A Case for Geometrically-Based Roundabout Capacity Equation Modeling

by

Mark T. Johnson, P.E.
MTJ Engineering, LLC
313 Price Place, Suite 5
Madison, WI 53705
Tel: 608-238-5000
mark@mtjengineering.com

and

David K. Hale, Ph.D.
Leidos, Inc.
11251 Roger Bacon Drive
Reston, VA 20190
Tel: 202-493-3296
david.k.hale@leidos.com

Word Count: 3,530 words text + 13 tables/figures x 250 words (each) = 6,780

(Original submission: November 5, 2014)

Revision submitted: March 19, 2015
ABSTRACT

Roundabout-controlled intersections have become prevalent in the United States, and the software models for analyzing such intersections are in heavy use across the country. Most of these software models rely on two basic analysis methods. One is a gap acceptance-based assessment of intersection performance, captured in the 2010 Highway Capacity Manual roundabout equation. The gap acceptance model was calibrated across a series of T-shaped intersections, offering only a limited set of geometric characteristics. The second basic analysis method is a statistically-derived model, developed in the UK, and based on the work of Kimber/Hollis. The resulting main equation utilizes six key geometric parameters for assessing roundabout performance. Engineers in the US have expressed concern over the likelihood of driver behavior differing substantially from country to country, and even from region to region. However, this paper synthesizes existing evidence to suggest that geometrics have a much stronger impact on capacity than geographic location, country of origin, or gaps that were calibrated under these assumptions. Moreover, there is significant data to indicate that the UK-developed Kimber equations are transferable to the US, and to other international conditions/drivers.

INTRODUCTION

This paper will provide an overview of (geometrically-based) roundabout capacity equation development in the UK. The overview will include traffic conditions and roundabout design principles that were in effect at the time of development, plus the significance of these conditions and principles. Additionally, we will focus in on the primary areas of disagreement by researchers, regarding the benefits of geometrically-based equations versus gap-based theoretical models.

A key component of this lasting debate is centered on the issue of Kimber’s findings; which show a continuous capacity increase caused by incrementally wider entry widths, often associated with ‘flared’ entries (single-lane approach flaring to a wider entry). This continually-increasing capacity contradicts conventional traffic theory and its equations. The conventional models were originally developed to analyze signalized intersections, but were later retrofitted for roundabout compatibility. In addition, these models must assume full lane widths to work correctly.

It appears that many researchers have not been able to discern capacity increases with sub-lane widening within the data they have available to collect, and have erroneously concluded that geometry-based methodologies and equations are not suitable for use outside the UK. The primary reason for this discrepancy is the UK roundabout design principles, which deliberately facilitated capacity increase due to sub-lane widening (via flared entries). These design principles are discussed later in the paper. Together with fully-saturated conditions, these design principles provided unique conditions from which to discover the geometric relationships. It is this perspective that researchers often lack – an understanding of the underlying geometric design attributes and traffic conditions necessary to replicate UK model predictions.
METHODS OF CAPACITY MEASUREMENT

Proper assessment or development of capacity equations begins with the method of capacity measurement. There are two common methods:

1. Direct Capacity Measurement

Capacity can be measured directly by repeatedly counting entering traffic and circulating traffic, under sufficiently congested conditions, typically for one- or two-minute intervals. A predictive relationship between capacity and circulating flow can then be derived by regression, without making theoretical assumptions about driver behavior. A large sample of data is necessary to determine slopes and distributions of regression lines.

2. Indirect Capacity Measurement

Predictive relationships between capacity and circulating flow can also be developed by measuring gaps, and using the gaps to estimate capacity. However, as flows and volume/capacity (v/c) ratios increase, or as geometrics change, the same methods will produce different gaps, and estimate different capacities. Methods of indirect capacity measurement require various model adjustments to prevent capacities from increasing unrealistically, in response to increasing congestion levels. Thus gap-based predictions must be diluted and assisted by other predictive mechanisms; including drivers modifying their headway due to encroaching entering vehicles, gap forcing, priority reversal, and merging.

KIMBER/HOLLIS CAPACITY EQUATION DERIVATION

The UK has been building roundabouts for over 70 years. The first UK roundabouts were similar to the US traffic circles built from the 1930s to the 1950s. Entering traffic had priority over circulating traffic, or had equal priority with traffic circulating in the adjacent lane. The two streams had to weave to either exit or proceed around the roundabout. Traffic circles were inefficient, with poor capacity and poor safety. Long weaving sections made them large and expensive. At high traffic flows they were in danger of locking, sometimes requiring police intervention before traffic movements could resume.

In 1966 the UK changed traffic regulations, giving priority to the circulating traffic. Yield lines were introduced. Locking ceased, and capacities dramatically increased. It was the introduction of this priority rule that led to the subsequent radical improvements in UK roundabout design, pioneered by Frank Blackmore (1). The resulting capacity increases were welcome, but some approach capacities were unexpectedly low. The reason for this was not clear; as geometry, flow levels, and turning proportions were similar. A new method for estimating yield-line capacities was needed, as the existing method was now obsolete. As the new roundabouts superficially appeared to be a series of T-junctions, it was natural for designers to use T-junction techniques to estimate yield-line capacity, as nothing better was available.
A mathematical model was developed to estimate T-junction capacity. It considered the probability of side-road vehicles entering main-road traffic gaps. This method, called ‘gap theory,’ was pioneered in 1962 by J.C. Tanner, a researcher at the UK’s government agency called the Transport and Road Research Laboratory (TRRL) (1). However, their gap-based procedures were found disappointingly unsuitable for estimating roundabout entry capacities. At saturated roundabouts, the comparison between gap-theory predictions and observed capacities revealed substantial disagreement. This was discouraging, as many congested roundabouts needed immediate improvement.

How to do this was uncertain, especially in urban locations where extra land was limited and expensive. To bypass this considerable uncertainty, the usual practice was to over-design and incur extra costs. It was not until several years later that over-design was discovered to substantially increase accidents. At this time, the UK had the unique advantage of having many saturated roundabouts, of different sizes and traffic flows. Thus, TRRL began an extensive research project into driver behavior and capacity at roundabouts.

Development of the UK Capacity Estimation

In the late 1970s, Professor Rod Kimber (then Head of Junction Design Section at TRRL) spearheaded a series of experiments on the TRRL test track, accompanied by extensive field measurements at UK road sites. On the TRRL test track, 35 different geometric parameters were tested for their effect on capacity. Six parameters per leg were found to be significant.

At 86 public road roundabouts, saturation capacity was observed, and detailed geometry was collected. This data, together with other data from various universities, was used to refine and calibrate the capacity model. Selected roundabouts operated at saturation capacity, and included a wide range for each of the six geometric parameters. A linear model was selected, as the non-linear model gave no better fit to the data. The field data included 11,000 minutes of at-capacity operation for approximately 500,000 vehicles. The statistically significant predictive variables are shown in Table 1 (2).

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Symbol</th>
<th>Data Range</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entry Width</td>
<td>e</td>
<td>3.6 - 16.5</td>
<td>Meters</td>
</tr>
<tr>
<td>Approach Half Width</td>
<td>v</td>
<td>1.9 - 12.5</td>
<td>Meters</td>
</tr>
<tr>
<td>Effective Flare Length</td>
<td>l&lt;sup&gt;1&lt;/sup&gt;</td>
<td>1.0 - ∞</td>
<td>Meters</td>
</tr>
<tr>
<td>Inscribed Circle Diameter</td>
<td>D</td>
<td>13.5 - 171.6</td>
<td>Meters</td>
</tr>
<tr>
<td>Entry Radius</td>
<td>r</td>
<td>3.4 - ∞</td>
<td>Meters</td>
</tr>
<tr>
<td>Entry Angle</td>
<td>ø</td>
<td>0.0 - 77</td>
<td>Degrees</td>
</tr>
<tr>
<td>Circulating Flow (Not Geometric)</td>
<td>Q&lt;sub&gt;c&lt;/sub&gt;</td>
<td>0 - 4,700</td>
<td>PCUs/hr</td>
</tr>
</tbody>
</table>
Calibration

Calibration of a model to local conditions is always desirable, but proper calibration is difficult and expensive. Given a population of roundabouts having identical geometries, all capacity models estimate mean capacities with a standard error of approximately ± 200 vehicles per hour. Proper calibration requires saturation capacity data at a sufficient number of sites.

For any given roundabout geometry, a distribution of entry capacities can be measured at different sites. Capacity measurement at a single roundabout only provides a random, unknown position on the capacity distribution. Until capacities are measured at a sufficient number of similar roundabouts, the mean capacity cannot be known with reasonable confidence. Even good data from a single (or a few) locations is insufficient for testing the accuracy of a capacity model; only the mean capacity line from a sufficiently large number of capacity lines is suitable for model comparison.

The 2010 Highway Capacity Manual (HCM) (3) roundabout capacity equation was derived from directly-measured capacity data, collected in the early 2000s. A non-linear (negative exponential) regression line was fitted to the data. The 2010 HCM recommends local calibration of capacity models to reflect local behavior. It specifies a calibration procedure using the observed critical headway and observed follow-up headway. The intercept and exponential coefficient can then be adjusted using gap parameters.

A primary concern with gap-based calibration procedures, even when done rigorously, is that the ‘calibration’ only applies to specific geometries, flows, and v/c ratios on the observed roundabouts. Therefore, those gaps will not accurately reflect a wide variety of differing geometries. As can be seen when calibrating the HCM 2010 procedure, small changes in gaps result in large changes to predictive capacity.

A recent U.S. case study (4) of the HCM 2010 calibration procedure illustrates this issue. The authors concluded that, based on their measured gaps, “further adjustment of the calibrated HCM model is likely needed.” They also stated that a “linear regression method provides a better fit to their collected data.” This better fit is illustrated below, in Figure 1. The study identified other difficulties with the HCM gap-based calibration procedure, including:

- It is difficult and expensive to collect enough data to analyze the critical headway.
- If one approach has much more data than other approaches, this approach may have an excessive impact on HCM model calibration.
- Critical headway and follow-up headway vary among the three study roundabouts, and even among the approaches at the same roundabout.
Thus while calibration of gap-based models may be necessary, it is a high-risk process. It must be carefully applied to avoid over-design; producing larger than necessary designs, which are less cost-effective and less safe. There is also a risk of under-design, which creates unexpected congestion too early in the facility’s design life.

Multiple peak-hour capacity lines derived from saturated conditions of the same entry are necessary to derive a statistically valid mean of the population of these capacity lines, as shown conceptually in Figure 2 below. With one line, or even a few lines, the population mean cannot be known. Moreover, location of the measured line in relation to this mean cannot be known.

![Single-Lane Roundabout Capacity for All Sites](image)

**FIGURE 1** Comparison of field data, linear regression model, and overall calibrated HCM model for all sites. (4)

![Multiple capacity lines of same entry.](image)

**FIGURE 2** Multiple capacity lines of same entry.
In the UK, “mean population” capacity lines were derived from a variety of geometries. Slopes of the lines are very robust. Calibration cannot be effective unless there is a sufficient population of capacity lines. The process is very expensive, and many saturated entries are required. However, the population of capacities has been previously derived in the UK. And once the US has sufficient amounts of roundabouts operating at saturated conditions, then the population mean can be compared with the UK population means, and a global calibration can be made if necessary.

TRANSFERABILITY TO OTHER COUNTRIES

Early US research erroneously concluded that the UK models over-predict US roundabout capacities. This raised questions about the transferability of non-US models to US conditions. Such questions included: Are UK drivers more experienced with roundabouts: Is their driver behavior different than US drivers? Does the different size of cars and trucks have an effect? Will such differences persist into the future? Some have assumed that UK field capacity measurements may not apply to the US because UK drivers are more experienced with roundabouts, and drive smaller vehicles. (The smaller vehicle size may have been true in the 1970s, but is not true in 2015.) However, a review of recent saturated US capacity data vs. UK capacity data shows very good correlation with Kimber’s equations (2), corroborating the transferability of these equations to US drivers (5).

As shown below in Figure 3, 4, and 5, US capacity data closely matches UK capacity data for both single-lane and dual-lane roundabouts.

![Figure 3 UK single-lane capacity predictions (2).](image-url)

**Line A:** Kimber’s prediction for single-lane closely matches US data.
Closely matches US data.

FIGURE 4 UK two-lane capacity predictions (6).

2012 US Data

FIGURE 5 Carmel, IN Linear and exponential models and maximum move-up data (5)
US application of Kimber/Hollis equations

The 2000 FHWA Roundabout Guide (7) and 2010 FHWA Roundabout Guide Second Edition (8) provide useful information and discussion on this topic. However, current US roadway and traffic planning standards and paradigms are predicated on signals. Below is an excerpt from the FHWA Roundabout Guide Second Edition (8) that provides a qualitative description:

1. FHWA Roundabout Guide 2010 Second Edition – 2.2.4 SPATIAL REQUIREMENTS (8)

   Roundabouts present opportunities to shape the cross section of a corridor in ways that are perhaps different from those afforded by signalized intersections.

   Signalized intersections operate most efficiently when they progress platoons of traffic, allowing the maximum number of vehicles to pass through on green without stopping. These platoons maximize the use of green time by promoting shorter headways. However, lane continuity between signals is needed to sustain these platoons through a series of signals, and the links tend to be underused between platoons.

   Roundabouts can be designed to accommodate node capacity, keeping the links between nodes more narrow. The resulting flow between roundabouts tends to be more random and makes more efficient use of the links between intersections.

Kimber’s equations analyze flared-entry roundabouts via the three primary capacity variables of approach width, entry width, and distance over which the entry is widened from the narrower approach width. This represents a significant change from conventional US traffic planning analysis, based on signalization and stop control procedures, which rely on adding or widening lanes to add capacity along the entire roadway section. Current US capacity research provides some qualitative guidance shown earlier; but provide no quantitative methods to correctly analyze a flared entry roundabout, or the correlation between its design and analysis.

Additionally, and as discussed earlier in this paper, there has been substantial confusion amongst traffic researchers and practitioners when evaluating what appears to be a flared entry. When the capacity effects are not present, researchers have erroneously concluded the flared entry does not help, and the analysis procedures that say otherwise are then also suspect. It was overlooked in these cases that in the US, single-lane roundabouts are not designed with the widened entry (shaped into a flare) to be utilized as a capacity mechanism. Instead, the entry has been widened only to accommodate large design vehicles, not for the operational effects that would be achieved with the dual-lane flared entry. This is a foundational missing component.

This issue is illustrated in the two pictures side-by-side in Figures 6 and 7 below. The southbound entry of the left-most roundabout is a single-lane design, with single circulating and exit lanes. The northbound entry for the right-most roundabout is clearly designed as a flared two-lane entry with effective geometrics, correct alignment entry to circulating and exit, plus corresponding striping and signing to illicit correct lane utilization of the entry.
Related to the above issue is the case when a two-lane entry is designed and analyzed for two lanes, but the geometric design and relationships are not conducive to entry utilization. This is illustrated in Figures 8 and 9; showing examples of designs that provided ineffective entry width (Figure 8), and effective entry width (Figure 9), respectively. The proper design facilitates effective use of the inside approach lane (9).

These examples illustrate the importance of the relationship of the geometric elements to achieve a close match between predicted and field measured operations. If the design elements and the associated visual messages to drivers aren’t congruent, this leads to ineffective lane utilization.

Figure 10 (design) and Figure 11 (aerial photo) illustrate a flared-entry design example. The single-lane approach roadway “flares” into two-lane entries, and two lanes circulating. Appropriate exit taper distances allow for the zippering or merging of traffic back to a single stream of traffic.
FIGURE 6 US Single-lane design – wide entry for trucks only. (Source: MTJ Engineering)

FIGURE 7 US Flared dual-lane entry design. (Source: MTJ Engineering)

FIGURE 8 Ineffective entry width. (Source: MTJ Engineering)
FIGURE 9  Proper entry design alignment.
(Source: WI DOT Facilities Development Manual) (9)

FIGURE 10  Effects of flared entry design geometry on operations.
(Source: MTJ Engineering)
This design was based on operational analysis from the software Rodel, which is predicated on Kimber’s equations for analysis of flared entries. It is designed for long-range traffic to meet acceptable operations on the 20-year traffic projections.

Video of the roundabout clearly shows how vehicles approach in the single-lane, stagger smoothly, and then double-up into each lane at the yield point. This roundabout exhibits the design elements and traffic distribution necessary for efficient operations. It will therefore produce a close match between model prediction, and measured field operations.

Design elements present include:
- Gradual smooth flaring from V to E
- Proper entry to circulating alignment
- Correct circulating to exit with sufficient receiving distance and taper length as a function of speeds and volumes
- Correct signing and pavement marking to elicit the splitting into two lanes

Figure 12 illustrates a screen shot from this video. It shows drivers naturally utilizing the flared entries, producing capacity that will closely match model predictions. This is due to the geometric elements, pavement markings, and signing present to elicit the inherent and natural driving behavior. The fact that this is the first roundabout in the small community of Waunakee (population 12,000 in Dane County Wisconsin) suggests that driver “experience” in different countries should not inhibit transferability of geometric-based modeling, as some researchers have implied.
Beyond the traffic operational benefits, installation of a flared-entry roundabout at this location helped to maintain a narrow cross-section through its historic downtown. This further allowed for on-street parking, bike lanes, wider pedestrian facilities, and other associated treatments, consistent with a Complete Streets design.

CONCLUSIONS

The geometrically-derived roundabout models represent a significant change from gap-derived signal and stop-control models, which were retrofitted to analyze roundabouts. Only Kimber’s equations account for the substantially different operational characteristics that roundabouts exhibit, due to natural driver behavior.

Kimber’s equations have proven to be statistically sound for approximately 25 years (10). This robustness can be attributed to the large statistical database on which the original research was performed. Some have said it will be difficult to replicate this original work in other countries; given the unique conditions available at the time, including widely-varying geometries and saturated demands (1, 6). But with improved understanding of the necessary data collection efforts, and stronger understanding of Kimber’s work, new research should be capable of substantiating and validating Kimber’s equations for drivers outside of the UK.

This paper showed that when roundabouts are designed with the essential elements for specific traffic distributions, flared entry roundabouts work and perform as discussed in the UK research literature that formed the basis for Kimber’s equations. This includes zippering/staggering, and doubling-up at entry. This occurs because the inherent and natural behavior present in drivers (irrespective of country of origin) is elicited with these design
elements present, and drivers respond accordingly as predicted by Kimber’s equations. When these design elements are not present, drivers react and behave differently.

More research and data on US saturated roundabout capacity is needed over a wide range of traffic flows and, importantly, geometry. It has become increasingly obvious that measuring capacity at non-saturated roundabouts without a good understanding of the issues presented in this paper can lead researchers to erroneous conclusions. Newer data, collected by FHWA during saturated conditions, shows US roundabouts operating very close to predicted UK capacities. These new results should motivate new efforts to transfer the extensive UK research to other countries, thus enabling accurate and robust designs across a wide range of conditions.
REFERENCES


