AN INVESTIGATION ON CAPACITY
REDUCTION DUE TO CONGESTION

by

Charles J. Robbins

Professional Mentor
Dave Roper
California Department of Transportation

CVEN 689
Transportation Information and
Control Systems Design

Department of Civil Engineering
Texas A&M University
College Station, Texas

August 1991
SUMMARY

This paper investigates a reduction in capacity due to congestion on freeway sections. The investigation consists of: (1) a literature review of the speed-flow-density relationships and (2) an analysis of traffic data containing volumes, speeds, and densities of a freeway section. The traffic data was collected during the peak hours of afternoon travel. The purpose of the investigation is to develop a supportive argument for the hypothesis. The hypothesis is congestion reduces the capacity or maximum hourly flow rate of a freeway section.

The literature reviewed suggests that the speed-flow-density relationships are discontinuous for data collected downstream of a bottleneck (or queue). The literature also suggests that the discontinuity occurs at the onset of congestion or at the formation of a queue. The discontinuity suggests that there are two distinct functions (uncongested and congested) to describe the data and indicates a reduction in flows at the onset of congestion.

The data were plotted to identify trends in the speed-flow-density relationships. The trends that were identified supported the literature reviewed. Additionally, trends in the speed-flow-density as a function of time of day plot were identified. The trends showed two reductions in maximum flow rates: (1) at the onset of congestion and (2) during congested conditions. The second reduction supports the hypothesis discussed. The flow rates followed a damped oscillation pattern, which began at the onset of congestion and continued until the second reduction in maximum flow rates occurred. Further research is recommended in this area to potentially develop a model to describe the damped oscillating pattern in the maximum flow rates. This model could aid the traffic system operators to prevent or minimize the loss in freeway section capacity due to congestion.
TABLE OF CONTENTS

INTRODUCTION .......................................................... J-1
  Problem Statement ............................................... J-1
  Research Objectives .............................................. J-1
  Scope of Research ................................................ J-3

BACKGROUND ................................................................ J-4

DATA COLLECTION AND REDUCTION ............................... J-14

RESULTS ................................................................. J-17
  Flow rate as a Function of Density ......................... J-17
  Speed as a Function of Flow Rate ......................... J-17
  Speed-Flow Rate-Density as a Function of Time of Day J-20
  Flow Rate as a Function of Time of Day ................... J-20

DISCUSSION OF FINDINGS ........................................... J-23

CONCLUSIONS ......................................................... J-24

RECOMMENDATIONS .................................................. J-25

ACKNOWLEDGEMENTS ............................................... J-25

REFERENCES ............................................................. J-26
INTRODUCTION

A critical part of analyzing traffic facilities is to identify the operational characteristics that determine the capacity of the facility. The knowledge of the capacity of a facility, along with current and predicted traffic demands enables a traffic engineer to design, plan, operate, and manage facilities that adequately satisfy society's transportation needs. Therefore, determining the capacity of a facility plays a vital role in the design, planning, operational, and management processes.

The 1985 Highway Capacity Manual (1985 HCM) contains the most widely accepted definition for the capacity of a facility. The 1985 HCM (1) defines the capacity of a facility as "the maximum hourly rate at which persons or vehicles can reasonably be expected to traverse a point or uniform segment of a lane or roadway during a given period under prevailing traffic, roadway, and control conditions." The 1985 HCM provides a process to theoretically determine the capacity of a freeway section. This process consists of: (1) assuming a capacity for ideal conditions and (2) reducing that value with factors that define the prevailing conditions for the freeway section. This will be referred to as the theoretical capacity of a freeway section.

Problem Statement

It is believed by the author and others that congestion reduces the theoretical capacity of a freeway section. This phenomenon is depicted in Figure 1. Figure 1 shows that as the traffic stream approaches an unstable state, the demand on a freeway section reaches its theoretical capacity. The unstable state is a result of the drivers' inability to adjust their speeds smoothly to the surrounding conditions. This causes the drivers to accelerate and decelerate which produces shockwaves within the traffic stream. These shockwaves cause more instability within the traffic stream. These series of events result in a progressive reduction in the maximum hourly rate (theoretical capacity) of vehicles able to traverse the section of freeway. This happens for the duration of congested flows. As the demands are reduced and the congestion is relieved, the freeway section begins to return to its theoretical capacity. This occurrence illustrates a temporary decrease in the maximum hourly rate for that section of freeway. Thus, there is a loss in capacity. There is a need to research the phenomenon in Figure 1. The purpose of the research is to determine how, why, and to what extent congestion reduces the theoretical capacity along a freeway section. The answers to these questions will enable the traffic engineer to develop techniques that may minimize and/or prevent the reduction in capacity. Minimizing and/or preventing the reduction in capacity due to congestion will provide safer and more efficient traveling conditions for drivers.

Research Objectives

The objective of the research is to investigate the reduction in capacity due to congestion as illustrated in Figure 1. The investigation is intended to open doors for future research in the area of congestion and how it affects the capacity of a freeway section. The specific objectives of the research are:
Figure 1. The loss of capacity due to congestion.
1. Review the background material that addresses the fundamental relationships of traffic flow on freeways. Note, the author has no intention of re-defining the fundamental relationships of traffic flow on freeways.

2. Obtain freeway traffic data (volumes, speeds, and densities) during the peak periods of travel.

3. Plot the speed-flow-density relationships of the data to: (1) determine any reduction in capacity and (2) support concepts in the literature reviewed that imply congestion affects the capacity of a freeway section.

4. Using the literature reviewed and plots of the data, develop an argument that supports the need to research the effect congestion has on capacity.

Scope of Research

The scope of the research is limited to the time allowed to collect the data for this study. The data collection required to confirm a reduction in capacity due to congestion is extensive. This perspective should be kept in mind when reviewing this study.

The data for this study consisted of volumes, speeds, and headways, obtained through electronic surveillance. Video tapes of the freeway section were obtained to verify the electronic surveillance data. Various plots were developed to analyze data.
BACKGROUND

The capacity of a freeway section is determined from the fundamental relationships between speed, flow, and density. Therefore, to understand the effects of congestion on capacity, one must first understand these fundamental relationships. The following background material will: (1) present how the 1985 HCM interprets the fundamental relationships between speed, flow, and density and (2) present additional research that expands on these fundamental relationships. The additional research also provides an overview of the traffic characteristics that help identify the fundamental relationship during the transition from uncongested conditions to congested conditions.

The 1985 HCM (1) illustrates the speed-flow, and density-flow relationships as an inverted U-shape, see Figure 2. Because of its shape, two speed or density values can exist for the same flow rate. The low density, high speed side of the curve (point A), depicts stable conditions on the freeway. The high density, low speed side of the curve (point B), depicts unstable conditions on the freeway. This unstable condition represents a breakdown in flow or forced flow (congested conditions) and is anticipated to begin when the freeway section reaches capacity flows. The capacity or maximum flow rate of the freeway section is reached when as a result of the increasing density and decreasing speed reduces maximum flow rates. The density that occurs at this time is called the critical density and the speed is called the critical speed, see Figure 2. At this time, there is no available space for vehicle maneuvering and any disturbance in the traffic stream can propagate along the freeway section. This occurrence can often lead to a breakdown in flow. Therefore, it is difficult to maintain a freeway at the theoretical capacity for a long period of time without a breakdown in flow.

The 1985 HCM (1) generated curves illustrating speed-flow and density-flow relationships for freeway sections under ideal conditions (see Figures 3,4). For the design speeds of 60 miles per hour (mph) and 70 mph, these curves depict the capacity to be 2,000 passenger cars per hour per lane (pcphpl), with a corresponding density in the range of 60 to 70 passenger cars per mile per lane (pcpmpml) and a speed in the range of 30 to 35 mph. The curves depict important characteristics that greatly influence the classification of operating conditions along the freeway section. These characteristics are:

1. Speeds are insensitive to increasing flows on the upper branch of the curve, see Figure 3. This occurs for a wide range of flows.

2. As flow approaches capacity, the curves in Figure 2 depict a dramatic decrease in speed and a rapid increase in density. Therefore, operating conditions change rapidly as flows approach capacity.

3. Both curves, speed-flow and density-flow, assume the transition from uncongested conditions to congested conditions to be a continuous function. However, the dashed line indicates that there is some uncertainty as to how the curve delineates after capacity has been reached.
Figure 2. Relationships among speed, flow, and density for freeways (1).
Figure 3. The speed-flow relationships under ideal conditions (1).
Figure 4. The density-flow relationship under ideal conditions (1).
These concepts provide a sound base for understanding the fundamental relationships of traffic flow theory. However, additional research has raised some different viewpoints on the configuration of the traditional speed-flow-density curves (see Figures 3, 4). The different viewpoints are directed towards conditions at and beyond the capacity of the freeway section and are supported by data that does not follow the traditional speed-flow-density curves. The particular part of data that does not follow the traditional curves is the discontinuity between the congested and uncongested condition. Following are some interpretations of this discontinuity.

* In the early 1970s, May (2) used 45 data sets to analyze single- and two-regime models. The single-regime model closely resembled the traditional curves. The two-regime model assumes completely separate functions describing congested and uncongested conditions. This model also shows a distinct discontinuity between the two conditions with the maximum flow rate of the congested condition lower than the maximum flow rate of the uncongested condition (see Figure 5). This discontinuity was illustrated on the flow-density relationship and ranged from a density of 50 to 60 vehicles per mile (vpm). The results of the analysis showed that the two-regime model provided a better representation of the data than the single-regime model at near capacity conditions in the flow-density relationship.

* Athol et al. (3) further investigated the flow-density relationship. Figure 6 depicts a flow-density criterion function that separates two distinct and discontinuous functions: (1) a linear function that was observed before congestion and (2) a nonlinear function that was observed after congestion. Figure 6 also shows that the flow-density relationship followed an inverted V-shape. Hall et al. (4, 5, 6) collected data that was similar to Athol's results (the inverted V-shape for flow-density relationship) and concluded that the inverted V-shape is a result of data collected at locations downstream of a queue.

* Hall et al. (4, 5, 6) related the discontinuity in the speed-flow-density relationships to incomplete data. The incompleteness of data occurs when it is collected at a location downstream of a bottleneck or represents a queue discharge. Instead of identifying the discontinuity in the incompleteness of data, Urbanik (7) illustrated that the traditional U-shape identifies queued traffic under forced flow (congested conditions) and that it does not describe capacity flows which are located downstream of a bottleneck.

* Hall (4, 5, 6) evaluated the freeway on a lane-by-lane basis and found that the parameters describing the speed-flow-density curves vary by lane and by location along the freeway. Hall and Hall (8) studied speed-flow relationships downstream of a queue and within a queue. The recommendations were to develop two different curves, each corresponding to a different location.
Figure 5. A comparison of the single- and two-regime models (2).
Figure 6. Flow-density relationship illustrated as two separate and distinct functions (3).
* Banks (9,10,11) recently studied the capacity of a freeway section after the formation of a queue (congested conditions) at a bottleneck. This research hypothesized a reduction in capacity at the onset of congested conditions, the "Two Capacity Phenomenon". The "Two Capacity Phenomenon" was supported using linear regression on 30-second counts for 12 minutes before and after the formation of an upstream queue. The results showed that congested conditions decreased flows by 3% across all lanes.

* Hall et al. (12,13) concluded that the speed-density-flow relationships may be continuous and discontinuous simultaneously. This was demonstrated by constructing a three-dimensional relationship. The three-dimensional relationship is derived from the three two-dimensional relationships, speed-flow, flow-density, and speed-density. This three-dimensional curve conveys a twisted V- or U-shape, as illustrated in Figure 7. Hall et al. (12,13) indicated that the scattering of the data can be explained in the three-dimensional curve. The scattering of data is observed due to the view of a three-dimensional curve in two dimensions, therefore when the data falls on a twisted surface it will look scattered in two-dimensions.

From the literature reviewed, there is evidence that there is some form of discontinuity in the speed-flow-density relationships which affects the capacity of a freeway section. The onset of congestion causes this discontinuity indicating a reduction in flow for the freeway section. Therefore, congestion may cause a reduction in capacity. To identify a reduction in capacity, it is important to understand the triggering factors that cause congestion. The remaining background material contains proposed explanations for the onset of congestion.

Banks (9) concluded that "the breakdown (congestion) appeared to be triggered by speed instability". This conclusion agrees with a study conducted by Drew et al. in the 1960s. Drew, Dudek, and Keese (14) studied acceleration noise, which is the standard deviation of vehicle accelerations. The acceleration noise of a vehicle was investigated as a parameter to identify levels of service for traffic flow. The study (14) showed that the acceleration noise of a vehicle increased with traffic interactions. Figure 8 shows the relationship between acceleration noise and speed. The relationship shows that after maximum output has been reached the acceleration noise increases rapidly as speeds decrease. Athol (15) suggested that congestion is triggered from a driver behavioral response to a threshold tolerance of other vehicles. These results also indicate that the instability in flow at maximum hourly rates is a result of the drivers' inability to adjust their speeds to the surrounding traffic conditions.

The information above provides a general overview of the fundamental relationships that are used to determine the capacity or maximum hourly rate of flow of a freeway section. The intent of providing this background material was to establish a basis for explaining any identified trends in the traffic data which was collected.
Figure 7. Three-dimensional interpretation of speed-flow-density relationship (12,13).
Figure 8. The relationship between speed and acceleration noise. (14).
DATA COLLECTION AND REDUCTION

The study site was located on the westbound lanes of US 290, just north of Houston, Texas. This section of freeway is located between the exits of Tidwell and Fairbanks (both exits lead to diamond interchanges) and is illustrated in Figure 9. US 290 is a six-lane divided (divided by transitway) freeway with eleven-foot lanes and an eight-foot outside shoulder and no shoulder for the median lane. The detectors were located 350 feet downstream of a 1060 foot merge lane, see Figure 10.

The volumes, speeds, and headways of the vehicles in each lane were measured using a set of six-by-eight foot inductive loop detectors. The detector data were collected in the afternoon between 2:50 p.m. and 6:30 p.m. Video tapes of the peak period traffic (4:00 p.m. to 6:00 p.m.) were used to validate the detector data. The video cameras recorded traffic conditions upstream of the detectors.

The volumes, speeds, and headways of the vehicles were broken down into one minute intervals. Banks (14) and others were concerned that larger intervals of time may neglect the sensitivity required to identify any trends in the data. Particular concerns were in the areas of data collection and reduction, which include averaging techniques which can lead to the misinterpretation of data. It was felt that the use of one minute intervals of time would help eliminate any misinterpretation of data due to these averaging techniques.

The data were then graphed to illustrate the various relationships between speed, flow, density, and time of day. These relationships were then compared to the relationships found in previous research to identify any trends in the functional operation of the freeway section studied. Additional evaluations were made to illustrate the hypothesis.
Figure 9. Site location.
Figure 10. Schematic of detector location.
RESULTS

The term "flow", for the traffic engineer, means the number of vehicles that traverse a point or section along the freeway in a given time period. The time period most often used is one hour. For this study, detectors were used to determine the number of vehicles traversing a point on the freeway over a one minute period. The volume data, vehicles per minute, were averaged across all lanes and then converted to an hourly rate and is expressed in the remainder of the paper as flow rate.

The results of the study are determined from trends observed in the speed-flow rate-density relationships. These trends are used to support comments in the literature reviewed which imply that congestion affects the capacity of a freeway section. Trends were found in the following plots: flow rate as a function of density, speed as a function of flow rate, speed-flow rate-density as a function of time of day, and flow rate as a function of time of day.

Flow Rate as a Function of Density (Figure 11)

Flow-density data was used by researchers (2,3,4) to show that there is discontinuity in the flow-density relationship. Athol et al. (3,4) concluded that there existed a linear function and a non-linear function in the flow-density relationship (see Figure 6). The linear function describes data collected under uncongested conditions and the non-linear function depicts data collected during congested conditions. Figure 11 is a plot of flow rate as a function of density. The trend in Figure 11 shows that the data falls into the two regions identified by Athol et al. (3,4). The two regions are described by a linear function and a non-linear function. Figure 11 also shows that the linear function peaks at higher flow rates than the non-linear function. Hence, the data supports the statement that a discontinuity exists in the flow-density relationship such that the maximum flow rates are higher in the linear function region than the non-linear function region.

The data form an inverted V-shape, which suggests that the data were collected downstream of a bottleneck (where a queue formed). The review of the video tapes showed that a queue did form upstream of the detectors, thus supporting the background material reviewed.

Speed as a Function of Flow Rate (Figure 12)

Figure 12 is a line-symbol plot of speed as a function of flow rate. The speed-flow rate plot depicts a gap in the data for flow rates of 1,800 vehicles per hour per lane (vphpl) and greater. The gap occurs for speeds in the 47 mph to 53 mph range. It can be presumed that the gap separates the two regions (uncongested and congested) identified in Figure 11. The arrows in Figure 12 identify where data traverses from the uncongested region to the congested region in a single interval. Keep in mind that the flow rate, expressed as an hourly rate, is based on minute intervals. Therefore, the data indicates that the freeway section experiences a rapid transition from uncongested conditions to congested conditions.
Figure 11. Flow rate-density relationship depicts two separate and distinct functions, similar to the results of Athol et al (3).
Figure 12. Speed as a function of flow rate illustrating discontinuity in the speed-flow rate relationship.
Speed-Flow Rate-Density as a Function of Time of Day (Figure 13)

Figure 13 is a good representation of how speeds, flow rates, and densities fluctuate over the time of day. The flow rate is displayed in vehicles per minute (vpm) for viewing purposes and represents the total flow of all three lanes. There are two trends recognized in Figure 13.

The first trend supports the 1985 HCM statement that when a decrease in speeds and an increase in densities result in a drop in flow rates, the freeway section has reached its capacity. Figure 13 illustrates how the speeds and the densities converge at a point when the maximum flow rates first drop. The flow rates drop approximately 5 vpm (100 vphpl) which constitutes approximately 4% of the maximum flow rates. A review of the video tapes indicated that this occurred at the onset of a queue formation upstream of the detectors. Hence, the data also supports Banks (9, 10, 11) research on the "Two Capacity Phenomenon".

The second trend is a second reduction in the maximum flow rates that occurs at approximately 5:20 pm. Figure 13 indicates that the speeds and densities diverge at the time of the second reduction in maximum flow rates. The second reduction occurs over a longer period of time than the first reduction and results in a greater reduction in maximum flow rates (approximately 15 vpm or 300 vphpl). The review of the video tapes showed that a small shockwave had passed over the detectors at this time. The shockwave signifies that the freeway section is in a congested condition and as a result the maximum flow rates are reduced. Therefore, the second reduction appears to support the hypothesis demonstrated in Figure 1. However, the reduction in maximum flow rates is not as severe as expected.

Flow Rate as a Function of Time of Day (Figure 14)

There is an interesting trend in the flow rates when they reach their maximum at approximately 4:30 pm. As illustrated in Figure 14, flow rates (shown as vphpl) begin to follow a damped oscillation pattern. In other words, the variance in flow rates (shown as X in Figure 14) dissipates over the time of day-from the time maximum flow rates are reached to the time the speeds and densities diverge and there is a reduction in the maximum flow rates. This is similar to the motion of a weight attached to a spring oscillating longitudinally.
Figure 13. Maximum flow rates drop in the flow rate-speed-density as a function of time of day plot.
Figure 14. Flow rate as a function of time of day depicts a damped oscillation pattern.
DISCUSSION OF FINDINGS

The following findings are based on the background material reviewed and the trends in the various plots of the available data. The trends in the data support concepts found in the literature and are listed below.

* The capacity of a freeway section is reached when the decreasing speeds and increasing densities result in a reduction in flows.

* The relationship between speed, flow, and density is discontinuous for locations downstream of a bottleneck.

* The discontinuity in the speed-flow-density relationship is caused by the formation of an upstream queue. The upstream queue is a sign of congestion. Therefore, congestion is the cause for the discontinuity in the relationship.

* Higher maximum flow rates occur before congestion is present, therefore, the capacity of a freeway section may change at the onset of congestion.

Conclusions reached by these concepts indicate that the capacity or maximum hourly rate of a freeway section is affected by congestion. The effect is signified by a rapid decrease in speeds and a rapid increase in densities, such that the maximum hourly rate for the freeway section is reduced.

The data depicted an additional and more severe reduction in maximum hourly rates. This reduction occurred after the formation of the upstream queue (see Figure 13). The reduction was a result of an small and brief shockwave that passed over the detectors. The shockwave may be the reason for a rapid increase in the speeds and a rapid decrease in the densities. The data also depicted a damped oscillation pattern in flow rates before the second reduction in maximum flow rates. If this pattern can be modelled, it may prove to be valuable to transportation management teams. This function can be used to provide the operators with a warning that a breakdown in operations is about to occur. The warning would indicate to the operators that the number of vehicles entering the freeway should be reduced. A reduction in vehicles entering the freeway, at this critical time, could ensure higher maximum hourly rates for that section of freeway.
CONCLUSIONS

The data suggests that the speed-flow-density relationship is discontinuous for data collected downstream of a queue. The flow rate-density relationship requires two separate functions to describe the data (uncongested and congested).

The data indicated two reductions in flow rates, one reduction (approximately 100 vphpl) occurred when the upstream bottleneck formed a queue, the second reduction (approximately 300 vphpl) occurred when the shockwave briefly passed over the detectors after the queue had formed upstream. The reduction in maximum flow rates was not as severe as expected, but the second reduction does support the hypothesis described in Figure 1. The data depicted an oscillating pattern in the flow rates prior to the second reduction in flow rates. This implies the potential development of an empirical model. This model could be used to alert the traffic system operators of breakdowns in the operation of the freeway section. The operators can use this alarm as an indication to reduce the flow rates entering the freeway section, thus preventing a breakdown in operations.
RECOMMENDATIONS

The trends in the data depicted a potential for the development of an empirical model to forecast a breakdown in freeway operations. The trend in Figure 14 represented a damped oscillation pattern prior to a reduction in capacity during congested conditions. Further research in this area is needed. The research would have to include the use of detectors to collect extensive data all along the freeway section. The target locations would be upstream, downstream, and at the beginning of the bottleneck. The freeway section should be video taped during the data collection to verify the detector data.

ACKNOWLEDGMENTS

This paper is based on the data supplied by the Texas Transportation Institute at Texas A&M University. This paper was developed through the a graduate class at Texas A&M University. The author would like to acknowledge the cooperation of Dr. Thomas Urbanik II and his staff for providing the data. The author would also like to acknowledge the professors who organized the class, Dr. Conrad Dudek and Dr. Carroll Messer. Special thanks to the mentors whose participation made the class possible, Dave Roper, Ed Rowe, Joe McDermott, Don Capelle, and Gary Trietsch.
REFERENCES


J-26


Charles J. Robbins received his B.E. in Civil Engineering in May 1990 from Manhattan College and is currently pursuing his M.S. in Civil Engineering at Texas A&M University. From May 1988 through September 1989, he was an Engineer for John Collins Engineer, P.C. His area of interest is traffic management systems.