This is the final version (unedited) of the paper published in ASCE Journal of Transportation Engineering

ANALYTICAL METHODS FOR DERIVING WORK ZONE CAPACITIES FROM FIELD DATA

Praveen Edara, Ph.D., P.E.¹*, Jalil Kianfar², Carlos Sun, Ph.D., P.E., J.D.³

¹Assistant Professor, E2509 Lafferre Hall, University of Missouri-Columbia, Columbia, Missouri 65201, edarap@missouri.edu, Tel: 573-882-1900, Fax: 573-882-4784

²Ph.D. Student, E1511 Lafferre Hall, University of Missouri-Columbia, Columbia, Missouri 65201, jkianfar@mizzou.edu, Fax: 573-882-4784

³Associate Professor, E2509 Lafferre Hall, University of Missouri-Columbia, Columbia, Missouri, 65201, sunc@missouri.edu, Tel: 573-884-6330, Fax: 573-882-4784

* Corresponding author

A revised paper submitted for publication in ASCE Journal of Transportation Engineering
ABSTRACT

Work zone capacity is a key variable in scheduling construction activity on roadways. Understanding the effect of different definitions of capacity on the values computed from field data will result in better planning and scheduling of lane closures. Capacity values obtained using three different definitions (and methods) of capacity: 1) maximum sustained flow, 2) re-scaled cumulative flow curves, and 3) 85th percentile flow were compared using field data from short-term work zones in Missouri.

The queue discharge flow (QDF) values were found to be the most conservative estimates of capacity. The 85th percentile flows were lower than the 15-min sustained flow values in all but one location. The pre-queue flow (PQF) values, indicative of near-constant flow prior to breakdown, did not occur in any of the four work zones. The four-site average capacities were 1149, 1267 and 1301 vphpl, for QDF, 85th percentile flow and 15-min sustained flow respectively. These capacity values were close to the 1240 vphpl HCM-based capacity value used in Missouri but lower than most values found in the nationwide survey.
INTRODUCTION

Work zone capacity is a key variable in scheduling construction activity on roadway facilities. Understanding the effect of different definitions of capacity on the computed values will result in better planning, scheduling of lane closures, and estimating the traffic impacts (e.g., capacity value is the key input to work zone user-cost analysis programs such as QuickZone and CA4PRS). Estimating traffic impacts and road user costs due to work zone lane closures is an area of high emphasis to both state departments of transportation (DOTs) and the federal highway administration (FHWA) (See Work Zone Safety and Mobility Rule (2004) for details on federal mandates on assessing mobility impacts at work zones for federal-aid projects). A considerable amount of research has been conducted in the area of work zone capacity estimation. Analytical (HCM 2000), regression (Kim et al. 2001), neural network (Karim and Adeli 2003), and other models have been developed using observed values of capacities at work zones. However, capacity was not uniquely defined across these studies; different definitions of capacity were used to obtain capacity values from field data. In this paper, capacity values obtained using different definitions of capacity were compared. Re-scaled cumulative flow curves (Banks 2009b), a popular technique for analyzing freeway bottlenecks, was used to evaluate work zone bottlenecks. Field data was collected at four work zones on two lane sections of I-70 in Missouri. The main objective of this paper is to present different methods based on different definitions of capacity to compute capacity from field data. The study goal is not to develop capacity estimation models (e.g., regression) but to compare different methods of extracting capacity information from the field data which will serve as the training (or estimating) data for capacity estimation models.

The research described in this paper makes the following new contributions to the existing body of knowledge on work zone capacity studies. First, past studies on work zone capacity focused on building models (regression, artificial intelligence, etc.) for estimating capacity at work zones using field data. Many of these papers assumed a certain definition of capacity (e.g., 15 minute sustained flow rate) without actually evaluating the effect of this definition on the capacity value. This is understandable
given their focus on building estimation models. The focus of our paper is to determine to what extent the definition of capacity at work zones influences the value computed from field data. For example, how does a capacity value obtained using a HCM capacity definition of maximum 15-minute sustained flow rate compare to a capacity value obtained using a 85th percentile definition, both obtained from the same field observations of traffic flow. This is important in two ways. First, the amount of capacity value variation that can be attributed to the difference in definition can be estimated. This information will lead to a fair comparison of the capacity values across different locations within a state or across different states. Second, the differences in the predicted values obtained from capacity models estimated using different capacity definitions will also be known. For example, when comparing the predicted capacity values using a regression model and a neural network model the difference in the predicted values that can be attributed to the definition of capacity can be ascertained.

Second, the state department of transportation (DOT) web-survey of 50 states administered in this research study provides information on the current work zone capacity practice in state DOTs in the US. Information on capacity definition, field measurement, estimation tools, factors influencing capacity, lane closure policy, and values adopted by different states are valuable to both traffic engineers and researchers. In recent literature, Sarasua et al. (2004), surveyed state DOTs for information on when lane closures were considered. Factors such as the use of traffic control plans, work zone information dissemination, allowable queue lengths, and threshold lane volumes were included in the survey. The survey conducted in the current study differs from this study in the following ways. First, the current survey is focused specifically on capacity definition and measurement, and factors influencing capacity. Since capacity is the key variable in the scheduling of lane closures, a synthesis of the current practice in state DOTs should be of value to practitioners. Second, the factors influencing work zone capacity was not part of the survey deployed by Sarasua et al. (2004). Obtaining this information from state DOTs will be beneficial to other state DOTs and to the HCM committee involved in the development of the list of factors deemed to influence work zone capacities. Last, the survey provides updated capacities being used in different state DOTs.
Third, the re-scaled cumulative flow curves method has been applied for the first time, in this paper, to evaluate work zone capacities. There are no studies to date that have evaluated this method for work zone capacity computations. The method has been proposed and validated at recurring bottlenecks [Banks, 2006, 2009] but has not been studied for non-recurring bottlenecks such as the work zone induced lane closures.

The paper is organized as follows. First, a review of the state of the practice in work zone capacity studies is presented. This includes a literature review and a state department of transportation (DOT) web survey that was administered to 50 states. Second, field studies are described along with details of data collection and processing. Third, the results of the application of three methods of capacity computation are shown. The first method is the maximum sustained flow method, the second is the new (to work zones) re-scaled cumulative flow curves method, and the third method is the 85th percentile method which was also recently proposed in the literature. Last, conclusions and future directions are presented.

REVIEW OF STATE OF THE PRACTICE

A review of state of the practice in work zone capacity was conducted. First, the literature on work zone capacity estimation was reviewed focusing on empirical studies and methods used to estimate capacity. Then, a questionnaire survey was prepared and sent to 50 state DOTs to obtain knowledge of their current practices in work zone capacity estimation.

Literature Review

Dudek and Richards (1982) reported that roadway capacities at work zones are lower than the capacities under normal operating conditions. Capacity value was computed as the maximum full-hour traffic count observed when queues existed upstream of the lane closure. They reported on six different lane closure scenarios and used data from 37 sites in Texas. These data were used to develop a chart showing the cumulative distribution of the work zone capacities. The Highway Capacity Manual (HCM) (1994) (and

Krammes and Lopez (1994) conducted research on work zones in major urban areas in Texas (Austin, Dallas, Houston, and San Antonio) where extensive frontage roads running parallel with the freeway function as an alternative to bypass the congested freeway conditions. Data were collected at 33 sites between 1987 and 1991 to update the capacity values for short-term freeway work zone lane closures. Capacity was computed as the mean-queue discharge flow rate averaged over consecutive hour volumes collected from a site. The researchers found that the new capacity values of short-term freeway work zone lane closures of 2 to 1 and 3 to 2 were significantly higher than the values reported in HCM 1994. HCM 2000 incorporated these findings. Unlike the capacity charts used in HCM 1994, a base capacity value of 1,600 pcphpl (passenger cars per hour per lane) is used for capacity computations in HCM 2000. This base value is adjusted through the application of adjustment factors such as the intensity of work activity, the effect of heavy vehicles, and the presence of ramps in close proximity to the work zone.

Dixon and Hummer (1998) conducted capacity studies at North Carolina short-term freeway work zones and found that North Carolina work zone capacities were higher than the HCM 1994 capacities by at least 10 percent. The 95th percentile 5-min flow value from all 5-min flow observations occurring after the onset of queue was computed for each site. This 5-min flow value was then converted to equivalent hourly flow rate which was reported as work zone capacity.

Al-Kaisy and Hall (2003) studied freeway capacities at six long-term work zones in Ontario, Canada. The data collection interval varied across locations – at two locations 20-second volume counts were collected, at two locations 5-minute counts were collected, and at the other two locations 15-minute counts were collected. Capacity was defined as the mean queue discharge flow rate as the authors consider it to reasonably represent the maximum sustainable flow during lane closure conditions. It was
found that all six sites had base capacity values lower than the HCM 2000 base capacity value. They developed a generic long-term work zone capacity model having a multiplicative form instead of an additive form, because it produced better estimates for the effect of heavy vehicles.

Karim and Adeli (2003) developed a neural network-based tool for the estimation of capacity and delay at work zones. The model considered 11 parameters in the estimation of capacity including number of lanes, number of open lanes, layout, percent trucks, grade, and intensity of work. The justification for using neural networks for this problem was that the functional form of the relationship between capacity and the identified independent variables was not known. Capacity values used in the study were generated using a lookup table developed by Ohio DOT and guidelines presented in the HCM for different highway and work zone capacities. The definition of capacity was not explicitly discussed in the paper. The developed estimation model was incorporated into an easy to use decision support system, IntelliZone (Jiang and Adeli 2004). IntelliZone estimates capacities and uses a deterministic queuing model to predict the queue length and delay.

Sarasua et al. (2004) conducted a study to determine the base capacity of short-term freeway work zones in South Carolina and eventually to determine the work zone capacity using equations derived from HCM 2000. Traffic volume, speed, and queue length data were collected at 22 sites on four interstates over a 1-year period. Hourly traffic flows were computed as the sum of 12 consecutive 5-minute flows. The maximum value of these hourly flows was defined as capacity. A straight line was fitted between speed and density based on linear regression. Using this equation along with the speed-flow-density relationship, the maximum value of flow, i.e., base capacity, was estimated. This base capacity value (1,460 pcp/h/l) was lower than the HCM 2000 base capacity value.

Schnell et al. (2002) evaluated traffic flow analysis tools applied to work zones. Highway Capacity Software (HCS), Synchro, CORSIM, NetSim, QUEWZ 92, and the Ohio DOT spreadsheet were used to estimate the capacity and queue length at four work zones on multilane freeways in Ohio. The
results were compared with the field data. From field data, capacity was computed as the average of hourly flow rates for the duration of queuing conditions. They reported that the simulation models could not be calibrated for oversaturated conditions that existed at the work zones, and even after calibration for other conditions, these models consistently under-predicted the queue lengths. They found that QUEWZ 92 was the most accurate in estimating the work zone capacity. When this capacity estimate was used in the Ohio DOT spreadsheet, it produced the most realistic estimates of queue lengths as compared to the estimates from other tools.

Chitturi and Benekohal (2004) compared the performance of QUEWZ 92, FRESIM, and QuickZone with field data at 11 freeway work zone locations in Illinois. Some of these work zones did not have queues. The results of the study showed that none of these models gave an accurate representation of real field conditions. QUEWZ 92 overestimated the capacity and underestimated the queue lengths, mainly because of its use of an outdated speed-flow relationship. FRESIM consistently overestimated the speeds under queuing conditions, overestimated the queue lengths for half of the cases, and underestimated the queue lengths for the other half of the cases. QuickZone consistently under-predicted the queue length and delay as compared to the field data. Benekohal et al. (2004) developed speed-flow curves for work zones based on data obtained from these 11 work zones and proposed estimating capacity using the operating speeds in a work zone. Recently, Chitturi et al. (2008) proposed a methodology for computing delay and user costs based on the speed-flow curves for work zones. In these studies, the capacity was computed from field data as follows. For sites with queuing, the top 15-minute interval that had the highest departure volumes was identified. The average headway was computed for this interval and the service capacity was computed as the inverse of average headway. For sites without queuing the top 5-minute was used instead of the 15-minute interval for service capacity computations.

Kim et al. (2001) developed a multiple regression model to estimate the capacity at work zones as a function of several key independent variables such as number of closed lanes, percentage of heavy vehicles, grade, and work intensity. To develop this model they collected data at 12 work zone sites in
Maryland. Work zones were operational after peak hours due to Maryland State Highway Agency’s policy of not closing lanes during peak hours. The method used to obtain capacity from field data was not specifically described in the paper. The authors reported that the developed regression model produced better capacity estimates as compared to the HCM model.

Sarasua et al. (2006) proposed two methods to estimate capacities at short-term work zones in South Carolina. The first method used curve fitting to establish speed-density-flow relationship in work zones and derived capacity as the maximum flow from the speed-flow curve. They reported that the parabolic speed-flow curve overestimated the capacity compared to observed capacities. To address this issue, they proposed an alternative method that used the 85th percentile volume as capacity. Xing et al. (2010) analyzed the work zone capacities of expressways in Japan. They used two definitions of capacity: 1) breakdown flow or the 15-minute flow rate immediately before the 5-minute space-mean speed decreased below 25 mph at a point immediately upstream of a bottleneck and 2) queue discharge flow or the average flow rate discharged during the congested conditions at a bottleneck. Nikolic et al. (2010) estimated the work zone capacities on single lane expressways outside Toronto. They used three methods: 1) direct estimation of capacity by identifying the maximum five-minute flows during queuing conditions, 2) estimation using time headways for different vehicle type combinations and proportions and 3) estimation from the VISSIM microscopic simulation model.

In summary, there were many studies that derived work zone capacities from field data. The primary focus of most studies was to develop a model that could adequately estimate work zone capacities without requiring the collection of actual flows. Each study assumed a certain definition of capacity in estimating its model. As previously mentioned, the main contribution of this paper is to illustrate the differences in capacity values obtained from different capacity definitions.
State-of-the-practice survey of state departments of transportation

A web survey was administered to 50 DOTs in the U.S and 29 DOTs responded. The survey inquired about the following: the definition of work zone capacity, the location where it is measured in the work zone, the factors affecting work zone capacity, the tools used by DOTs to estimate capacity and the capacity values for different lane configurations.

Survey responses

Work Zone Capacity Definition: Sixty-two percent of the respondent DOTs defined work zone capacity as the maximum observed hourly flow during pre-queue conditions. Only Texas, Maine, Washington and Washington, D.C. considered mean queue discharge flow as the work zone capacity. Few states such as Oregon and Colorado defined work zone capacity as the maximum observed 15 minute flow rate irrespective of the existence of a queue.

Work Zone Capacity Measurement in Field: Fifty-nine percent of the respondents indicated that they collect field data to estimate work zone capacity. Out of these respondents, 87% collect traffic volume data, 47% collect queue lengths, and 60% collect speed data in the work zone to estimate capacity. A few states such as Minnesota, Washington DC, Montana and Oregon also measure average headways of the vehicles in the work zone. Survey results suggested that 50% of the respondents measure work zone capacity well into the work zone area (i.e. near activity area). A few states such as Iowa, Oregon and Wisconsin measure capacity at the beginning of the taper whereas Massachusetts measures it at the end of the taper area. Remaining DOTs (17%) use pre-defined work zone capacity values from past experience and simulation studies rather than field measurements.

Tools to Estimate Work Zone Capacity: Sixty percent of the respondents follow the HCM procedure to estimate work zone capacity. A few states such as West Virginia, Texas and Washington use analytical tools such as QUEWZ to estimate capacity values, while New York and Rhode Island use micro-
simulation tools such as VISSIM and CORSIM for estimating work zone capacities. Florida DOT has developed a custom work lane closure policy document based on past empirical studies.

**Factors Influencing Work Zone Capacity:** State DOTs specified the factors they used for estimating work zone capacities as shown in Table 1. The list essentially repeated the factors documented in HCM 2000. Over fifty percent of the respondents included work zone configuration, work zone length, lane width, heavy vehicle proportion and ramp presence. All DOTs were asked to specify the adjustment values associated with the chosen factors, but only Wisconsin provided them. They are:

- Urban work zone capacity exceed rural work zone capacity by 200 pcphpl
- Base capacity is reduced by a factor of 0.97 when the shoulder width is less than 6 ft
- Base capacity value is reduced by the hourly ramp volume up to a maximum of 600 vph when a ramp is within 1500 ft of the work zone
- One truck is equivalent to 2 passenger cars
- Long term work zones capacities are up to 150 pcphpl higher than short term work zones
- Multiply base capacity by 0.97 if lane width is 11 ft, and by 0.95 if lane width is 10.5 ft
- Crossover work zone capacity can be 200 pcphpl less than non-crossover work zone capacity, especially in rural areas

**Work Zone Capacity Values Adopted by State DOTs:** The capacity values adopted by the state DOTs that responded to the survey are shown in Table 2. These values are later compared with the values obtained from the field studies.

**Work Zone Lane Closure Policy:** Ninety-three percent of survey respondents indicated that they have a policy of closing lanes either during night time or off-peak hours depending on the type of roadway (e.g., highway, secondary roadway, etc).
DATA COLLECTION AND PROCESSING

Field measurements were taken from four short-term maintenance work zones on Interstate 70 (I-70). Due to the heavy daytime traffic volumes on I-70 inside the urban area of Columbia, lanes were not closed during the day (MoDOT 2010a). Accordingly, all work zones involved nighttime lane closures that still had significant traffic volumes. Data was collected at the following work zones (see Fig. 1 for locations on the map):

Case 1. On June 22, 2009 data was collected at mile marker 125.7 on I-70. Approximate times of collection were between 7:00 PM and 9:30 PM. Westbound traffic was monitored.

Case 2. On June 23, 2009 data was collected at mile marker 125.7 on I-70. Approximate times of collection were between 7:00 PM and 10:00 PM. Westbound traffic was monitored.

Case 3. On June 28, 2009 data was collected at mile marker 124.7 on I-70. Approximate times of collection were between 7:00 PM and 10:00 PM. Westbound traffic was monitored.

Case 4. On June 29, 2009 data was collected at mile marker 124.7 on I-70. Approximate times of collection were between 7:00 PM and 10:00 PM. Westbound traffic was monitored.

All work zones involved a right lane closure with the passing lane open (2 to 1 work zone). The traffic was filmed using two cameras – one camera looking at vehicles approaching the taper, and the second camera looking at vehicles receding from the taper. The details of the standard MUTCD-type work zone setup can be found in the MoDOT policy on temporary traffic control (MoDOT 2004).

The cameras were deployed on surveillance trailers with 30 foot masts. The trailers were located on the side of the freeway beyond the clear zone. Even though the cameras were not directly overhead, the pan-tilt-zoom capability produced a very good field-of-view for video processing purposes. Since the work zones were two lanes to one, only the open lane at the taper area was monitored. Thus there were no occlusion problems. Image processing software, Autoscope, was used to automate data extraction using
virtual count detectors. The accuracy of counts obtained from Autoscope was verified by manually counting for 10 minutes in each tape. Data was collected in the summer when the daylight hours were long. Although, data collection period extended until it was dark, data was only processed for daylight hours and not for twilight transitions and night time conditions.

The speeds of vehicles approaching the work zone were captured using three methods: 1) a radar speed gun, 2) speeds derived from video processing and 3) speeds derived from virtual speed traps. Due to accuracy concerns, the virtual speed trap became the main method and is described as follows. A virtual speed trap was implemented using two known locations on the video field-of-view. The physical distances within the video field-of-view were recorded using a measuring wheel. These included the distance between reference delineators deployed by the researchers, the critical distances between other research equipment and distances between centerline dashes. The centerline dash spacing was verified to be at a 40 foot interval as specified in the MoDOT policy, which conforms to the MUTCD. Additionally, 5 minutes of speed radar data for each footage was also used for calibrating speeds. The calibration tweaked the speed trap distance to minimize the average absolute difference of speeds. Automated video processing software was used in detecting vehicles at the virtual speed traps. This elaborate system proved to be more accurate than the speeds derived directly from video processing software.

**COMPUTATION OF CAPACITY FROM FIELD DATA**

**Method 1: Sustained flow**

Chapter 13 of the HCM (2000) defines freeway capacity as “the maximum sustained 15-min flow rate, expressed in passenger cars per hour per lane that can be accommodated by a uniform freeway segment under prevailing traffic and roadway conditions in one direction of flow.” In addition to the maximum sustained 15-min flow rate, maximum sustained 10-min and 5-min flow rates were also computed in this research. Moving time windows of 15-min, 10-min, and 5-min were obtained by grouping 1-min traffic counts over the size of the respective time window. The maximum sustained flow rate was then obtained
by aggregating counts within a group. To be consistent with Missouri DOT’s units for work zone capacities (MoDOT 2010b) the reported capacities are in vehicles per hour per lane (vphpl). The truck percentages observed at each site were used in conjunction with passenger car equivalents (PCE) to convert capacities to equivalent passenger cars per hour per lane. All sites were on level terrain, therefore a PCE value of 1.5 was selected for trucks as recommended by the HCM.

Method 2: Automated re-scaling of cumulative flow curves

The Method 1 capacity definition of the maximum sustained flow could have occurred before or after the formation of queues. Past studies have demonstrated that there are two different types of maximum flow observed at freeway bottlenecks: queue discharge flow (QDF) and pre-queue flow (PQF); see e.g., Banks (1991; 2006; 2007), Hall and Agyemang-Duah (1991), Persaud et al. (2001). The QDF is the traffic flow discharged from an active bottleneck and PQF is the near-constant flow observed before the breakdown of flow (Banks, 2006). Persaud and Hurdle (1991), Ringert and Urbanik (1993) have conducted freeway bottleneck capacity studies using the mean discharge flow rate as a good proxy for the maximum sustained flow rate.

This section applies the concept of QDF and PQF to work zones. Cassidy and Bertini (1999) proposed the use of re-scaled cumulative vehicle arrival curves (re-scaled N-curves) to study the traffic features at two freeway bottlenecks in Toronto, Canada. The re-scaling procedure has been used in other bottleneck studies (Banks and Amin 2003; Kurada et al. 2007). First, an N-curve is built by piece-wise linear approximation of vehicle arrivals in discrete time intervals. A background cumulative vehicle count is then subtracted from the N-curve to form a rescaled N-curve. The re-scaling procedure assists in the visual identification of changes in flow. Using this methodology, the QDF, PQF and flow patterns observed during queue discharge (stationary versus non-stationary patterns) can be identified.

In a recent study Banks presented an alternative to the visual identification of changes in flow. Banks (2009b) explains, “The cumulative curves eliminate most of the noise in the data without obscuring
rapid changes in the average value of the time series; however, the identification of periods of nearly
canstant flow requires visual identification of segments of the cumulative curve that are nearly straight.
There were no definite criteria (other than the analyst’s subjective impression) for deciding what
segments were ‘nearly’ straight”. In other words, the visual identification of periods of constant flow or
straight lines in the curve depends on analyst judgment and could vary from one researcher to another. To
alleviate this inconsistency, Banks (2009b) suggested an automated method for freeway bottleneck
analysis. The method consists of using software routines to identify periods of breakdown and pre-queue
flow. He used this method to identify PQF and QDF in 21 non-work zone freeway sites in several
metropolitan areas in United States. In this paper, we apply this automated flow analysis method to
identify PQF and QDF in work zones. For all work zone applications studied in this research, the flow
breakdown was a result of an active bottleneck being present at the work zone location and not due to a
downstream queue spillback.

The automated flow analysis method starts with a piecewise linear approximation of the
cumulative vehicle arrivals. The points of greatest change in the average flow rate are then identified.
This is accomplished by locating points that maximize the absolute value of the difference between the
cumulative curve and a straight line representing its average value. Identification of these points is an
iterative procedure and continues until a desired level of linear approximation is achieved. These points
divide the time series into periods or segments that are further analyzed. The procedure is as follows
(Banks 2009a):

Assume $v(t)$ is the value of traffic flow time series during the time period $t_0$ to $t_n$. The cumulative flow,
$C(t)$, at time $t$ is defined as:

$$C(t) = \sum_{t_0}^{t} v(t)$$

The mean value of cumulative flow between $t_0$ and $t_n$ is defined as:
\[ \bar{C}(t_0, t_n) = \frac{C(t_n) - C(t_0)}{t_n - t_0} \]

The difference between cumulative flow at time \( t \) and a straight line representing the mean value of cumulative flow between \( t_0 \) to \( t_n \) is:

\[ \Delta(t) = C(t) - \bar{C}(t_0, t_n) \]

The point of greatest change in the mean cumulative flow, \( t' \), is selected such that

\[ |\Delta(t')| = \max(|\Delta(t)|), \text{ for } t_0 \leq t \leq t_n \]

The change in mean cumulative flow at time \( t' \) is considered significant if

\[
\begin{align*}
|\Delta(t')| &\geq \Delta_c \\
|\bar{C}(t_0, t') - \bar{C}(t', t_n)| &\geq \Delta_x
\end{align*}
\]

Where, \( \Delta_c \) represents the minimum difference between cumulative flow and the mean of cumulative flow that is considered significant, and \( \Delta_x \) represents the minimum change in the mean of the cumulative flow that is considered significant.

Periods of congestion are identified by the average speed. If the average speed of a period is below a threshold value, \( u_{\text{crit}} \), and the period lasts five minutes or longer, the period is considered to be congested. QDF exists whenever there is congestion upstream from the work zone, but there is no congestion downstream that has spilled back to the work zone. Periods before congestion are examined to identify PQF by fitting a regression line to the flow time series before the occurrence of breakdown. The slope of regression line shows increasing, decreasing, or constant flow. By definition, PQF is a period of nearly constant flow. Nearly constant flow is verified based on three conditions: (i) slope of the regression line is less than 100 vehicles per hour per lane per hour (vphplph), (ii) a t-test at 85% confidence verifies that the slope is statistically significant, and (iii) an additional test to verify that a small slope is not a result of fluctuations in flow.
The routine to find PQF is as follows. Assume $t_q$ is the breakdown time. Consider the period of piecewise linear approximation of cumulative flow to be between $t_{q-1}$ and $t_q$. A linear regression line is fitted to the flow time series between $t_{q-1}$ and $t_q$:

$$\dot{q} = bt + q_0$$

where, $\dot{q}$ represents the linear estimate of flow $q$ at time $t$, $b$ indicates the change of flow rate and $q_0$ is the flow at beginning of the period $t_{q-1}$. This process is repeated for periods starting at $t_{q-2}$, $t_{q-3}$ until the beginning of analysis period or another congestion period. The result is a set of regression slopes $\{b_1, b_2, \ldots, b_i, \ldots, b_n\}$, where $b_i$ represents the slope of regression line for the segment $t_{q-i}$ to $t_q$. The starting time $t_p$ of PQF is then obtained as:

$$b_p = \min(|b_i|) \quad \forall i : t_{q-n} < t_{q-i} < t_q$$

As described previously, the following three conditions need to be satisfied:

1. $$|b_p| \leq 100 \text{ vphplph}$$  (Condition 1)

2. $$\alpha_p \geq 0.85$$, where $\alpha_p$ is the significance level of $b_p$ (Condition 2)

A nearly zero $b_p$ value could be a result of a period with highly fluctuating flow, where decreasing and increasing flows could cancel out each other. To eliminate such a possibility in PQF computation, a period is only considered if

$$\begin{cases}
q_j - \frac{C(t_q) - C(t_p)}{t_q - t_p} (t_j - t_p) \leq 32 \text{ vpl} & \forall t_j, t_p < t_j < t_q \\
|\bar{q}_p - \bar{q}_{q-1}| < 200 \text{ vphpl} & \text{(Condition 3)}
\end{cases}$$

where, $C(t)$ represents the cumulative flow at time $t$ and $\bar{q}_p$ and $\bar{q}_{q-1}$ are average flow between $t_p$, $t_q$ and $t_q$, $t_{q-1}$, respectively. Condition 3 consists of two sub-conditions. The first sub-condition verifies if for
each point \( j \) the vertical intercept between the cumulative flow plot and piecewise linear approximation plot is less than 32 \( vpl \). The second sub-condition verifies if the difference between average flow of the two segments is less than 200 \( vphpl \).

Data from the camera looking upstream of taper was used for building the cumulative flow curves. This field-of-view was used in order to capture the onset of the queue at the taper area. The beginning of the taper is where the capacity is measured. Banks’ (2009a) method was applied for four work zones to identify PQF and QDF values. The average speed threshold \( u_{crit} \) was assumed to be 30 mph in the work zone, \( \Delta_c \) to be 10 \( vpl \), and \( \Delta_r \) to be 2 \( vpl \). These parameters, \( \Delta_c \) and \( \Delta_r \), were selected so that the time period of shortest segment at least sustains for five minutes. This was accomplished by assuming high initial values for these parameters and then gradually decreasing them until the shortest segment was five minutes long. The segmented cumulative vehicle arrival plot and the regression lines fitted to the corresponding segments of the flow time series plot are shown, as an example, in Fig. 2 for case 1 data. The QDF values are reported in Table 4. The first and second conditions were not satisfied in all four cases studied when applying the automated flow analysis method for finding PQF. This meant that there was no period of PQF preceding QDF in any of the cases, so no PQF values were included in Table 4.

**Method 3: 85th percentile traffic flow**

As previously discussed, Sarasua et al. (2006) proposed the use of 85th percentile traffic flow as capacity at work zones. They explained that the 85th percentile value is extensively used as a threshold in transportation and statistical areas and would be a suitable estimate of capacity. The 85th percentile traffic flow values were computed for the four work zones using the five-minute flow values. To do this, the five-minute flows were sorted in ascending order and cumulative percentiles were computed. For example, say there are 20 five-minute flows, the least flow would represent the 5th percentile \((1/20 \times 100)\) and the highest flow would represent the 100th percentile \((20/20 \times 100)\). The percentile for \( k \)th flow value would be computed as \((k/20 \times 100)\). The 85th percentile flow was determined from the plot of cumulative...
percentiles (see Fig. 3). The 85th percentile flow values are shown in Table 5. Case 3 was unusual as compared to the other cases because of a large jump in flow after the 85 percentile. Case 3 also had the lowest 85th percentile flow.

**DISCUSSION OF RESULTS**

The maximum sustained flows decreased as the aggregation interval increased from 5 to 15 minutes at all work zone sites. The capacity values obtained using the three methods (see Fig. 4) show that the QDF values were the most conservative estimates of capacity in all four cases. It was also found that the 15 minute sustained flows occurred before the formation of queue in the first three case studies. In case study 4, the demand was so high that the queue starting forming as soon as the work zone taper was set. In this case, the sustained 15-min flow occurred immediately after the queue formation. The times of occurrence of 85th percentile flow were after the sustained 15-min flow time intervals in all case studies. The QDF values, by definition, occurred after the queue formation with durations shown in Table 4. The lower QDF values reflect hysteresis in traffic flow caused by the onset of congestion (Kuhne and Michalopoulos 2009). Also, the difference between the 15 minute flow and QDF values decreased with the increase in the duration of QDF.

The PQF values did not occur in any of the four work zones. This means that there was no period of near-constant flow prior to the breakdown of flow. This was expected for case study 4 since the queue started forming immediately after the taper was set (not enough time for gradual flow buildup before breakdown). However, the non-existence of PQF in the other three case studies (where queues did not form until well after the taper was set) raises an important question about the existence of PQF conditions at short-term maintenance work zones. This finding has implications on the design of traffic control methods to alleviate work zone congestion. The existence of PQF and its correlation with probability of flow breakdown provide guidance on delaying or eliminating flow breakdown using traffic control methods such as ramp metering. More research is needed to measure the PQF at short-term work zones.
This is critical since even for normal (non-work zone) conditions on high-volume urban freeways the existence of PQF is not certain. For example, Banks (2009b) found that it occurred about 50% of the time for initial breakdowns and only 30% of the time for non-initial breakdowns (those that occurred following recovery from congestion).

The 85th percentile flows were lower than the 15-min sustained flow values in all cases except Case 1 for which it was slightly higher by 4vphpl. The 85th percentile flow is heavily influenced by the demand. Therefore, a more frequently congested location will have a higher 85th percentile flow than a less frequently congested location even though both locations could theoretically have the same 15 minute sustained flows. In addition to being influenced by the demand, the 85th percentile flow also depends on the duration of the analysis period. The effect of duration on the 85th percentile flow value, however, can be positive or negative, meaning the capacity value can either increase or decrease with the increase of duration. The dependency of 85th percentile flow on demand and the duration of analysis period can be perceived as disadvantages when using it as capacity. The capacity values in passenger cars per hour per lane (pcphpl), that were computed using passenger car equivalents for trucks, were higher on the second day of work zone deployment (case 2 had higher capacity than case 1 and case 4 had higher capacity than case 3). The type of work activity, location, traffic control plan, remained the same across days. This means that the driver population adapted better to the work zone traffic conditions on the second day of deployment. The average capacity value obtained by averaging all four cases were 1301, 1267, 1149 vphpl, for 15-min sustained flow, 85th percentile flow, and QDF, respectively. These values were close to the HCM-based capacity value of 1240 vphpl currently used by Missouri DOT for a 2 to 1 lane work zone (MoDOT 2004).

When comparing different methods using field data it is desirable to perform statistical analysis to verify if the differences in results produced by these methods are statistically significant. Performing statistical analysis for short-term work zone capacity is challenging. One reason for this is the limited capacity data that can be collected at short-term work zones. Capacity value, unlike other traffic
parameters such as speeds, typically occurs only once during an observation period (a day or a peak period). Short-term maintenance work zones like the ones studied in this paper exist only for one or two days at the same location. Therefore, collecting a dataset that is large enough to perform statistical comparisons may not be possible. In addition, for short-term work zone locations many times researchers need to deploy their own traffic monitoring equipment as there may not be any DOT sensors near the work zone taper area. Capacity studies for recurring congestion locations (non-work zones) do not have this problem as the locations of interest (e.g., a merge/diverge area on a freeway) are instrumented with DOT sensors continuously providing data on a daily basis. Therefore, obtaining a sufficiently large dataset (for statistical analysis) covering multiple days of data at recurring congestion locations is practical and not as challenging as the non-recurring short-term work zones.

**CONCLUSIONS**

The 2 to 1 work zone capacity values found in the DOT surveys were higher than the field values observed in Missouri. On the high end were Florida (1800 vph) and New York (1800 pcphpl). Several states ranged around 1450 to 1600 vphpl or pcphpl including Wisconsin, Nevada, Massachusetts, Hawaii and Iowa. Unfortunately truck factors were not available to translate all reported capacity values into a uniform unit of pcphpl.

The DOT surveys also revealed a diversity of practice among DOTs in treating work zone capacity. Some DOTs used simulation with historical data, some used analytical tools such as QUEWZ and some measured field capacities. The field capacities were not measured uniformly but were measured at different locations such as the beginning of the taper, the activity area or the end of the taper area.

The variability of capacity values as a function of the methodology was illustrated using field studies in Missouri. QDF might be a truer measure of what is sustainable in work zones but is harder to derive than either 85th percentile or the 15-min sustained flow. The 15-min sustained flow might be preferable to the 85th percentile flow, because the latter is dependent on traffic demand. The QDF values
were also the most conservative estimates of capacity in all four case studies. One other advantage of the QDF is that it is the relevant capacity measure when planners seek to predict queue lengths in work zones. The maximum sustained flow and 85th percentile definitions do not explicitly take into account the traffic characteristics induced by the lane closure. The cumulative curve method on the other hand, specifically tracks the traffic speeds before, during, and after the formation of queue at the work zone in determining the PQF and QDF values. In other words, the cumulative curve method is based on the accepted traffic flow theory of analyzing active bottlenecks. This method is also more statistically sound than the other two methods as it involves hypothesis testing for the flow regression line slopes. However, collecting traffic speed data could be considered as an additional data requirement of the cumulative curve method compared to the maximum sustained flow and 85th percentile methods. Methodologically, one challenging aspect of the cumulative curve method is the selection of threshold values for the parameters used in the method.

The use of maximum 15-min sustained flow has the benefit of familiarity of traffic engineers and practitioners to HCM definition of capacity. Similarly, the 85th percentile measure is popularly used in traffic studies (although mainly for speeds). The 85th percentile flow, however, is heavily influenced by demand. Therefore, a more frequently congested location will have a higher 85th percentile flow than a less frequently congested location even though both locations could theoretically have the same 15 minute sustained flows. In addition to being influenced by demand, the 85th percentile flow also depends on the duration of the analysis period. The effect of duration on the 85th percentile flow value, however, can be positive or negative, meaning the capacity value can either increase or decrease with the increase of duration. The dependency of 85th percentile flow on demand and the duration of analysis period can be perceived as disadvantages when using it as capacity.

ACKNOWLEDGMENTS

This research project was funded by the Federal Highway Administration’s Smart Work Zone Deployment Initiative Pooled Fund. The authors are thankful for the assistance provided by MoDOT.
traffic engineers Ken Strube, Erik Menenga and Dan Smith for coordinating field data collection sites. The authors wish to acknowledge the contributions of Amit Dhatrak, Jordan Freborg and Kyle Ervin, who helped with data collection and analysis, and Indrajit Chatterjee, who assisted with the state of practice survey.

REFERENCES


24. Missouri Department of Transportation (MoDOT). (b). “MoDOT work zone safety and mobility policy.”


LIST OF FIGURE CAPTIONS

Figure 1. Location of work zones field sites on I-70 (Google, 2010)

Figure 2. (a) Segmented cumulative vehicle arrival plot, (b) Time series of flow showing regression lines fitted to segments for case 1.

Figure 3. Plots of flows sorted in ascending order

Figure 4. Comparison of capacity values obtained from different definitions
Table 1. Factors influencing work zone capacity estimation

<table>
<thead>
<tr>
<th>Factor</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work zone configuration (number of open and closed lanes)</td>
<td>74%</td>
</tr>
<tr>
<td>Length of work zone</td>
<td>74%</td>
</tr>
<tr>
<td>Lane width</td>
<td>70%</td>
</tr>
<tr>
<td>Proportion of heavy vehicles</td>
<td>65%</td>
</tr>
<tr>
<td>Presence of ramp near work zone</td>
<td>52%</td>
</tr>
<tr>
<td>Posted speed duration</td>
<td>44%</td>
</tr>
<tr>
<td>Partial lane closure versus cross over</td>
<td>35%</td>
</tr>
<tr>
<td>Day time versus night time</td>
<td>35%</td>
</tr>
<tr>
<td>Intensity of the work zone</td>
<td>30%</td>
</tr>
<tr>
<td>Work duration factor</td>
<td>30%</td>
</tr>
<tr>
<td>Location of closed lane (left, middle, or right)</td>
<td>26%</td>
</tr>
<tr>
<td>Work zone grade</td>
<td>26%</td>
</tr>
<tr>
<td>Short term versus long term</td>
<td>26%</td>
</tr>
<tr>
<td>Type of lane delineation</td>
<td>22%</td>
</tr>
<tr>
<td>Horizontal alignment</td>
<td>13%</td>
</tr>
<tr>
<td>Driver familiarity</td>
<td>9%</td>
</tr>
</tbody>
</table>
### Table 2. Work zone capacity values adopted by state dots

<table>
<thead>
<tr>
<th>State</th>
<th>2 to 1</th>
<th>3 to 1</th>
<th>3 to 2</th>
<th>Two way one lane (TWOL)</th>
<th>TWOL (with median crossover)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Florida</td>
<td>1800 vph</td>
<td>1500 vph</td>
<td>3600 vph</td>
<td>1400 vph</td>
<td></td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1500 pcphpl</td>
<td>1500 vph</td>
<td>1500 pcphpl</td>
<td>1500-1600 pcphpl</td>
<td>1500-1600 pcphpl</td>
</tr>
<tr>
<td>Nevada</td>
<td>1500-1600 pcphpl</td>
<td>1500-1600 pcphpl</td>
<td>1500-1600 pcphpl</td>
<td>1500-1600 pcphpl</td>
<td>1500-1600 pcphpl</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>1500 vph</td>
<td>1500 vph</td>
<td>3000 vph</td>
<td>850-1100 vph</td>
<td></td>
</tr>
<tr>
<td>Hawaii</td>
<td>1600 pcphpl</td>
<td>1600 pcphpl</td>
<td>1600 pcphpl</td>
<td>600-800 pcphpl</td>
<td></td>
</tr>
<tr>
<td>Iowa</td>
<td>1450 vphpl</td>
<td>1450 vphpl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York</td>
<td>1800 pcphpl</td>
<td>1600 pcphpl</td>
<td>1700 pcphpl</td>
<td>1800 pcphpl</td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>1300-1400 vphpl</td>
<td>1200-1300 pcphpl</td>
<td>3000-3200 vphpl</td>
<td>600-750 vphpl</td>
<td>1200-1500 vphpl</td>
</tr>
</tbody>
</table>

Note: vph means vehicle per hour, pcphpl means passenger car per hour per lane.
<table>
<thead>
<tr>
<th>Case</th>
<th>15-min</th>
<th>10-min</th>
<th>5-min</th>
<th>Percentage of Trucks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(vphpl)</td>
<td>(pcphpl)</td>
<td>(vphpl)</td>
<td>(pcphpl)</td>
</tr>
<tr>
<td>Case 1</td>
<td>1256</td>
<td>1356</td>
<td>1314</td>
<td>1418</td>
</tr>
<tr>
<td>Case 2</td>
<td>1408</td>
<td>1563</td>
<td>1410</td>
<td>1565</td>
</tr>
<tr>
<td>Case 3</td>
<td>1192</td>
<td>1267</td>
<td>1290</td>
<td>1373</td>
</tr>
<tr>
<td>Case 4</td>
<td>1348</td>
<td>1448</td>
<td>1386</td>
<td>1488</td>
</tr>
<tr>
<td>Case</td>
<td>Flow (vphpl)</td>
<td>Flow (pcphpl)</td>
<td>Duration (min)</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>--------------</td>
<td>---------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Case 1</td>
<td>1026</td>
<td>1108</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td>1223</td>
<td>1358</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Case 3</td>
<td>1079</td>
<td>1148</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Case 4</td>
<td>1268</td>
<td>1362</td>
<td>33</td>
<td></td>
</tr>
</tbody>
</table>
Table 5. 85th percentile flows for all case studies

<table>
<thead>
<tr>
<th>Case</th>
<th>85th percentile flow (vphpl)</th>
<th>(pcphpl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>1260</td>
<td>1360</td>
</tr>
<tr>
<td>Case 2</td>
<td>1356</td>
<td>1505</td>
</tr>
<tr>
<td>Case 3</td>
<td>1110</td>
<td>1181</td>
</tr>
<tr>
<td>Case 4</td>
<td>1343</td>
<td>1442</td>
</tr>
</tbody>
</table>
Figure 1. Location of work zones field sites on I-70 (Google, 2010)
Figure 2. (a) Segmented cumulative vehicle arrival plot, (b) Time series of flow showing regression lines fitted to segments for case 1.
Figure 3. Plots of flows sorted in ascending order
Figure 4. Comparison of capacity values obtained from different definitions