

A Statistical Comparison of Three Goodness-of-Fit Criteria Used in Modelling Distances

ROBERT F. LOVE[†] AND HALIT ÜSTER

*Management Science/Information Systems Area
Michael G. DeGroot School of Business
McMaster University, Hamilton
Ontario, Canada*

Abstract. Distance predicting functions may be used in a variety of applications for estimating travel distances between points. To evaluate the accuracy of a distance predicting function and to determine its parameters, a goodness-of-fit criteria is employed. *AD* (Absolute Deviations), *SD* (Squared Deviations) and *NAD* (Normalized Absolute Deviations) are the three criteria that are mostly employed in practice. In the literature some assumptions have been made about the properties of each criterion. In this paper, we present statistical analyses performed to compare the three criteria from different perspectives. For this purpose, we employ the $\ell_{k,p,\theta}$ -norm as the distance predicting function, and statistically compare the three criteria by using normalized absolute prediction error distributions in seventeen geographical regions. We find that there exist no significant differences between the criteria. However, since the criterion *SD* has desirable properties in terms of distance modelling procedures, we suggest its use in practice.

Keywords: Transportation, Accuracy in Distance Modelling, Goodness-of-Fit Criteria.

1. Introduction

When objects in space, such as different cities in a geographic region, activity centers in a plant, or computer terminals of a LAN, can be represented by points, a distance predicting function may be used to transform point coordinate differences of two points into an estimate of the distance between the points. Thus, distance predicting functions have a number of uses. Some of these uses are discussed below.

For validating the accuracy of actual road network distance data, distance predicting functions can be used as suggested by Ginsburgh and Hansen [8].

To determine the optimal mix of trunking and tramping of a truck transportation network for the movement of finished goods and raw materials

[†] Requests for reprints should be sent to Robert Love, MSIS Area, Michael G. DeGroot School of Business, McMaster University, Hamilton, Ontario, Canada.

among national distribution centers, regional depots, and producers, a distance predicting function was utilized by Westwood [25] to obtain estimates of the travel distances between possible links in the network. In some distribution problems for which only the demands and the general location of customers are known [7], a distance predicting function may be employed to calculate a predicted travel distance between the depot and the general area.

Distance predicting functions can also be used in models that determine the response time of emergency vehicles to calls such as the model proposed by Kolesar [10] for calculating the response time of fire engines to fires.

Klein [9] suggests that distance predicting functions which reflect the nature of a geographic region's road network should be used for constructing Voronoi diagrams of the region. A Voronoi diagram subdivides a region into a number of subregions with each subregion being formed around a point belonging to a set of points. For example, the set of points may be the region's police stations, fire halls, or hospitals. Once the location of a query point is determined, the appropriate point of the set is notified to respond to the call by looking at the Voronoi diagram.

Distance predicting functions appear within the context of larger models such as facilities

location problems [6], [13]. Distance predicting functions in these models obviate the need for determining actual distances between the new facilities and the existing facilities. In addition, by using distance predicting functions which have empirical parameters that reflect the nature of a region's road network, more accurate cost structures should be obtained than if an assumed distance function is used by an analyst.

Presently, a distance predicting function is being utilized in software packages TruckStops2 [20] and Roadnet [21]. When an analyst provides data regarding the customer demands, customer locations, and truck types for a transportation network, TruckStops2 assigns customers to different trucks and determines the routes for the trucks.

Distance predicting functions may be used for calculating distances in a Geographic Information System (GIS). As Star and Estes [22] state, distance measurements are of value in many geographic circumstances. Some of these circumstances are planning an irrigation channel between a pond and a field, locating a site for a fire tower in a forest, and calculating the distances among different geographic regions. To calculate distance measurements, a distance predicting function may be incorporated into a GIS.

In order to evaluate the accuracy of a distance predicting function, a criterion is required. The criterion not only provides a numerical value so that different distance predicting functions can be compared but also

provides the means for determining any empirical parameters of a distance predicting function. Researchers are presently using three goodness-of-fit criteria [1], [2], [3], [5], [11], [12], [18], [23], [24]:

1. Sum of Absolute Deviations (*AD*),
2. Sum of Squared Deviations (*SD*),
3. Sum of Normalized Absolute Deviations (*NAD*),

In addition, *AD* and *SD* have been used by Love and Morris [11], [12] to develop tests for statistically comparing the accuracy of different distance predicting functions.

There are several motivations for conducting the study presented in this paper. Love, Walker and Tiku [18] describe a procedure to find the confidence intervals for a fitted distance. The procedure utilizes the statistical properties of the errors produced when a distance predicting function is fitted to a particular geographic region. Since different criteria could lead to different statistical properties of the fitting errors, we do statistical analyses of these errors for the three fitting criteria.

Secondly, in the literature the three criteria were assumed to have different properties in terms of predicting distances. For example, it has been assumed that if the *AD* criterion is used, the weighted ℓ_p -norm will predict long distances more accurately than short distances. The *SD* criterion has been characterized as having prediction errors with better statistical properties but still being similar to the *AD* criterion in terms of its accuracy in predicting long distances [11], [12]. The *NAD* criterion, on the other hand, has been assumed to predict short distances as accurately as long distances [3], [5], [15], [18].

In this paper, we present statistical properties of the fitting errors and a comparison of the above mentioned criteria. Statistical analyses are applied to seventeen different geographic regions using the $\ell_{kp\theta}$ -norm as the distance predicting function. In Section 2, the three criteria and the distance predicting function are described. In Section 3, the statistical test procedures and comparison results are presented. Finally, in Section 4, conclusions based on our analyses of these results are discussed.

2. The Distance Function and the Criteria

The weighted ℓ_p -norm ($\ell_{kp\theta}$) was employed as the distance predicting function. With the $\ell_{kp\theta}$ -norm the travel distance between points $\mathbf{x} = (x_1, x_2)$

and $\mathbf{y} = (y_1, y_2)$ is given by

$$\ell_{kp\theta}(\mathbf{x}', \mathbf{y}') = k (|x'_1 - y'_1|^p + |x'_2 - y'_2|^p)^{1/p}, \quad k > 0, \quad p \geq 1 \quad (1)$$

where

$$\begin{aligned} \mathbf{x}' &= (x'_1, x'_2), \quad \mathbf{y}' = (y'_1, y'_2), \quad \theta \in [0, 90^\circ] \quad \text{and} \\ \begin{pmatrix} x'_1 & x'_2 \\ y'_1 & y'_2 \end{pmatrix} &= \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \end{aligned}$$

This norm was selected because insights into the peculiarities of road networks are provided by the empirical parameters k , p , and θ of the norm when the empirical parameters are determined for a sample of road distances from a geographic region. The parameter p measures the rectangular bias of the road network. The angle θ is a rotation parameter which ensures that the coordinate axes are rotated counterclockwise from the analyst's defined coordinate axes until the road network is in phase with the rotated coordinate axes [5]. The parameter k is an inflation factor which accounts for the hills, valleys and other types of noise in the road networks.

A criterion is used to measure the accuracy of a distance predicting function and also to determine its optimal parameters. We next describe the general methodology for fitting the distance predicting function to a given geographic region. A random sample of points within the geographic region is chosen. Based on an arbitrary coordinate system, cartesian coordinates for each point are assigned and the actual distances between each pair of points are measured or read from distance charts. Then the parameters (k , p and θ) of the distance predicting function are computed to minimize the value of the selected criterion.

Let $\ell_{kp\theta}(\mathbf{a}_i, \mathbf{a}_j)$ be the predicted distance between points \mathbf{a}_i and \mathbf{a}_j and $A(\mathbf{a}_i, \mathbf{a}_j)$ be the actual distance between \mathbf{a}_i and \mathbf{a}_j , and n be the number of points in the data set. Then the mathematical expressions for three goodness-of-fit criteria that will be analysed in this paper are the minimizations of the following sums:

$$AD = \sum_{i=1}^{n-1} \sum_{j=i+1}^n |\ell_{kp\theta}(\mathbf{a}_i, \mathbf{a}_j) - A(\mathbf{a}_i, \mathbf{a}_j)| \quad (2)$$

$$SD = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{(\ell_{kp\theta}(\mathbf{a}_i, \mathbf{a}_j) - A(\mathbf{a}_i, \mathbf{a}_j))^2}{A(\mathbf{a}_i, \mathbf{a}_j)} \quad (3)$$

$$NAD = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{|\ell_{kp\theta}(\mathbf{a}_i, \mathbf{a}_j) - A(\mathbf{a}_i, \mathbf{a}_j)|}{A(\mathbf{a}_i, \mathbf{a}_j)} \quad (4)$$

The first criterion AD is the minimization of the sum of absolute deviations. Since the terms in AD are not the weighted ones but only the absolute errors for each pair, it has been described as a criterion which should estimate long distances more accurately than short distances. The second criterion SD is the minimization of the sum of squared deviations where each squared error term is weighted by $1/A(\mathbf{a}_i, \mathbf{a}_j)$. Squared deviations and the division by actual distance provide the criterion with certain desirable statistical properties [11], [12]. However, the assumption has still been made that the difference in the accuracy of predictions involving long and short distances in a region will favour the long distances [1], [2], [11], [12], [23], [24]. The last criterion NAD is relatively new in the literature [3], [5], [15], [18]. With the NAD criterion, a sum of normalized absolute deviations is minimized and the basic premise is that equal accuracy in predicting long and short distances in a region will result. Normalization is realized by dividing the absolute deviation by the actual distance between each pair. In this way both long and short distances are treated on the same relative basis.

Besides their above-mentioned structures, the three criteria also differ from each other by the computational procedures performed to determine the optimal parameter values of the distance predicting function. The computational procedures for fitting the AD and the SD criteria are given by Brimberg and Love [4]. For the NAD criterion it is shown that the computational procedure is identical to that of AD [15]. In general, the best θ and p values are determined by using a two-stage incremental search procedure and a four-stage incremental search procedure, respectively. To find the best k value some properties of the criteria are used. It is known that AD is a convex function of k , and SD is a strictly convex function of k [4]. NAD was shown to be a convex function of k by Love and Walker [15]. When using the AD and NAD criteria it is necessary to employ an algorithm to find the optimal k for a given (θ, p) pair. The optimal k for the SD criterion is calculated with a simple closed-form formula derived by Brimberg and Love [4]. The property of having a closed-form formula to find the best value of k makes the application of the SD computationally more efficient than using either the AD or the NAD criterion.

To model the parameters of the $\ell_{kp\theta}$ -norm Love and Walker [16] collected sample data from seventeen geographic regions. The sample data for each geographic region included 15 points (locations) based on random selection of *point coordinates* on the map. These 15 points provided 105 actual distances to be modelled by the distance predicting function $\ell_{kp\theta}$ using each criterion.

The empirical parameter values for the $\ell_{kp\theta}$ -norm and the corresponding minimum criterion values for seventeen geographic regions are computed

by Love and Walker [17] for the *AD* and *SD* criteria, and by Love and Walker [15] for the *NAD* criterion. In Table 1 we present the best parameter values of the $\ell_{kp\theta}$ function used in this study.

3. Statistical Comparison of the Criteria

The purpose of this section is to conduct statistical comparisons of the three criteria by adopting the normalized absolute prediction error as the random variable. For our work on road distances, the errors are the differences between actual distance and fitted distance pairs. Using the previously defined notation the model that we used to determine the relationship between the fitted distance and the actual distance is given by

$$A(\mathbf{x}_i, \mathbf{x}_j) = \ell_{kp\theta}(\mathbf{x}_i, \mathbf{x}_j) + e(\mathbf{x}_i, \mathbf{x}_j), \quad (5)$$

where $e(\mathbf{x}_i, \mathbf{x}_j)$ is the error term for the $\mathbf{x}_i, \mathbf{x}_j$ pair. From the random sample of points for a geographic region, the point estimates of the empirical distance predicting function parameters are calculated. Substituting these point estimates into the empirical distance predicting function, an estimate of the actual distance, $\ell_{kp\theta}(\mathbf{x}_i, \mathbf{x}_j)$, is obtained. The error term for any pair of points embodies errors that may arise in determining the fitted distance for that pair of points. For empirical distance functions which utilize point coordinate differences, these errors may arise from point coordinate measurements, inaccurate instrument calibrations, and road network peculiarities that are not captured by the distance model.

In order to compare the three criteria, we use a transformed random variable given as $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$. There are three reasons for using this transformation. First, the new random variable frees the error terms from their directions so that the absolute errors are to be compared. Second, since each criterion produces errors in different units the division of each error term by its actual distance provides the comparison to be performed on the same basis for the criteria. Finally, the accuracy in predicting long and short distances in a given region can be compared on the same basis by this new random variable.

3.1. Comparison of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions

In order to compare the normalized absolute errors, $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$, their distribution for each criterion was first checked for normality. For that purpose and also to present the descriptive statistics for each distribution, Table 2, which includes the means (\bar{x}), variances (s^2), skewness

Table 1. Best Parameter Values of $\ell_{kp\theta}$, $1 \leq p \leq 2$

No.	Region	AD	θ	k	p
1	Australia	13106.91	3	1.1176	1.6848
2	BC Province	6369.31	71	1.2737	1.6322
3	Canada	8029.42	86	1.1772	1.4705
4	France	1592.26	68	1.0468	1.7734
5	Great Britain	2311.36	38	1.1185	1.8124
6	NY State	1637.44	86	1.1035	1.6946
7	Pennsylvania	1207.34	3	1.0671	1.6274
8	United States	6516.24	0	1.0817	1.7290
9	Brussels	46.46	46	1.0488	1.7660
10	London Central	61.45	18	1.0697	1.7524
11	London North	28.49	15	1.0638	1.6505
12	Los Angeles	111.89	89	1.0626	1.5699
13	NY City	124.29	5	1.0674	1.5822
14	Paris	48.71	75	1.0704	1.7859
15	Sydney	13.11	7	1.1048	1.4061
16	Tokyo	28.91	59	1.0961	1.8591
17	Toronto	65.98	87	1.0118	1.1333
No.	Region	SD	θ	k	p
1	Australia	1163.597	0	1.1460	1.8585
2	BC Province	1038.727	68	1.2495	1.5609
3	Canada	565.617	83	1.1715	1.4849
4	France	92.327	70	1.0609	1.8430
5	Great Britain	219.429	40	1.1095	1.7895
6	NY State	159.805	86	1.0794	1.5823
7	Pennsylvania	107.061	4	1.0611	1.6244
8	United States	342.684	0	1.0792	1.6641
9	Brussels	3.55	47	1.0549	1.8180
10	London Central	16.53	72	1.1182	1.9241
11	London North	1.78	11	1.0599	1.6456
12	Los Angeles	15.50	2	1.0721	1.5734
13	NY City	13.58	6	1.1069	1.7340
14	Paris	6.52	86	1.0613	1.8189
15	Sydney	1.35	8	1.1266	1.4719
16	Tokyo	2.30	58	1.0963	1.8252
17	Toronto	5.10	88	1.0279	1.1863
No.	Region	NAD	θ	k	p
1	Australia	6.261	1	1.0959	1.7545
2	BC Province	11.068	69	1.2701	1.7080
3	Canada	5.234	85	1.1732	1.4584
4	France	3.611	71	1.0396	1.7417
5	Great Britain	5.981	0	1.1032	1.8826
6	NY State	6.004	87	1.0308	1.4950
7	Pennsylvania	6.486	7	1.0673	1.6958
8	United States	3.295	0	1.0825	1.7427
9	Brussels	4.68	45	1.0495	1.7802
10	London Central	8.75	8	1.0495	1.7802
11	London North	4.70	14	1.0591	1.6171
12	Los Angeles	7.90	87	1.0672	1.5684
13	NY City	6.90	4	1.0737	1.6975
14	Paris	8.41	11	1.0635	1.6649
15	Sydney	6.50	6	1.0991	1.3940
16	Tokyo	4.30	64	1.0963	1.8901
17	Toronto	4.45	87	1.0121	1.1261

($\hat{\mu}_3$) and kurtosis (in ($\hat{\mu}_4 - 3$) form), is constructed. In Table 2, we observe

that skewness and kurtosis values for the distributions are different enough from zero that we cannot conclude the distributions of $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ are from normal distributions for each criterion in the regions. The normal probability plots and histograms for the United States and Toronto can be found in Love and Üster [14]. These plots also support the non-normality of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions. Therefore, a nonparametric test was applied to determine if the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions for each criterion were significantly different from each other in a given region. The Friedman Test [19], which is used for multiple matched samples, was employed as the main effect test to compare the three $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions at the 5% significance level. The p-values for seventeen geographic regions are listed in Table 3. Since the p-values are well above 0.05, no pair of criteria is significantly different at the 5% significance level. In Table 4 the mean absolute percent errors are reported for each criterion and region. Based on the figures in Table 4, it can be said that the average percent absolute errors for a given region are very close to each other for the criteria, and in general, they are small enough to conclude that the predicted distances are close approximations of actual distances. For example, in Brussels the percent absolute errors in predicting distances are 4.46%, 4.47% and 4.46% for the *AD*, *SD*, and *NAD* criteria, respectively.

We next test the long and short distance distributions of the random variable $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ for normality. Non-normality of this variable's distribution formed by 105 pairs in a given region does not guarantee that a subset of these 105 pairs which is non-randomly formed by 35 pairs, also comes from a non-normal distribution. There are six different distributions (two for each criterion) used in the comparisons to identify the differences between the three criteria for predicting long actual distances and short actual distances. If we find that in these six distributions there is at least one non-normal distribution for each criterion (either long or short distance distribution, but not all three are the same), then we need to use nonparametric tests for the following parts of this section. Therefore, the skewness and kurtosis values of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions for long distances using the *AD* and *NAD* criteria, and for short distances using the *SD* criterion are reported for each region in Table 5. The skewness and kurtosis values in Table 5 are sufficiently different from zero (the distributions are always skewed right, and are generally more peaked than the normal distribution) to provide evidence that the distributions come from non-normal populations. This is also supported by the normal probability plots and histograms which can be found in Love and Üster [14]. There-

Table 2. Sample Statistics of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$
 (Row 1:AD; Row 2:SD; Row 3:NAD)

No.	Region	\bar{x}	s^2	$\hat{\mu}_3$	$\hat{\mu}_4 - 3$
1	Australia	0.0634	0.0027	1.5524	2.9709
		0.0636	0.0025	1.5999	3.3357
		0.0593	0.0027	1.2950	1.892
2	BC Province	0.1069	0.0071	0.7062	-0.1670
		0.1063	0.0069	0.7309	-0.0870
		0.1052	0.0067	0.7595	0.0682
3	Canada	0.0499	0.0028	1.747	3.2356
		0.0504	0.0025	1.5718	2.1566
		0.0497	0.0028	1.8150	3.3933
4	France	0.0335	0.0009	1.7694	4.1584
		0.0341	0.0008	1.7285	4.1728
		0.0334	0.0010	1.9857	5.1221
5	Great Britain	0.0591	0.0025	1.1108	1.3866
		0.0587	0.0024	1.2424	1.9158
		0.0570	0.0027	1.6056	3.2751
6	NY State	0.0623	0.0023	0.8010	0.1435
		0.0612	0.0021	1.0104	0.5856
		0.0626	0.0026	1.2077	1.0528
7	Pennsylvania	0.0617	0.0033	1.6327	3.7136
		0.0623	0.0033	1.7010	3.7626
		0.0608	0.0031	1.5613	3.5665
8	United States	0.0314	0.0009	1.8475	4.8344
		0.0326	0.0008	1.8210	4.1774
		0.0314	0.0009	1.8510	4.9158
9	Brussels	0.0446	0.0017	1.6735	3.7924
		0.0447	0.0017	1.6599	3.8810
		0.0446	0.0017	1.6728	3.7951
10	London Central	0.0833	0.0212	4.0173	19.1818
		0.0872	0.0195	4.2164	20.9644
		0.0832	0.021	4.1062	19.9933
11	London North	0.0446	0.0013	1.2163	1.5499
		0.0442	0.0012	1.2857	1.8672
		0.0440	0.0012	1.2825	1.8599
12	Los Angeles	0.0751	0.0138	6.7183	57.5431
		0.0757	0.0131	7.0910	62.3922
		0.0748	0.0139	6.9068	59.9512
13	NY City	0.0660	0.0039	1.7232	2.6192
		0.0682	0.0031	1.5462	2.3413
		0.0653	0.0039	1.6307	2.4097
14	Paris	0.0815	0.0068	3.0579	13.457
		0.0827	0.0063	2.9045	12.1986
		0.0801	0.0069	2.4590	9.0103
15	Sydney	0.0621	0.0033	1.4150	2.0364
		0.0620	0.0031	1.3346	1.8110
		0.0617	0.0033	1.4393	2.1473
16	Tokyo	0.0410	0.0021	3.9404	23.7253
		0.0414	0.0021	3.8788	23.1547
		0.0409	0.0022	3.9262	23.5744
17	Toronto	0.0425	0.0020	2.2888	6.7672
		0.0440	0.0017	1.9516	5.0684
		0.0424	0.0021	2.2782	6.6037

fore, nonparametric tests should be used for unbiased comparisons of the criteria involving the six distributions.

Table 3. The p-values for the Friedman Test of $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions

No.	Region	p-value	No.	Region	p-value
1	Australia	0.3732	9	Brussels	0.9765
2	BC Province	0.3750	10	London Central	0.6386
3	Canada	0.2315	11	London North	0.6839
4	France	0.4806	12	Los Angeles	0.3724
5	Great Britain	0.9355	13	NY City	0.7062
6	NY State	0.4047	14	Paris	0.1038
7	Pennsylvania	0.7280	15	Sydney	0.7606
8	United States	0.8889	16	Tokyo	0.8816
			17	Toronto	0.2450

Table 4. Mean absolute percent errors for criteria

No.	Region	<i>AD</i>	<i>SD</i>	<i>NAD</i>
1	Australia	6.34	6.36	5.93
2	BC Province	10.69	10.63	10.52
3	Canada	4.99	5.04	4.97
4	France	3.35	3.41	3.34
5	Great Britain	5.91	5.87	5.70
6	NY State	6.23	6.12	6.26
7	Pennsylvania	6.17	6.23	6.08
8	United States	3.14	3.26	3.14
9	Brussels	4.46	4.47	4.46
10	London Central	8.33	8.72	8.33
11	London North	4.46	4.42	4.40
12	Los Angeles	7.51	7.57	7.48
13	NY City	6.60	6.82	6.53
14	Paris	7.82	8.27	8.01
15	Sydney	6.21	6.20	6.17
16	Tokyo	4.10	4.14	4.09
17	Toronto	4.25	4.40	4.24

3.2. Accuracy of the criteria in predicting long and short distances

It has already been shown that we must employ a nonparametric test to compare the accuracy of the three criteria in predicting long and short distances. For that purpose, two Friedman tests (long and short distance distributions) for the matched triples are performed for each region. A p-value less than 0.05 is supposed to indicate the existence of a significantly different pair among the criteria for the given region. The p-values of these main effect tests are provided in Table 6. The significance levels listed in Table 6 can be interpreted as follows: In general, the accuracy in predicting long or short distances is not significantly different for the criteria. However, for long actual distances, in five of the eight large geographic

Table 5. Skewness and Kurtosis Values for the Normality of Three Distributions.

No.	Region	Long Distance <i>AD</i>		Short Distance <i>SD</i>		Long Distance <i>NAD</i>	
		Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
1	Australia	0.4881	-0.6541	1.2536	1.3505	0.4131	-0.9182
2	British Columbia	0.1244	1.0772	0.3501	-0.1516	1.3306	1.6137
3	Canada	1.3520	1.6556	0.6961	-0.5406	1.3277	1.8534
4	France	2.6315	9.8096	2.0758	5.8926	2.7389	10.1983
5	Great Britain	1.4349	1.3900	0.5296	-0.2207	1.1993	0.9675
6	NY State	0.8914	-0.2543	0.9814	1.0950	0.7519	-0.7202
7	Pennsylvania	0.5275	-1.0158	1.2628	1.0579	0.4775	-1.0018
8	United States	1.6883	4.4469	1.4650	2.5399	1.6432	4.1495
9	Brussels	1.5948	2.4744	1.8813	5.0096	1.6090	2.6409
10	London Central	3.3989	12.8349	3.8639	15.9083	3.4378	13.7044
11	London North	0.6037	-0.7204	0.9873	0.3199	0.8531	0.2911
12	Los Angeles	1.4745	2.5833	4.8098	26.0766	1.4541	2.7294
13	NY City	2.2495	6.4738	0.9857	-0.2267	2.1354	5.8431
14	Paris	0.8056	-0.0447	2.0891	4.8714	0.7644	-0.5594
15	Sydney	1.9071	4.9922	0.5940	-0.3163	2.0189	5.6586
16	Tokyo	1.7651	5.3646	3.3773	15.1482	1.9370	6.3196
17	Toronto	1.4835	2.2125	1.1633	1.5794	1.2894	1.4425

regions; Australia, British Columbia, Great Britain, New York State, and the United States, and for the short actual distances, in four regions; Australia, New York State, Los Angeles and Paris, there is at least one pair of criteria with significantly different distance prediction accuracy.

Table 6. The p-values of Friedman test comparing long and short distance distributions

No.	Region	Short D.	Long D.
1	Australia	0.0000	0.0002
2	BC Province	0.1242	0.0080
3	Canada	0.4624	0.9107
4	France	0.6897	0.7553
5	Great Britain	0.2466	0.0073
6	NY State	0.0187	0.0116
7	Pennsylvania	0.0527	0.1242
8	United States	0.1066	0.0300
9	Brussels	0.9749	0.9491
10	London Central	0.3213	0.1870
11	London North	0.1342	0.2684
12	Los Angeles	0.0083	0.4389
13	NY City	0.1048	0.2355
14	Paris	0.0316	0.6175
15	Sydney	0.8534	0.2388
16	Tokyo	0.9767	0.4296
17	Toronto	0.1691	0.6897

In order to identify which criterion is more accurate in predicting long or short distances in the above exceptional regions, multiple comparisons are performed by using nonparametric Wilcoxon matched pairs tests. However,

instead of reporting the results of this test, average percent absolute errors ($100 * E[|e(\mathbf{x}_i, \mathbf{x}_j)| / A(\mathbf{x}_i, \mathbf{x}_j)]$) for predicting long and short distances for the criteria are presented in Table 7. Inspecting the average percent absolute

Table 7. Average percent absolute errors in predicting long and short distances

No.	Region	Short Distance			Long Distance		
		<i>AD</i>	<i>SD</i>	<i>NAD</i>	<i>AD</i>	<i>SD</i>	<i>NAD</i>
1	Australia	8.45	8.37	6.23	5.03	5.02	6.42
2	BC Province	13.85	13.53	12.67	7.79	7.90	8.40
3	Canada	8.92	8.80	8.76	1.89	1.97	1.96
4	France	3.93	4.03	3.81	2.56	2.51	2.67
5	Great Britain	6.73	6.61	5.99	3.60	3.76	5.08
6	NY State	7.29	6.91	6.03	5.14	5.44	6.82
7	Pennsylvania	8.32	8.59	8.00	4.13	4.10	4.44
8	United States	4.01	4.22	3.97	2.06	2.02	2.09
9	Brussels	5.30	5.31	5.29	3.25	3.26	3.24
10	London Central	11.18	11.00	11.03	4.60	5.34	4.70
11	London North	6.22	5.98	5.87	2.77	2.87	2.92
12	Los Angeles	11.60	11.39	11.33	4.15	4.70	4.34
13	NY City	7.71	7.82	7.17	5.75	5.41	5.84
14	Paris	11.61	11.73	10.68	6.18	6.27	6.91
15	Sydney	7.97	7.86	7.85	5.61	5.60	5.66
16	Tokyo	5.63	5.64	5.64	2.65	2.62	2.77
17	Toronto	5.59	5.70	5.52	2.51	2.39	2.57

errors of the first five exceptional regions listed above, it is observed that the *AD* and *SD* criteria generate less average percent absolute error than the *NAD* criterion in predicting long distances. For example, in Australia the average percent absolute errors for the *AD* and *SD* criteria are 5.02% and 5.03%, respectively, whereas for the *NAD* criterion, it is 6.42%. Inspecting Table 7 for short actual distances, we see that the *NAD* criterion provides better prediction accuracy than either of the *AD* and *SD* criteria for the four exceptional regions. For example, in Australia the *NAD* criterion generates a 6.23% absolute error in predicting short distances. However, the *AD* and *SD* criteria provide 8.45% and 8.37% absolute errors for the same region, respectively.

3.3. Accuracy of each criterion in predicting the long versus short distances

The purpose of this section is to compare the accuracy of predicting long distances versus short distances in a region. First, in order to determine whether the variance of the $|e(\mathbf{x}_i, \mathbf{x}_j)| / A(\mathbf{x}_i, \mathbf{x}_j)$ distribution is constant

for a given criterion in a region, Levene tests are conducted for each criterion in seventeen regions. Hence if the p-value of the Levene test for a criterion is significant, then we can conclude that the variance of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distribution for that criterion in that region is constant and vice versa. The 2-tail p-values for the Levene tests are presented in Table 8. Based on this table, the variances of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions are not homoscedastic in 10 regions for AD , again in 10 regions for SD , and in 11 regions for NAD . However, note that the significance levels for the urban centers are not as strong as the significance levels for the large geographical regions.

Table 8. p-values of the Levene test for $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$

No.	Region	AD	SD	NAD
1	Australia	0.084	0.031	0.383
2	BC Province	0.725	0.617	0.758
3	Canada	0.000	0.000	0.000
4	France	0.285	0.351	0.252
5	Great Britain	0.159	0.234	0.466
6	NY State	0.741	0.925	0.014
7	Pennsylvania	0.002	0.000	0.005
8	United States	0.013	0.001	0.018
9	Brussels	0.084	0.093	0.090
10	London City	0.066	0.039	0.066
11	London North	0.001	0.003	0.007
12	Los Angeles	0.017	0.024	0.018
13	NY City	0.016	0.054	0.011
14	Paris	0.006	0.004	0.022
15	Sydney	0.043	0.082	0.048
16	Tokyo	0.025	0.019	0.030
17	Toronto	0.000	0.001	0.000

In order to see the general pattern of differences in variances for the three criteria we inspect Table 9 where the variances of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions resulting from long and short distance predictions are reported. In general, it can be said that the variance of the distribution of $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ for long distances is less than the variance for short distances for each criterion in each region. But this conclusion does not always hold at the 5% significance level as the Levene tests suggest in Table 8. To see the converging funnels formed by the difference in variances the scatter plots of $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ are inspected. The plots for United States and Toronto are shown in Love and Üster [14]. 5

The accuracy of each criterion in predicting the long versus short distances in a given region is examined by using the nonparametric 2-tailed

Table 9. Variances of $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ in predicting long and short distances

No.	Region	Short Distance			Long Distance		
		<i>AD</i>	<i>SD</i>	<i>NAD</i>	<i>AD</i>	<i>SD</i>	<i>NAD</i>
1	Australia	0.0036	0.0035	0.0030	0.0015	0.0015	0.0019
2	BC Province	0.0072	0.0073	0.0070	0.0058	0.0057	0.0064
3	Canada	0.0043	0.0039	0.0047	0.0003	0.0002	0.0003
4	France	0.0011	0.0010	0.0012	0.0007	0.0006	0.0007
5	Great Britain	0.0022	0.0021	0.0025	0.0014	0.0014	0.0021
6	NY State	0.0025	0.0024	0.0021	0.0025	0.0023	0.0034
7	Pennsylvania	0.0058	0.0060	0.0056	0.0012	0.0010	0.0014
8	United States	0.0012	0.0013	0.0013	0.0004	0.0003	0.0004
9	Brussels	0.0023	0.0023	0.0023	0.0009	0.0009	0.0009
10	London City	0.0364	0.0354	0.0367	0.0060	0.0045	0.0058
11	London North	0.0020	0.0019	0.0020	0.0005	0.0005	0.0006
12	Los Angeles	0.0343	0.0335	0.0352	0.0013	0.0012	0.0012
13	NY City	0.0053	0.0038	0.0053	0.0027	0.0023	0.0026
14	Paris	0.0145	0.0134	0.0136	0.0019	0.0018	0.0033
15	Sydney	0.0042	0.0039	0.0042	0.0026	0.0025	0.0027
16	Tokyo	0.0041	0.0040	0.0042	0.0006	0.0005	0.0006
17	Toronto	0.0025	0.0019	0.0026	0.0004	0.0004	0.0004

Mann-Whitney Test with a 5% significance level. The 2-tailed p-values of the tests for three criteria are presented in Table 10. In this table, a p-value less than 0.05 indicates that there is a significant difference between the long distance and short distance $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions of the geographical region. For the *AD* criterion only in New York City and Sydney and for the *SD* criterion only in New York State, New York City and Sydney, do the long distance and short distance $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions apparently come from the populations having the same distributions. For the *NAD* criterion the same conclusion holds in the regions of Australia, France, Great Britain, New York State, Pennsylvania, New York City, Paris and Sydney. As also seen in Table 7, the average percent absolute errors in a given region are considerably different for long and short distances and indeed the accuracy in predicting long distances is higher than the accuracy in predicting short distances in a given region.

4. Summary of Results and Conclusions

The three goodness-of-fit criteria, *AD*, *SD* and *NAD*, were compared with different perspectives for seventeen geographical regions including nine large geographical regions and eight urban centers. Several generalized conclusions, based on these seventeen regions, can be drawn from the above statistical tests.

Table 10. p-values of the Mann-Whitney Test for long and short distance $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$

No.	Region	<i>AD</i>	<i>SD</i>	<i>NAD</i>
1	Australia	0.0118	0.0197	0.5145
2	BC Province	0.0024	0.0044	0.0285
3	Canada	0.0000	0.0000	0.0000
4	France	0.0203	0.0070	0.0821
5	Great Britain	0.0042	0.0057	0.4920
6	NY State	0.0446	0.1622	0.9205
7	Pennsylvania	0.0124	0.0118	0.0548
8	United States	0.0061	0.0038	0.0093
9	Brussels	0.0340	0.0360	0.0307
10	London Central	0.0012	0.0316	0.0026
11	London North	0.0004	0.0008	0.0021
12	Los Angeles	0.0039	0.0093	0.0103
13	NY City	0.5413	0.0942	0.9859
14	Paris	0.0350	0.0177	0.2401
15	Sydney	0.1638	0.1501	0.1943
16	Tokyo	0.0061	0.0060	0.0132
17	Toronto	0.0061	0.0002	0.0123

- The $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ populations are non-normal for each criterion. The histograms are highly peaked with more occurrences close to zero and skewed right.
- There are generally no pairs of the criteria for which the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions are significantly different.
- In terms of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$, the three criteria are not significantly different in predicting either long distances or short distances. However, each criterion has a higher accuracy in predicting relatively long distances than in predicting relatively short distances.
- The variance of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distribution for long distances is significantly different than the variance of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distribution for short distances. The former is smaller than the latter, and hence the scatter plots of the $|e(\mathbf{x}_i, \mathbf{x}_j)|/A(\mathbf{x}_i, \mathbf{x}_j)$ distributions form a converging funnel as the actual distance between the sample points increases.

Finally, we can say that because of the computational efficiency provided by the closed form formula to determine the best value of parameter k when fitting the $k\ell_p$ -norm, and since the $e(\mathbf{x}_i, \mathbf{x}_j)$ have an expected value of zero without any exceptions in all the regions, it would seem to be advantageous to use the *SD* criterion in practice.

References

1. W. Berens. The suitability of the weighted ℓ_p -norm in estimating actual road distances. *European Journal of Operational Research*, 34:39–43, 1988.
2. W. Berens and F. J. Körling. Estimating road distances by mathematical functions. *European Journal of Operational Research*, 21:54–56, 1985.
3. J. Brimberg, P. D. Dowling, and R. F. Love. Estimating the parameters of the weighted ℓ_p -norm by linear regression. *IIE Transactions*, 28:363–367, 1996.
4. J. Brimberg and R. F. Love. Estimating travel distances by the weighted ℓ_p -norm. *Naval Research Logistics*, 38:241–259, 1991.
5. J. Brimberg, R. F. Love, and J. H. Walker. The effect of axis rotation on distance estimation. *European Journal of Operational Research*, 80:357–364, 1995.
6. Z. Drezner, editor. *Facility Location, A Survey of Applications and Methods*. Springer-Verlag, New York, 1995.
7. S. Eilon, C. D. T. Watson-Gandy, and N. Christofides. *Distribution Management: Mathematical Modeling and Practical Analysis*. Hafner Publishing Company, New York, 1971.
8. V. Ginsburgh and P. Hansen. Procedures for the reduction of errors in road network data. *Operational Research Quarterly*, 25:321–322, 1974.
9. R. Klein. Voronoi diagrams in the Moscow metric. In *Graph-Theoretic Concepts in Computer Science*, pages 434–441, 1988.
10. P. Kolesar, W. Walker, and J. Hausner. Determining the relation between fire engine travel times and travel distances in New York City. *Operations Research*, 23:614–627, 1975.
11. R. F. Love and J. G. Morris. Modeling inter-city road distances by mathematical models. *Operational Research Quarterly*, 23:61–71, 1972.
12. R. F. Love and J. G. Morris. Mathematical models of road travel distances. *Management Science*, 25:130–139, 1979.
13. R. F. Love, J. G. Morris, and G. O. Wesolowsky. *Facilities Location – Models and Methods*. North-Holland, New York, 1988.
14. R. F. Love and H. Üster. Comparison of the properties and the performance of the criteria used to evaluate the accuracy of distance predicting functions. Technical Report 419, McMaster University, Faculty of Business, Hamilton, Ontario, Canada, 1996.
15. R. F. Love and J. H. Walker. A new criterion for evaluating the accuracy of distance predicting functions. Technical Report 370, McMaster University, Faculty of Business, Hamilton, Ontario, Canada, 1991.
16. R. F. Love and J. H. Walker. Distance data for eighteen geographic regions. Technical Report 383, McMaster University, Faculty of Business, Hamilton, Ontario, Canada, 1993.
17. R. F. Love and J. H. Walker. An empirical comparison of block and round norms for modeling actual distances. *Location Science*, 2(1):21–43, 1994.
18. R. F. Love, J. H. Walker, and M. L. Tiku. Confidence intervals for $\ell_{k,p,\theta}$ distances. *Transportation Science*, 29:93–100, 1995.
19. M. M. Hollander and d. A. Wolfe. *Nonparametric Statistical Methods*. John Wiley and Sons, New York, 1973.
20. MicroAnalytics. *TruckStops 2*. Suite One, 2045 North 15th Street, Arlington, Virginia and 1986 Queen Street East, Toronto, Ontario, 1993.
21. Roadnet-Technologies. *ROADNET*. 10540 York Rd., Huntvalley, Maryland, 1993.
22. J. Star and J. Estes. *Geographic Information Systems – An Introduction*. Prentice-Hall, New Jersey, 1990.

23. J. E. Ward and R. E. Wendell. A new norm for measuring distance which yields linear location models. *Operations Research*, 28:836–844, 1980.
24. J. E. Ward and R. E. Wendell. Using block norms for location modeling. *Operations Research*, 33:1074–1090, 1985.
25. J. B. Westwood. A transport planning model for primary distribution. *Interfaces*, 8:1–10, 1977.