greater detail, the total pulses column in Table 3 tallies the number of pulses recorded by the detector during the period of concurrent video data collection at the given station, while the total vehicles column tallies the corresponding number of vehicles that traveled in the lane as seen in the video. The total number of actual splashover events is reported for the given lane where the vehicle was incorrectly detected (i.e., the target lane) as well as the subtotals for unique splashover and combined splashover. The total pulses do not always correspond to the sum of total vehicles and unique splashover events, e.g., from the totals in the last row of the table, 3,756 - (3,177 + 473) leaves 106 extra pulses. The concurrent video showed that these remaining pulses are due to vehicles changing lanes and being counted in both lanes.[6] The source and target lanes giving rise to the splashover are shown in the second to the last column. The final column shows the ASR. For instance, target lane 3

[6] One might view these lane change maneuver errors as a milder manifestation of the splashover error. Furthermore, none of the four data sets with splashover exhibited any other extra non-vehicle pulses during the period of study, e.g., pulse breakup events, though some of the 17 other data sets did.

at station 38 WB has 115 unique splashover events and 2 combined splashover events, all of which are caused by vehicles traveling in source lane 2. So this lane pair has an ASR of 117/242, or 48.3%, as shown in the table.

Lane 1 at station 104 EB has the highest ASR (91.1%), and lane 2 at station 38 WB has the lowest non-zero ASR (1.2%). There were a total of 537 actual splashover events from the four detector stations, 473 of which (88%) are unique splashover. So 473 out of 3,756 total pulses resulted from vehicles being counted in a second lane due to actual splashover at these four stations, i.e., 12.6% over-counting due to splashover.[7] As previously mentioned, while the combined splashover events do not impact flow, both types of splashover will lead to an erroneous increase in occupancy.

"Place Table 3 about here"

4.1 Application and results

The final columns of Table 4 present the adjusted rate of suspected splashover, ARSS, from the splashover detection algorithm (i.e., Equation 2) applied to the loop detector data for the periods with ground truth. Of the 94 adjacent lane pairs from 68 loop detectors at 15 loop detector stations, a total of seven lane pairs returned a positive ARSS. Five of these seven lane pairs exhibited actual splashover events (as per Table 3) while two did not, and the algorithm failed to identify two lane pairs with actual splashover events. Employing the ground truth data, the seven lane pairs that exhibited actual splashover are shaded in Table 4. The two lane pairs with actual splashover events that the algorithm failed to identify are target lane 2 from source lane 1 at both station 38 WB and station 41 EB. Looking at these two lane pairs in Table 3, they had a relatively small ASR, i.e., 1.2% in target lane 2 at station 38 WB and 3.3% in target lane 2 at station 41 EB. In both cases the total pulses in the target lane was significantly greater than the total pulses in the source lane, i.e., the number of pulses in target lane 2 is more than 140% of source lane 1 in the station 38 WB data set, and more than 150% in the station 41 EB data set, thereby increasing the chance of finding non-splashover events that contribute to TRFP. In the end, the

[7] The total over-counting rate rises to 15.6% if one also includes the 106 vehicles counted in two lanes during lane change maneuvers.

RSS at each of these two lane pairs was exceeded by TRFP. These two cases represent the threshold of "chronic splashover" that the algorithm can detect. In contrast, target lane 3 at station 56 WB also has a low ASR (3.0%) but the algorithm correctly identifies this lane pair because the number of pulses in target lane 3 is about 19% of source lane 2 and thus, the TRFP is small in this case.

Moreover, the algorithm correctly classified all but two detector pairs without splashover, the two errors being target lane 4 from source lane 3 at both station 2 NB and station 31 NB. Both of these failures occurred immediately upstream of two different lane drops where lane 4 ends. The nature of the traffic patterns upstream of a lane drop exasperates conditions for our algorithm. Relatively few drivers use a lane immediately upstream of when it drops, reducing TRFP. While those drivers that remain in lane 4 are seeking gaps to move over to lane 3 in anticipation of the pending lane drop. Such a vehicle is more likely to still be in lane 4 as they pass the detector station if there is a concurrent lane 3 vehicle precluding a convenient gap, i.e., violating our underlying assumption of independent arrivals in adjacent lanes and increasing RSS in the absence of splashover events.

"Place Table 4 about here"

4.2 A comparison of algorithm performance

To illustrate the fact that splashover may elude the existing error detection methodologies, we compared the performance of three earlier error detection methodologies (as reviewed in Section 1) against our algorithm using the same detector data sets that were used to generate Table 4. At each detector we implemented the following tests: Chen and May (1987), [C&M], which tabulates the percent of individual actuations with an off-time under 15/60 seconds; Jacobson et al. (1990), [JNB], which tabulates the percent of macroscopic data (20 sec samples) outside of the combined *acceptable thresholds* on flow and occupancy; Turochy and Smith (2000), [T&S], which tabulates the percent of macroscopic data (30 sec samples) with flow greater than 3,100 vehicles/hr; and our method from Section 3 (on each adjacent lane pair), [L&C]. For each of the preceding tests we used the parameters and settings as given in the respective article. The earlier error detection methodologies were not specifically designed to identify a

loop detector with splashover problems, but in theory the previous tests are capable of catching problems that arise from splashover errors. For example, the resulting double count due to unique splashover leads to inaccurate measurements, i.e., higher flow and occupancy, potentially allowing JNB or T&S to identify a loop detector with measurement errors via the macroscopic measurements. Meanwhile, C&M explicitly note that they found detectors exhibiting splashover. A unique splashover between valid pulses can lead to unexpected short off-time, potentially allowing C&M to identify a loop detector with splashover errors using their short off-time threshold.

We applied each of the three earlier error detection methodologies to all of the detectors with ground truth, thus effectively generating a separate table similar to Table 4 for each of the three earlier methodologies showing the results for each detector. For brevity these tables are not shown[8], instead, we segregate the detectors into two groups: the 7 with actual splashover (now simply termed "splashover"), and the 61 without (termed "non-splashover"). Within each of these two groups (i.e., splashover and non-splashover detectors), for a given test we calculate the minimum (min), maximum (max), mean, and median values that were found by the test across the detectors in the group. These summary statistics are shown on Fig. 6. To illustrate this process in the case of L&C, the statistics simply summarize Table 4, e.g., the mean splashover value is the mean of the seven shaded cells, the mean non-splashover comes from the 61 non-shaded cells, and so forth.

"Place Fig.6 about here"

One should not compare absolute values between methodologies since each test measures different features; rather, compare the relative values between splashover and non-splashover for a given methodology as a gauge of the test's ability to differentiate between the two conditions. Only T&S and L&C have a near zero mean or median for the non-splashover detectors. However, the difference between the splashover and non-splashover detectors is small for T&S across all four statistics. All four tests exhibit identical performance in terms of min. On the remaining three statistics, among the four tests

[8] The respective tables for the other methods are shown in the appendices of Lee (forthcoming)

L&C exhibits the largest relative difference between the splashover and non-splashover conditions. In fact, for mean and max, among the four tests L&C also exhibits the largest total (positive) difference between splashover and non-splashover. For the median JNB exhibits a slightly larger total difference between the two conditions compared to L&C. However, because the median of the non-splashover detectors from JNB is so much larger, L&C has a larger relative difference for the median values. Unfortunately, JNB exhibits the undesirable feature that the max and mean from the non-splashover detectors is larger than the corresponding statistics of the splashover detectors, the only test to exhibit such an inversion. The general result that our method, L&C, offers the best performance at differentiating between the splashover and non-splashover detectors is not surprising since as noted previously the other tests were not specifically designed to identify splashover. However, this example also illustrates the fact that splashover errors can easily go undetected using the existing suite of error detection methodologies.

4.3 A brief before and after study of loop detector sensitivity

This section seeks to illustrate the diagnostic power of the test, namely whether it can be used to eliminate splashover problems. Klein et al. (2006) note that splashover usually occurs when the sensitivity level of a loop detector sensor card is set too high or the physical loop is too close to the lane line. The latter splashover errors require restriping the roadway or cutting new loop detectors to solve the problem. The former splashover errors due to the sensor card setting, however, can be easily resolved by reducing the sensitivity level of the sensor card. After completing Table 4, it was not known whether the actual splashover errors were due to the sensor setting or the loop placement in the lane. Two stations were selected for further investigation: station 104 EB with the highest ASR in Table 3 and station 56 westbound with the lowest ASR in Table 3 that yielded a positive ARSS (i.e., that was detected by the algorithm, as shown in Table 4).

Our team asked the operating agency (the Ohio Department of Transportation, ODOT) to reduce the sensitivity setting in lane 3 at station 56 WB and lane 1 at station 104 EB. ODOT completed the changes on June 10, 2009 and reported that the sensitivity level in lane 3 at station 56 WB was reduced

from “Normal” to “Low” and lane 1 at station 104 EB was reduced from “High” to “Low”. A second round of ground truth data was collected for each station in free flow conditions after the change: 30 min (09:33 ~ 10:03) on June 17, 2009 from station 56 WB and 15 min (13:14 ~ 13:29) on June 26, 2009 from station 104 EB. No actual splashover events were found in the ground truth data at either station after the change and the algorithm labeled all of the lanes at those stations as being non-splashover.[9]

5. Conclusions

This paper developed and tested an algorithm for identifying loop detectors with chronic splashover problems. The examination of both loop detector data and concurrent video data found that the errors are manifest as unique splashover and combined splashover. If not caught, the resulting errors from splashover will lead to inaccurate measurements, e.g., higher flow from the double count in unique splashover, and both forms of splashover will lead to higher occupancy. At the crux of our algorithm is the fact that the pulse from a unique splashover in the target lane is usually bounded by a pulse of the valid vehicle detection in the source lane. The task is complicated by the fact that two concurrent vehicles in adjacent lanes can yield the same pattern in the time series detector data, even though it is not an error and it is not splashover. So the algorithm calculates the RSS and balances it with the TRFP as a function of the observed traffic conditions to find loop detectors suspected of chronic splashover problems.

On a trial of 94 adjacent lane pairs from 68 loop detectors at 15 loop detector stations, concurrent video was used to generate ground truth data for all of the passing vehicles. The algorithm correctly classified five out of the seven loop detectors that exhibited actual splashover. While the remaining two loop detectors that did not yield a positive response were below the threshold of the test, having a low ASR and a high TRFP. Two out of 61 loop detectors that did not have actual splashover were erroneously labeled as having splashover. Both of these failures were just upstream of where the given lane drops. In each case the false positive arose due to a combination of low flow in the target lane and the fact that any

[9] Obviously, a reduction in detector sensitivity should reduce the occurrence of splashover, while increasing the likelihood of pulse breakup and other errors. We explicitly studied pulse breakup and it is the topic of Lee and Coifman (2012b).

driver that remains in the dropped lane past the detector is typically doing so because there is a concurrent vehicle in the through lane, thus violating our assumption that arrivals are independent in adjacent lanes.

The algorithm was compared against three existing error detection methodologies and it exhibited the best performance differentiating between detectors with and without actual splashover problems. Although the tests developed herein do not attempt to differentiate between the various factors that may have caused a splashover problem, splashover often arises when the sensitivity level of a loop detector is set too high (Klein et al., 2006). In conventional practice it is difficult to know the physical loop's responsiveness, which makes it difficult to select the appropriate sensitivity level on the sensor. This work offers a viable tool to ensure the sensitivity level is set correctly on a loop detector, e.g., based on the results of the splashover test, two of the detectors identified as exhibiting actual splashover were subsequently adjusted and the change was shown to eliminate the splashover problem.

Although reducing the detector sensitivity solved the splashover problem for the two lane pairs we tested, in general the splashover problems do not appear to be correlated with measures of *over-sensitive* detectors. Lee and Coifman (2012a) used the daily median on-time to measure detector sensitivity and examined many of the same detectors studied herein. Among the stations common to the two studies there does not appear to be a correlation between the daily median on-time results (Fig 1 in Lee and Coifman, 2012a) and splashover (Table 4, herein). In fact at station 104 EB the daily median on-time for lane 1 (with splashover) falls below that of lane 2 and 3 (without splashover). The stations that exhibited the highest (station 4 SB) and lowest (stations 2 NB and 15 NB) daily median on-times in that study did not exhibit splashover in this study.

The algorithm is compatible with the existing loop detector infrastructure. However, because the microscopic data are usually discarded by the controller after it calculates the macroscopic measures, either these data need to be preserved (e.g., as is done in Columbus, Ohio- see Coifman, 2006b for details) or the test needs to be incorporated into the controller software.[10] In either case the algorithm

[10] Our algorithm is simple enough that it is a viable option to incorporate it into the traffic monitoring software of most traffic controller models.

would monitor the performance of existing detector stations in situ. Finally, with only minor modifications, the algorithm should also be applicable to other detector technologies. For example the out-of-pavement, wayside mounted Remote Traffic Microwave Sensor (RTMS) exhibited similar "splashover-like" errors due to apparent difficulties segmenting vehicles into lanes (Coifman, 2006a).

Acknowledgements

This material is based upon work supported in part by NEXTRANS the USDOT Region V Regional University Transportation Center and by the California PATH (Partners for Advanced Highways and Transit) Program of the University of California, in cooperation with the State of California Business, Transportation and Housing Agency, Department of Transportation. The Contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data and results presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification or regulation.

The authors are grateful for the help of the Ohio Department of Transportation and in particular Matt Graf, in facilitating this research as well as the constructive feedback from the reviewers of this paper.

References

Caltrans, 2007. VideoSync, http://www.dot.ca.gov/research/operations/videosync/index.htm, accessed on June 5, 2007.

Cheevarunothai, P., Wang, Y., Nihan, N. L., 2007. Using dual-loop event data to enhance truck data accuracy. *Transportation Research Record 1993*, TRB, Washington, DC, pp. 131-137.

Chen, C., Kwon, J., Rice, J., Skabardonis, A., Varaiya, P., 2003. Detecting errors and imputing missing data for single loop surveillance systems. *Transportation Research Record 1855*, TRB, Washington, DC, pp. 160-167.

Chen, L., May, A.D., 1987. Traffic detector errors and diagnostics. *Transportation Research Record 1132*, TRB, Washington, DC, pp. 82-93.

Cleghorn, D., Hall, F., Garbuio, D, 1991. Improved data screening techniques for freeway traffic management systems. *Transportation Research Record 1320*, TRB, Washington, DC, pp. 17-23.

Coifman, B., 1999. Using dual loop speed traps to identify detector errors. *Transportation Research Record 1683*, TRB, Washington, DC, pp. 47-58.

Coifman, B., 2006a. Vehicle level evaluation of loop detectors and the remote traffic microwave sensor. *Journal of Transportation Engineering*, 132 (3), pp. 213-226.

Coifman, B., 2006b. *The Columbus Metropolitan Freeway Management System (CMFMS) effectiveness study: part 2 – the after study*. The Ohio Department of Transportation.

Coifman, B., Dhoorjaty, S., 2004. Event data based traffic detector validation tests. *Journal of Transportation Engineering*, 130 (3), pp. 313-321.

Coifman, B., Kim, S., 2009. Speed estimation and length based vehicle classification from freeway single loop detectors. *Transportation Research Part C*, 17 (4), pp. 349-364.

Coifman, B., Krishnamurthy, S., 2007. Vehicle reidentification and travel time measurement across freeway junctions using the existing detector infrastructure. *Transportation Research Part C*, 15 (3), pp. 135-153.

Federal Highway Administration (FHWA), 2001. *Traffic Monitoring Guide*. May, 2001, U.S. DOT, Office of Highway Policy Information.

Hourdakis, J., Michalopoulos. P. G., 2002. Evaluation of ramp control effectiveness in two twin cities freeways. *Transportation Research Record 1811*, TRB, Washington, DC, pp. 21-29.

Jacobson, L., Nihan, N., Bender, J., 1990. Detecting erroneous loop detector data in a freeway traffic management system. *Transportation Research Record 1287*, TRB, Washington, DC, pp. 151-166.

Klein, L. A., Mills, M. K., Gibson, D. R. P., 2006. *Traffic detector handbook*, 3rd edition, Federal Highway Administration, U.S., Department of Transportation.

Kwon, J., Coifman, B., Bickel, P., 2000. Day-to-Day travel time trends and travel time prediction from loop detector data. *Transportation Research Record 1717*, TRB, Washington, DC, pp. 120-129.

Lang, L, Coifman, B., 2006. Identifying lane mapping errors at freeway detector stations. *Transportation Research Record 1945*, TRB, Washington, DC, pp. 89-99.

Lee, H., Coifman, B., 2012a. Quantifying loop detector sensitivity and correcting detection problems on freeways. [in press] *Journal of Transportation Engineering*

Lee, H., Coifman, B., 2012b. Identifying and correcting pulse-breakup errors from freeway loop detectors. [in press] *Transportation Research Record*.

Lee, H., (forthcoming), *Algorithms to identify splashover errors and pulse breakup errors in freeway loop detector data*. PhD Dissertation, The Ohio State University.

Papageorgiou, M., Hadj-Salem, H., Middleham, F., 1997. ALINEA local ramp metering summary of field results. *Transportation Research Record 1603*, TRB, Washington, DC, pp. 90-98.

Payne, H. J., Thompson, S. M., 1997. *Development and testing of operational incident detection algorithms: technical report*. U.S. Federal Highway Administration.

Payne, H. J., Tignor, S. C., 1978. Freeway Incident-detection algorithms based on decision trees with states. *Transportation Research Record 682*, TRB, Washington, DC, pp. 30-37.

Turochy, R. E., Smith, B. L., 2000. New procedure for detector data screening in traffic management systems. *Transportation Research Record 1727*, TRB, Washington, DC, pp. 127-131.

Williams, B.M., Guin, A., 2007. Traffic management center use of incident detection algorithms: findings of a nationwide survey. *IEEE Transactions on Intelligent Transportation Systems*, 8 (2), pp. 351-358.

LIST OF FIGURES

Fig. 1. A hypothetical example of a splashover event from source lane 2 to target lane 1.

Fig. 2. (a) A plot of transition pulses, and the corresponding video image at station 104 EB showing successively for the first through third lane 1 pulses: (b) unique splashover, (c) combined splashover at the head end, and (d) combined splashover at the tail end.

Fig. 3. (a) A scatter plot of on-times in unique splashover and combined splashover at station 104 EB, target lane 1 and source lane 2, (b) CDF of the difference of on-time in unique splashover and (c) combined splashover.

Fig. 4. (a) The combined relationship between the difference in falling transition times and the difference in rising transition times: (lane 1 – lane 2), (b) Scatter plot of the difference in rising and falling transition time for the splashover events from the ground truth data at station 104 EB, target lane 1 and source lane 2. Within each region the brackets tally the total number of observations of [unique splashover, combined splashover].

Fig. 5. The splashover detection algorithm to select (a) a suspected splashover event and (b) the non-splashover events that contribute to the expected rate of false positives.

Fig. 6. Bar charts comparing the min, max, mean, and median results for detectors with splashover and non-splashover from the four error detection methods, as applied to the data sets with ground truth data underlying Table 4. Note that the vertical scales differ among the four plots.

LIST OF TABLES

Table 1, Application of the splashover detection algorithm to station 104 EB.

Table 2, Various factors contributing to nSS, EnFP, and TnFP.

Table 3, Summary of the ground truth data with splashover in free flow.

Table 4, Percentage of adjusted suspected splashover relative to source lane. Shaded cells indicate a loop detector with splashover verified from the ground truth data, and all of the non-shaded cells represent detectors that did not exhibit splashover in the ground truth data.

LIST OF KEYWORDS

Loop Detectors;

Error Detection;

Splashover;

Freeway Traffic Monitoring

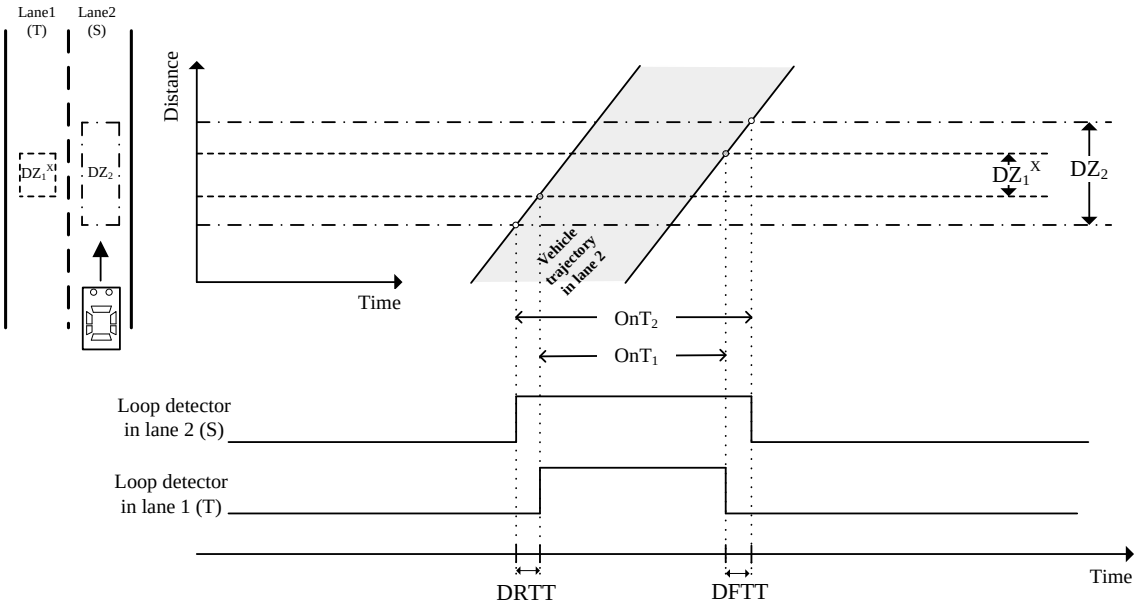

Fig. 1. A hypothetical example of a splashover event from source lane 2 to target lane 1.

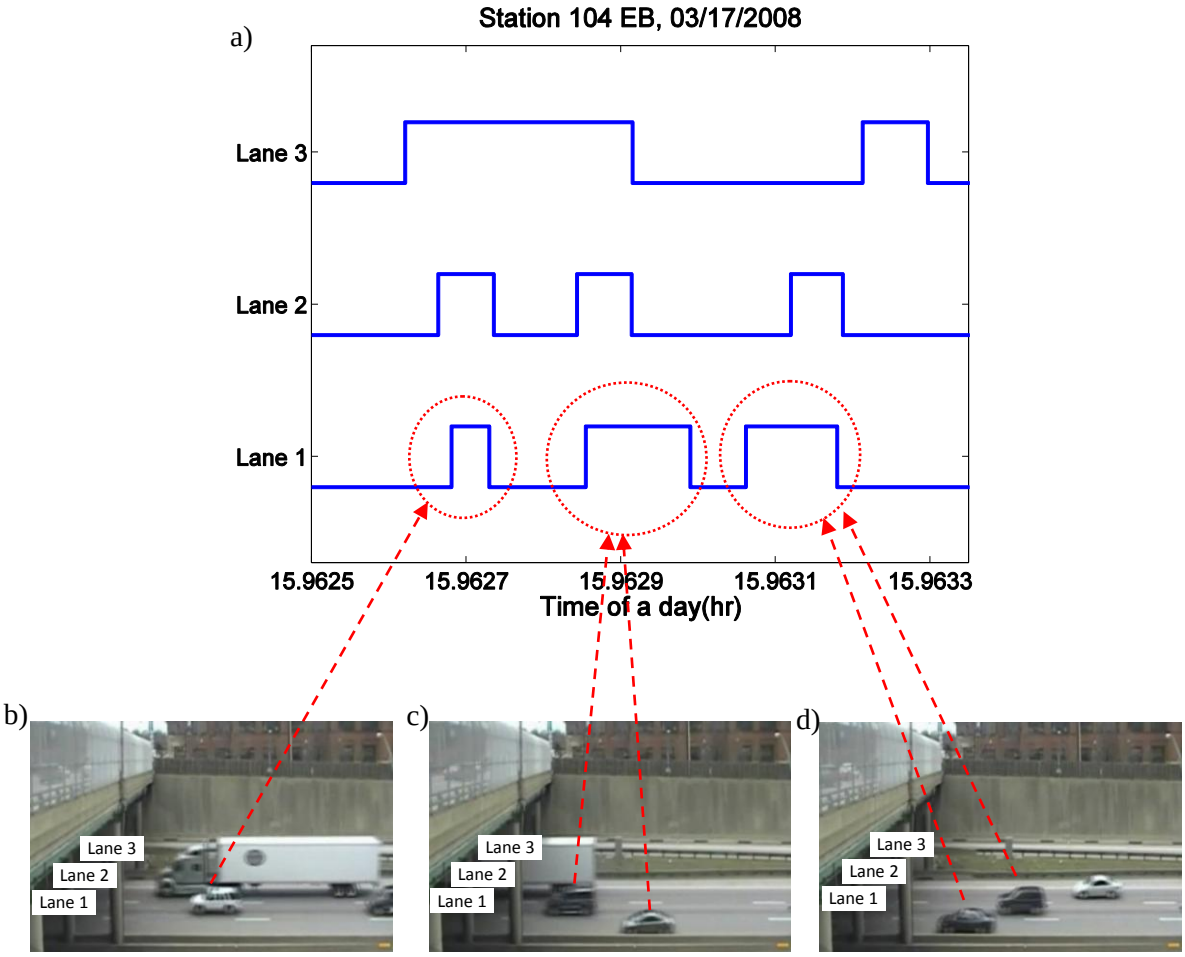

Fig. 2. (a) A plot of transition pulses, and the corresponding video image at station 104 EB showing successively for the first through third lane 1 pulses: (b) unique splashover, (c) combined splashover at the head end, and (d) combined splashover at the tail end.

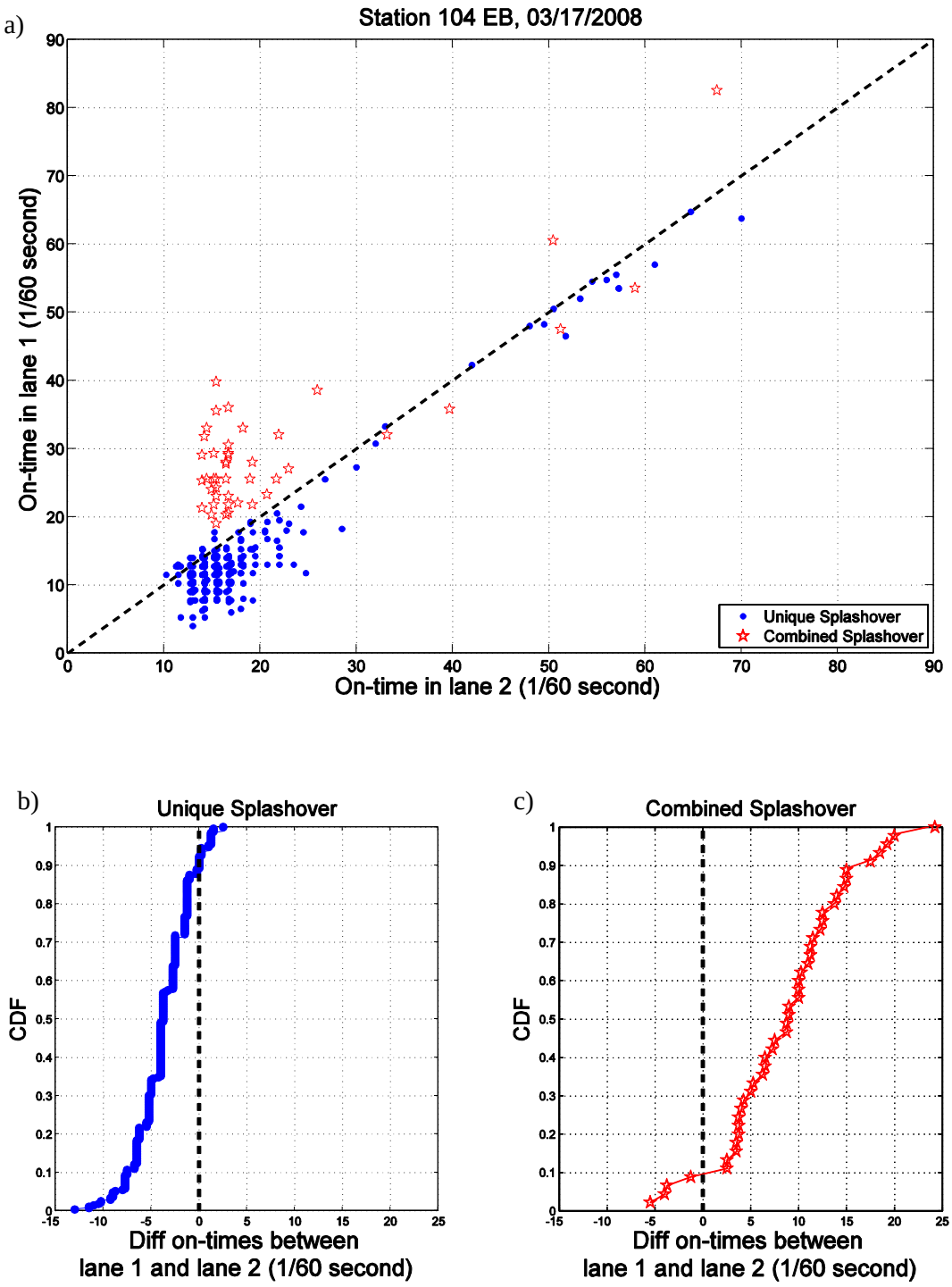

Fig. 3. (a) A scatter plot of on-times in unique splashover and combined splashover at station 104 EB, target lane 1 and source lane 2, (b) CDF of the difference of on-time in unique splashover and (c) combined splashover.

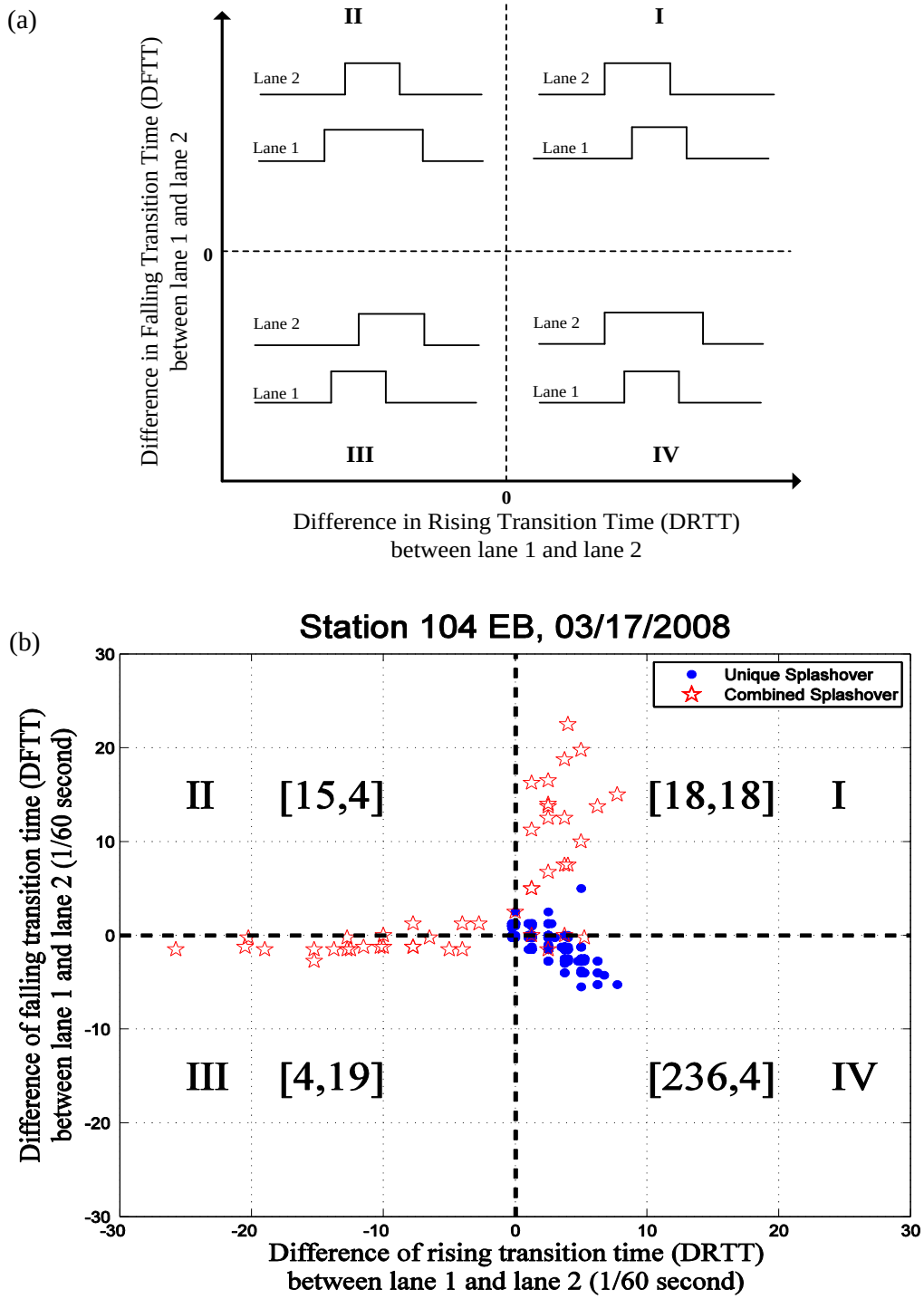

Fig. 4. (a) The combined relationship between the difference in falling transition times and the difference in rising transition times: (lane 1 – lane 2), (b) Scatter plot of the difference in rising and falling transition time for the splashover events from the ground truth data at station 104 EB, target lane 1 and source lane 2. Within each region the brackets tally the total number of observations of [unique splashover, combined splashover].

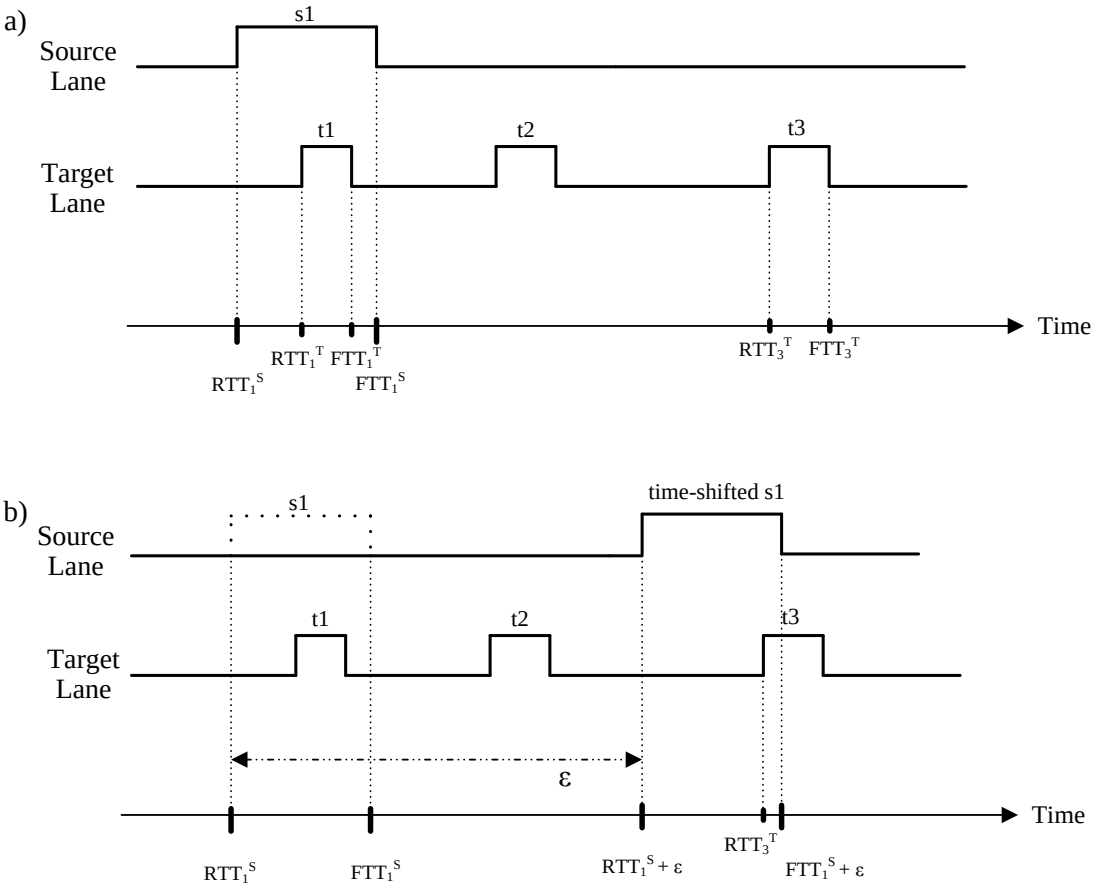

Fig. 5. The splashover detection algorithm to select (a) a suspected splashover event and (b) the non-splashover events that contribute to the expected rate of false positives.

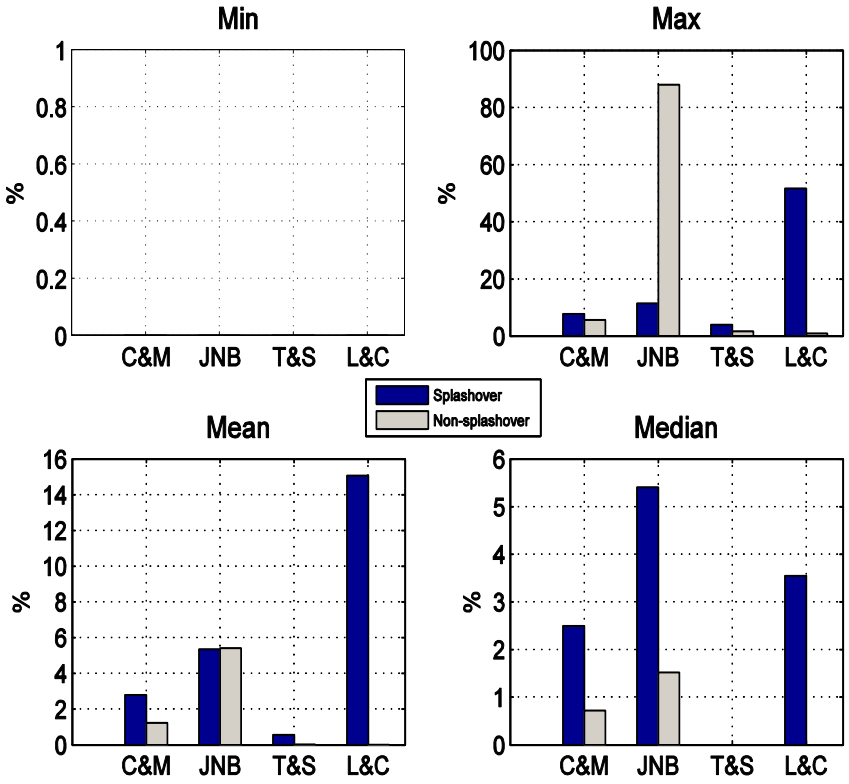

Fig. 6. Bar charts comparing the min, max, mean, and median results for detectors with splashover and non-splashover from the four error detection methods, as applied to the data sets with ground truth data underlying Table 3. Note that the vertical scales differ among the four plots.

Table 1
Application of the splashover detection algorithm to station 104 EB.

St 104 EB	[**Source lane** → Target lane]			
	Lane1 → Lane 2	Lane 2 → Lane 1	Lane 2 → Lane 3	Lane 3 → Lane 2
Actual splashover number (ASn)	0	318	0	0
Number of suspected splashover (nSS)	32	240	7	35
Estimated number of false positives (EnFP)	10	14	6	24
nSS – EnFP	22	226	1	11
Dynamic threshold number of false positives (TnFP)	61	60	47	74
nSS - TnFP	-29	180	-40	-39

Table 2
Various factors contributing to nSS, EnFP, and TnFP.

Factors	nSS & RSS	EnFP & ERFP	TnFP & TRFP
Valid concurrent actuation	√	√	√
Driver interaction	√		
Actual splashover	√		
Lane changing maneuver	√		
Demand responsive adjustment factor			√

Table 3
Summary of the ground truth data with splashover in free flow.

Station number (Direction)	Lane number	Total pulses	Total vehicles	Actual splashover number (ASn)			[Source lane → Target lane]	% actual splashover rate (ASR)
				Total	Unique splash-over	Combined splash-over		
38 (WB)	1	172	172	0	0	0	-	-
	2	242	235	2	2	0	Lane 1 → Lane 2	1.2%
	3	206	90	117	115	2	Lane 2 → Lane 3	48.3%
	4	56	39	17	17	0	Lane 3 → Lane 4	8.3%
41 (EB)	1	336	274	53	39	14	Lane 2 → Lane 1	10.5%
	2	506	475	11	8	3	Lane 1 → Lane 2	3.3%
56 (WB)	1	345	340	0	0	0	-	-
	2	632	610	0	0	0	-	-
	3	121	84	19	19	0	Lane 2 → Lane 3	3.0%
104 (EB)	1	441	164	318	273	45	Lane 2 → Lane 1	91.1%
	2	349	347	0	0	0	-	-
	3	350	347	0	0	0	-	-
Total		3,756	3,177	537	473	64		

Table 4
Percentage of adjusted suspected splashover relative to source lane. Shaded cells indicate a loop detector with splashover verified from the ground truth data, and all of the non-shaded cells represent detectors that did not exhibit splashover in the ground truth data.

Condition	Station number	Dire-ction	Date	Time	Adjusted rate of suspected splashover, ARSS [**Source lane** → Target lane]					
					Lane 1→ Lane 2	Lane 2→ Lane 1	Lane 2→ Lane 3	Lane 3→ Lane 2	Lane 3→ Lane 4	Lane 4→ Lane 3
Splashover	38	WB	09/09/2008	12:05 ~ 12:25 (20min)	0%	0%	41.3%	0%	6.8%	0%
	41	EB	09/09/2008	11:00 ~ 11:35 (35min)	0%	3.6%	-	-	-	-
	56	WB	11/21/2008	09:00 ~ 09:40 (40min)	0%	0%	2.2%	0%	-	-
	104	EB	03/17/2008	16:00 ~ 16:10 (10min)	0%	51.6%	0%	0%	-	-
Non-splashover	2	NB	03/09/2009	17:21 ~ 17:50 (29min)	0%	0%	0%	0%	0.6%	0%
	3	NB	03/17/2008	10:57 ~ 11:20 (23min)	0%	0%	0%	0%	0%	0%
	3	SB	04/18/2008	15:55 ~ 16:55 (60min)	0%	0%	0%	0%	0%	0%
	4	SB	03/17/2008	10:15 ~ 10:35 (20min)	0%	0%	0%	0%	0%	0%
	6	NB	04/18/2008	15:55 ~ 16:55 (60min)	0%	0%	0%	0%	-	-
	9	NB	06/05/2006	12:20 ~ 14:20 (120min)	0%	0%	0%	0%	-	-
	9	SB	06/05/2006	12:20 ~ 14:20 (120min)	0%	0%	0%	0%	-	-
	15	NB	03/10/2009	17:18 ~ 17:47 (29min)	0%	0%	0%	0%	-	-
	18	NB	03/09/2009	08:24 ~ 08:57 (33min)	0%	0%	0%	0%	-	-
	19	NB	03/17/2008	09:25 ~ 09:40 (15min)	0%	0%	0%	0%	-	-
	31	NB	11/21/2008	10:35 ~ 11:05 (30min)	0%	0%	0%	0%	0.9%	0%
	38	EB	08/29/2008	15:05 ~ 15:25 (20min)	0%	0%	0%	0%	-	-
	43	EB	09/02/2008	08:50 ~ 09:15 (25min)	0%	0%	0%	0%	-	-
	43	WB	09/02/2008	08:50 ~ 09:15 (25min)	0%	0%	0%	0%	-	-
	56	EB	09/03/2008	16:40 ~ 17:25 (45min)	0%	0%	0%	0%	-	-
	102	EB	03/10/2009	17:05 ~ 17:20 (15min)	0%	0%	0%	0%	-	-
	104	WB	03/12/2009	17:00 ~ 17:18 (18min)	0%	0%	0%	0%	-	-