McMASTER
UNIVERSITY

MICHAEL G. DeGROOTE
SCHOOL OF BUSINESS

RESEARCH AND WORKING PAPER SERIES

COMPARISON OF THE PROPERTIES AND THE PERFORMANCE OF THE CRITERIA USED TO EVALUATE THE ACCURACY OF DISTANCE PREDICTING FUNCTIONS

By

Robert F. Love and **Halit Üster**

Department of Management Science/systems
McMaster University,
Hamilton, Ontario,
Canada L8S 4M4

Working Paper # 419

December, 1996

TER UNIVERSITY
Street West
Ontario, Canada L8S 4M4
(905) 525-9140

COMPARISON OF THE PROPERTIES AND THE PERFORMANCE OF THE CRITERIA USED TO EVALUATE THE ACCURACY OF DISTANCE PREDICTING FUNCTIONS

By

Robert F. Love and **Halit Üster**

Department of Management Science/systems
McMaster University,
Hamilton, Ontario,
Canada L8S 4M4

Working Paper # 419

December, 1996

Comparison of the Properties and the Performance of the Criteria Used to Evaluate the Accuracy of Distance Predicting Functions

ROBERT F. LOVE and HALİT ÜSTER

Department of Management Science/Systems

McMaster University, Hamilton, Ontario, L8S 4M4

Abstract: Distance predicting functions may be used in a variety of applications for estimating travel distances between points. To evaluate the accuracy of distance predicting functions, goodness-of-fit criteria are employed. AD_f (Absolute Deviations), SD_f (Squared Deviations) and NAD_f (Normalized Absolute Deviations) are the three criteria that are mostly employed for modelling distances. 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 the $\ell_{k,p,\theta}$ norm was employed as the distance predicting function. First we analyse statistical properties of the prediction errors, and then we statistically compare the three criteria by using absolute normalized error distributions in seventeen geographical regions.

When objects in space, such as different cities in a geographic region, activity centres 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 among national distribution centres, 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 (see Eilon et al. [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 et al. [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 (see e.g. Love, Morris and Wesolowsky [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 by MicroAnalytics in TruckStops2 [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 [19] 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. Sum of Absolute Deviations (AD_f),
2. Sum of Squared Deviations (SD_f),
3. Sum of Normalized Absolute Deviations (NAD_f),

(see, e.g. Berens [1]; Berens and Körling [2]; Brimberg, Dowling and Love [4]; Brimberg, Love and Walker [6]; Love and Morris [11, 12]; Love, Walker and Tiku [17]; and Ward and Wendell [22, 23]). In addition, AD_f and SD_f 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 [17] 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_f criterion is used, the $\ell_{k,p,\theta}$ norm will predict long distances more accurately than short distances. The SD_f criterion has been characterized as having prediction errors with better statistical properties but still being similar to the AD_f criterion in terms of its accuracy in predicting long distances (see Love and Morris [11, 12]).

The NAD_f criterion, on the other hand, has been assumed to predict short distances as accurately as long distances (Love and Walker [14]; Brimberg, Love and Walker [6]; Love, Walker and Tiku [17]; and Brimberg, Dowling and Love [4]).

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 $\ell_{k,p,\theta}$ norm as the distance predicting function. In section two, the three criteria and the distance predicting function are described. In section three, the statistical test procedures and results are presented. Finally, in section four, conclusions based on our analyses of these results are discussed.

THE DISTANCE PREDICTING FUNCTION AND THE GOODNESS-OF-FIT CRITERIA

The weighted ℓ_p norm ($\ell_{k,p,\theta}$) was employed as the distance predicting function. For the $\ell_{k,p,\theta}$ norm, the travel distance between the points $x_1 = (x_{11}, x_{12})$ and $x_2 = (x_{21}, x_{22})$ is given by

$$\ell_{k,p,\theta} = k\left[|x_{11}' - x_{21}'|^p + |x_{12}' - x_{22}'|^p\right]^{1/p}$$

where
$$\begin{pmatrix} x_{11}' & x_{12}' \\ x_{21}' & x_{22}' \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

$$k \in R^+,\ p \in [1, 2],\ \text{and}\ \theta \in [0, 90]\ .$$

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 (see Brimberg, Love and Walker [6]). 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 optimal 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 of the distance predicting function are computed to minimize the value of the selected criterion. The three goodness-of-fit criteria that will be analysed in this paper are the minimizations of the following sums:

$$AD_f = \sum_{i=1}^{n-1} \sum_{j=1+1}^{n} \left| d_f(x_i,x_j) - A(x_i,x_j) \right|,$$

$$SD_f = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{(d_f(x_i,x_j) - A(x_i,x_j))^2}{A(x_i,x_j)}$$

$$NAD_f = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\left| d_f(x_i,x_j) - A(x_i,x_j) \right|}{A(x_i,x_j)}$$

where n = the number of points in a data set,
$A(x_i,x_j)$ is the actual distance between x_i and x_j,
and $d_f(x_i,x_j)$ is the predicted distance between points x_i and x_j using distance predicting function f.

The first criterion, AD_f, is the minimization of the sum of absolute deviations. Since the terms in AD_f 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_f, is the minimization of the sum of squared deviations where each squared error term is weighted by $1/A(x_i,x_j)$. Squared deviations and the division by actual distance provide the criterion with certain desirable statistical properties (see Love and Morris [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 (Berens [1], Berens and Körling [2], Love and Morris [11, 12] and Ward and Wendel [22, 23]). The last criterion, NAD_f, is relatively new in the literature. It has been utilized by Love and Walker [14]; Brimberg, Love, Walker [6]; Love, Walker and Tiku [17]; and Brimberg, Dowling and Love [4]. With the NAD_f 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_f and the SD_f criteria are given by Brimberg and Love [5]. For the NAD_f criterion the computational procedure is identical to that of AD_f (Love and Walker [14]). In general, the best θ and p values are determined by using an incremental search procedure and a four-stage incremental search procedure, respectively. In order to find the best k value some properties of the criteria are used. It is known that AD_f is a convex function of k, and SD_f is a strictly convex function of k (Brimberg and Love [5]). NAD_f was shown to be a convex function of k by Love and Walker [14].

Therefore, when using the AD_f and NAD_f 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_f criteria is calculated with a simple closed formula derived by Brimberg and Love [5]. The property of having a closed-form formula to find the best value of parameter k makes the application of the SD_f criterion computationally more efficient than using either the AD_f or the NAD_f criterion.

In order to model the parameters of the $\ell_{k,p,\theta}$ norm Love and Walker [15] collected sample data from seventeen geographic regions. For each geographic region, 15 points were randomly chosen. These 15 points provided 105 actual distances to be modelled by the distance predicting function $\ell_{k,p,\theta}$ using each criterion. The actual distance data and point coordinates from the seventeen geographic regions are presented in Love and Walker [15]. The empirical parameter values for the $\ell_{k,p,\theta}$ norm and the corresponding minimum criterion values for seventeen geographic regions computed by Love and Walker [16] for the AD_f and SD_f criteria, and by Love and Walker [14] for the NAD_f criterion are given here in Tables 1and 2. Table 1 includes the parameter values and the criteria values corresponding to the minimum practical criteria values. Therefore it also includes parameter p values greater than 2. Table 2, on the other hand, reports the same information corresponding to the p values in the (0,2] interval. As stated by Brimberg and Love [4], two sets of parameter values are theoretically the same for a region. For example, for the NAD_f criterion in Australia in Table 1, p is 2.3281 and in Table 2 it is 1.7545, and the corresponding criterion values are 6.23 and 6.26 respectively. Although the two sets of parameters are theoretically same, in our analysis we have chosen the parameter values in Table 1 with very slightly smaller criterion values. It should be kept in mind that while searching for the parameter values of the $\ell_{k,p,\theta}$ norm in a region other than the ones included in this study, it is enough to search for a p value in the (0,2] interval.

	AD_f				SD_f				NAD_f			
Region	Criterion	θ	k	p	Criterion	θ	k	p	Criterion	θ	k	p
Australia	13106.91	3	1.1176	1.6848	1158.04	42	1.1827	2.1897	6.23	45	1.1605	2.3281
BC Province	6362.35	26	1.3912	2.6182	1027.59	23	1.3894	2.7290	11.04	16	1.3502	2.3744
Canada	8029.42	86	1.1772	1.4705	557.06	36	1.3384	3.1066	5.21	38	1.3621	3.4717
France	1592.26	68	1.0468	1.7734	92.32	70	1.0609	1.8430	3.61	71	1.0396	1.7417
Great Britain	2311.36	38	1.1185	1.8124	219.42	40	1.1095	1.7895	5.98	0	1.1032	1.8826
NY State	1637.44	86	1.1035	1.6946	159.80	86	1.0794	1.5823	6.00	87	1.0308	1.4950
Pennsylvania	1206.24	54	1.1544	2.5539	106.71	50	1.1573	2.5760	6.38	54	1.1392	2.4360
United States	6516.24	0	1.0817	1.7290	342.68	0	1.0792	1.6641	3.29	0	1.0825	1.7427
Brussels	46.46	46	1.0488	1.7660	3.55	47	1.0549	1.8180	4.68	45	1.0495	1.7802
London City	61.06	63	1.1328	2.3358	16.53	27	1.1354	2.0821	8.74	61	1.1359	2.3117
London North	27.88	61	1.1474	2.4789	1.73	57	1.1528	2.5676	4.62	56	1.1582	2.6086
Los Angeles	110.66	43	1.1760	2.7970	15.13	48	1.1909	2.7704	7.85	42	1.1757	2.6790
NY City	122.75	51	1.1741	2.6915	13.49	50	1.1751	2.3716	6.86	49	1.1510	2.3996
Paris	48.58	36	1.1204	2.2501	6.45	39	1.1066	2.2835	8.41	11	1.0635	1.6649
Sydney	13.11	7	1.1048	1.4061	1.35	8	1.1266	1.4719	6.50	6	1.0991	1.3940
Tokyo	28.89	15	1.1328	2.2059	2.29	13	1.1389	2.2262	4.30	20	1.1244	2.1492
Toronto	65.98	87	1.0118	1.1333	5.07	42	1.3140	5.176	4.45	87	1.0121	1.1261

Table 1. Optimal parameter values of $\ell_{k,p,\theta}$ for the criteria for $p \geq 1$

STATISTICAL TESTS AND THE PRESENTATION OF RESULTS

The purpose of this section is two-fold. First, the statistical properties of the errors in predicting distances are examined for each criterion, and for each region. Second, the statistical comparisons of the three criteria are conducted by adopting the absolute normalized error as the random variable.

Statistical Properties of Errors

For our work on road distances, the errors are the differences between actual distance and fitted distance pairs. The model that determines the relationship between the fitted distance and the actual distance is given by

$$A(x_i,x_j) = d_f(x_i,x_j) + e(x_i,x_j)$$

	AD_f				SD_f				NAD_f			
Region	Criterion	θ	k	p	Criterion	θ	k	p	Criterion	θ	k	p
Australia	13106.91	3	1.1176	1.6848	1163.59	0	1.1460	1.8585	6.26	1	1.0959	1.7545
BC Province	6369.31	71	1.2737	1.6322	1038.72	68	1.2495	1.5609	11.06	69	1.2701	1.7080
Canada	8029.42	86	1.1772	1.4705	565.61	83	1.1715	1.4849	5.23	85	1.1732	1.4584
France	1592.26	68	1.0468	1.7734	92.32	70	1.0609	1.8430	3.61	71	1.0396	1.7417
Great Britain	2311.36	38	1.1185	1.8124	219.42	40	1.1095	1.7895	5.98	0	1.1032	1.8826
NY State	1637.44	86	1.1035	1.6946	159.80	86	1.0794	1.5823	6.00	87	1.0308	1.4950
Pennsylvania	1207.34	3	1.0671	1.6274	107.06	4	1.0611	1.6244	6.48	7	1.0673	1.6958
United States	6516.24	0	1.0817	1.7290	342.68	0	1.0792	1.6641	3.29	0	1.0825	1.7427
Brussels	46.46	46	1.0488	1.7660	3.55	47	1.0549	1.8180	4.68	45	1.0495	1.7802
London City	61.45	18	1.0697	1.7524	16.53	72	1.1182	1.9241	8.75	8	1.0495	1.7802
London North	28.49	15	1.0638	1.6505	1.78	11	1.0599	1.6456	4.70	14	1.0591	1.6171
Los Angeles	111.89	89	1.0626	1.5699	15.50	2	1.0721	1.5734	7.90	87	1.0672	1.5684
NY City	124.29	5	1.0674	1.5822	13.58	6	1.1069	1.7340	6.90	4	1.0737	1.6975
Paris	48.71	75	1.0704	1.7859	6.52	86	1.0613	1.8189	8.41	11	1.0635	1.6649
Sydney	13.11	7	1.1048	1.4061	1.35	8	1.1266	1.4719	6.50	6	1.0991	1.3940
Tokyo	28.91	59	1.0961	1.8591	2.30	58	1.0963	1.8252	4.30	64	1.0963	1.8901
Toronto	65.98	87	1.0118	1.1333	5.10	88	1.0279	1.1863	4.45	87	1.0121	1.1261

Table 2. Optimal parameter values of $\ell_{k,p,\theta}$ for the criteria for $1 \le p \le 2$.

where $A(x_i,x_j)$ is the actual distance between points x_i and x_j, $d_f(x_i,x_j)$ is the predicted distance, and $e(x_i,x_j)$ is the error term for the x_i, 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, $d_f(x_i,x_j)$, 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.

The error term, $e(x_i,x_j)$, is a continuous random variable since it is a purely random part of the actual distance, $A(x_i,x_j)$, that cannot be explained by the model. It is assumed that the errors for different pairs of points in a region are independent, i.e., the errors of $d_f(x_i,x_j)$ about $A(x_i,x_j)$ are not related to the errors of $d_f(x_k,x_l)$ about $A(x_k,x_l)$ for the points i, j, k, l in a geographic region. In order to examine the statistical properties of the error term populations, the $e(x_i,x_j)$'s which form a random sample of 105 observations for each region were used to calculate the sample statistics for the population parameters. The estimates of the population mean and variance are the sample mean $(\bar{x})$ and the sample variance (s^2), respectively. The estimates of the population's Pearson coefficients of skewness and kurtosis, $\sqrt{\beta_1}$ and β_2, are the sample Pearson coefficients which are denoted by $\sqrt{b_1}$ and b_2. For a large sample $(n \geq 100)$, $\sqrt{b_1}$ and b_2 are unbiased estimators of $\sqrt{\beta_1}$ and β_2 (Stuart and Ord [20]). The $\bar{x}$, s^2, $\sqrt{b_1}$ and $(b_2 - 3)$ values are presented in Table 3 for the seventeen regions. In this table kurtosis is given as $(b_2 - 3)$ because the SPSS [18] reports it in this way for convenient comparison purposes with the normal distribution. In order to determine whether the $e(x_i,x_j)$'s are from a normal distribution, we first check the sample Pearson coefficients. The sample Pearson coefficients not only indicate how skewed and peaked the samples are, but also provide an indication of how skewed and peaked the populations from which the samples were drawn are. If $\sqrt{b_1}$ is less than zero, then the sample is skewed left, and if $\sqrt{b_1}$ is greater than zero, then the sample is skewed right. A $\sqrt{b_1}$ value of zero indicates that the sample is symmetric around its mean. A (b_2-3) value which is less than zero (greater than zero) indicates that the sample is less peaked (more peaked) than a sample from a normal population which would have a (b_2-3) value of zero. The sample Pearson coefficients for the different geographic regions confirm that the populations are non-normal.

However, the degree of non-normality varies from region to region. In most regions the distributions are skewed right and are more peaked than the normal distribution.

Besides the sample Pearson coefficients, Normal Probability plots and histograms were examined. The related graphs are given in Figures 1 and 2 for the United States and Toronto respectively. On normal probability plots, a linear relation is expected between the observed cumulative probabilities and the expected cumulative probabilities for a sample distribution to be from a normally distributed population. The histograms are expected to have a symmetric bell-shaped appearance with no violations at the tails. The normal probability plots and histograms also confirm that there is enough evidence to assume non-normal distributions of errors for the seventeen geographic regions.

In order to test the equality of the variances for the three criteria the Levene test (using a 5% significance level) was conducted for each region. Levene's test is a powerful test when the data come from continuous, but not necessarily normal distributions. The p-values for the two-tail significance test are listed for each geographic region in Table 4. Since the p-values for the 2-tailed Levene test are greater than 0.10, it is confirmed that the $e(x_i,x_j)$ distributions of AD_f, SD_f and NAD_f have the same variance at the 5% significance level in all the regions.

Region	$\bar{x}$	S^2	$\sqrt{b_1}$	$(b_2 - 3)$	Urban Center	$\bar{x}$	S^2	$\sqrt{b_1}$	$(b_2 - 3)$	Criteria
	19.229	27908.360	1.154	1.023		0.029	0.377	-0.219	1.911	AD_f
Australia	11.091	26810.490	1.032	0.936	Brussels	0.033	0.375	-0.192	1.704	SD_f
	76.609	30706.590	1.107	0.652		0.040	0.376	-0.181	1.840	NAD_f
	6.136	6920.109	0.842	1.649		0.326	1.097	2.245	8.365	AD_f
BC Province	9.781	6869.654	0.770	1.581	London City	0.158	1.125	2.150	7.674	SD_f
	18.070	7263.024	1.008	1.896		0.290	1.103	2.245	8.363	NAD_f
	4.397	10377.210	0.468	1.174		0.005	0.112	-0.329	-0.094	AD_f
Canada	5.260	9960.011	0.455	0.695	London North	0.017	0.112	-0.379	0.004	SD_f
	7.382	10320.270	0.474	1.184		0.001	0.116	-0.419	0.128	NAD_f
	2.890	415.358	1.074	2.110		0.373	1.833	0.248	0.536	AD_f
France	0.673	409.954	1.035	2.104	Los Angeles	0.144	1.911	0.068	0.219	SD_f
	4.022	422.440	1.162	2.525		0.288	1.883	0.294	0.461	NAD_f
	0.242	949.459	1.099	1.914		0.442	2.559	1.294	2.649	AD_f
Great Britain	2.104	957.126	1.329	1.837	NY City	0.128	2.584	1.085	2.086	SD_f
	10.496	1168.140	1.075	2.030		0.543	2.528	1.239	2.294	NAD_f
	0.040	539.481	1.658	3.669		-0.034	0.373	-0.794	1.237	AD_f
NY State	1.523	542.542	1.606	3.353	Paris	0.062	0.361	-0.767	1.346	SD_f
	9.030	693.586	1.477	1.905		-0.087	0.432	-0.902	1.245	NAD_f
	0.845	230.726	0.366	0.010		0.022	0.029	0.701	2.128	AD_f
Pennsylvania	1.021	223.838	0.318	-0.107	Sydney	0.013	0.029	0.504	1.879	SD_f
	2.516	241.910	0.481	0.074		0.026	0.030	0.762	2.256	NAD_f
	17.026	7452.193	1.089	1.678		0.051	0.148	1.316	3.653	AD_f
United States	3.367	7120.392	0.992	1.606	Tokyo	0.022	0.148	1.267	3.598	SD_f
	19.085	7556.449	1.098	1.670		0.073	0.150	1.409	3.759	NAD_f
						-0.052	0.775	-0.791	2.704	AD_f
					Toronto	0.049	0.756	-0.729	2.388	SD_f
						-0.108	0.787	-0.813	2.614	NAD_f

Table 3. Sample statistics of $e(x_i,x_j)$ for seventeen regions

Region	p-value	Urban Center	p-value
Australia	0.668	Brussels	0.999
BC Province	0.973	London City	0.995
Canada	0.993	London North	0.996
France	0.997	Los Angeles	0.925
Great Britain	0.524	NY City	0.994
NY State	0.284	Paris	0.840
Pennsylvania	0.983	Sydney	0.999
United States	0.998	Tokyo	0.995
		Toronto	0.993

Table 4. Two-tail values for the Levene test (equality of variances) of the three criteria

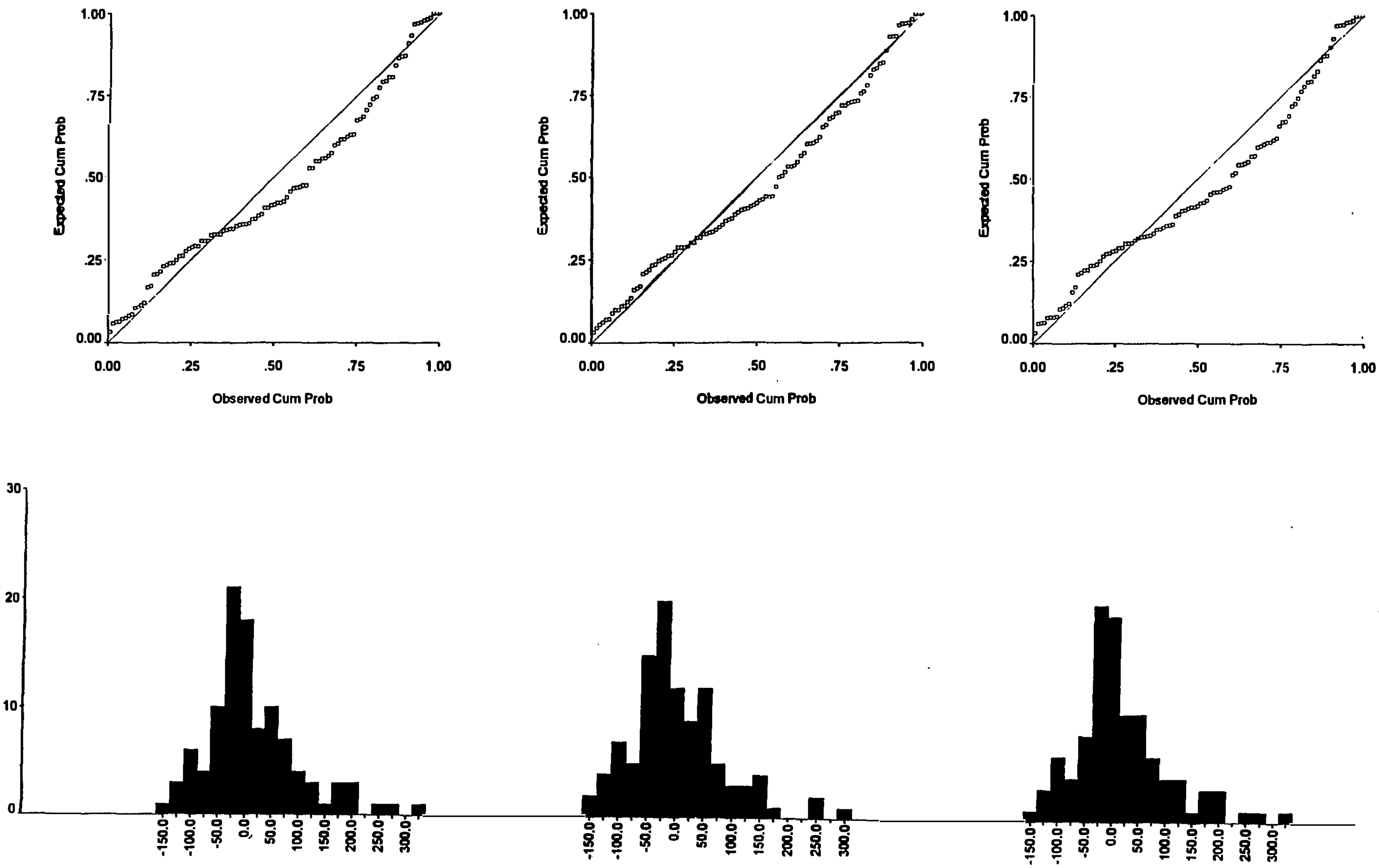


Figure 1. Normal Probability Plots and Histograms of $e(x_i,x_j)$ distributions for United States- left to right AD_f, SD_f, NAD_f criterion

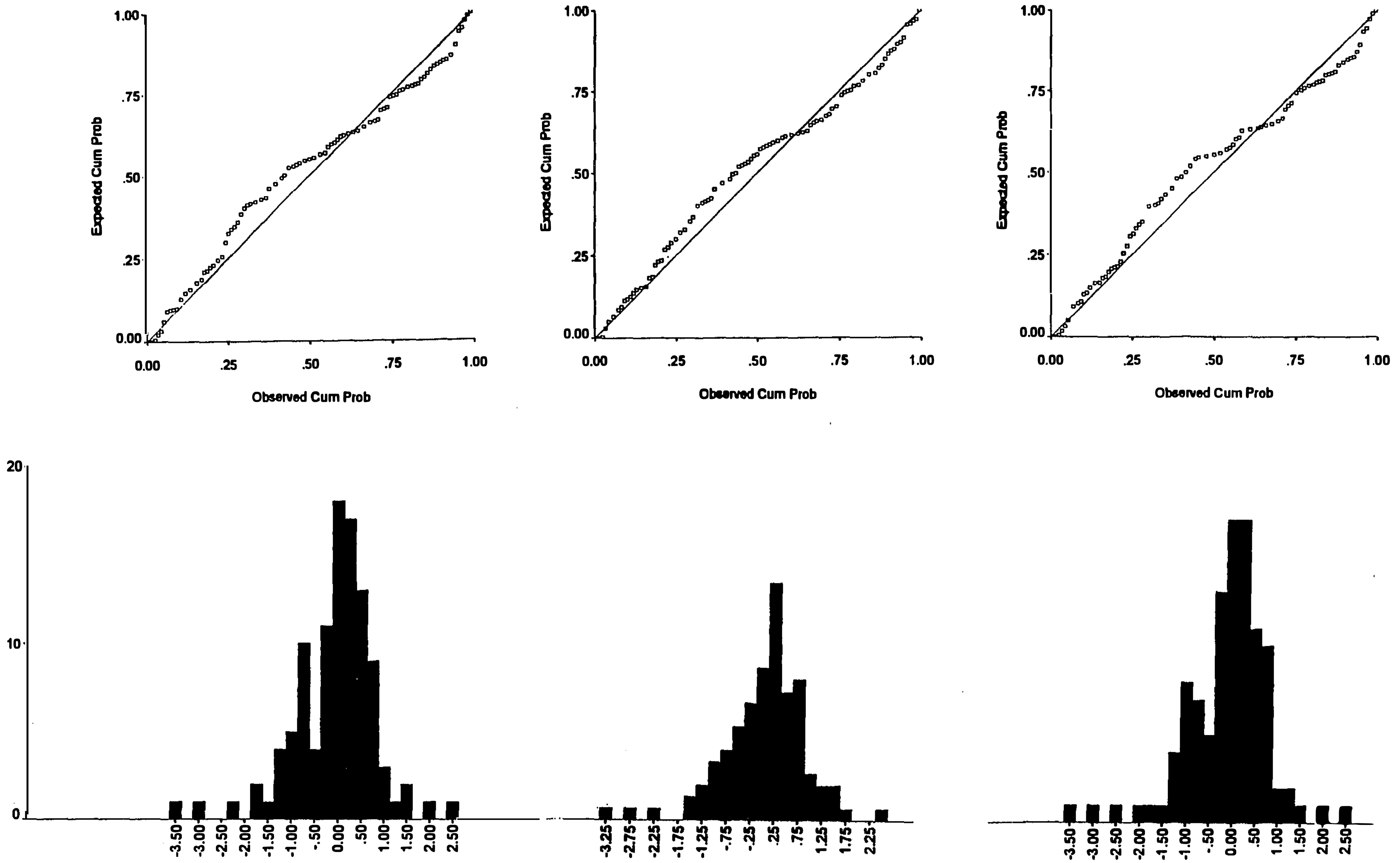


Figure 2. Normal Probability Plots and Histograms of $e(x_i,x_j)$ distributions for Toronto- left to right AD_f , SD_f , NAD_f criterion

To examine the homoscedasticity (see Wesolowsky [24]) for each criterion, the sample sets of 105 pairs are divided into three groups after they are ordered in their increasing order of actual distances for each geographic region. The first and the third groups, the 35 short actual distance pairs and the 35 long actual distance pairs respectively, are extracted to use in testing the homoscedasticity of the $e(x_i,x_j)$'s for each criterion. In order to clarify what is meant by long and short actual distances, Table 5 was constructed. The means of the long actual distance and short actual distance distributions, and also the ratio of the former to the latter are listed in Table 5. The ratios are not too much different for all regions except Canada which has a relatively large ratio of mean long actual distances to mean short actual distances of 5.605.

Region	Long D.	Short D.	Ratio	Urban Center	Long D.	Short D.	Ratio
Australia	3567.42	1072.83	3.325	Brussels	16.22	5.83	2.782
BC Province	1017.98	299.65	3.397	London City	12.50	4.39	2.847
Canada	4289.72	765.33	5.605	London North	11.08	3.58	3.095
France	727.06	258.17	2.816	Los Angeles	25.85	9.45	2.714
Great Britain	701.80	193.54	3.626	NY City	27.93	10.29	2.714
NY State	428.72	117.36	3.653	Paris	9.69	3.39	2.858
Pennsylvania	373.41	102.37	3.647	Sydney	3.32	1.25	2.656
United States	3596.99	1078.15	3.336	Tokyo	11.22	4.28	2.621
				Toronto	26.11	9.01	2.898

Table 5. The means of $A(x_i,x_j)$ for long distance and short distance distributions

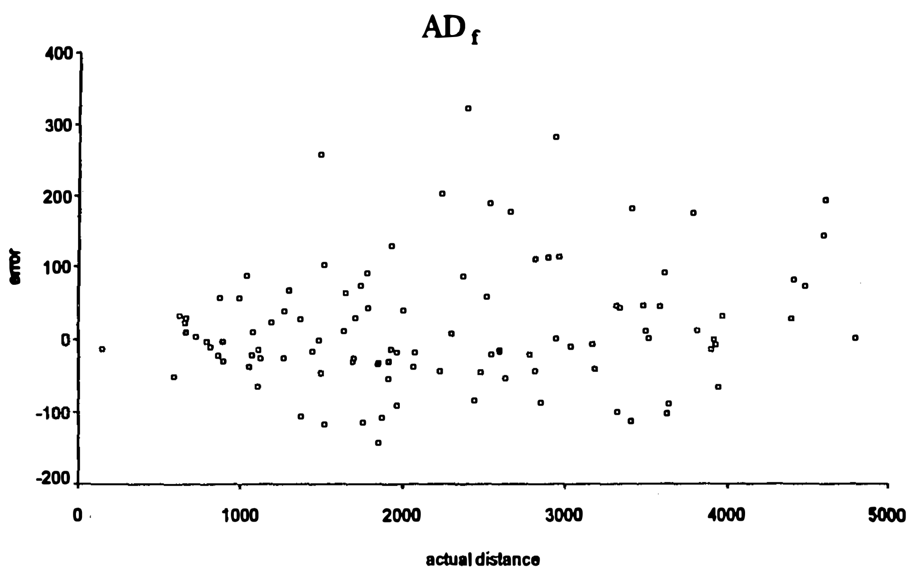
Levene tests for equality of variances of the prediction error distributions for the long and short distances for the three criterion in each region are conducted. The standard deviation (σ) of the $e(x_i,x_j)$ distributions for the long and short 35 pairs, and the 2-tail p-values of the Levene test are presented for each criterion in Table 6.

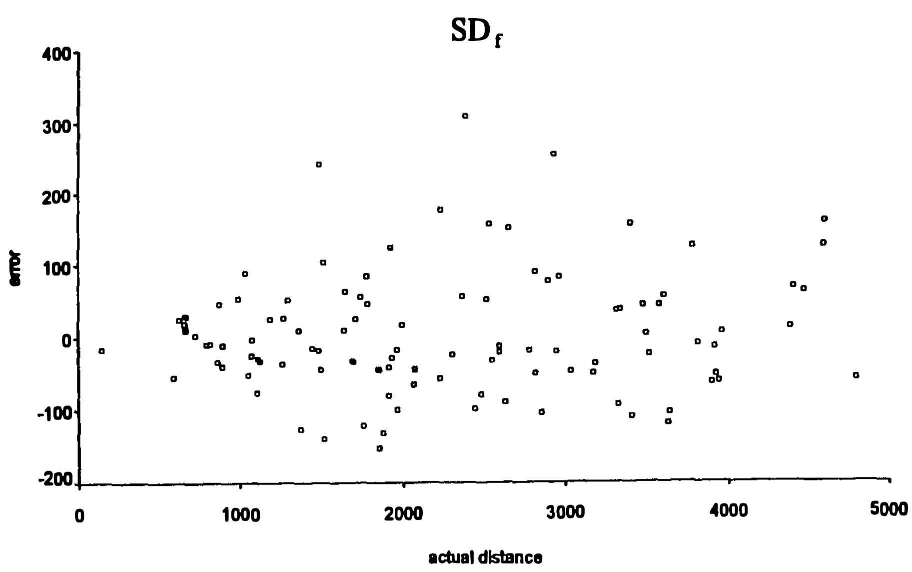
The p-values, based on a 5% significance level, suggest that the $e(x_i,x_j)$'s for each criterion are heteroscedastic except possibly in the five regions of Canada, London City, Los Angeles, Tokyo and Toronto out of the seventeen geographic regions. The standard deviations are always higher for the long distance pairs. To illustrate this, the scatter plots of $e(x_i,x_j)$'s for the United States and Toronto are presented in Figures 3 and 4 respectively. These scatter plots also confirm that the $e(x_i,x_j)$'s increase as the actual distance between the pairs increases.

Finally, we were interested in the expected values of the $e(x_i,x_j)$'s for each criterion. Since we already have enough evidence for the non-normality and the heteroscedasticity of the $e(x_i,x_j)$ distributions, a nonparametric test, the Wilcoxon Signed Rank test with 5% significance level, was performed to see if the $E[e(x_i,x_j)]$'s for each criterion is equal to zero. The results of the test are given in Table 7. The p-values, which are greater than 0.05, present enough evidence to conclude that the $e(x_i,x_j)$'s for the three criteria have an expected value of zero. The possible exceptions are 3 regions for the AD_f criterion (London City, Los Angeles and New York City), 5 regions for the NAD_f criterion (Australia, Great Britain, New York State, London City, New York City) and none for the SD_f criterion.

Statistical Comparison of the Three Criteria

In order to compare the three criteria, we used a transformed random variable given as $|e(x_i,x_j)|/A(x_i,x_j)$. There are three reasons for using this transformation. First, the new random variable frees the error terms, the $e(x_i,x_j)$'s, from their directions so that the absolute errors are to be compared. Second, since each criteria 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 each criterion. 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 ($|e(x_i,x_j)|A(x_i,x_j)$).





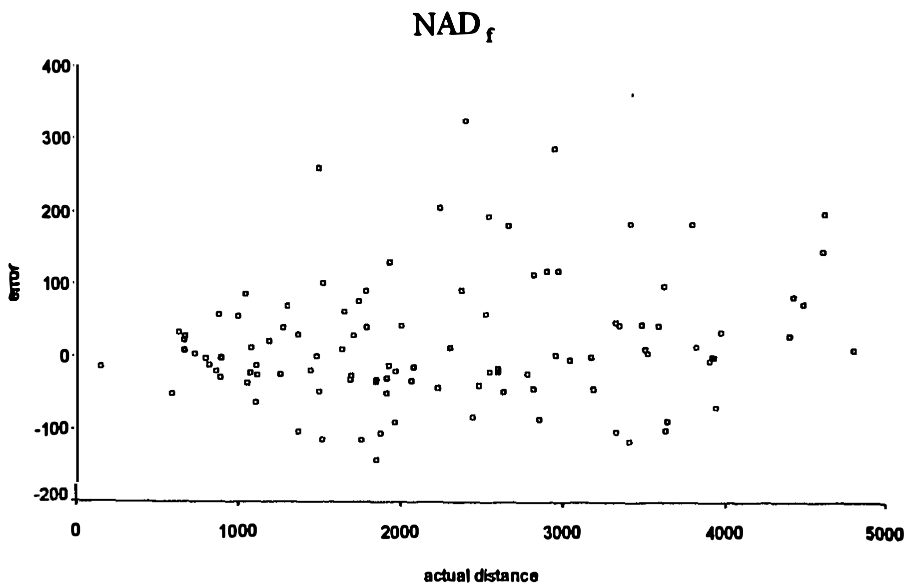
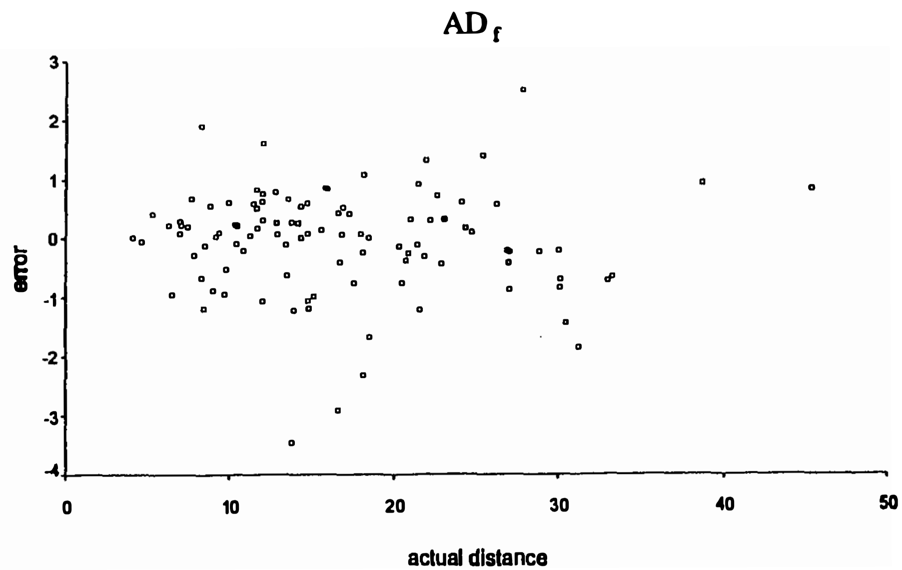
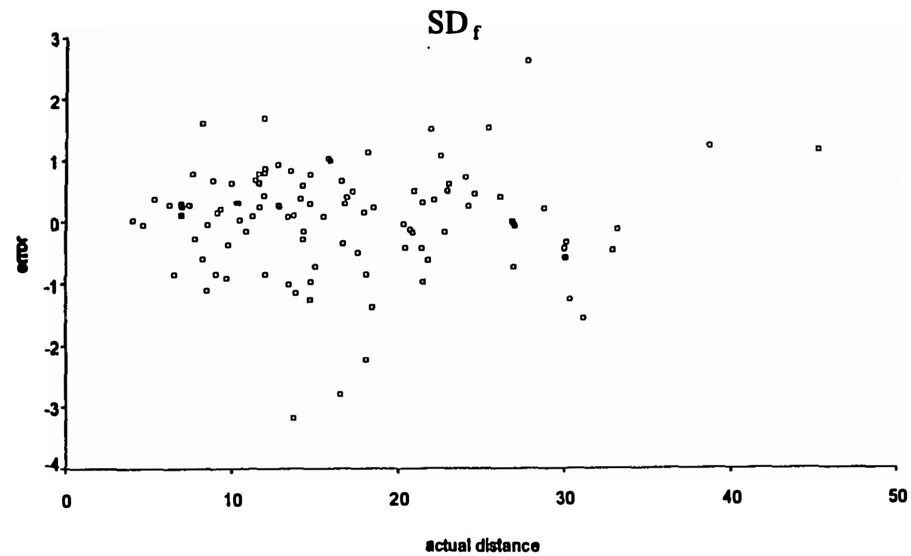


Figure 3.Scatter Plots of $e(x_i,x_j)$ distribution for United States





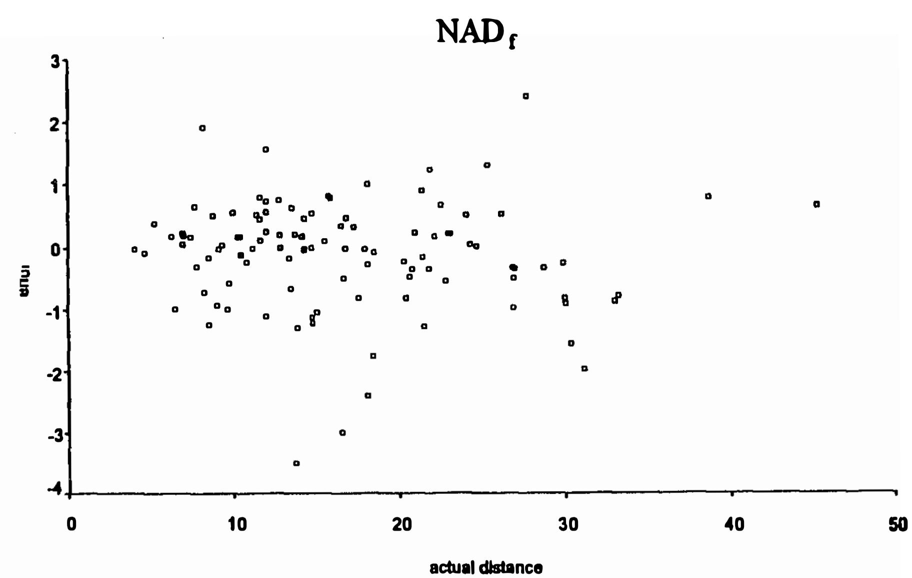


Figure 4.Scatter Plots of $e(x_i,x_j)$ distributions for Toronto

Geographical Regions	AD_f			SD_f			NAD_f		
	σ_S	σ_L	p-value	σ_S	σ_L	p-value	σ_S	σ_L	p-value
Australia	72.950	198.770	0.000	71.580	199.250	0.000	69.820	192.570	0.000
British Columbia	47.950	105.740	0.001	47.660	104.730	0.001	46.260	109.510	0.000
Canada	80.900	110.600	0.184	78.160	109.850	0.125	79.550	107.610	0.235
France	11.250	23.490	0.001	11.280	23.820	0.001	11.050	23.670	0.001
Great Britain	16.550	34.510	0.010	16.460	34.410	0.010	17.280	39.290	0.000
NY State	8.240	31.690	0.000	8.270	31.310	0.000	9.070	33.660	0.000
Pennsylvania	8.570	17.730	0.000	8.800	16.620	0.001	8.390	17.970	0.000
United States	64.620	92.190	0.030	66.250	87.720	0.033	64.390	93.430	0.031
Brussels	0.407	0.684	0.022	0.408	0.682	0.022	0.406	0.680	0.021
London City	0.839	0.982	0.294	0.843	0.979	0.306	0.838	0.987	0.280
London North	0.248	0.357	0.020	0.238	0.378	0.009	0.236	0.388	0.007
Los Angeles	1.232	1.409	0.363	1.221	1.465	0.255	1.222	1.434	0.284
NY City	0.981	2.005	0.006	0.953	2.007	0.005	0.932	1.959	0.006
Paris	0.579	0.714	0.040	0.565	0.709	0.038	0.551	0.840	0.005
Sydney	0.117	0.233	0.003	0.114	0.232	0.002	0.116	0.235	0.002
Tokyo	0.380	0.390	0.392	0.381	0.407	0.300	0.381	0.386	0.427
Toronto	0.678	0.856	0.163	0.648	0.840	0.162	0.681	0.817	0.168

Table 6. Standard deviations of the $e(x_i,x_j)$ distributions, and the 2-tail p-values of Levene Test for homoscedasticity

Large Geo. Region	p-value AD_f	p-value SD_f	p-value NAD_f	Urban Center	p-value AD_f	p-value SD_f	p-value NAD_f
Australia	0.815	0.654	0.003	Brussels	0.598	0.548	0.490
BC Province	0.944	0.480	0.134	London City	0.007	0.964	0.048
Canada	0.862	0.801	0.565	London North	0.658	0.381	0.681
France	0.555	0.621	0.246	Los Angeles	0.012	0.364	0.086
Great Britain	0.269	0.740	0.042	NY City	0.039	0.983	0.007
NY State	0.060	0.278	0.035	Paris	0.801	0.080	0.747
Pennsylvania	0.996	0.854	0.291	Sydney	0.295	0.509	0.207
United States	0.309	0.649	0.202	Tokyo	0.486	0.916	0.235
				Toronto	0.978	0.208	0.492

Table 7. Wilcoxon signed rank test results for $E[e(x_i,x_j)] = 0$

This section is presented in three sets of comparisons: (i) the comparison of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for 105 pairs of each geographic region, (ii) the comparison of the accuracy of the three criteria in predicting the 35 pairs of long distances and 35 pairs of short distances in a given region, and (iii) the comparison of the accuracy for each criteria in predicting the long distances versus short distances in a given region.

(i) In order to compare the absolute normalized errors, $|e(x_i,x_j)|/A(x_i,x_j)$, their distribution for each criterion was first checked for normality. For that purpose and also to present the descriptive statistics for each distribution, Table 8, which includes the means $(\bar{x})$, variances (s^2), skewness (b_1) and kurtosis (in(b_2-3) form), is constructed. Furthermore, the normal probability plots and histograms for each criterion and region were constructed. Two of the normal probability plots and histograms for the United States and Toronto are presented in Figures 5 and 6, respectively.

In Table 8, we observe that skewness and kurtosis values for the distributions are different enough from zero that we cannot conclude the distributions of $|e(x_i,x_j)|/A(x_i,x_j)$ are from normal distributions for each criteria in the regions. The normal probability plots and histograms in Figures 5 and 6 also support the non-normality of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions. Therefore a nonparametric test was applied to determine if the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for each criterion were significantly different from each other in a given region. The Friedman Test, which is used for multiple matched samples, was employed as the main effect test to compare the three $|e(x_i,x_j)|/A(x_i,x_j)$ distributions at the 5% significance level.The p-values for seventeen geographic regions are listed in Table 9. Since the p-values in Table 9 are well above 0.05,no pair of criteria is significantly different at the 5% significance level. The mean absolute errors in percentages are reported for each criterion and region in Table 10. Based on the figures in Table 10, 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_f, SD_f, and NAD_f criteria, respectively.

Region	$\bar{x}$	S^2	$\sqrt{b_1}$	$(b_2 - 3)$	Urban Center	$\bar{x}$	S^2	$\sqrt{b_1}$	$(b_2 - 3)$	Criteria
	0.0634	0.0027	1.5524	2.9709		0.0446	0.0017	1.6735	3.7924	AD_f
Australia	0.0636	0.0025	1.5999	3.3357	Brussels	0.0447	0.0017	1.6599	3.8810	SD_f
	0.0593	0.0027	1.2950	1.892		0.0446	0.0017	1.6728	3.7951	NAD_f
	0.1069	0.0071	0.7062	-0.1670		0.0833	0.0212	4.0173	19.1818	AD_f
BC Province	0.1063	0.0069	0.7309	-0.0870	London City	0.0872	0.0195	4.2164	20.9644	SD_f
	0.1052	0.0067	0.7595	0.0682		0.0832	0.021	4.1062	19.9933	NAD_f
	0.0499	0.0028	1.747	3.2356		0.0446	0.0013	1.2163	1.5499	AD_f
Canada	0.0504	0.0025	1.5718	2.1566	London North	0.0442	0.0012	1.2857	1.8672	SD_f
	0.0497	0.0028	1.8150	3.3933		0.0440	0.0012	1.2825	1.8599	NAD_f
	0.0335	0.0009	1.7694	4.1584		0.0751	0.0138	6.7183	57.5431	AD_f
France	0.0341	0.0008	1.7285	4.1728	Los Angeles	0.0757	0.0131	7.0910	62.3922	SD_f
	0.0334	0.0010	1.9857	5.1221		0.0748	0.0139	6.9068	59.9512	NAD_f
	0.0591	0.0025	1.1108	1.3866		0.0660	0.0039	1.7232	2.6192	AD_f
Great Britain	0.0587	0.0024	1.2424	1.9158	NY City	0.0682	0.0031	1.5462	2.3413	SD_f
	0.0570	0.0027	1.6056	3.2751		0.0653	0.0039	1.6307	2.4097	NAD_f
	0.0623	0.0023	0.8010	0.1435		0.0815	0.0068	3.0579	13.457	AD_f
NY State	0.0612	0.0021	1.0104	0.5856	Paris	0.0827	0.0063	2.9045	12.1986	SD_f
	0.0626	0.0026	1.2077	1.0528		0.0801	0.0069	2.4590	9.0103	NAD_f
	0.0617	0.0033	1.6327	3.7136		0.0621	0.0033	1.4150	2.0364	AD_f
Pennsylvania	0.0623	0.0033	1.7010	3.7626	Sydney	0.0620	0.0031	1.3346	1.8110	SD_f
	0.0608	0.0031	1.5613	3.5665		0.0617	0.0033	1.4393	2.1473	NAD_f
	0.0314	0.0009	1.8475	4.8344		0.0410	0.0021	3.9404	23.7253	AD_f
United States	0.0326	0.0008	1.8210	4.1774	Tokyo	0.0414	0.0021	3.8788	23.1547	SD_f
	0.0314	0.0009	1.8510	4.9158		0.0409	0.0022	3.9262	23.5744	NAD_f
						0.0425	0.0020	2.2888	6.7672	AD_f
					Toronto	0.0440	0.0017	1.9516	5.0684	SD_f
						0.0424	0.0021	2.2782	6.6037	NAD_f

Table 8. Descriptive statistics of the $|e(x_i,x_j)|/A(x_i, x_j)$ for geographic regions

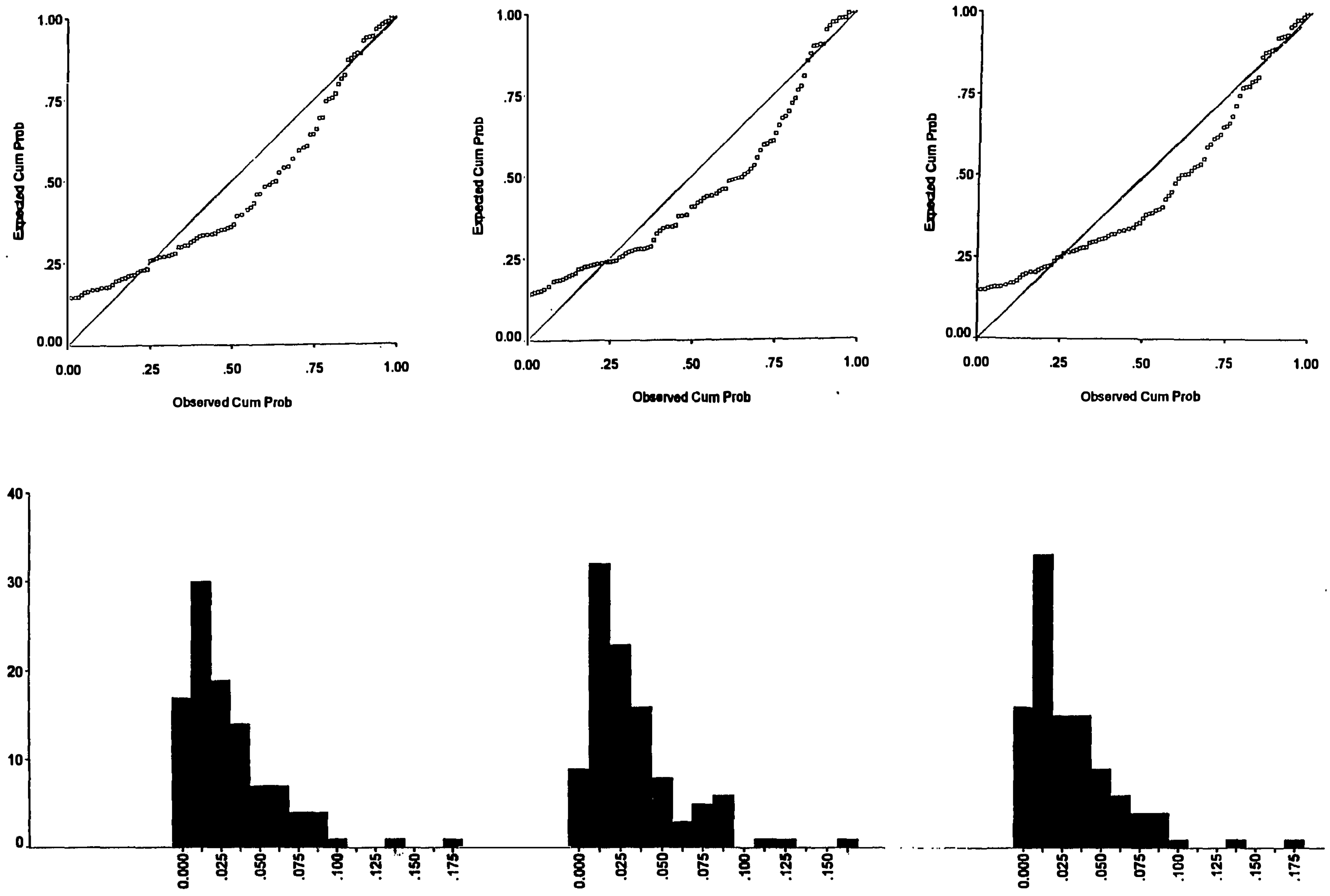


Figure 5. Normal Probability Plots and Histograms of $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for United States- left to right AD_f, SD_f, NAD_f criterion.

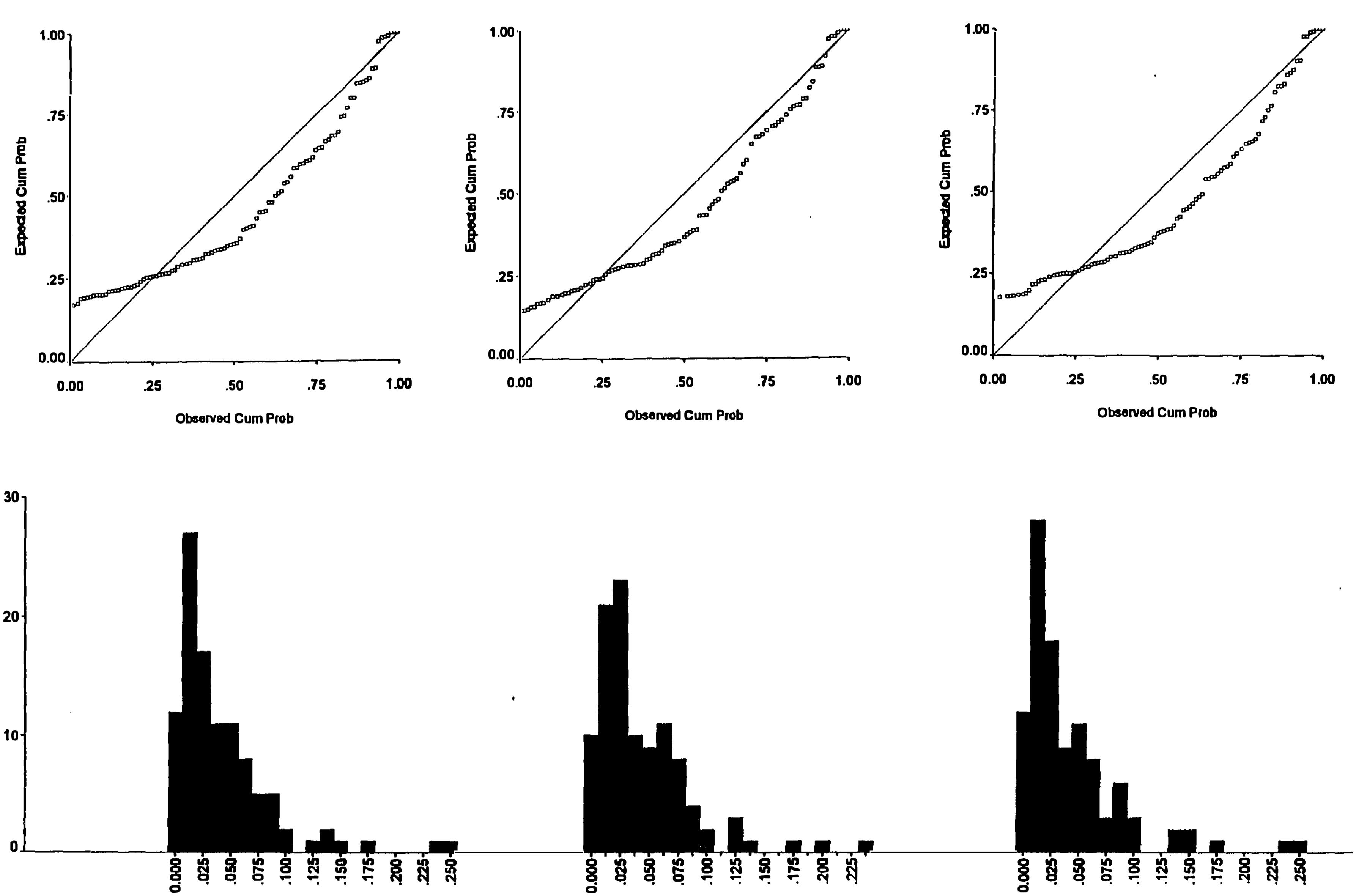


Figure 6. Normal Probability Plots and Histograms of $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for Toronto - left to right AD_f, SD_f, NAD_f criterion.

Region	p-value	Urban Center	p-value
Australia	0.3732	Brussels	0.9765
British Columbia	0.3750	London City	0.6386
Canada	0.2315	London North	0.6839
France	0.4806	Los Angeles	0.3724
Great Britain	0.9355	New York City	0.7062
New York State	0.4047	Paris	0.1038
Pennsylvania	0.7280	Sydney	0.7606
United States	0.8889	Tokyo	0.8816
		Toronto	0.2450

Table 9. The p-values for the Friedman Test of $|e(x_i,x_j)|/A(x_i,x_j)$ distributions of the criteria

Region	AD_f	SD_f	NAD_f	Urban Center	AD_f	SD_f	NAD_f
Australia	6.34	6.36	5.93	Brussels	4.46	4.47	4.46
British Columbia	10.69	10.63	10.52	London City	8.33	8.72	8.33
Canada	4.99	5.04	4.97	London North	4.46	4.42	4.40
France	3.35	3.41	3.34	Los Angeles	7.51	7.57	7.48
Great Britain	5.91	5.87	5.70	NY City	6.60	6.82	6.53
NY State	6.23	6.12	6.26	Paris	8.15	8.27	8.01
Pennsylvania	6.17	6.23	6.08	Sydney	6.21	6.20	6.17
United States	3.14	3.26	3.14	Tokyo	4.10	4.14	4.09
				Toronto	4.25	4.40	4.24

Table 10. Absolute percent errors in predicting distances for the criteria in geographic regions.

We next test the distributions of the random variable $|e(x_i,x_j)|/A(x_i,x_j)$ (corresponding to long and short actual distances) for normality. Non-normality of the $|e(x_i,x_j)|/A(x_i,x_j)$ distribution formed by 105 pairs of 15 points in a given region does not guarantee that a subset of these 105 pairs which is nonrandomly formed by 35 pairs of points, also does not come from a non-normal distribution. There are six different distributions used in the comparisons to identify the differences between the three criteria for predicting long actual distances and short actual distances. The six distributions are sketched in Figure 7.

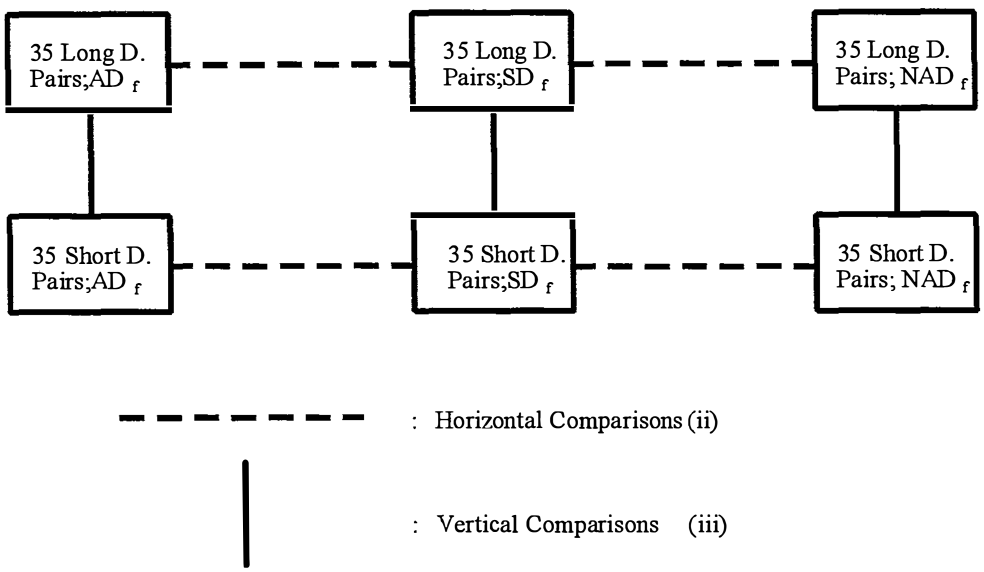


Figure 7 :Six distributions for long and short distance pairs comparisons.

If at least two of six distributions, one from each row, are not from a normal distribution, then a nonparametric test should be used for the main effect test of horizontal comparisons (i.e., for part (ii)). If at least three, each from a different criterion, of six distributions are not from a normal distribution, then vertical comparisons (i.e., for part (iii)) should be performed by using a nonparametric test. Therefore, the skewness, $\sqrt{b_1}$, and kurtosis, (b_2-3), values of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for long distances using the AD_f criterion, for short distances using the SD_f criterion and for long distances using the NAD_f criterion are reported for each region in Table 11. If we guarantee that in the six distributions of Figure 7 there is at least one distribution in each row and one distribution in each column coming from non-normal populations, then we need to use nonparametric tests for the following parts of the section. In addition, for each criterion the normal probability plots and histograms of the above mentioned distributions for the United States and

Toronto are represented in Figures 8 and 9, respectively. The skewness and kurtosis values in Table 11 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 in Figures 8 and 9. Therefore, nonparametric tests should be used for unbiased comparisons of the criteria involving the six distributions outlined in Figure 7.

Geographical Region	Long Distance- AD_f		Short Distance- SD_f		Long Distance- NAD_f	
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Australia	0.4881	-0.6541	1.2536	1.3505	0.4131	-0.9182
British Columbia	0.1244	1.0772	0.3501	-0.1516	1.3306	1.6137
Canada	1.3520	1.6556	0.6961	-0.5406	1.3277	1.8534
France	2.6315	9.8096	2.0758	5.8926	2.7389	10.1983
Great Britain	1.4349	1.3900	0.5296	-0.2207	1.1993	0.9675
New York State	0.8914	-0.2543	0.9814	1.0950	0.7519	-0.7202
Pennsylvania	0.5275	-1.0158	1.2628	1.0579	0.4775	-1.0018
United States	1.6883	4.4469	1.4650	2.5399	1.6432	4.1495
Brussels	1.5948	2.4744	1.8813	5.0096	1.6090	2.6409
London City	3.3989	12.8349	3.8639	15.9083	3.4378	13.7044
London North	0.6037	-0.7204	0.9873	0.3199	0.8531	0.2911
Los Angeles	1.4745	2.5833	4.8098	26.0766	1.4541	2.7294
New York City	2.2495	6.4738	0.9857	-0.2267	2.1354	5.8431
Paris	0.8056	-0.0447	2.0891	4.8714	0.7644	-0.5594
Sydney	1.9071	4.9922	0.5940	-0.3163	2.0189	5.6586
Tokyo	1.7651	5.3646	3.3773	15.1482	1.9370	6.3196
Toronto	1.4835	2.2125	1.1633	1.5794	1.2894	1.4425

Table 11. Skewness ($\sqrt{b_1}$) and kurtosis (b_2-3) values for the normality of three distributions.

(ii) 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. However, in order to obtain more comparisons between the criteria, tests can be applied to check the variances for equality. There are two sets of horizontal comparisons (long distances and short distances) and each set includes three distributions formed by 35 pairs of points in a region (see Figure 7).

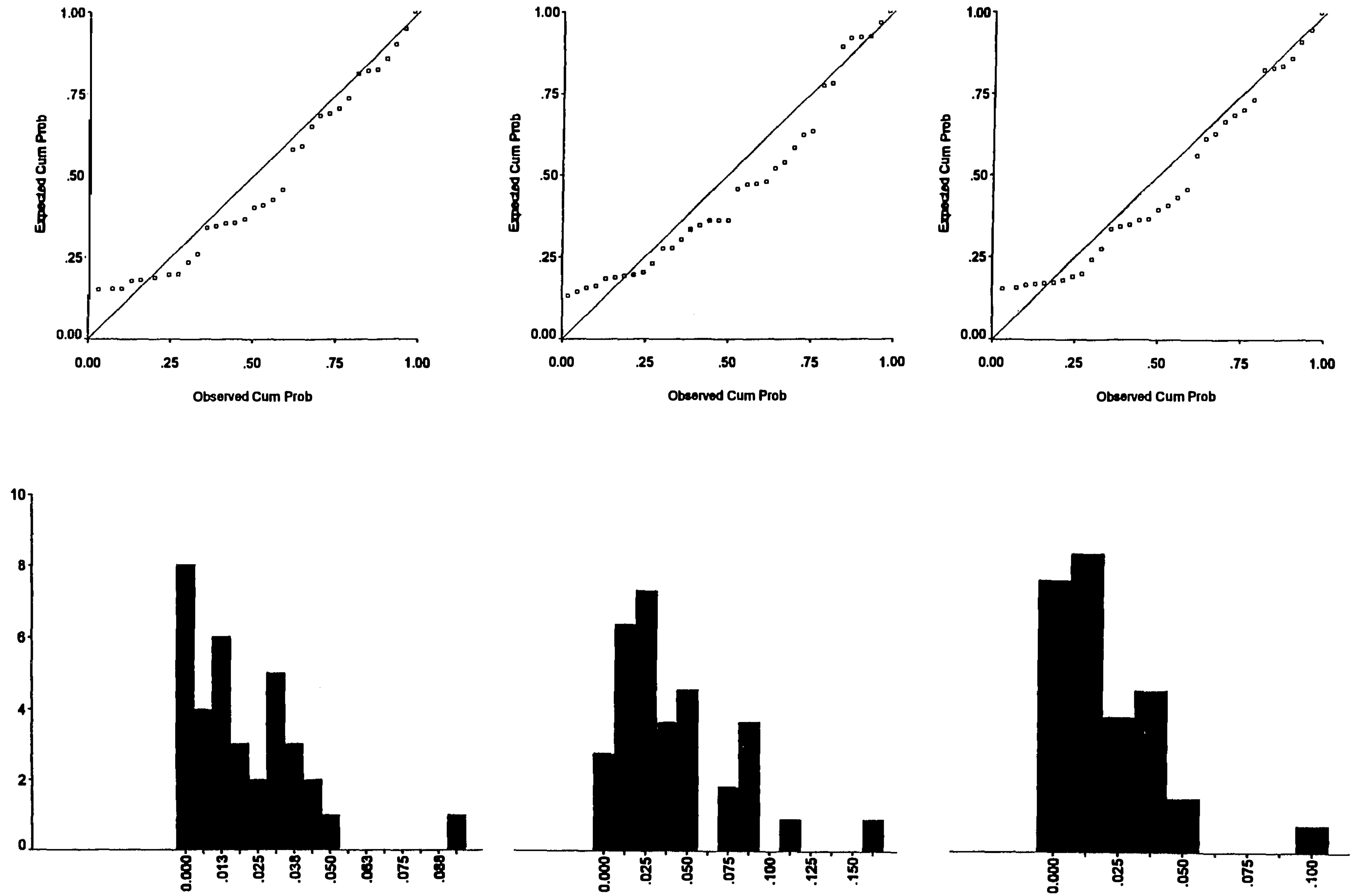


Figure 8. Normal Probability Plots and Histograms of th selected $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for United States- left to right: AD_f; Long Distance, SD_f; Short Distance, NAD_f; Long Distance.

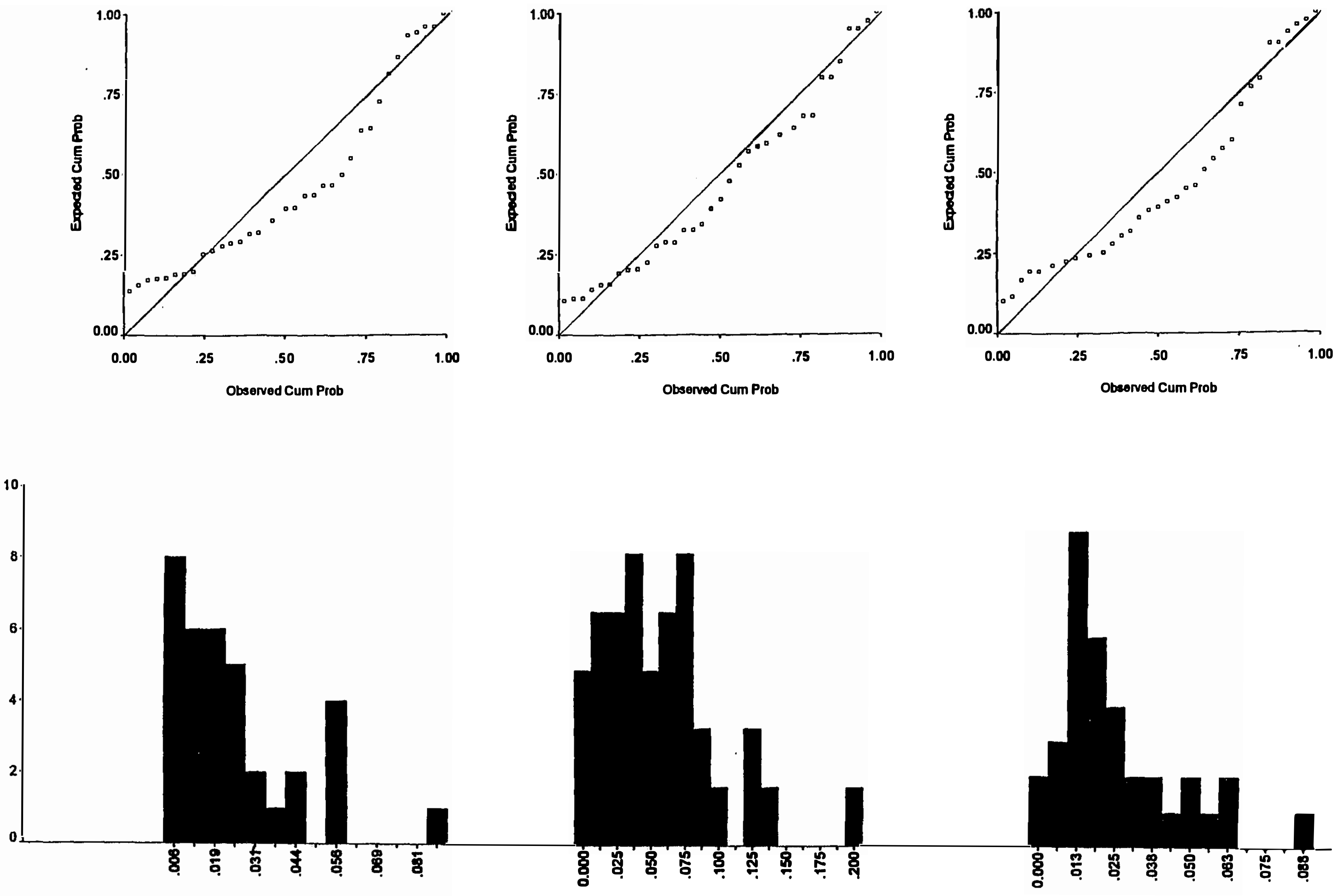


Figure 9. Normal Probability Plots and Histograms of th selected $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for Toronto - left to right: AD_f; Long Distance, SD_f; Short Distance, NAD_f; Long Distance.

A Levene equality-of-variance test at a significance level of 5% is performed for each set. The 2-tail p-values of the Levene test for the long and short distance distributions are given for each region in Table 12. Since the p-values in Table 12 are well above 0.10 (the one exception is the long distance distribution in Paris with a p-value of 0.06), they do not provide enough evidence at the 5% significance level to reject the hypothesis that the variances of long distances and short distances are significantly different.

Region	Short D. p-value	Long D. p-value	Urban Center	Short D. p-value	Long D. p-value
Australia	0.945	0.335	Brussels	0.999	0.999
British Columbia	0.997	0.991	London City	1.000	0.912
Canada	0.875	0.937	London North	0.975	0.911
France	0.871	0.963	Los Angeles	0.995	0.960
Great Britain	0.808	0.448	NY City	0.589	0.997
NY State	0.412	0.250	Paris	0.976	0.060
Pennsylvania	0.984	0.412	Sydney	0.947	0.991
United States	1.000	0.369	Tokyo	0.998	0.923
			Toronto	0.815	0.992

Table 12. 2-tail p-values for equality of variances of long and short distance distributions for the three criteria in each region

We next turn our attention to the comparisons of the accuracy of the three criteria in predicting long distances and short distances in a given region. For that purpose, two Friedman tests for the matched triples of both horizontal sets (see Figure 7) are performed. A p-value less than 0.05 is supposed to indicate the existence of a significantly different pair from the three criteria for the given region. The p-values of these main effect tests are provided in Table 13.

The significance levels listed in Table 13 can be interpreted as follows. In general, the accuracy in predicting distances is not significantly different for the three criteria. However, for long actual distances, in five of the eight large geographic regions; Australia, British Columbia, Great

Britain, New York State, and the United States, and for the short actual distances, in three regions; Australia, New York State, Paris, there is at least a pair of criteria with significantly different distance prediction accuracy.

Region	Long D. p-value	Short D. p-value	Urban Center	Short D. p-value	Long D. p-value
Australia	0.0002	0.0000	Brussels	0.9491	0.9749
British Columbia	0.0080	0.1242	London City	0.1870	0.3213
Canada	0.9107	0.4624	London North	0.2684	0.1342
France	0.7553	0.6897	Los Angeles	0.4389	0.0083
Great Britain	0.0073	0.2466	NY City	0.2355	0.1048
NY State	0.0116	0.0187	Paris	0.6175	0.0316
Pennsylvania	0.1242	0.0527	Sydney	0.2388	0.8534
United States	0.0300	0.1066	Tokyo	0.4296	0.9767
			Toronto	0.6897	0.1691

Table 13. The p-values of Friedman test comparing long and short distance distributions for the criteria.

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 ($100*E[|e(x_i,x_j)|/A(x_i,x_j)]$) for predicting long and short distances for the criteria are presented in Table 14.

By inspecting the average percent absolute errors of the first five exceptional regions listed above, it can be observed that the AD_f and SD_f criteria generate less average percent absolute error than the NAD_f criterion in predicting long distances. For example, in Australia the average percent absolute errors for the AD_f and SD_f criteria are 5.02% and 5.03%, respectively, whereas for the NAD_f criterion, it is 6.42%. However, the United States is exceptional in those five regions since the SD_f criterion has less average percent absolute error, 2.02% than the AD_f and NAD_f criteria which have

2.06% and 2.09%, respectively. By inspecting Table 14 for short actual distances, we see that the NAD_f criterion provides better prediction accuracy than the either of the AD_f and SD_f criteria for the three exceptional regions. For example, in Australia the NAD_f criterion generates a 6.23% absolute error in predicting short distances. However the AD_f and SD_f criteria provide 8.45% and 8.37% absolute errors for the same region, respectively.

Geographical Region	Short Distance			Long Distance		
	AD_f	SD_f	NAD_f	AD_f	SD_f	NAD_f
Australia	8.45	8.37	6.23	5.03	5.02	6.42
British Columbia	13.85	13.53	12.67	7.79	7.90	8.40
Canada	8.92	8.80	8.76	1.89	1.97	1.96
France	3.93	4.03	3.81	2.56	2.51	2.67
Great Britain	6.73	6.61	5.99	3.60	3.76	5.08
New York State	7.29	6.91	6.03	5.14	5.44	6.82
Pennsylvania	8.32	8.59	8.00	4.13	4.10	4.44
United States	4.01	4.22	3.97	2.06	2.02	2.09
Brussels	5.30	5.31	5.29	3.25	3.26	3.24
London City	11.18	11.00	11.03	4.60	5.34	4.70
London North	6.22	5.98	5.87	2.77	2.87	2.92
Los Angeles	11.60	11.39	11.33	4.15	4.70	4.34
New York City	7.71	7.82	7.17	5.75	5.41	5.84
Paris	11.61	11.73	10.68	6.18	6.27	6.91
Sydney	7.97	7.86	7.85	5.61	5.60	5.66
Tokyo	5.63	5.64	5.64	2.65	2.62	2.77
Toronto	5.59	5.70	5.52	2.51	2.39	2.57

Table 14. Average percent absolute errors in predicting long and short distances of the criteria.

(iii) The purpose of this part is to compare the accuracy of predicting long distances versus short distances in a region. Therefore, three sets of vertical comparisons, one for each criterion (see Figure.7), are performed. First, in order to determine whether the variance of the $|e(x_i,x_j)|/A(x_i,x_j)$ 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 (i.e., greater than 0.10 significance level for the two-tail test) then we can conclude that the variance of the $|e(x_i,x_j)|/A(x_i,x_j)$ 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 15. Based on this table, the variances of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for all the urban centers are not equal since the test values are less than 0.10 for all the criteria. Additionally, there are four exceptions for each criterion in the large geographic regions; for the AD_f and SD_f criteria, the variance of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions are homoescedastic in British Columbia, France, Great Britain, NY State, and for the NAD_f criterion, the $|e(x_i,x_j)|/A(x_i,x_j)$ distribution is homoscedastic in Australia, British Columbia, France and Paris.

Region	AD_f	SD_f	NAD_f	Urban center	AD_f	SD_f	NAD_f
Australia	0.084	0.031	0.383	Brussels	0.084	0.093	0.090
BC Province	0.725	0.617	0.758	London City	0.066	0.039	0.066
Canada	0.000	0.000	0.000	London North	0.001	0.003	0.007
France	0.285	0.351	0.252	Los Angeles	0.017	0.024	0.018
Great Britain	0.159	0.234	0.466	NY City	0.016	0.054	0.011
NY State	0.741	0.925	0.014	Paris	0.006	0.004	0.022
Pennsylvania	0.002	0.000	0.005	Sydney	0.043	0.082	0.048
United States	0.013	0.001	0.018	Tokyo	0.025	0.019	0.030
				Toronto	0.000	0.001	0.000

Table 15. The 2-tail p-values for the equality of variance of $|e(x_i,x_j)|/A(x_i,x_j)$ for each criterion.

In order to see the general pattern of differences in variances for the three criteria we inspect Table 16 where the variances of the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions resulting from long and short distance predictions are reported. In general, it can be said that the variance of the distribution of $|e(x_i,x_j)|/A(x_i,x_j)$ for long distances is less than the variance for short distances for each criteria in each region. But this conclusion does not always hold at the 5% significance level as the Levene tests suggest in Table 15. In order to represent the converging funnels formed by the variances plotted against increasing actual distances in a region, the scatter plots of $|e(x_i,x_j)|/A(x_i,x_j)$ for United States and Toronto are shown in Figures 10 and 11, respectively.

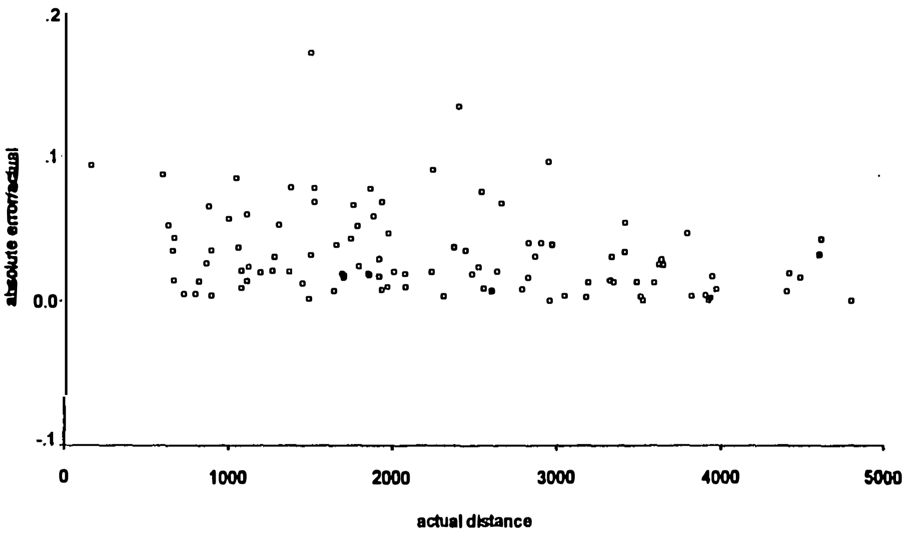
Geographical Region	Short Distances			Long Distances		
	AD_f	SD_f	NAD_f	AD_f	SD_f	NAD_f
Australia	0.0036	0.0035	0.0030	0.0015	0.0015	0.0019
BC Province	0.0072	0.0073	0.0070	0.0058	0.0057	0.0064
Canada	0.0043	0.0039	0.0047	0.0003	0.0002	0.0003
France	0.0011	0.0010	0.0012	0.0007	0.0006	0.0007
Great Britain	0.0022	0.0021	0.0025	0.0014	0.0014	0.0021
NY State	0.0025	0.0024	0.0021	0.0025	0.0023	0.0034
Pennsylvania	0.0058	0.0060	0.0056	0.0012	0.0010	0.0014
United States	0.0012	0.0013	0.0013	0.0004	0.0003	0.0004
Brussels	0.0023	0.0023	0.0023	0.0009	0.0009	0.0009
London City	0.0364	0.0354	0.0367	0.0060	0.0045	0.0058
London North	0.0020	0.0019	0.0020	0.0005	0.0005	0.0006
Los Angeles	0.0343	0.0335	0.0352	0.0013	0.0012	0.0012
NY City	0.0053	0.0038	0.0053	0.0027	0.0023	0.0026
Paris	0.0145	0.0134	0.0136	0.0019	0.0018	0.0033
Sydney	0.0042	0.0039	0.0042	0.0026	0.0025	0.0027
Tokyo	0.0041	0.0040	0.0042	0.0006	0.0005	0.0006
Toronto	0.0025	0.0019	0.0026	0.0004	0.0004	0.0004

Table 16. Variances of $|e(x_i,x_j)|/A(x_i,x_j)$ distributions in predicting long and short distances of the criteria.

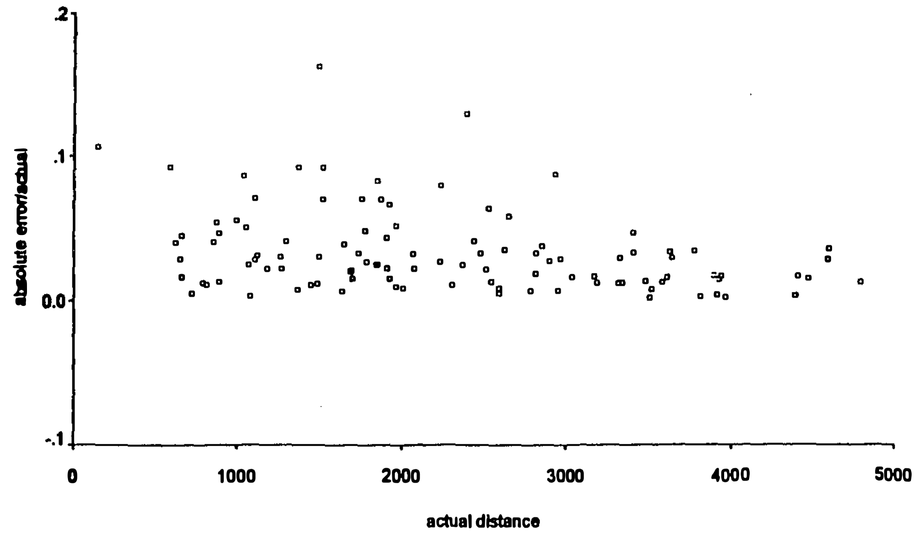
The accuracy of each criteria in predicting the long versus short distances in a given region is examined by using the nonparametric 2-tailed Mann-Whitney Test (see figure 7, vertical comparisons) with a 5% significance level. The 2-tailed p-values of the tests for three criteria are presented in Table 17. In this table, a p-value less than 0.10 indicates that there is a significant difference between the long distance and short distance $|e(x_i,x_j)|/A(x_i,x_j)$ distributions of the geographical region.

Therefore, for the AD_f criterion only in New York city and Sydney and for the SD_f criterion only in New York State and Sydney, do the long distance and short distance $|e(x_i,x_j)|/A(x_i,x_j)$ distributions apparently come from the populations having the same distributions. For the NAD_f

AD_f



SD_f



NAD_f

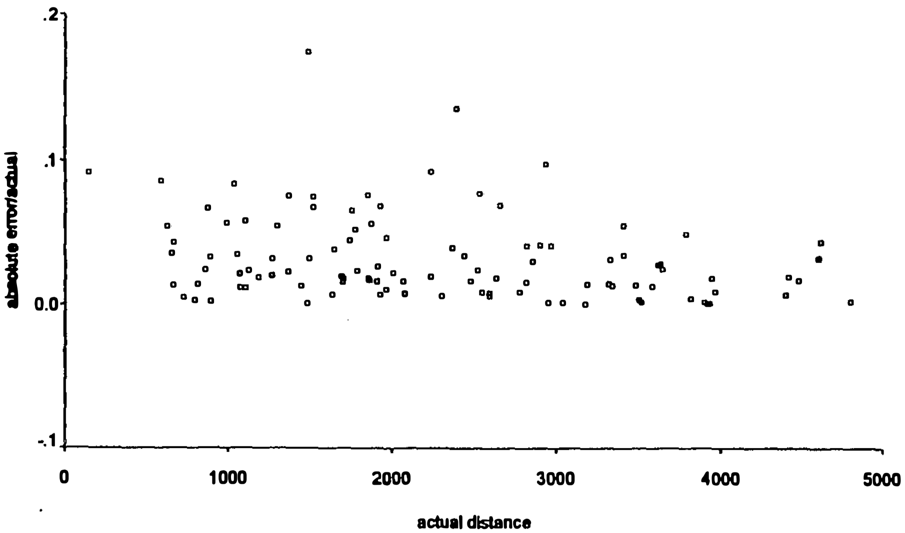


Figure 10. Scatter Plots of $|e(x_i,x_j)|/A(x_i,x_j)$ distribution for United States

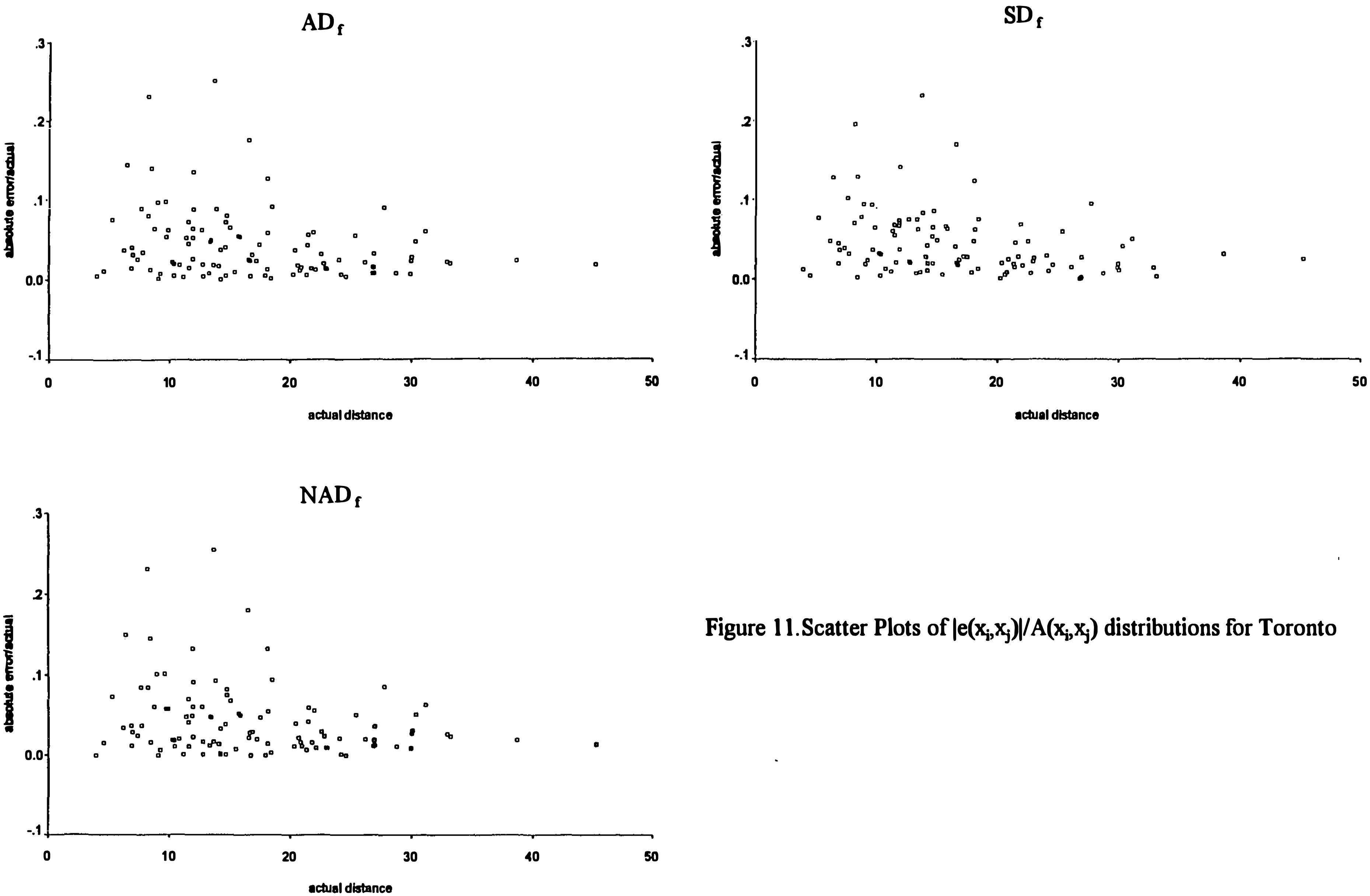


Figure 11.Scatter Plots of $|e(x_i,x_j)|/A(x_i,x_j)$ distributions for Toronto

criterion the same conclusion holds in the regions of Australia, Great Britain, New York State, New York City, Paris and Sydney. As also seen in Table 16, 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.

Region	AD_f	SD_f	NAD_f	Urban Center	AD_f	SD_f	NAD_f
Australia	0.0118	0.0197	0.5145	Brussels	0.0340	0.0360	0.0307
BC Province	0.0024	0.0044	0.0285	London City	0.0012	0.0316	0.0026
Canada	0.0000	0.0000	0.0000	London North	0.0004	0.0008	0.0021
France	0.0203	0.0070	0.0821	Los Angeles	0.0039	0.0093	0.0103
Great Britain	0.0042	0.0057	0.4920	NY City	0.5413	0.0942	0.9859
NY State	0.0446	0.1622	0.9205	Paris	0.0350	0.0177	0.2401
Pennsylvania	0.0124	0.0118	0.0548	Sydney	0.1638	0.1501	0.1943
United States	0.0061	0.0038	0.0093	Tokyo	0.0061	0.0060	0.0132
				Toronto	0.0061	0.0002	0.0123

Table 17. Two-tailed Mann-Whitney Test p-values for $|e(x_i,x_j)|/A(x_i,x_j)$ distributions in predicting long and short distances of the criteria.

SUMMARY OF THE RESULTS AND CONCLUSIONS

The three goodness-of-fit criteria, AD_f, SD_f and NAD_f, 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 a variety of statistical tests.

Several conclusions regarding the properties of the $e(x_i,x_j)$ distributions for the three criteria can be drawn.

1. The $e(x_i,x_j)$ populations are non-normal, generally highly peaked and skewed right, for each criterion.
2. For the three criteria, the variances of the $e(x_i,x_j)$ distributions are not significantly different.

3. For each criterion, the variance of the $e(x_i,x_j)$ distribution for long distances is significantly different, and indeed greater than the variance of the $e(x_i,x_j)$ distribution for short distances. Hence, in general, the scatter plots for $e(x_i,x_j)$ forms a diverging funnel as the actual distance increases for the sample pairs.

4. The expected values of the $e(x_i,x_j)$ distributions are zero for the geographical regions with a few exceptions for the AD_f and NAD_f criteria. There were no exceptions for the SD_f criterion.

The following conclusions regarding the comparison of the three criteria in terms of their accuracy in predicting relatively long distances and short distances can be drawn.

1. The $|e(x_i,x_j)|/A(x_i,x_j)$ populations are non-normal for each criterion. The histograms are highly peaked with more occurrences close to zero and skewed right.

2. There are generally no pairs of the criteria for which the $|e(x_i,x_j)|/A(x_i,x_j)$ distributions are significantly different.

3. In terms of the $|e(x_i,x_j)|/A(x_i,x_j)$'s, 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.

4. The variance of the $|e(x_i,x_j)|/A(x_i,x_j)$ distribution for long distances is significantly different than the variance of the $|e(x_i,x_j)|/A(x_i,x_j)$ distribution for short distances. The former is smaller than the latter, and hence the scatter plots of the $|e(x_i,x_j)|/A(x_i,x_j)$ 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 $\ell_{k,p,\theta}$ norm, and since the $e(x_i,x_j)$'s have an expected value of zero without any exceptions in all the regions, it would seem to be advantageous to use the SD_f criterion in practice.

REFERENCES

[1] Berens, W., "The suitability of the weighted ℓ_p-norm in estimating actual road distances", *European Journal of Operational Research,* 34, 39-43 (1988).

[2] Berens, W., and Körling, F.J., "Estimating Road Distances by Mathematical Functions", *European Journal of Operational Research,* 21, 54-56 (1985).

[3] Brimberg, J., Dowling, P.D., Love, R.F., "Estimating the parameters of the weighted ℓ_p-norm by linear regression", *IIE Transactions,* 28, 363-367 (1996).

[4] Brimberg, J. and Love, R.F., "Directional Bias of the ℓ_p Norm", *European Journal of Operational Research,* 67, 287-294 (1993)

[5] Brimberg, J., and Love, R.F., "Estimating Travel Distances by the Weighted ℓ_p Norm", *Naval Research Logistics*, 38, 241-259 (1991).

[6] Brimberg, J., Love, R.F., Walker, J.H., "The effect of axis rotation on distance estimation", *European Journal of Operational Research*, 80, 357-364 (1995).

[7] Eilon, S., Watson-Gandy, C.D.T., and Christofides, N., *Distribution Management: Mathematical Modelling and Practical Analysis*, Hafner Publishing Company, New York (1971).

[8] Ginsburgh, V. and Hansen, P., "Procedures for the Reduction of Errors in Road Network Data", *Operational Research Quarterly*, 25, 321-322 (1974).

[9] Klein, R., "Voronoi Diagrams in the Moscow Metric", Proc. Graph-Theoretic Concepts in Computer Science (WG '88), Amsterdam, 434-441 (1988).

[10] Kolesar, P., Walker, W., and Hausner, J., "Determining the Relation between Fire Engine Travel Times and Travel Distances in New York City", *Operations Research*, 23, 614-627 (1975).

[11] Love, R.F., and Morris, J.G., "Modelling Inter-City Road Distances by Mathematical Models", *Operational Research Quarterly*, 23, 61-71 (1972).

[12] Love, R.F., and Morris, J.G., "Mathematical Models of Road Travel Distances", *Management Science*, 25, 130-139 (1979).

[13] Love, R.F., Morris, J.G., and Wesolowsky, G.O., *Facilities Location - Models & Methods*, North-Holland, New York (1988).

[14] Love, R.F. and Walker, J.H., "A New Criterion for Evaluating the Accuracy of Distance Predicting Functions", Faculty of Business Working Paper No. 370, McMaster University, Hamilton, Ontario, Canada (1991).

[15] Love, R.F., and Walker, J.H., "Distance Data for Eighteen Geographic Regions", Faculty of Business Working Paper No. 383, McMaster University, Hamilton, Ontario, Canada (1993).

[16] Love, R.F., Walker, T.H., "An Empirical Comparison of Block and Round Norms for Modelling Actual Distances", *Location Science*, 2 No. 1, 21-43 (1994).

[17] Love, R.F., Walker, J.H, Tiku, M.L., "Confidence Intervals for $\ell_{k,p,\theta}$ Distances", *Transportation Science*, 29, No. 1, 93-100 (1995).

[18] SPSS for Windows, Release 6.1.3., SPSS Inc. 444 North Michigan Avenue, Chocago, IL 60611 USA, (1995)

[19] Star, J. and Estes, J., *Geographic Information Systems – An Introduction*, Prentice-Hall, New Jersey (1990).

[20] Stuart, A., and Ord, J.K., *Kendall's Advanced Theory of Statistics, Volume I - Distribution Theory*, Oxford University Press, New York (1987).

[21] TruckStops 2, MicroAnalytics, Suite One, 2045 North 15th Street, Arlington, Virginia and 1986 Queen Street East, Toronto, Ontario (1990).

[22] Ward, J.E., and Wendell, R.E., "A New Norm for Measuring Distance Which Yields Linear Location Models", *Operations Research*, 28, 836–844 (1980).

[23] Ward, J.E., and Wendell, R.E., "Using Block Norms for Location Modeling", *Operations Research*, 33, 1074–1090 (1985).

[24] Wesolowsky, G.O., *Multiple Regression and Analysis of Variance*, John Wiley and Sons, Toronto (1976).

[25] Westwood, J.B., "A Transport Planning Model for Primary Distribution", *Interfaces*, 8, 1–10 (1977).

Faculty of Business
McMaster University

WORKING PAPERS - RECENT RELEASES

395. Roy Adams, "A Pernicious Euphoria: 50 Years of Wagnerism in Canada", November, 1994.

396. Roy Adams, "The Determination of Trade Union Representativeness in Canada and the United States", November, 1994.

397. Shouhong Wang, Norman P. Archer, "A Fuzzy Decision Making Model", December, 1994.

398. Kalyan Moy Gupta and Ali R. Montazemi, "A Methodology for Evaluating the Retrieval Performance of Case-Based Reasoning Systems", December, 1994.

399. Jiang Chen and George Steiner, "Lot Streaming with Detached Setups in Three-Machine Flow Shops", December, 1994.

400. Jiang Chen and George Steiner, "Lot Streaming with Attached Setups in Three-Machine Flow Shops", December, 1994.

401. Ali R. Montazemi and Feng Wang, "An Empirical Investigation of CAI in Support of Mastery Learning", February, 1995.

402. Kalyan Moy Gupta and Ali Reza Montazemi, "Retrieval in Case-Based Reasoning Systems with Modified Cosine Matching Function", February, 1995.

403. Kalyan Moy Gupta and Ali Reza Montazemi, "A Connectionist Approach for Similarity Assessment in Case-Based Reasoning Systems", March, 1995.

404. John W. Medcof, "Selecting Alliance and Network Partners - Strategic Issues", March 1995.

405. Jiang Chen and George Steiner, "Discrete Lot Streaming in Two-Machine Flow Shops", March, 1995.

406. Harish C. Jain and S. Muthuchidambaram, "Strike Replacement Ban in Ontario and Its Relevance to U.S. Labor Law Reform", May, 1995.

407. Ali R. Montazemi and Kalyan Moy Gupta, "A Framework for Retrieval in Case-Based

Reasoning Systems", June, 1995.

408. Ali R. Montazemi and Kalyan Moy Gupta, "An Adaptive Agent for Case Description in Diagnostic CBR Systems", June, 1995.

409. Roy J. Adams, Noel Cowell and Gangaram Singh, "The Making of Industrial Relations in the Commonwealth Caribbean", June, 1995.

410. Jiang Chen and George Steiner, "Approximation Methods for Discrete Lot Streamingin Flow Shops", June, 1995.

411. Harish C. Jain and S. Muthuchidambaram, "Bill 40 Amendments to Ontario Labour Relations Act: An Overview and Evaluation", June, 1995.

412. Jiang Chan and George Steiner, "Discrete Lot Streaming in Three-Machine Flow Shops", July, 1995.

413. J. Brimberg, A. Mehrez and G.O. Wesolowsky, "Allocation of Queueing Facilities Using a Minimax Criterion", January, 1996.

414. Isik Zeytinoglu and Jeanne Norris, "Global Diversity in Employment Relationships: A Typology of Flexible Employment", March, 1996.

415. N. Archer, "Characterizing World Wide Web Search Strategies", April, 1996.

416. J. Rose, "Immediacy and Saliency in Remedying Employer Opposition To Union Organizing Campaigns", July, 1996.

417. Roy J. Adams and Parbudyal Singh, "Worker Rights Under Nafta: Experience with the North American Agreement on Labor Cooperation", September, 1996.

418. George Steiner and Paul Stephenson, "Subset-Restricted Interchange for Dynamic Min-Max Scheduling Problems", September, 1996.