Street Condition Ratings: Their Use among North Carolina Cities

David N. Ammons and T. Dwane Brinson

Many cities across the nation report a single statistic or sometimes a small cluster of statistics to reflect the condition of the community’s roadways. Some report a statistic called the Pavement Condition Index (PCI), while others declare their Pavement Condition Rating (PCR), Pavement Service Rating (PSR), Overall Condition Index (OCI), or any of several similar ratings (Yu, Chou, and Luo 2007). Typically, a city will report a single average score for all streets, the percentage of streets with ratings exceeding or falling below a given score, the percentage of streets deemed to be in acceptable condition, or a pair of these statistics in combination.

This bulletin examines the prevalence of street condition ratings among municipalities in North Carolina, the frequency of condition assessments, the tendency to perform ratings in-house or by contract, and the uses for ratings. It also examines the practice of some cities to report summary statistics among their street maintenance performance measures and considers the validity of comparisons across cities.

Street Condition Ratings

A review of the literature on street condition rating systems reveals an almost bewildering array of rating systems. More than two dozen different systems exist, with labels such as Pavement Condition Index (PCI), Pavement Condition Rating (PCR), International Roughness Index, Pavement Distress Index (PDI), Pavement Serviceability Rating (PSR), and Surface Distress.
Index (SDI). The PCI is one of the most popular rating systems nationally. It was created by the U.S. Army Corps of Engineers and focuses on structural integrity, structural capacity, roughness, skid resistance, and rate of distress (Zhou et al. 2010). Among North Carolina cities, however, the PCR is the predominant choice.

Although the various rating systems differ from one another in ways large or small, broadly speaking their purposes are similar. Each is designed to assess systematically the condition of roadways or a particular aspect of road condition. Typically, scores are assigned (or points deducted) based on specified guidelines. For instance, one rating system focuses on eighteen different road condition variables, including potholes/debonding and patching. In this system, scores for the potholes/debonding variable distinguish between street sections having fewer than five potholes per mile, sections with five to ten potholes per mile, and sections with more than ten potholes per mile. Similarly, scores for patching distinguish between street sections having fewer than ten patches per mile, sections with ten to twenty patches per mile, and sections having more than twenty patches per mile (Highway Preservation Systems 2001).

The guidelines for scoring street conditions are explicit. Efforts are made to minimize the degree of subjectivity in each of these rating systems.

Research Design
Public officials—primarily public works directors—in the 106 North Carolina cities and towns with populations of 5,000 or greater were surveyed electronically by the authors in August 2010 regarding street condition ratings. They were asked about rating practices, rating uses and usefulness, and any tendencies to compare their own rating with the street condition ratings of other communities. After a series of follow ups, responses were received from eighty-six municipalities, yielding a response rate of 81 percent.

Findings among North Carolina Cities
Among cities and towns (hereafter referred to simply as “cities”) responding to the August 2010 survey, 83 percent periodically compile a comprehensive assessment of the condition of their streets (i.e., they rate their streets), while the other 17 percent do not. Although more populous cities are slightly more likely than less populous cities to have street condition ratings ($r = 0.16$), some very small cities are among the group that compiles street condition ratings (e.g., Whiteville [pop. 5,148] and Morrisville [pop. 5,208]), while some fairly large cities are not (e.g., New Bern [pop. 23,128] and Lumberton [pop. 20,795]). Nevertheless, the influence of city size is apparent, with the street-rating group among all cities with 5,000 residents or more having a median population of 41,229, and the non-rating group having a median population of 11,379.

---

1. Among the various rating systems are these: Pavement Condition Index (PCI), Pavement Condition Rating (PCR), Profile Index, Roughness Index (RI), Cracking Index (CI), Critical Condition Index (CCI), Distress Index (DI), International Roughness Index, Mean Panel Rating, Overall Pavement Condition (OPC), Overall Pavement Index (OPI), Pavement Condition Evaluation System (PACES), Pavement Condition Survey (PCS), Pavement Distress Index (PDI), Pavement Quality Index (PQI), Pavement Serviceability Rating (PSR), Present Serviceability Index (PSI), Present Serviceability Rating, Public Ride Perception Index (PRPI), Remaining Service Life (RSL), Ride Number, Ride Quality Index (RQI), Sufficiency Rating (SR), Surface Distress Index (SDI), and Surface Rating (SR).
Among the fifteen responding cities that did not systematically rate the condition of their streets, more than half indicated that insufficient staff time or staff capacity was at least part of the reason for their decision, though some also cited the expense of street condition ratings or questioned their value (Figure 1).

**Contracting Out or Performing Ratings In-House**

Most of the cities with street condition ratings contract out the rating work, with only 26 percent performing the work in-house using municipal staff. Contracting cities either work with the Institute for Transportation Research and Education (ITRE) at North Carolina State University in partnership with an engineering consultant (30 percent of the cities that have street condition ratings do this) or contract directly with an engineering consulting firm (44 percent).²

ITRE has a long history of work in roadway condition ratings. When approached by several municipal officials in 1983 to assist their cities with street condition ratings, ITRE was already doing similar work with the North Carolina Department of Transportation (NCDOT) on state highways (ITRE n.d.). Only slight modifications were needed to convert the NCDOT’s PCR system for municipal purposes. The resulting assessment measured eight types of pavement distress for each section of roadway and produced the pavement condition rating, reported conditions warranting particular attention, prescribed maintenance activities, and estimated the repair costs for each street. Altogether, ITRE has worked with more than 160 North Carolina municipalities and conducted more than 400 pavement management surveys or updates (ITRE n.d.).

---

². Because the contracts for ITRE street rating projects are formally between the municipality and ITRE’s partnering engineering consultant, it is possible that the percentage of ITRE projects is understated in these responses and actually constitutes a larger portion of the 74 percent of street ratings that are contracted out.
Frequency of Ratings
Three-fourths of the cities that compile street condition ratings report doing so at least every five years; 45 percent report a frequency rate of three years or less; and 16 percent report compiling new street condition ratings annually. There is only a modest tendency for cities that rate their streets frequently to be more populous (r = 0.111, nonsignificant) but a much stronger tendency for those that update their ratings regularly to handle the rating duties in-house rather than by contract (r = 0.529, p < .05) (Figure 2).

One of the items reported by cities responding to the School of Government survey was the date of their most recent street condition rating. Even reporting at mid-year (August), 15 percent of the cities had already rated their streets during the current year and 42 percent had done so since the beginning of 2009—that is, a rating had been completed within the previous nineteen months (Figure 3). This pace is slightly ahead of the normal frequency of ratings declared by respondents.3

Some cities base their ratings on the condition of a sample of streets, while others include all streets in their review. Most North Carolina survey respondents (88 percent) reported that all streets were included in their most recent assessments.

Rating System Choice and Techniques
As was previously mentioned, among North Carolina cities that systematically rate their street inventories, the Pavement Condition Rating (PCR) is the overwhelming choice of assessment tool. Eighty-five percent of the School of Government survey respondents use the PCR, sometimes in combination with other systems, compared to 6 percent who use the Pavement Condition Index (PCI) and 19 percent divided among a variety of other rating systems.4

The rating scales for the PCR and PCI both range from 0 to 100. Score interpretations, however, differ. For the PCR, 91–100 is considered very good; 81–90, good; 66–80, fair; 51–65, poor; and below 51, very poor (B&F Consulting 2010).5 For the PCI, 86–100 is considered excellent; 71–85, very good; 56–70, good; 41–55, fair; 26–40, poor; 11–25, very poor; and 0–10, failure. Individual cities often are found to depart from these prescriptions as they set their own thresholds of acceptable road condition scores (Xie and Levinson 2007).

Most North Carolina cities use a 100-point rating scale, but considerable variation exists regarding the score that different cities consider satisfactory. The median satisfactory score among survey respondents with 100-point scales was 80. While some of these cities considered a score as low as 50 to be satisfactory, a few others set their threshold score at 90.

The literature on street condition rating methods lists a wide variety of techniques and equipment used by raters in arriving at their assessments, including profiling devices (contact and noncontact), falling weight deflectometers, ground penetrating radar, Automated Road Analyzer (ARAN) systems with Wisecrax software, roughometers, and dipsticks (e.g., Han and Lund 2007). Altogether, sixteen of these techniques were mentioned to the officials surveyed, but only

---

3. Being ahead or behind the declared pace of ratings could simply be a matter of the survey’s timing. If several cities that rate their streets only every four to ten years recently completed their ratings, this could cause the group as a whole to appear to be ahead of the normal pace.

4. The percentages for different rating systems sum to more than 100 percent because some cities reported using more than one system.

5. Variation can exist even in the official interpretation of a single system’s scale. For instance, a different source explains PCR scoring as 91–100, very good; 76–90, good; 66–75, fair; 56–65, fair to poor; 41–55, poor; and 0–40, very poor (Highway Preservation Systems 2001).
six techniques were reported by two or more respondents as being used in the rating of their streets. Visual inspection by raters was reported as the most-used method overall, as stated by 64 of the 70 officials (91 percent) responding to a survey question on this point, and seven officials (10 percent) reported visual inspection of images by raters as their predominant method. Video photologs are used in two of the responding cities; automated distress collection vans in four cities; digital photo records in eleven; and hand odometer, straightedge, and ruler in nine of the cities. A city’s population had little apparent effect on choice of equipment or technique.
Complexity
Rating a street is a more technical task than most laypersons might realize. Among the street characteristics to be rated in major systems are the following: alligator cracking, bleeding, corrugations, edge cracks, longitudinal cracking, patching, potholes/debonding, raveling and weathering, rutting and shoving, settlement, and traverse cracking. Because so many North Carolina cities choose to contract out their rating duties, survey respondents—mostly public works directors—were asked about the perceived difficulty of rating each of these characteristics. The category labeled “potholes/debonding” was the only one not regarded by any of the respondents as difficult to rate. “Potholes/debonding” and “patching” were regarded as easy to rate by the largest numbers of respondents (86 and 81 percent, respectively), while “raveling and weathering” and “bleeding” were deemed difficult to rate by more respondents (19 and 18 percent, respectively) than the other conditions. This coincides with the findings of a study conducted in New Mexico, where comparisons of ratings by expert and beginning raters showed “bleeding” to be the most difficult type of distress to assess (Bogus et al. 2010).

Did a tendency for some respondents to regard more of the conditions as difficult to rate lead their cities to contract out the rating duties? Apparently not. No significant correlation was detected supporting such a relationship.

Rating Scale Uses and Usefulness
Local officials report a variety of uses for their street condition ratings (Figure 4). The most frequent uses were for establishing maintenance priorities, supporting budgetary and capital improvement requests, determining the maintenance needs of a given street, capital planning, and reporting on street conditions to the governing body.

The actual value of street condition ratings sometimes meets expectations for intended uses and sometimes falls short (Figure 5). North Carolina officials in cities with street ratings overwhelmingly found their ratings helpful in establishing maintenance priorities and were only slightly less enthusiastic in their endorsement of ratings for several other uses.

Preferred Street Condition Rating Statistics
When street condition ratings are reported among performance measures in local government budgets, they most often appear as a single score—typically the mean or average score for all streets. Occasionally, the percentage of streets with ratings exceeding a given score is reported, and sometimes the percentage with ratings falling below a given score is reported. What do public works directors and other survey respondents say is the most helpful way to depict the results of a street condition rating? The ranking of choices depends only slightly on whether the purpose is managing or reporting (Table 1). The first choice for the most helpful statistic in either case is a street-by-street listing of condition scores.

Whether for managing street maintenance or reporting street condition to others, respondents were less enthusiastic about the value of a grand mean of all street conditions than the frequent reporting of this statistic would imply. Even for reporting the community’s street condition to others, the mean is rated no better than a fifth-place choice, falling behind a listing of individual streets by their condition scores, the percentage of streets falling below a given score, the percentage exceeding a given score, and even the range of scores. It is likely that public works directors had in mind, as they responded to the questionnaire, the value of street-by-street ratings when called upon to explain to disgruntled citizens and elected representatives why another street is given priority for repairs over the one they favor. Individual reporting of
Street Condition Ratings: Their Use among North Carolina Cities

this kind is one thing, however; summary reporting is quite another. While the tabulation of ratings for each street would be unwieldy for most reporting documents, reporting the percentage of streets falling short of or exceeding a desired score—the second and third choices among respondents—would be a relatively simple matter. Apparently, the common tendency to report the mean is either a choice made by street officials without reflecting upon the possibility of better choices or it is simply the preference of the office compiling the reporting document.

Most lay recipients of street condition rating statistics—including mayors, city councils, city managers, citizens, and members of the news media—have little, if any, basis for judging the
numbers. Even if the public works director says that the city considers 70 to be a satisfactory score and 80 to be pretty good, an average score of, say, 76 leaves a great deal of information unrevealed. If the street-by-street scores yielding a mean of 76 are tightly clustered around the mean, then most of the streets might well exceed the threshold score of 70; if they are not so tightly clustered, a large portion of the street inventory might be sub-par. Perhaps this is why so many survey respondents preferred reporting the percentages above or below a specified target score.

### Suitability of Ratings for Comparison with Others

With the public works director’s declaration of a threshold score denoting satisfactory street condition in a given city, citizens and community decision makers can compare the condition of local streets to the desired score as they make their own assessments of the street inventory. Some citizens and officials, however, might desire a context broader than that provided by the public works director’s prescribed score—perhaps a comparison with the street condition ratings of other communities. Are such comparisons actually made? Would such comparisons be valid or fair?

Among respondents to the School of Government survey, 57 percent reported that they or others sometimes compared their own street ratings with others’ ratings, although only 13 percent of the respondents said that they did so “often.” The other 43 percent of respondents reported “never” comparing scores with other municipalities.

A perusal of city budgets and other reporting documents from municipalities across the nation shows many cities reporting the average PCI (Pavement Condition Index), PCR (Pavement Condition Rating), or whatever rating system is used for all city streets or the percentage of streets with ratings above (or below) a given score, without any explicit comparison to corresponding statistics for other communities. Occasionally, however, direct comparisons are made

### Table 1. Average Rank-Order Preference of Selected Street Condition Rating Statistics among Responding Public Officials

<table>
<thead>
<tr>
<th>Best Statistic for Use in Managing Street Maintenance</th>
<th>Best Statistic for Reporting Community’s Street Condition to Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>A listing of streets by their condition scores</td>
<td>A listing of streets by their condition scores</td>
</tr>
<tr>
<td>Percentage of streets falling below a given score</td>
<td>Percentage of streets falling below a given score</td>
</tr>
<tr>
<td>Range of scores (highest to lowest)</td>
<td>Percentage of streets exceeding a given score</td>
</tr>
<tr>
<td>Percentage of streets exceeding a given score</td>
<td>Range of scores (highest to lowest)</td>
</tr>
<tr>
<td>Community’s average (mean) score</td>
<td>Community’s average (mean) score</td>
</tr>
<tr>
<td>Community’s median score</td>
<td>Community’s median score</td>
</tr>
</tbody>
</table>

**Source:** David N. Ammons and T. Dwane Brinson, Electronic Survey of N.C. Public Officials, August 2010. School of Government, University of North Carolina at Chapel Hill.

**Note:** Average rankings were based on respondent rankings, with “most helpful” scored as 6, “second-most-helpful” scored as 5, and so forth. Averages for usefulness in managing street maintenance ranged from 1.49 to 5.55. Averages for usefulness in reporting to others ranged from 1.97 to 4.81.
in budgets or other reports. In an especially notable case, one fairly large southwestern city was even willing to compare across different rating systems, comparing its own average PCI (using a 0-to-100 scale) to the average Pavement Quality Index (PQI, using a 0-to-10 scale adjusted to achieve presumed comparability) of streets in a nearby city. Even if cross-jurisdictional comparisons might be considered reasonable when both jurisdictions use the same rating system, would ratings made under different systems hold sufficient comparability to warrant such comparisons?

It is important to recognize differences in components and rating guidelines designed by the originators of different rating systems. Various street defects are weighted differently by different rating systems (Chan, Whitman, and Elioff 2006). Although a perfect score might be 100 for several of these systems, a street inventory meriting an average score of 85 in one might not earn an 85 in another.

Recent studies examining the comparability of rating systems used by different state highway departments and by different cities in California’s Los Angeles County are instructive. The study of condition ratings of state roadways casts doubt on the viability of interstate comparisons primarily because of wide disparities in rating systems (Papagiannakis et al. 2009). Not only do rating methods, scoring computations, and rating scales differ among the states (for instance, some use a 100-point scale while others use a 5-point scale), but even those using a 100-point scale differ in their designation of a “good” score. For example, 75 is the lowest “good” score in some states’ systems, while a score of 50 makes the “good” range in other systems.

An effort in Los Angeles County to develop a comprehensive assessment of roadway conditions and the maintenance needs of eighty-nine separate jurisdictions within the county using twenty different rating systems required analysts to correlate the various indices used (Chan, Whitman, and Elioff 2006). Simple, direct comparisons across different rating systems were not considered reasonable.

Even if comparisons across different rating systems are problematic, what about comparison of the street condition of two jurisdictions, both using the same rating system—for example, both using the PCI or both using the PCR? The reasonableness of comparison in this case is more difficult to dismiss.

Clearly, a PCI-to-PCI or PCR-to-PCR comparison is not absolutely problem-free. Whenever ratings are prepared by different raters, questions of inter-rater reliability may arise. In fairness, however, it should be noted that reliability questions could be raised even within a single city when the ratings of different streets are compared—whether different raters assessed the condition of different streets or even when the same rater assessed streets at different times. Yet these potential reliability problems are considered insufficient to seriously restrict intracity comparison. No rating system is perfect. Even factors like pavement temperature and moisture, direction and angle of sunlight, and direction of observation can affect interrater reliability (Smith, Freeman, and Pendleton 1998). Modest imperfections in rating precision, however, do not impair the ability of local officials to assess the condition of streets overall or the relative condition of individual streets.

Scholars have noted the influence of rater experience and subjectivity on condition ratings, despite the presence of well-defined guidelines or criteria, and have developed methods to detect problems needing corrective action (Bianchini, Bandini, and Smith 2010). On the whole, much of the rating process is straightforward for experienced raters. Survey respondents tended to regard few street surface characteristics as especially difficult to rate. The same rationale that recognizes modest inter-rater reliability issues but regards these problems as insufficient to rule
out street-to-street comparisons within a single jurisdiction would seem to open the way to comparisons across cities as well.

In the case of North Carolina cities, most street condition ratings are performed by engineering firms following thoroughly specified rating guidelines. Professionalism and the desire to maintain a good reputation may be expected to exert the same influence for objectivity on these firms as their counterpart consulting professionals in accounting and law feel on projects in their fields of professional expertise. Furthermore, any desire to provide inflated ratings that could be touted by client cities in a form of civic boosterism is mitigated by the fact that a primary use of these ratings is to support capital planning and budgeting requests—where negative scores might be more helpful than positive scores.

The keys to the viability of comparisons of street condition ratings across different cities lie in the comparability of rating systems and the conscientiousness and integrity of raters. The comparability issue is best laid to rest by making comparisons only with cities using the same rating system—that is, comparing PCI scores only with other PCI-using cities and comparing PCR scores only with other PCR-using cities. The interrater reliability issue is perhaps best addressed by considering comparison not as absolute but only as approximate. For instance, the overall condition of streets in Denver, Colorado, with an average PCI reported at 72 in 2008, might fairly be considered approximately the same as the overall condition of streets in San Antonio, Texas, and Wichita, Kansas, with average PCIs of 70.2 and 72.12, respectively, but not as good as the condition of streets in Lake Havasu City, Arizona, or Shawnee and Leawood, Kansas, with PCIs of 78, 89.5, and 91.0, respectively. Judgments based on small differences in ratings—even within the same rating system—would assume too much precision; judgments based on larger differences and a willingness to accept close scores as approximately the same would not.

Comparisons of street condition ratings across cities would not be the first municipal service to be usefully compared despite modest vulnerability to concerns about subjectivity and inter-rater reliability. Comparisons across cities are made, for instance, of citizen ratings of park appearance and the perceived safety of neighborhoods; customer satisfaction with custodial services and fleet maintenance services; arson incidents per 10,000 population; hours billed as a percentage of mechanic hours available; and crime statistics—all good measures whose comparisons with other municipalities may be valuable to managers, but all of which are at least as vulnerable to reliability concerns as are comparisons of street condition ratings made under a single rating system. In fact, the elaborate rating guidelines of street rating systems and the tendency of some contractors to specialize in street ratings and perform ratings for several municipalities may place street comparison reliability in a more favorable position than many other often-compared services.

Conclusions
The systematic rating of street conditions is a widespread practice among cities nationally. Among North Carolina cities with populations of 5,000 or greater responding to the 2010 School of Government survey, 83 percent periodically compile a comprehensive assessment of the condition of their streets, often using contractors to perform the actual ratings. The

---

6. Each of the measures noted here is collected by the Center for Performance Measurement, International City/County Management Association (ICMA) and compared for participating local governments in a national benchmarking project (ICMA 2008).
frequencies at which these ratings are compiled differ from city to city, but most re-rate their streets at least every two or three years, and some do so annually.

Many different street condition rating systems exist. The systems reviewed for this bulletin differ somewhat in conditions examined, scales used, scoring instructions, and weighting of scores, but a characteristic common to all of the systems used is the reliance of each on a detailed set of scoring instructions, thereby reducing the degree of subjectivity involved.

Cities that rate their streets often list the overall average PCI (Pavement Condition Index) or PCR (Pavement Condition Rating) among their street maintenance performance measures. The usefulness of this choice is drawn into question by the responses of surveyed officials, most of whom favor other statistics not only for managing street maintenance but also for reporting street condition. Reporting the percentage of streets falling below a specified score might be a better summary statistic.

Comparing street condition ratings across cities using different rating systems could be problematic and is probably inappropriate in most instances. We conclude, however, that comparisons of street condition ratings across cities are viable when each city uses the same rating system and when the ratings are considered approximate rather than absolute.

References
City and County of Denver (CO). *Mayor’s Proposed 2010 Budget*.
City of Lake Havasu City (AZ). *Fiscal Year 2009–2010 Annual Budget*.
City of Leawood (KS). *2010 Annual Budget*.
City of San Antonio (TX). *Adopted Annual Operating and Capital Budget: Fiscal Year 2009*.
City of Shawnee (KS). *2009 Annual Budget*.
Institute for Transportation Research and Education (ITRE). “Pavement Management Systems in North Carolina” (Raleigh, NC: North Carolina State University, undated).


