Ramp Metering and the Two-Capacity Phenomenon in Freeway Operations

TIAN Zong-zhong

(Department of Civil & Environmental Engineering University of Nevada, Reno 89557, USA)

Abstract: Previous studies have demonstrated a unique operational feature for freeway operations, which is often referred to as the two-capacity phenomenon. The two-capacity phenomenon suggests that freeways can obtain higher flows under freely-flowing (unqueued) conditions than under queued conditions. At an isolated ramp metering location, it is due to the existence of the two-capacity phenomenon that signifies the effectiveness of ramp metering in reducing overall delays for both the freeway and the ramp traffic. This point is first illustrated using a numerical example. The paper further explores whether microscopic simulation models (VISSIM in this case) could be calibrated to replicate the two-capacity phenomenon as well as to analyze some characteristics associated with the phenomenon. It is important to note that microscopic simulation models, when used for evaluating ramp-metering operations with a focus on delay measures, have the capability of replicating the two-capacity phenomenon. Otherwise, erroneous and invalid conclusions might result. From this study, we found that most microscopic simulation models, such as VISSIM, use traditional car-following theory, which does not automatically exhibit the two-capacity phenomenon. However, with a special coding and parameter selection, the two-capacity phenomenon can be properly replicated in VISSIM. Using the VISSIM model, we conducted analysis on the characteristics of ramp metering and the two capacity phenomenon. We found that ramp metering diminishes the difference between the two capacity values, especially when queue flush policies are adopted. While the freeway capacity under freely-flowing conditions proves to be difficult to estimate, it should have a value no less than the capacity under queue-discharge conditions.

Key words: two-capacity phenomenon; freeway operations; ramp metering; simulation

CLC number: U412.366

在高速公路中的匝道控制与双能力现象

田宗忠

（内华达大学雷诺分校土木与环境工程系，雷诺 89557, 美国）

摘要：已有研究已经论证高速公路独有的运营特点，包括双能力现象。双能力现象表明，高速公路在自由运行（非排队式）条件下比在排队条件下具有更大能力。在孤立的匝道区，双能力现象的存在，决定了匝道控制在缓和高速公路和匝道交通延迟的效率。本文进一步探讨了微观仿真模型（VISSIM）是否能够在分析现象有关特征的基础上，校准复制双能力现象。本文指出，微观仿真模型用于研究匝道控制延时评价时，必须具有再现双能力现象的能力，否则就会得到错误的结论。本文发现大多数微观仿真模型，例如 VISSIM，尽管使用了传统的跟车理论，但不能自动展示双能力现象。然而，如果在 VISSIM 中使用一组特殊的编码参数就能再现双能力现象。利用 VISSIM 模型，本文分析了匝道控制和双能力现象的特征，发现匝道控制减小了双能力值的差距，尤其是采用排队政策时。由于高速公路在自由运行状态下的能力难以评估，本文建议可用有排队状态下的能力值来代替。

关键词：双能力现象；高速公路管理；匝道控制；仿真

中图分类号：U412.366
0 Introduction

Unlike other traffic facilities, freeways have a unique operational feature often referred to as the two-capacity phenomenon, which is characterized by the ability to obtain higher flows under freely-flowing (unqueued) conditions than under queued (i.e., congested) conditions. The transition from the freely-flowing condition to the congested condition is often referred to as a freeway breakdown, characterized by a sudden drop in speed, an increase in density, and perhaps a drop in flow (Persaud et al., 2001). Because freeway capacities are usually estimated based on the measured flows, the drop in flow also results in capacity reduction once breakdown occurs, which is perhaps the origin of the two-capacity terminology.

While the majority of the studies have supported the two-capacity phenomenon hypothesis (Hall and Agyenmang-Duah, 1991; Cassidy and Bertini, 1999; Persaud et al., 1998; Lorenz and Elefteriadou, 2001; Zhang and Levinson, 2004), there has been disagreement on the level of capacity drop when breakdown occurs. Some studies have indicated that the queue-discharge capacity might be even lower than the free-flow capacity (Banks, 1991; Ringert and Urbanik, 1994). One of the major objectives of this paper is to further explore the nature of the two-capacity phenomenon from new perspectives. Most importantly, the paper emphasizes the importance of the two-capacity phenomenon in evaluating ramp metering operations. We want to demonstrate that ramp metering would be effective in reducing system delays only if the two-capacity phenomenon existed. Another issue to be discussed is related to the application of microscopic traffic simulation models. It is important to emphasize that a microscopic simulation model to be used for studying freeway and ramp metering operations should have the capability of replicating the two-capacity phenomenon. Otherwise, erroneous and invalid conclusions might result.

The paper is organized as follows. First, a comprehensive literature review is provided on the research findings related to the two-capacity phenomenon. Data collected from the field is also used to support the findings. Using a sample calculation, the freeway and ramp delays are estimated to demonstrate the importance of the two-capacity phenomenon in evaluating ramp-metering operations. The study then focuses on testing whether the VISSIM microscopic simulation model can replicate the two-capacity phenomenon. Furthermore, the characteristics of the two-capacity phenomenon associated with ramp metering are investigated using VISSIM. Finally, a section of summary and conclusions is provided.

1 Literature Review

A number of publications have been devoted to studying the two-capacity phenomenon in freeway operations. Freeway capacities are usually estimated based on the measured flows over a sustained period, such as a 15-minute period as defined in the Highway Capacity Manual (HCM) (Transportation Research Board, 2000). The critical elements related to measuring the flows include the measurement location, the measurement period and the presence or absence of queues both upstream and downstream of the measurement location. The flows for estimating queue-discharge capacity must be measured at an active bottleneck location (Hall and Agyenmang-Duah, 1991). An active bottleneck, as defined by Daganzo (Daganzo, 1997) is a location characterized by the presence of an upstream queue (to guarantee that vehicles are flowing at a maximum rate) and the absence of queuing downstream, such that vehicles are not influenced by another bottleneck located further downstream. In measuring flows for estimating the free-flow capacity, it should be restricted to the period prior to breakdown when demand is approaching the maximum (Hall and Agyenmang-Duah, 1991). Otherwise, the capacity would be underestimated. At freeway merge sections, the flow measurement location should be somewhere downstream of the freeway merge, so that both freeway and ramp flows are counted.

Based on review of the existing literature, the majority of the studies support the two-capacity phenomenon hypothesis. (Hall and Agyenmang-Duah, 1991; Cassidy and Bertini, 1999; Persaud et al., 1998; Lorenz and Elefteriadou, 2001; Zhang and Levinson, 2004). However, studies have shown different results in terms of the difference between the two capacities. These studies
have reported different ranges of capacity drops once breakdown occurs, somewhere between 2% and 16%.

On the other hand, literature is also found to dispute the two-capacity phenomenon. In a study by Banks (1991), he studied the two-capacity issue using the data collected at four sites in San Diego, California. Although he supported the existence of the two-capacity phenomenon while examining the flows on individual lanes, he concluded that there is no evidence, when looking at traffic across all the lanes, that the flows are significantly different before and after breakdown. He actually found that the flows increased after breakdown occurred in three out of the four sites examined. Ringert and Urbanik (1994) studied the freeway breakdown issue at three freeway merge locations in Texas. They found that the queue-discharge flow was higher at one site, lower at one site, and no difference at the other site.

By examining the methodologies used in the two studies by Banks and by Ringert and Urbanik, a few factors are suspected to have led to their conclusions. For example, Banks used a data collection point upstream of the merge (i.e., a location that does not include the ramp flows), which may not have reflected the true freeway capacity. The measured lower flows under free-flow conditions may be due to the fact that the pre-breakdown flows did not reach its maximum demand level. Strictly speaking, the true free-flow capacity should not be less than the queue-discharge capacity.

Nonetheless, all the literature tends to agree that freeway breakdown is probabilistic in nature, i.e., freeway breakdown could occur at different flow levels (Persaud et al., 2001; 1998; Lorenz and Elefteriadou, 2001; Elefteriadou et al., 1995). Random variations exist for the flows under both free-flow and queue-discharge-flow conditions. However, the variations on queue-discharge flows are generally smaller compared to that on the free flows.

2 Field Evidence of the Two-capacity Phenomenon

In this study, we used the detector data collected by the Ministry of Transportation of Ontario, Canada, at a freeway merge location near the Cawthra Rd. Queen Elizabeth Way interchange (Rajcoomar, 2003) to verify the two-capacity phenomenon. This site is far from the next downstream freeway merge, thus is free from influence by downstream queues. Two days of data were collected: one was on a Wednesday and one was on a Sunday. Each day included 24-hr continuous data from the freeway detectors.

The raw detector data included speed, occupancy, and flow based on 20-sec intervals. Figure 1 shows the flow and speed data (Wednesday) aggregated by 15-min moving average from the raw 20-sec interval data. Therefore, each point in the figure represents the average value from the previous 15 minutes. Using a longer time interval (e.g., 15 min) provides smoother lines; however, the details of flow-regime transition may not well be captured. Using a 15-min interval as shown in Figure 1 is consistent with the freeway capacity definition as given in the HCM for sustained flows.

![Time series flow-speed diagram](image)

As can be seen from Figure 1, the freeway experienced a sudden drop in speed at about 6:10 a.m., indicating the start of congestion and a possible start of a breakdown. While the speed continued to drop, the flows (demand) continued to rise until about 6:20 a.m. than started to drop. The speed dropped to below 60 km/hr and sustained until about 9:20 a.m., indicating the period of breakdown. When the data is plotted with 15-min moving average shown in Figure 1, the difference between the maximum flows during breakdown and before breakdown can be clearly seen. A lower flow rate (averaged about 6100 vph) under breakdown can be seen compared to the flow rate of about 6700 vph before
the breakdown. While much higher flows could be achieved within a short time interval (e.g., 20 second, 1 minute), the capacity flow is often measured at a much longer time interval, for example, a 15-minute period as defined in the Highway Capacity Manual. It is noted that the queue-discharge (i.e., under breakdown) flows were measured over a sustained longer period, thus a more reliable estimate on the queue-discharge capacity was achieved. However, close-to-capacity flows before breakdown only existed for a shorter period (less than 3 minutes). Such a short period generally results in an underestimation of the free-flow capacity. Examining on the Sunday data revealed similar trends to the Wednesday data. Based on the data shown in Figure 1, it clearly indicates that higher flows and higher capacity exist under the free-flow condition than under the congested condition, which supports the two-capacity hypothesis.

3 Ramp Metering and the Two-capacity Phenomenon

One of the major purposes of some ramp metering strategies is to maintain freeway in free-flow conditions by controlling the total traffic flow not to exceed the freeway capacity. In fact, it is the two-capacity phenomenon that determines the significance of ramp-metering applications. If no such a two-capacity phenomenon existed and freeway capacity were a single value, ramp metering itself would not be effective in reducing overall system delays. A numerical example is presented next to illustrate this point.

Before presenting the numerical calculations, the so-called cumulative arrival and departure method (McShane et al, 1998; Lawson et al., 1997) is first introduced. This method is used later to calculate the freeway and ramp delays in the examples. Equations 1 through 4 provide a complete description of the cumulative arrival and departure method in discrete form. In fact, this method can be used to estimate delay at any traffic facilities.

\[
\frac{\Delta A}{\Delta t} = \frac{A(t + \Delta t) - A(t)}{\Delta t} = V(t) \quad (1)
\]

\[
\frac{\Delta D}{\Delta t} = \frac{D(t + \Delta t) - D(t)}{\Delta t} = Q(t)
\]

\[
q(t) = \begin{cases} 
    c, & \text{if } A(t) > D(t) \\
    V(t), & \text{otherwise}
\end{cases}
\]

\[
q(t) = A(t) - D(t) \quad (3)
\]

\[
TD(t) = \frac{q(t - 1) + q(t)}{2} \Delta t \quad (4)
\]

where

- \(\Delta t\) = time step, sec or min or hr
- \(A(t)\) = cumulative number of vehicle arrivals at time \(t\), veh
- \(c\) = capacity of the facility, vps or vpm or vph
- \(D(t)\) = cumulative number of vehicle departures at time \(t\), veh
- \(A(t)\) = throughput flow rate at time \(t\), vph
- \(q(t)\) = queue length at time \(t\), veh
- \(V(t)\) = traffic demand at time \(t\), vps or vpm or vph
- \(TD(t)\) = total vehicle delays during the period \((t - 1, t)\) in veh-sec or veh-min or veh-hr depending on the unit of \(\Delta t\)

Equation 1 defines the traffic demand at time \(t\), which is a derivative of the cumulative arrival function, \(A(t)\). Equation 2 defines the throughput of the facility, which would equal to the capacity if the cumulative arrival is greater than the cumulative departure (i.e., there is a queue at the facility as calculated from Equation 3). Otherwise, the throughput would equal to the demand. Equation 4 calculates the total delay during the period of \(t - 1\) and \(t\). The cumulative arrival and departure method described in Equations 1 through 4 calculates delays at a traffic facility in a deterministic manner. When applied to analyze ramp metering operations, it resembles a fixed-rate metering operation.

Table 1 and Table 2 illustrate the delay calculations using the cumulative arrival and departure method for the two cases: 1) with the two-capacity phenomenon, and 2) with the single-capacity phenomenon. In the examples, a fixed ramp-metering rate of 1200 vph is assumed. The traffic demand profile is created in such a way that at the beginning of the analysis period, the freeway mainline demand, \(V(t)\), is less than the free-flow capacity, \(c_F\), but the total freeway mainline demand and the ramp demand exceeds \(c_F\). Such a demand profile represents a situation where the freeway would
break down without ramp metering, but would remain at free flow with ramp metering at an appropriate rate.

For the calculations in Table 1, the free-flow capacity, $c_f$, is assumed to be 6600 vph, and the queue-discharge capacity, $c_d$, is assumed to be 6000 vph. For each 5-min interval ($t = 5\text{min}$), the freeway demand, $V_t(t)$, and the ramp traffic demand, $R(t)$, are given in vph, and the vehicle queues and delays are calculated. For the calculations in Table 2, the capacity is assumed to be 6500 vph.

For the No Metering scenario in Table 1, the freeway has a lower capacity of $c_d = 6\ 000$ vph at the beginning interval (0–5) of the analysis period, because the total freeway and ramp demand exceeds the free-flow capacity. It represents a case of freeway breakdown, so the lower queue-discharge capacity is used. The queue length at the end of the first interval is calculated at $(7\ 000 - 6\ 000) 60 \times 5 = 83$ vph, and the delay is calculated at $(0 + 83) 2 \times 5 = 60 = 3.5$ veh-hr. For the second interval (5–10), the queue reaches 166 veh (83 from the previous interval and 83 from the current interval). For the case of ramp metering, however, the freeway is able to maintain at a free-flow condition, thus has a higher capacity (6 600 vph) at the beginning interval (0–5) of the analysis period. In this case, freeway traffic has zero queues and delays, but the ramp traffic has a queue length of $(2\ 000 + 1\ 200) 60 \times 5 = 67$ vph, and delays of $(0 + 67) 2 \times 5 = 60 = 2.8$ veh-hr. Overall, ramp metering is able to reduce the system delays from 315 veh-hr to 289 veh-hr with the assumption of the two-capacity phenomenon. Similar calculations are carried out in Table 2 assuming a single capacity of 6500 vph. In this case, ramp metering actually increases the system delays from 168 veh-hr to 198 veh-hr.

The above sample calculations proved an earlier statement that ramp metering would be effective in reducing system delays only if the two-capacity phenomenon existed.

Table 1 Delays with without ramp metering: single-capacity phenomenon

<table>
<thead>
<tr>
<th>$t$ (min)</th>
<th>$V_t(t)$ vph</th>
<th>$R(t)$ vph</th>
<th>$c(t)$ vph</th>
<th>$q_f(t)$ vph</th>
<th>$q_d(t)$ vph</th>
<th>$TD_f$ veh-hr</th>
<th>$TD_d$ veh-hr</th>
<th>$O_f(t)$ veh</th>
<th>$O_d(t)$ veh</th>
<th>$V_r(t)$ vph</th>
<th>$R(t)$ vph</th>
<th>$c(t)$ vph</th>
<th>$q_f(t)$ vph</th>
<th>$q_d(t)$ vph</th>
<th>$TD_f$ veh-hr</th>
<th>$TD_d$ veh-hr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–5</td>
<td>5 000</td>
<td>2 000</td>
<td>6 000</td>
<td>83</td>
<td>3.5</td>
<td>1 200</td>
<td>67</td>
<td>6 200</td>
<td>6 600</td>
<td>0</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–10</td>
<td>5 000</td>
<td>2 000</td>
<td>6 000</td>
<td>166</td>
<td>10.4</td>
<td>1 200</td>
<td>134</td>
<td>6 200</td>
<td>6 600</td>
<td>0</td>
<td>8.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–15</td>
<td>8 000</td>
<td>2 200</td>
<td>6 000</td>
<td>517</td>
<td>28.5</td>
<td>1 200</td>
<td>217</td>
<td>9 200</td>
<td>6 000</td>
<td>267</td>
<td>25.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15–20</td>
<td>8 000</td>
<td>1 000</td>
<td>6 000</td>
<td>767</td>
<td>35.3</td>
<td>1 200</td>
<td>200</td>
<td>9 200</td>
<td>6 000</td>
<td>533</td>
<td>50.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–25</td>
<td>4 000</td>
<td>800</td>
<td>6 000</td>
<td>667</td>
<td>59.7</td>
<td>1 200</td>
<td>167</td>
<td>5 200</td>
<td>6 000</td>
<td>467</td>
<td>56.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25–30</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>550</td>
<td>50.7</td>
<td>1 200</td>
<td>117</td>
<td>5 200</td>
<td>6 000</td>
<td>400</td>
<td>47.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–35</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>433</td>
<td>41.0</td>
<td>1 200</td>
<td>67</td>
<td>5 200</td>
<td>6 000</td>
<td>333</td>
<td>38.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35–40</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>317</td>
<td>31.3</td>
<td>1 200</td>
<td>17</td>
<td>5 200</td>
<td>6 000</td>
<td>267</td>
<td>28.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40–45</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>200</td>
<td>21.5</td>
<td>800</td>
<td>0</td>
<td>4 800</td>
<td>6 000</td>
<td>167</td>
<td>18.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45–50</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>83</td>
<td>11.8</td>
<td>600</td>
<td>0</td>
<td>4 600</td>
<td>6 000</td>
<td>50</td>
<td>9.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50–55</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>0</td>
<td>3.5</td>
<td>600</td>
<td>0</td>
<td>4 600</td>
<td>6 000</td>
<td>0</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55–60</td>
<td>4 000</td>
<td>600</td>
<td>6 000</td>
<td>0</td>
<td>0</td>
<td>600</td>
<td>0</td>
<td>4 600</td>
<td>6 000</td>
<td>0</td>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: 1. $V_t(t)$ is the freeway demand at $t$; $R(t)$ is the ramp demand at $t$; $c(t)$ is the capacity at $t$; $q_f(t)$ and $q_d(t)$ are queue length at $t$ on freeway and ramp; $TD_f$ and $TD_d$ are total delays on freeway and ramp; $O_f(t)$ is the throughput flow from the ramp.

2. $c(t)$ is equal to $c_f$ if $q_d(t - 1) = 0$ and $V_r(t) + O_d(t) \leq c_f$; otherwise, $c(t)$ is equal to $c_d$.

3. For the case of No Metering, queues and delays are calculated together for freeway and ramp.

4. $c_f = 6\ 000$ vph, $c_d = 6\ 000$ vph.

### 4 The Two-capacity Phenomenon in Traffic Simulation Models

As earlier discussions indicate that the two-capacity phenomenon is an important aspect in modeling ramp-metering and freeway operations. Therefore, any simulation model, when used for studying ramp-metering related issues, should have the capability of producing the two-capacity phenomenon.

While microscopic simulation models have been
widely used in evaluating ramp-metering algorithms (Hasan et al., 2002; Zhang et al., 2001; Chu et al., 2002), no literature has been found to document whether the simulation models in their studies were able to produce this freeway operational feature. In fact, microscopic simulation models (e.g. VISSIM) that are developed based on traditional car-following theories do not automatically yield the two-capacity phenomenon (ITC Private Communication, 2004).

Table 2 Delays with and without ramp metering: single-capacity phenomenon

<table>
<thead>
<tr>
<th>t (min)</th>
<th>V_r (t) (vph)</th>
<th>R(t) (vph)</th>
<th>( q_r(t) ) (vph)</th>
<th>( q_r(t) ) (veh)</th>
<th>TD_r (veh-hr)</th>
<th>Q_d(t) (vph)</th>
<th>Q_d(t) (veh)</th>
<th>TD_d (veh-hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>5000</td>
<td>2000</td>
<td>6500</td>
<td>42</td>
<td>1.7</td>
<td>1200</td>
<td>67</td>
<td>6200</td>
</tr>
<tr>
<td>5-10</td>
<td>5000</td>
<td>2000</td>
<td>6500</td>
<td>83</td>
<td>5.2</td>
<td>1200</td>
<td>133</td>
<td>6600</td>
</tr>
<tr>
<td>10-15</td>
<td>8000</td>
<td>2000</td>
<td>6500</td>
<td>192</td>
<td>19.8</td>
<td>1200</td>
<td>217</td>
<td>9200</td>
</tr>
<tr>
<td>15-20</td>
<td>8000</td>
<td>1000</td>
<td>6500</td>
<td>392</td>
<td>41.3</td>
<td>1200</td>
<td>200</td>
<td>9200</td>
</tr>
<tr>
<td>20-25</td>
<td>4000</td>
<td>800</td>
<td>6500</td>
<td>458</td>
<td>44.1</td>
<td>1200</td>
<td>167</td>
<td>5200</td>
</tr>
<tr>
<td>25-30</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>300</td>
<td>31.6</td>
<td>1200</td>
<td>117</td>
<td>5200</td>
</tr>
<tr>
<td>30-35</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>142</td>
<td>18.4</td>
<td>1200</td>
<td>67</td>
<td>5200</td>
</tr>
<tr>
<td>35-40</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>59</td>
<td>5.9</td>
<td>1200</td>
<td>17</td>
<td>5200</td>
</tr>
<tr>
<td>40-45</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>0</td>
<td>4800</td>
</tr>
<tr>
<td>45-50</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>0</td>
<td>0</td>
<td>600</td>
<td>0</td>
<td>4600</td>
</tr>
<tr>
<td>50-55</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>0</td>
<td>0</td>
<td>600</td>
<td>0</td>
<td>4600</td>
</tr>
<tr>
<td>55-60</td>
<td>4000</td>
<td>600</td>
<td>6500</td>
<td>0</td>
<td>0</td>
<td>600</td>
<td>0</td>
<td>4600</td>
</tr>
</tbody>
</table>

Total Delay = 108 veh-hr

Note: 1. \( V_r(t) \) is the freeway demand at \( t \); \( R(t) \) is the ramp demand at \( t \); \( q_r(t) \) is the capacity at \( t \); \( q_r(t) \) and \( q_r(t) \) are queue lengths at \( t \) on freeway and ramp; \( TD_r \) and \( TD_d \) are total delays on freeway and ramp; \( Q_d(t) \) is the throughput flow from the ramp.
2. \( q_r(t) \) has single value of 0, 500 vph.
3. For the case of No Metering, queues and delays are calculated separately for freeway and ramp.

An effort was made to see whether VISSIM can replicate the two-capacity phenomenon by modifying some of the coding and model parameters. It was found that only with careful selection of the model parameters can the two-capacity phenomenon be produced in VISSIM. The following summarizes the specific model coding requirements in order to achieve the two-capacity phenomenon.

It is noted, however, that these coding requirements only applied to VISSIM Version 3.60. With further development and enhancement, later versions of the software may not exhibit exactly the same characteristics. The required coding process in order to produce the two-capacity phenomenon includes:

- Code the freeway links as urban motorized road instead of freeway link.
- Use the Wiedemann-74 car-following model instead of the Wiedemann-99 model.
- Code the merging section no more than 400 feet.

In order to replicate the two-capacity phenomenon using simulation, an appropriate traffic demand profile must also be created, such as the one shown in Figure 2.

In Figure 2, the free-flow capacity and the queue-discharge capacity of the freeway bottleneck were assumed to be known as shown (note that the actual values need to be estimated from VISSIM). The hypothetical traffic demand profile was created to ensure freeway
breakdown occurrence by including a period during which demand exceeded the presumed free-flow capacity. On the other hand, a reasonable period of free-flow conditions should also exist with sufficient demand so that the free-flow capacity can be estimated. As pointed out earlier, while the queue-discharge capacity can normally be observed over a prolonged period, the free-flow capacity is more difficult to obtain because the flows for estimating it are more unstable. Estimating the free-flow capacity using measured flows would always result in an underestimation because the freeway demand has to be lower than the free-flow capacity in order to maintain free-flow conditions. At a demand level near the free-flow capacity, free-flow conditions could only last for a short period before breakdown would occur. Of course, in a microsimulation environment, it may require a number of trial-and-error runs in order to achieve the above demand profile and a good estimate of the free-flow capacity.

Based on a generic ramp merge site with three lanes on the freeway and the hypothetical traffic demand profile shown in Figure 2, an investigation on the two-capacity phenomenon in VISSIM was conducted under two conditions: without ramp metering and with ramp metering. Figure 3 illustrates the flow-speed plot obtained from VISSIM based on 20-sec intervals when ramp metering was not present. The 15-min moving average flow and speed lines are also shown. The data were collected in simulation at a detector location near the end of the merge lane, where the queue formation was observed while still capturing both the freeway traffic and the ramp traffic. As can be seen, the plot closely resembles the field data shown in Figure 1, where the two flow regimes can be clearly identified. It is noted that calibration of the simulation model to any field data was not attempted. From the figure, the free-flow capacity was estimated at \( C_F = 7220 \) vph, and the queue-discharge capacity was estimated at \( C_Q = 6320 \). The data presented in this figure were based on a particular simulation run with a random seed of 100. Also noted from the figure is that much higher flows are observed with the shorter 20-sec intervals than with the 15-min intervals. The capacity values were determined based on the 15-min moving average flow values. The free-flow capacity was estimated based on the highest 15-min moving average flow rate before breakdown, while the queue-discharge capacity was estimated based on the average flows during the period of breakdown. As pointed out earlier, the free-flow capacity may have been underestimated because the 15-min flows may have included flows that had not reached the highest. Under free-flow conditions, the average speed on the freeway was above 100 km hr. Under apparent breakdown conditions, the average speed was slightly above 60 km hr.

![Fig. 3 Speed-flow diagram from VISSIM](image)

Figure 4 illustrates the time series flow-speed plots from different simulation runs based on different random seeds. The figure clearly indicates the stochastic nature of freeway breakdown and its maximum flows, i.e., even with similar average traffic demand levels, freeway breakdown could occur under different conditions with different maximum flows.

Figure 5 illustrates the estimated capacity values from 10 different simulation runs with different random seeds. The average free-flow capacity was estimated at 7040 vph with a standard deviation of 110 vph, and the queue-discharge capacity was estimated at 6500 vph with a standard deviation of 55 vph. The variation in the queue-discharge capacity under breakdown conditions was smaller than that under the free-flow conditions, which appeared to be consistent with the findings from previous studies (Cassidy and Bertini, 1999; Persaud et al., 1998).

The simulation results presented above indicate that the two-capacity phenomenon was adequately reflected in VISSIM when ramp metering was not present. However,
traffic flow characteristics were found to be different when ramp metering was present. Figure 6 through Figure 8 illustrate the flow-speed diagrams from three simulation runs with the same random seed but with different ramp controls. Figure 6 shows the case when ramp metering was not turned on; Figure 7 shows the case when ramp metering was on but without queue flush; and Figure 8 shows the case when the ramp metering was on but with queue flush. Queue flush is a type of ramp-metering policy where the ramp metering operation is suspended whenever a large ramp queue is detected. Queue flush has been used as means of controlling queue spillback onto surface-street system; however, it would significantly diminish the efficiency of ramp metering operations (Tian, 2004).

Fig. 4 | Speed-flow diagrams from different VISSIM runs

![Speed-flow diagrams from different VISSIM runs](image1)

Fig. 5 | Capacity values from 10 VISSIM runs

![Capacity values from 10 VISSIM runs](image2)

Fig. 6 | Flow-speed diagram from VISSIM without ramp metering

![Flow-speed diagram from VISSIM without ramp metering](image3)

In simulation runs, the free-flow capacity without ramp metering was about 7260 vph, while its queue-discharge capacity was about 6290 vph (see Figure 6). With ramp metering and no queue flush, the two capacities were about 7130 vph and 6760 vph, respectively (see Figure...
7), indicating a drop in the free-flow capacity, but an increase in the queue-discharge capacity. The drop in free-flow capacity with ramp metering was due to the ramp traffic, which involved acceleration to the desired speed. The acceleration process resulted in reduced flows. Furthermore, with ramp metering and queue flush, it is no longer easily differentiate the two capacity regimes (see Figure 8). The lower estimated free-flow capacity in Figure 8 was simply due to the low demand during the counting period.

![Flow-speed diagram from VISSIM with ramp metering and no queue flush](image)

Fig. 7 Flow-speed diagram from VISSIM with ramp metering and no queue flush

With ramp metering and queue flush (results shown in Figure 8), the freeway at the merge location seemed to frequently transition between breakdown (lower speed) and free-flow (higher speed), which was visually observed in VISSIM as a persistent shockwave movement on the freeway. The average flows under this traffic condition could be seen at approximately 6880 vph.

Similar observations can also be found for the speeds. With ramp metering (see Figure 7), the speeds under free-flow conditions were lower than those without ramp metering. This was due to the ramp traffic accelerating to its desired speed after entering the freeway. The influence of ramp metering operations on both flows and speeds suggests that ramp metering does show some effectiveness in minimizing the capacity drop after breakdown, which agrees with the findings by Zhang and Levinson (2004) based on field studies, i.e., ramp metering can generally yield higher freeway flows under breakdown conditions than without ramp metering.

![Flow-speed diagram from VISSIM with ramp metering and queue flush](image)

Fig. 8 Flow-speed diagram from VISSIM with ramp metering and queue flush

pomenon from the perspective of microsimulation experiments. The study specifically emphasizes the importance of the two-capacity phenomenon in evaluating freeway and ramp-metering operations. Using the VISSIM traffic simulation model, some traffic features related to freeway breakdown and the two-capacity phenomenon are investigated. Major conclusions of this study are summarized below:

The study findings seem consistent with the two-capacity hypothesis in freeway operations. Although the level of capacity drop after breakdown may vary, the free-flow capacity should at least be equal to or greater than the queue-discharge capacity.

In general, maximum flows for capacity estimation under free-flow conditions are more difficult to obtain than those under queue-discharge conditions. As a result, high variations on the free-flow capacity usually exist and an underestimation of the capacity generally results. In addition, the location to measure the flows for capacity estimation must be downstream of the merge and free from downstream congestion and queueing.

With adequate calibration of the model parameters, VISSIM can well produce the two-capacity phenomenon. Results from the simulation revealed that the two-capacity phenomenon is more evident when there is no ramp metering. With ramp metering and queue flush operation, freeway could remain in constant transition between free flow and breakdown, thus making it difficult to observe the two-capacity phenomenon. However, this feature may need further validation based on field studies.

5 Summary and Conclusions

This study addresses the freeway two-capacity phe-
References


