Inter and intra buffer variability: A case study using scale

S Dawn Haynie
Georgia State University
shaynie@gsu.edu

Abstract

Many morphological studies select sample sets to explore neighbourhoods of interest, particularly in terms of their structural street properties, measures of scale or density, and proximity to the metropolitan center (Cervero and Gorham, 1995; Crane and Crepeau, 1998; French and Scoppa, 2007; Handy et al., 2003; Jacobs, 1993; Peponis et al., 2007; Southworth and Owens, 1993); yet beyond the established and distinctive structures of these neighborhoods, few have analyzed, in depth, the variability in their measures. This study randomly samples 4,321 localities from the 24 largest American metropolitan areas and describes a method using the measures of length and area to evaluate the variability both between and within these localities. Calculated as the standard deviation of mean scale, Inter Buffer Variability is introduced to describe the variation between these localities while Intra Buffer Variability describes the variation, or consistency, within these localities. How varied then are the measures of scale, and are the measures for some MSAs more varied than others? As will be shown, the MSA Inter Buffer Variability for both length and area are broad, which is expected given both the urban and suburban localities captured across each MSA; and yet, the MSA Intra Buffer Variability is also broad suggesting more variation within these localities than originally suggested by the samples illustrated within the literature. Comparatively for each measure of length and area, both Inter and Intra Buffer Variability are graphed one in relation to the other with their associated means used to delineate those trending higher or lower than average. Interestingly, four quadrants emerge distinctively delineating the measures of scale for these MSAs.

Keywords

Urban morphology, urban design, variability, neighbourhood scale.

1. Introduction

Beyond the differences demonstrated between neighborhoods and their associated measures of scale, density, or connectivity, few have been able to analyze, in depth, variability in the sprawling landscape of the metropolitan city. Methodologically, many studies have extracted representative samples from larger metropolitan areas to illustrate extremes between various neighborhoods or localities of particular interest (Cervero and Gorham, 1995; Crane and Crepeau, 1998; Doxiadis, 1968; Frank et al., 2007; Handy et al., 2003; Hess, 1997; Jacobs, 1993; Jo, 1998; Ozbil and Peponis, 2007; Peponis et al., 2007; Siskna, 1997; Southworth and Owens, 1993). These neighborhoods were often chosen for their purity of structural type, planning history, demographics, or primary mode of transportation. Statistical inferences were drawn, and frequently, these results were generalized to describe the consistency, density, and pedestrian-oriented context of the city center in juxtaposition to the varied, sprawling, auto-oriented context of the peripheral suburb.
In a global context, Doxiadis (1968) illustrated variations in scale across different metropolitan areas to describe patterns of growth, as did Abler and Adams (1976), Passonneau and Wurman (1966), and Adams in his discussion of the New York Regional Plan (1929). In a local context, Cervero and Gorham (1995) illustrated differences in scale between neighborhoods with prevailing modes of choice for transit; Handy (2003) illustrated differences between neighborhoods of specific structural interest; Jacobs (1993) illustrated differences in the measures between neighborhoods of significant historical interest; and Peponis et al. (2007) demonstrated differences between localities influenced by various planning policies and urban design initiatives.

In each of these cases, the measures of road segments and blocks were analyzed to offer a fundamental sense of the scale or size of urban elements that combine to form the texture of the urban fabric; and yet, given the method of selective sampling, these neighborhoods are not necessarily representative of the city in its entirety, essentially ignoring the variability experienced between or within these illustrated extremes.

If neighborhoods, or localities, were sampled randomly, with equitable probability, and for a population size that yields statistically significant results, would the inferences in the measures, as established in the literature, persist? How varied are the average measures between localities, how varied are the measures within each locality, and perhaps more fundamental, how can the variability in scale be described analytically?

2. Defining Measures of Variability

To discuss variability, two distinctions are suggested. First, a measure to describe differences between the average measures of each locality is defined to capture, as an example, differences between the average lengths of road segments for a city center in comparison to the average for a remote suburb. Second, an alternate measure to describe differences within the measures within each locality is defined to capture, as an example, the consistency in the lengths of road segments for city centers with a strong planning initiative.

**Inter Buffer Variability:**

describes differences between the average measures of neighborhoods; thus, it is calculated as the standard deviation for a set of neighborhood means. For example, the Inter Buffer Variability of Length is calculated by taking the standard deviation of mean road length for a sampled set of neighborhoods.

**Intra Buffer Variability:**

describes differences within the measures of any particular neighborhood, thus, it is calculated as the mean for a set of neighborhood standard deviations, which is calculated for any set of elements within that neighborhood. For example, the Intra Buffer Variability of Length is calculated as the mean of the standard deviation of road length for roads within a sampled set of neighborhoods.

3. Method for constructing a Randomly Sampled Set

To ensure an equitable distribution of sampled areas, a framework was established. From a defined point of center, rings radiated outward at a distance relevant to the scale of the maps studied; and a coordinate system, fixed by the point of center, was superimposed and rotated 45 degrees to define rather than divide the quadrants of North, South, East and West. From each section of this established framework, x and y coordinates were randomly selected at a particular distance and degree from the designated point of center. Included was a provision for eliminating the potential of overlapping areas such that all selected buffers were complete and distinct from one another. These randomly selected...
coordinates, along with each point of center, were imported into ESRI software to create circular buffers at a radius relevant to the variables being studied and then used to extract spatial information to describe smaller, more local areas within the maps.

4. Case Study analyzing the variability of scale in American Cities

To study the urban and suburban conditions found across the United States, 24 of the largest, most populated cities were selected for random sampling. These MSAs included: Atlanta, Baltimore, Boston, Chicago, Cincinnati, Cleveland, Dallas, Denver, Detroit, Houston, Los Angeles, Miami, Minneapolis, New York City, Philadelphia, Phoenix, Portland, San Diego, San Francisco, Seattle, St. Louis, Tampa, and Washington D.C. For each MSA, GIS based vector data was compiled from the Street Map and County databases released by ESRI in their ArcMap software. Data was decompressed, exported, converted into various shape files, dissolved, clipped and then eventually merged into a single shape file representative of each MSA.

Initially, each city was defined simply by the legal boundary of its larger Metropolitan Statistical Area (MSA); yet in several cases, the overall density and development of the city was continuous across the landscape from one MSA to another. In these cases, the two MSAs were combined into a single area for analysis to complete the overall morphology of the city and to reduce any possible distortions in the measures due to ‘edge effects.’ These combinations include the union of Cleveland with Akron, Denver with Boulder, Los Angeles with Riverside and Ventura, Philadelphia with Trenton, and San Francisco with San Jose. In comparison to the metropolitan areas originally evaluated by Abler and Adams (1976), the MSAs selected here captured 5% of the total land area held within the contiguous U.S., and they represented 49.5% of the population.

In this selected set of MSAs, the point of center was established by the position of the original City Hall and/or a similarly associated, politically significant building in the MSA. Rings radiated outward from the point of center at 5, 15, 30, and 60 mile intervals (Figure 1). From each section of this established framework, coordinates were randomly selected using a script programmed in Java. From these randomly selected coordinates, along with each point of center, circular buffers measuring 2 miles in diameter were established.

---

Intending to capture 10% of the total land area, 363 coordinate pairs were identified for each of the selected MSAs. Included was a provision for eliminating the potential of overlapping areas; and thus, all selected buffers were complete and distinct from one another. If all coordinate pairs captured development, the Randomly Sampled Set would have contained 8,712 buffers for study; but inevitably, many fell either outside the political boundary of the MSA or in rural, undeveloped areas (Figure 2).
To test this methodology and to access the effectiveness of the proposed definitions for variability, the following measures were extracted from the randomly sampled set:

**Length**
- the distance (feet) between two choice intersections, with road segments drawn as street centerlines

**Buffer Length**
- the sum length of road segments (feet) in a defined buffer divided by the number of road segments captured by that buffer; i.e. the mean length of road segments for a defined buffer

**Area**
- the landmass (acres) of a block bounded by a continuous set of road segments, with road segments drawn as street centerlines

**Buffer Area**
- the sum area of blocks (acres) in a defined buffer divided by the number of blocks captured by that buffer; i.e. the mean area of blocks for a defined buffer

In addition to those buffers capturing undeveloped areas, road segments and blocks of extreme scale, both large and small, were identified and excluded. To prevent potential distortion from extremes within the database, the work of Thomas Jefferson and his influence on the Land Ordinance of 1785 (Rashid, 1996) was assessed in conjunction with the work of Doxiadis (1965) and Leon Krier (1976) to set parameters for pragmatically defining and removing extremes. Blocks more than 640 acres in area or less than 0.12 acres were excluded. Similarly, road segments more than 1 mile in length or less than 72 feet were excluded.

5. Limitations of the Database

While there were no significant ‘edge effects’, the measures were subject to a number of other limitations. Area calculations were accurate, subject to the projections necessary for GIS to represent a spherical model in a single plane of two dimensions. Block area included not only the sum area of its parcels but a portion of its defining streets as well because the vectors were drawn to represent the centerline of each street without consideration of street width. In addition, the calculations of available street length were accurate, subject to the reliability of the available GIS data.

In consequence to the defined method of sampling, several complications were encountered, which limit the inferences that can be made of the resulting Randomly Sampled Set. First, the automated process for sampling allowed for the creation of a larger set of samples than would otherwise be possible; however, the size of the Randomly Sample Buffers then made it impossible to examine the data to correct errors.

---

2 The area calculation for each MSA was calculated from the .shp files of the ESRI Database in a NAD 83 projection.

3 According to ESRI, their databases were created and compiled from individual TIGER files and subsequently corrected in house to remove traditional issues of alignment and varying coordinate systems when connecting larger networks. The accuracy of this data remains subject to the individual error declared by ESRI. The individual line segments of each road network have not been verified or corrected here within this study, but the error within each overall network system chosen for analysis is presumed to be minimal given the scale of the analysis and thus, should have little impact on the results.
As a result, the Randomly Sampled Set of Buffers was only as good as the original set from which the data was drawn, without interventions to correct inaccuracies. Second, many of the Randomly Sampled Buffers captured significantly more area or length than was initially intended, given that the blocks and road segments intersecting each buffer were included along with those completely contained within (Figure 3). As a result, the Randomly Sampled Sets of blocks and road segments were not complete, congruent and comprehensive sets. For this reason, statistical correlations between the measures of the Randomly Sampled Buffers should be considered carefully.

Figure 3: Illustrations of the associated blocks (shown in a darker grey) and road segments (shown in black) captured within each of the Randomly Sampled Buffers

6. Analyzing Inter and Intra Buffer Variability

Measures for the Randomly Sampled Buffers were studied to assess the variations between them. Calculated as the standard deviation of Buffer Length or Buffer Area, the Inter Buffer Variability described the variability or consistency found between the measures of Buffer Length or Buffer Area for the Randomly Sampled Set. Subsequently, the MSA Inter Buffer Variability of Length or Area described the average variation or consistency found between the measures of buffer scale within a particular MSA.

In addition to analyzing the variation in the measures of scale between buffers, or the differences in Buffer Scale for these Randomly Sampled Buffers, variations in the measures of Length and Area were also considered within each buffer. Calculated as a mean for the standard deviation of Buffer Scale, Intra Buffer Variability of Length and Area described the variability or consistency found within the measures of Length or Area for the Randomly Sampled Buffers. Subsequently, the MSA Intra Buffer Variability of Length and Area described the average variation or consistency found among the measures of the Randomly Sampled Road Segments or Blocks within each buffer.

For the Randomly Sampled Buffers, Inter Buffer Variability of Length measures 855.12 feet (Table 1). When calculated for each metropolitan area, the MSA Inter Buffer Variability of Length varies from 458.71 in Boston to 1070.88 feet in Phoenix. Likewise, Inter Buffer Variability of Area measures 144.36 acres (Table 2), and when calculated for each metropolitan area, the MSA Inter Buffer Variability of Area varies from 67.67 acres in Boston to 191.55 acres in Minneapolis – St. Paul.

Notably, Atlanta, Boston, New York City, and Washington D.C. illustrate a lower MSA Inter Buffer Variability of Length and Area, which suggests greater consistency between the measures of their buffers. Los Angeles – Riverside – Ventura, Phoenix, San Diego, San Francisco – San Jose, and Seattle exhibit a higher MSA Inter Buffer Variability of Length with a lower MSA Inter Buffer Variability of Area, which suggests relative consistency between the measures of Buffer Area despite variability between the
measures of Buffer Length. Lastly, Cincinnati, Dallas, and St. Louis exhibit a lower MSA Inter Buffer Variability of Length with a higher MSA Inter Buffer Variability of Area, which suggests relative consistency between the measures of Buffer Length despite variability between the measures of Buffer Area.

Table 1: Statistical Measures of Buffer Length for the Randomly Sampled Buffers

<table>
<thead>
<tr>
<th>MSA</th>
<th>Randomly Sampled Buffers</th>
<th>MSA Mean of Buffer Length (feet)</th>
<th>MSA Median of Buffer Length</th>
<th>MSA Intra Buffer Variability</th>
<th>MSA Inter Buffer Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td>all MSAs</td>
<td>4321</td>
<td>1382.6562</td>
<td>1158.4521</td>
<td>968.4598</td>
<td>855.12</td>
</tr>
<tr>
<td>Atlanta</td>
<td>291</td>
<td>1279.3100</td>
<td>999.0812</td>
<td>992.8728</td>
<td>639.97</td>
</tr>
<tr>
<td>Baltimore</td>
<td>90</td>
<td>1145.9438</td>
<td>927.0117</td>
<td>874.3102</td>
<td>627.74</td>
</tr>
<tr>
<td>Boston</td>
<td>125</td>
<td>921.0272</td>
<td>674.9106</td>
<td>785.2839</td>
<td>458.71</td>
</tr>
<tr>
<td>Chicago</td>
<td>213</td>
<td>1328.2620</td>
<td>1115.2535</td>
<td>884.1587</td>
<td>1023.93</td>
</tr>
<tr>
<td>Cincinnati</td>
<td>155</td>
<td>1516.6831</td>
<td>1259.1298</td>
<td>1094.8248</td>
<td>669.08</td>
</tr>
<tr>
<td>Cleveland - Akron</td>
<td>101</td>
<td>1528.7196</td>
<td>1292.0128</td>
<td>1036.9714</td>
<td>985.69</td>
</tr>
<tr>
<td>Dallas</td>
<td>258</td>
<td>1450.6091</td>
<td>1200.6352</td>
<td>1038.7708</td>
<td>734.94</td>
</tr>
<tr>
<td>Denver - Boulder</td>
<td>190</td>
<td>1790.6753</td>
<td>1578.7486</td>
<td>1117.2205</td>
<td>1058.15</td>
</tr>
<tr>
<td>Detroit</td>
<td>134</td>
<td>1323.7026</td>
<td>1098.7564</td>
<td>934.3410</td>
<td>853.93</td>
</tr>
<tr>
<td>Houston</td>
<td>280</td>
<td>1423.4125</td>
<td>1186.7527</td>
<td>998.2042</td>
<td>856.64</td>
</tr>
<tr>
<td>Los Angeles – Riverside - Ventura</td>
<td>212</td>
<td>1195.4960</td>
<td>1006.4543</td>
<td>822.9983</td>
<td>856.85</td>
</tr>
<tr>
<td>Miami</td>
<td>68</td>
<td>926.6310</td>
<td>744.7042</td>
<td>756.9758</td>
<td>693.21</td>
</tr>
<tr>
<td>Minneapolis - St. Paul</td>
<td>225</td>
<td>1671.2101</td>
<td>1441.4431</td>
<td>1086.7686</td>
<td>965.85</td>
</tr>
<tr>
<td>New York City</td>
<td>209</td>
<td>894.8752</td>
<td>683.1129</td>
<td>724.4346</td>
<td>566.01</td>
</tr>
<tr>
<td>Philadelphia - Trenton</td>
<td>188</td>
<td>1171.5391</td>
<td>931.5364</td>
<td>895.3749</td>
<td>700.21</td>
</tr>
<tr>
<td>Phoenix</td>
<td>209</td>
<td>1751.4890</td>
<td>1594.6980</td>
<td>1048.7428</td>
<td>1070.88</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>189</td>
<td>1259.9701</td>
<td>987.0297</td>
<td>1017.4978</td>
<td>662.22</td>
</tr>
<tr>
<td>Portland</td>
<td>211</td>
<td>1532.9129</td>
<td>1365.6079</td>
<td>1036.3200</td>
<td>800.23</td>
</tr>
<tr>
<td>San Diego</td>
<td>114</td>
<td>1365.8487</td>
<td>1165.2267</td>
<td>969.0424</td>
<td>918.59</td>
</tr>
<tr>
<td>San Francisco - San Jose</td>
<td>181</td>
<td>1526.1984</td>
<td>1303.6005</td>
<td>1000.4448</td>
<td>1042.18</td>
</tr>
<tr>
<td>Seattle</td>
<td>150</td>
<td>1569.4133</td>
<td>1390.3612</td>
<td>971.0437</td>
<td>930.79</td>
</tr>
<tr>
<td>St. Louis</td>
<td>266</td>
<td>1568.2355</td>
<td>1332.2371</td>
<td>1081.8066</td>
<td>739.21</td>
</tr>
<tr>
<td>Tampa – St. Petersburg</td>
<td>96</td>
<td>1186.9352</td>
<td>964.4266</td>
<td>840.5445</td>
<td>769.50</td>
</tr>
<tr>
<td>Washington D.C.</td>
<td>166</td>
<td>1176.1295</td>
<td>940.7750</td>
<td>890.5396</td>
<td>651.96</td>
</tr>
</tbody>
</table>

Alternately, for Length within these Randomly Sampled Buffers, Intra Buffer Variability of Length measures 968.46 feet (Table 1). When calculated for each metropolitan area, the MSA Intra Buffer Variability of Length varies from 724.43 feet in New York City to 1117.22 feet in Denver – Boulder. Likewise for Area, the Intra Buffer Variability of Area measures 118.89 feet (Table 2). When calculated for each metropolitan area, the MSA Intra Buffer Variability of Area varies from 71.31 feet in Tampa – St. Petersburg to 156.99 feet in Pittsburgh.
Like many of the neighborhoods sampled and studied by Jacobs (1993), the Intra Buffer Variability of both length and area is also broad in this case study. Miami, Los Angeles – Riverside – Ventura, New York City, and Tampa – St. Petersburg illustrate a lower MSA Intra Buffer Variability of Length and Area, which suggests greater consistency between the measures within their buffers. Denver - Boulder, Phoenix, San Diego, and San Francisco – San Jose exhibit a higher MSA Intra Buffer Variability of Length with a lower MSA Intra Buffer Variability of Area, which suggests relative consistency among the measures of area despite variability between the measures of length. Boston, Baltimore, Detroit, and Philadelphia - Trenton exhibit a lower MSA Intra Buffer Variability of Length with a higher MSA Intra Buffer Variability of Area, which suggests relative consistency among the measures of length despite variability between the measures of area.
When the measures of scale are averaged by buffer and the variability between them is evaluated, the measures differ considerably. Given that buffers are sampled from the metropolitan center as well as the periphery of each MSA, not surprisingly, the Inter Buffer Variability of Scale, or the variability in the average scale between buffers, is broad for both Buffer Length and Buffer Area; but unexpectedly, the Intra Buffer Variability, or the variability in the scale of road segments and blocks within each of the buffers, is also broad.

As a comparison, Inter Buffer Variability is graphed in relation to the Intra Buffer Variability of Scale, with their associated means used to delineate those trending higher or lower than average.

Figure 4: Graph of MSA Mean of Inter Buffer Variability for Scale in relation to MSA Mean of Intra Buffer Variability for Scale, for the Randomly Sampled Set

In considering the variability of length, both between and within each buffer, the graph illustrates four distinct quadrants for these selected MSAs (Figure 4A). Denver – Boulder exhibits a higher MSA Inter Buffer Variability of Length and a higher MSA Intra Buffer Variability of Length, with Cleveland – Akron, Minneapolis – St. Paul, Phoenix, San Diego, San Francisco – San Jose, and Seattle similar. Contrastingly, Boston exhibits a lower MSA Inter Buffer Variability of Length with a lower MSA Intra Buffer Variability of Length with Baltimore, Miami, New York City, Philadelphia – Trenton, Tampa – St. Petersburg, and Washington D.C. similar. Atlanta, Cincinnati, Dallas, Houston, Pittsburgh, Portland, and St. Louis exhibit a lower Inter Buffer Variability of Length with a higher Intra Buffer Variability of Length. Interestingly, only Chicago exhibits high variability in the average measures of length between buffers with more consistency of length within its buffers.

In considering the variability of Area, both between and within each buffer, the graph again illustrates four distinct quadrants (Figure 4B). For these selected MSAs, Chicago, Cincinnati, Cleveland – Akron, Dallas, Detroit, Minneapolis – St. Paul, and St. Louis exhibit a higher MSA Inter Buffer Variability of Area and a higher MSA Intra Buffer Variability of Area, illustrating variability both between and within the measures of the buffers; in contrast, Los Angeles – Riverside – Ventura, Miami, New York City, Phoenix, San Diego, San Francisco – San Jose, Seattle, Tampa – St. Petersburg, and Washington D.C. exhibit a lower MSA Inter Buffer Variability of Area and a lower MSA Intra Buffer Variability, illustrating greater consistency both between and within the measures of the buffers. Atlanta, Baltimore, Boston, Philadelphia – Trenton, Pittsburgh, and Portland exhibit a higher MSA Intra Buffer Variability with a lower
Inter Buffer Variability of Area. Interestingly, only Denver – Boulder and Houston exhibit high variability in the average measures of area between buffers with more consistency of area within them.

7. Statistical Inferences

Measures for the Inter Buffer Variability, as calculated from these randomly sampled neighborhoods, capture the expected variation within this selected set of MSAs. Miami exhibits consistency among the average measures of its neighborhoods while Denver – Boulder exhibits variation, and neither is surprising given the gridded street structure of Miami juxtaposed to the varied geography of Denver – Boulder.

Measures for the Intra Buffer Variability capture far more variation and unpredictability in the measures of scale than originally anticipated, particularly given the examples often analytically studied as extremes in the literature. Chicago exhibits consistency in the length of road segments within its neighborhoods though there is greater variation in its block size despite its regularized plan. Contrastingly, Minneapolis – St. Paul exhibits variation in its road segment length and its block area.

When Inter Buffer Variability is plotted against Intra Buffer Variability for both length and area, several probabilities are suggested. First, the measures for length and area operate independently in several MSAs, despite their strong correlation (Peponis et al. 2007). Second, only a few of the 24 MSAs demonstrate a higher Inter Buffer Variability with a lower Intra Buffer Variability so the chance of encountering greater variability in the measures of scale between buffers and yet not within them is quite low. Lastly, almost twice as many MSAs exhibit low variability between their buffers while the variability within each is relatively evenly distributed around the mean. In summary, many buffers within these MSAs behave as originally perceived, despite perhaps the extraordinary variation amid the scale of the road segments and blocks within them.

8. Implications

As introduced, Inter and Intra Buffer Variability can be used to capture the consistency and/or variation in any number of measures. Inter Buffer Variability describes the differences between the average measures of each locality while Intra Buffer Variability describes the differences in the measures within each locality, regardless of the variable analyzed. Furthermore, the method introduced for sampling localities randomly ensures statistical significance in the results and can be utilized in other analyses.
References


Passonneau, J. and Wurman, R. (1966), Urban atlas: 20 American cities; a communication study notating selected urban data at a scale of 1:48,000, Cambridge, MIT Press.


