Transportation 14: 175–192 (1987)

The reliability of using the gravity model for forecasting trip distribution

LEONNIE N. DUFFUS[1], ATTAHIRU SULE ALFA[2] & AFIFI H. SOLIMAN[1]

[1]*Dept. of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2;* [2]*Transport Institute and Dept. of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2; (requests for offprints)*

Key words: gravity model; trip distribution; forecasting

Abstract. The main purpose of this study is to assess the forecasting capability of the gravity model and to investigate the merit of including K-factors when using the model. Peak hour trip data was obtained for four study year periods 1962, 1971, 1976 and 1981 for the City of Winnipeg. Analysis of the calibration results indicated that the F-factors for the twenty year period were stable within a range of values. In general, however, the K-factors were found to be inconsistent from one prediction period to the next, and when used in forecasting trips they resulted in larger errors than without their use. The validity of using K-factors or the method which has been used to determine them is questionable. It was concluded that while K-factors are very meaningful in theory (as defined), they are not appropriate for use in predicting O-D matrices based on the method by which they are currently estimated (i.e. as a simple ratio). Further study is needed to investigate an alternative method of calibrating the gravity model such as the cell-by-cell regression method.

Introduction

The gravity model is a popular model for use in trip distribution. Its calibration process has been well documented. In fact, most analysts believe that research into aspects of this model is completed and hence well understood. But there is an unanswered question: "How reliable is the use of gravity models for prediction of future trips?" This question has not been properly addressed in the literature due to the lack of available data, and perhaps, because it has not been asked.

In general, predictions of future travel are based on a number of assumptions, one of which is that people will, in the future, respond to a particular set of travel conditions that are measured in terms of travel costs. In order to make these predictions, planners use transportation demand models. The most commonly used demand model is a four stage model consisting of: trip generation, trip distribution, modal split, and trip assignment. Trip distribution is the main focus of this paper.

The purpose of trip distribution models is essentially to estimate the trip interchanges between zone pairs in a study area. As input, it uses a set of trip productions and attractions for each zone and tries to estimate the trip interchanges between the zone pairs. Trip distribution is assumed to be based on the type and extent of the transportation facilities available, the pattern and intensity of land use, and the socio-economic characteristics of the population in the area (Sosslau et al. 1978). Hence the allocation of trips is based on these assumptions.

There are many trip distribution methods that have been developed to describe and forecast trip interchanges. To date the most widely used trip distribution model, for medium to large urban centres, is the gravity model. It is simple in concept and has been well documented.

Since its formulation, the gravity model has been calibrated by many agencies and researchers for base years. However, it has not been shown how well the calibrated model predicts O-D trips for horizon years. Little has been done to check the accuracy of the forecasts provided by the models against observations made in subsequent years. Emmerson (1982) tested the predicting capability of several trip distribution models, one of which was a simple exponential gravity model, and studied the errors generated by the predictions. The analysis was restricted to the use of this simple model and other models such as the Furness and Growth Factor Methods. Since only two sets of data years were used for the comparison in that study, the analysis could not establish trends that may be present in the predictions or in the resulting errors.

In the calibration of the gravity model, it is generally accepted that the model is a good fit when the calibrated frequency distribution is statistically close to the observed frequency distribution for the base year trips. Comparison of trip length frequency distributions does not necessarily provide enough information for assessing how well the trip interchanges match. Hence a "good fit" as defined above does not guarantee accurate predictions of future zone-to-zone trip interchanges. Because the predicted trips will be used in the final phase of the demand model, i.e. trip assignment, it is important to have a model that produces reliable and consistent origin-destination (O-D) trip interchanges.

The objective of this study, therefore, is to examine the process of calibrating the gravity model and to assess the accuracy with which the calibrated model is able to predict future O-D trips at the zone-to-zone level. Based on the observed trends, the study will determine the most reliable F-factors to be used for prediction purposes and establish the validity of the use of an adjustment factor (K-factors) with respect to future predictions. The City of Winnipeg has been used as the case study. The gravity model is calibrated for the AM peak work trip. These trips are seen to be generally stable and therefore easier to predict. Four sets of observed O-D data over a 20-year period (1962, 1971, 1976 and 1981) were used to establish the respective trends.

Approaches to gravity model calibration

The gravity model is based on the assumption that all trips starting from a given area or zone are attracted to other zones in direct proportion to the attractive power of the zone, and inversely proportional to the travel impedance between the two zones (Sosslau et al. 1978). The gravity model can be stated generally as:

$$T_{ij} = \frac{P_i A_j F_{ij} K_{ij}}{\sum_v A_v F_{iv} K_{iv}} \tag{1}$$

T_{ij} = trips between zone i and zone j,
P_i = number of trips produced by zone i,
A_j = number of trips attracted by zone j,
F_{ij} = empirically derived travel time factor which expresses the average area wide effect of spatial separation on trip interchange between zones i and j – referred to as F-factor,
K_{ij} = a specific zone-to-zone adjustment factor (for zones i and j) to allow for the incorporation of the effect of socio-economic variables on travel patterns that are not otherwise accounted for in the model – referred to as K-factor.

Calibration of the gravity model involves estimating the values of F_{ij}, (F-factor) and if necessary, K_{ij} (K-factor), that will reproduce the observed trips at the base year. The zone-to-zone travel time (or travel cost), and the total trip productions and attractions of each zone are used as inputs in the model calibration. The shape of the F-factor curve is usually that of an inverse exponential function that represents the impedence to travel between zone pairs.

Bureau of public roads method of calibration

The most widely used method for calibrating the gravity model is that developed by the Bureau of Public Roads (BPR) (Stopher & Meyburg 1975). In the BPR method, the purpose of the calibration procedure is to establish the relationship between the F-factors and travel cost (usually travel time) between zone pairs for the base year. The calibration is deemed reasonable when the observed frequency distribution for the trips is sufficiently close to the calibrated frequency distribution, based on a chosen closure criterion.

After the F-factors have been estimated from the model calibration, it is common practice to introduce an adjustment factor, K-factor, at the zone-to-zone level to account for the lack of fit between the observed and estimated trip interchanges. The K-factor is supposed to be a reflection of the socio-economic characteristics of each zone pair, (i.e. producing and attracting zones, i and j), and of other factors not incorporated in the model parameters, but said to affect tripmaking. The K-factors are computed as the ratio of the observed trips, to the estimated trips, between each zone pair. These K-factors are then used for future forecasts.

Calculating the K-factors as a simple ratio seems inconsistent with the way in which they are defined. This method of estimating K-factors does not attempt to model some of the complex socio-economic relationship that exists between a pair of zones. There are also problems associated with these factors when later used in the prediction process. It has been emphasized by Stopher & Meyburg (1975), that unless the K-factors can be related to specific socio-economic characteristics, in some statistical or mathematical form, the K-factor value obtained for each zone pair must otherwise be assumed to remain the same throughout the forecast period. It is clear that since socio-economic parameters are expected to change, this assumption can introduce serious problems in the usefulness of the model for prediction purposes. Also, there can be difficulties in estimating K-factors for new zone pairs that did not exist when calibration was carried out. For the above reasons, it is difficult to justify the use of K-factors, particularly based on the manner in which they are estimated. There does not seem to be a consensus among planners as to whether K-factors should be used or not.

Like most models, the calibration process of the gravity model by the BPR method has resulted in a number of disadvantages (Dickey 1983). First, it is assumed that the F-factors found through the model calibration, will remain the same over the prediction period. It is unlikely that the travel-time factors would remain truly constant throughout the urban area, or to the horizon period, since travel conditions do change. This aspect will be investigated in this paper. Second, the model has been found to overestimate near trips and underestimate far trips. The closure criteria used in the calibration process is often judged on the total trip length distribution of each time interval. This distribution may exhibit a rapid decay and satisfactory closure will be frequently obtained, for example when near trips are fitted well at the expense of longer trips. It has also been noted that the frequency trip length distribution may be an insensitive parameter in determining F-factors (Todes 1978). It was found that there were only small differences between the F-factors calibrated from different sets of data, despite the fact that there were significant differences between the frequency distributions of the data. Finally, if the model shows only an approximate agreement with field data, K-factors are introduced in an attempt to cause each interzonal trips to be reproduced at the base year. Since socio-economic parameters may be expected to change over any forecast period, the use of K-factors is likely to introduce problems in the prediction process.

Cell-by-cell regression method of calibration

A calibration process that tries to eliminate some of the disadvantages associated with the BPR method is a cell-by-cell (zone-to-zone) regression method suggested by Todes (1978).

This approach tends to develop a more disaggregate form of the F-factors that are to be used in the gravity model. Unlike the BPR method, the cell-by-cell regression approach does not require a series of iteration to obtain the area-wide F-factors. It is based on obtaining a specific F-factor for each O-D pair, and later relating these O-D pair specific factors to a number of variables that affect trip-making. This is done by regression techniques. The pair-specific factors are calculated as the ratio of the observed zonal trips between i and j, to the calculated trips, due to only the production at i and attraction to j. The calculated trips, $\bar{t}_{ij}$, are obtained from the expression:

$$\bar{t}_{ij} = \frac{P_i A_j}{\sum_v A_v} \tag{2}$$

F_{ij} is then estimated as the ratio $t_{ij}/\bar{t}_{ij}$ where t_{ij} is the observed trips between i and j.

The cell-by-cell F-factors are said to represent several characteristics of the study area, including the resistance to travel between zones, accessibility, socio-economic characteristics, and other factors affecting the trip distribution. Since each zone pair has its own specific factor, their values, when used in the gravity model, will cause the calculated zone-to-zone trips to equal the observed. At this stage of the regression method, the pair-specific factors can be considered as "ideal" distribution factors. These ideal factors must now be represented, by some mathematical function, to the variables that affect trip making. Regression analysis is used to relate the O-D pair specific factors to the trip-making variables. The benefit of using this technique is that through regression, the resulting area-wide F-factor function can incorporate several variables such as a measure of accessibility, income, etc., as well as travel costs or time. In addition, this approach does not need to estimate K-factors because the effects of socio-economic characteristics will be incorporated into the area wide F-factors.

The quality of the predictions, using the model calibrated by the regression method, is assessed by a comparison of the zone-to-zone simulated and observed trips. This overcomes the problem associated with using the trip length frequency distributions as a means of assessing the performance of the calibrated gravity model, as in the BPR method. The closeness of the distributions does not necessarily mean the observed trip matrix will be reproduced at the zone-to-zone level.

Accuracy measure

To assess the performance of the gravity model, several statistical parameters were chosen to measure the differences between the estimated and observed O-D trips. The chosen indicators were the absolute mean error, the root-mean-square error (rmse), and the D-statistics or Gini-coefficient. The absolute mean error expresses the average absolute differences between corresponding observed and estimated O-D trips, while the rmse is a measure of the standard deviation of the differences between observed and estimated trips. The rmse is more sensitive to large deviations than the absolute mean error. The D-statistics or Gini-coefficient measures the percent of the zone-to-zone trips in the estimated O-D matrix that have to be reassigned, in order for all the values in that matrix to be equal to the observed (Emmerson 1982). It compares the differences between the zone-to-zone trips in each cell, as a percentage of the total trips. Other indicators such as the correlation coefficient and chi-square statistics were not used. The correlation coefficient measures the association between variables rather than their accuracy, and the chi-square statistics tends to give unrealistic results for cells with small or near zero observations (Wilson 1976). To make direct comparisons between the years, each of the chosen accuracy indicators was expressed as a percentage of the average observed inter-zonal trips.

Case study: City of Winnipeg

Winnipeg is a medium size city with a population of about 600,000 people. The first area wide travel survey in Winnipeg was conducted in 1962 (City of Winnipeg Travel & Demographic Trends, 1982). In 1962 all homes (100%) within the city were surveyed and data collected on all work trips originating in each household. Following the completion of the 1962 transportation study, travel surveys have been conducted every five years, starting in 1971 and coinciding with the national census. In 1971 and 1976, the basic data collection procedure was repeated using a 20% sample of all homes in the metropolitan area. In 1981, a mail survey was conducted and the final return rate resulted in a 10% sample size. A statistical analysis of the sample results undertaken by the city planning department found that a sample size ranging from between 10 – 20% was considered cost effective and statistically reliable.

The City of Winnipeg has experienced no major population changes over the 21 year period from 1962 to 1981. Population growth has been steady at about 1.0% per annum. While the growth rate has remained relatively unchanged, there has been significant people movement from the downtown to the outer areas of the city. The average income per dwelling unit however has been increasing at 10.8% per annum between 1971 and 1981.

The travel trends in Winnipeg have also not seen much variations throughout the 20 year period. The total observed a.m. peak hour work trips have been increasing by 1.9% every year. In 1962 the a.m. peak period work trips were about 73,000 and by 1981, they were in the order of 104,000. The a.m. peak trips alone make up about 45% of the total daily work trips. There are two predominant modes of travel in the City of Winnipeg; public transit and automobile. Over the years, an average of 24% of the work trips have been made by transit.

Eight areas of concentrated employment are in Winnipeg. These areas attract the majority of work trip makers. The downtown area in Winnipeg has the largest concentration of employment, at 26%, and is the largest attractor of work trips.

Data acquisition

Winnipeg's observed O-D trip and travel time data sets for 1962, 1971, 1976 and 1981 were obtained from the Research and Development Branch of the City of Winnipeg's Streets and Transportation Department. The area is divided into 139 traffic zones, each having its own socio-economic characteristics and travel patterns.

To make computations easier, the traffic zones were aggregated into thirty-six established superzones as shown in Fig. 1. The corresponding origin-destination trips were aggregated and tabulated for each superzone. Based on the observed travel time between the smaller traffic zones, an average travel time between each pair of the superzones was estimated. Only auto travel times were available and used to distribute trips among the zones.

Analysis of results

The BPR and the cell-by-cell regression techniques were both used to calibrate the gravity model. The results from the regression method however were not satisfactory. A power function fitted to the zone-to-zone F-factors, using travel time alone as the independent variable, resulted in r^2 values varying from 0.40 to 0.26. Using an exponential function produced equally poor results with r^2 varying from 0.36 to 0.21. The regression equations and the results of the regression fit are shown in Table 1. The poor fit is due in part to the fact that only travel time was used as the independent variable. Had other variables such an income, and/or an accessibility index been used in addition to travel time, it is expected that the regression fit could have been improved. Unfortunately, these additional variables were not available.

Despite the poor regression fit, the resulting area-wide F-factors, derived from

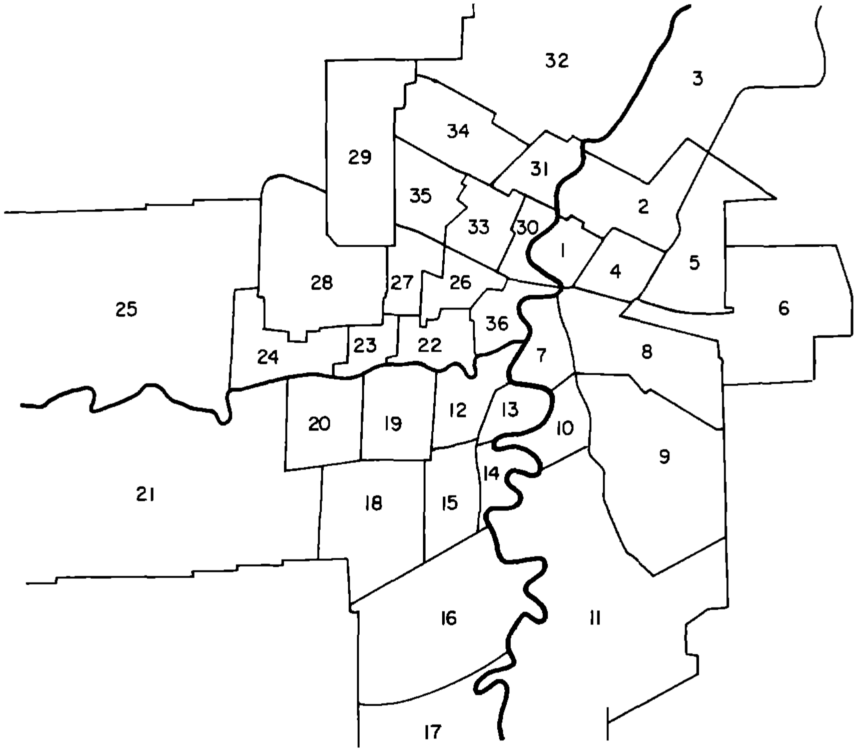

Fig. 1. City of Winnipeg superzones.

Table 1. Results of the regression fit to cell-by-cell F-factors.

Power function: $F = kt^{-a}$		
Year	Equation	R^2
1981	$F = 77.04\ t^{-1.506}$	0.40
1976	$F = 22.34\ t^{-1.141}$	0.37
1971	$F = 32.38\ t^{-1.266}$	0.33
1962	$F = 52.30\ t^{-1.467}$	0.26
Exponential function: $F = ke^{-at}$		
Year	Equation	R^2
1981	$F = 4.17\ e^{-0.075t}$	0.36
1976	$F = 2.77\ e^{-0.061t}$	0.32
1971	$F = 3.40\ e^{-0.067t}$	0.27
1962	$F = 3.40\ e^{-0.021t}$	0.21

Table 2. Goodness of fit to observed zone-to-zone base years trips.

BPR method				
Base year	Calibration accuracy			
	Observed study area mean zone-to-zone trips	Absolute mean error (% of study area mean trips)	RMSE (% of study area mean trips)	D-stat (%)
1962	56	13 (23%)	38 (67%)	12
1971	72	20 (28%)	46 (64%)	14
1976	77	23 (30%)	55 (72%)	15
1981	80	24 (30%)	50 (63%)	15
Regression method				
Base year	Calibration accuracy			
	Observed study area mean zone-to-zone trips	Absolute mean error (% of study a area mean trips)	RMSE (% of study area mean trips)	D-stat (%)
1962	56	15 (27%)	43 (77%)	13
1971	72	22 (31%)	51 (71%)	15
1976	77	25 (32%)	59 (77%)	16
1981	80	25 (31%)	52 (65%)	16

the power function, were used to estimate the base year zone-to-zone trips. The observed and estimated zone-to-zone trips were compared and the error measures are shown in Table 2. A comparison of the error measures for the regression and BPR method at the base year (Table 2), indicates that the regression method resulted in slightly higher errors. As a percent of the study area mean interzonal trips, the absolute mean error, the rmse, and the D-statistics across all of the years were higher than the BPR method on the average 2.5%, 6% and 1% respectively.

Because the regression method resulted in such a poor fit of the zone-to-zone F-factors, and because it produced slightly higher errors than the BPR method in estimating the base year trip interchanges, it was not used for forecasting. It was expected that the procedure would have provided a basis for the comparison of future forecasts, against those made by the conventional BPR method. Instead, the F-factors generated through the conventional BPR method alone were analysed for calibration and forecasting accuracy.

Calibration results

The trip length frequency distributions obtained from the BPR method, for each

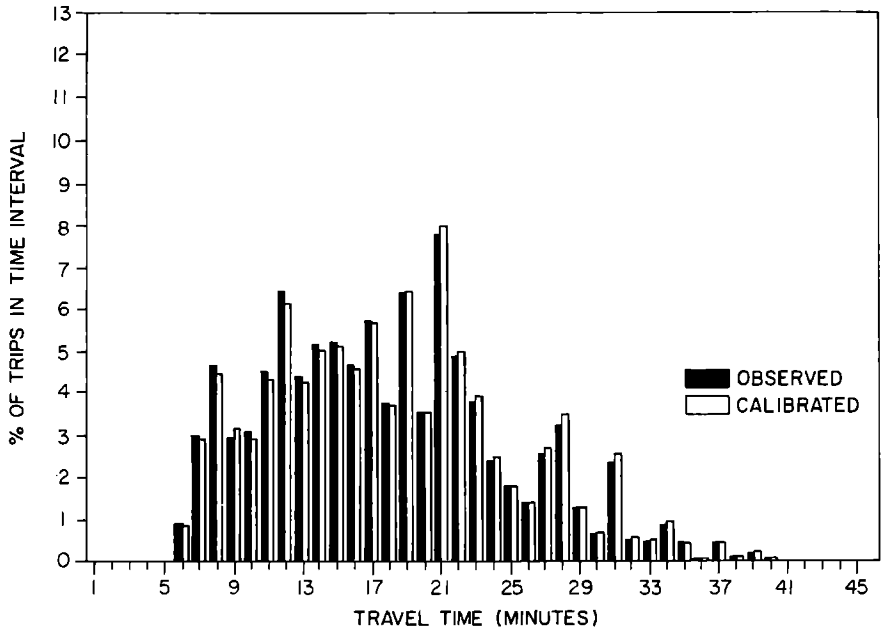

Fig. 2a. 1981 Trip length frequency distribution.

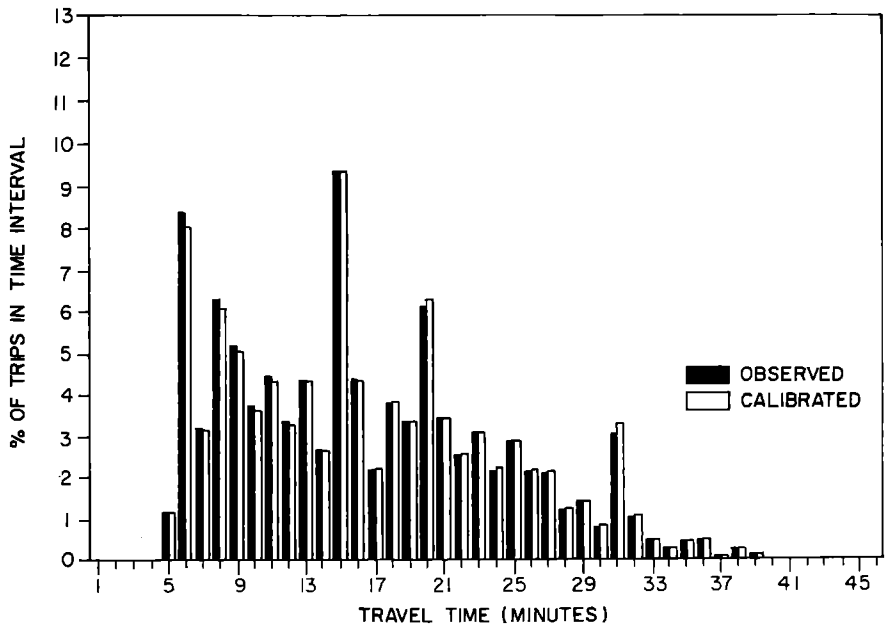

Fig. 2b. 1976 Trip length frequency distribution.

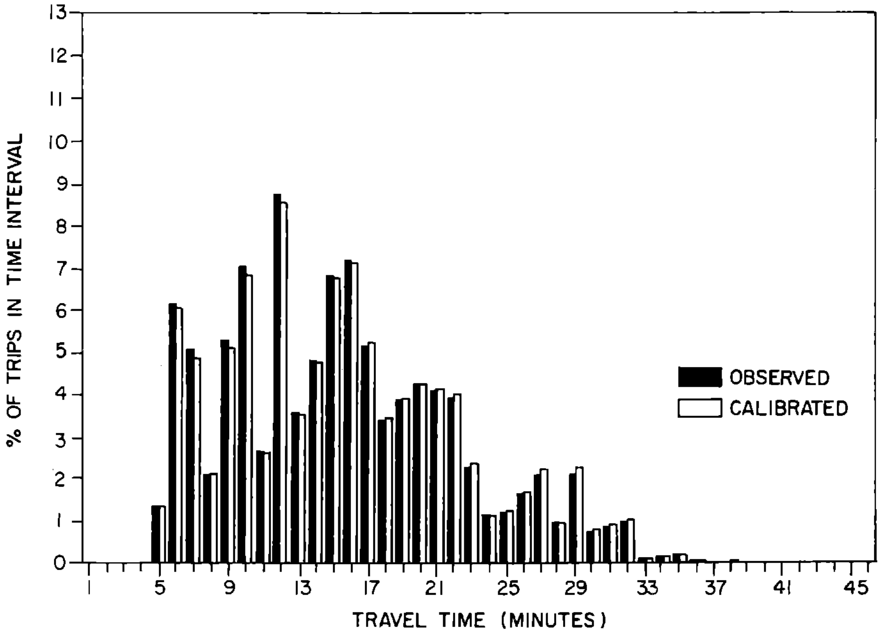

Fig. 2c. 1971 Trip length frequency distribution.

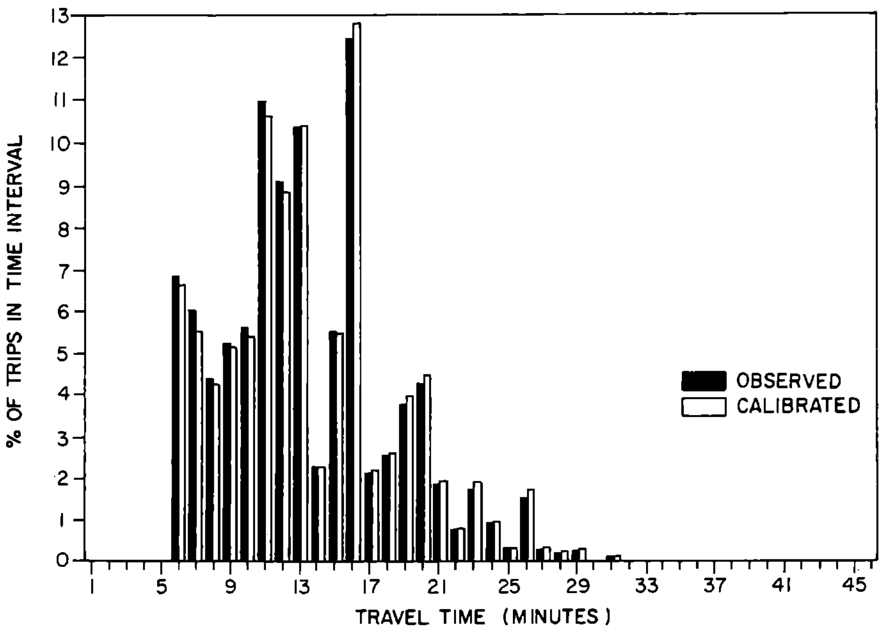

Fig. 2d. 1962 Trip length frequency distribution.

year, are shown in Figs. 2a – d. The calibrated trip length frequency distributions are comparable with the observed distributions for each year. The differences between the observed and calibrated trip length frequency distributions, by travel time interval, are less than 7% for all years. From Figs. 2a – d, it can be seen that the distributions have been shifting over time and the trend is towards longer average trip lengths. The average trip lengths and the standard deviations of the distributions about the mean are shown in Table 3. The average trip lengths have seen an increase from 13.4 min in 1962 to 18 min in 1981. The increase in average trip length is not surprising since the city has had substantial suburban growth, and travellers are willing to accept longer travel time to their place of work. The coefficient of variability (CV), the ratio of the standard deviation to the mean trip, is used to compare the relative variability of the distributions (Friedman 1972). The magnitude of the CV values for all of the distributions are between 0.37 and 0.47 (Table 3). The distributions appear to have similar variability, and indicate a proportionate increase in the variability of the distributions as the average trip length increases.

While the BPR calibrated model closely reproduced the observed frequency distributions, as expected it did not replicate the observed trip matrix as closely at the zone-to-zone level. The BPR calibrated gravity model was used to estimate the zone-to-zone base year trips. The observed and estimated base year trips were compared. The error measures for the differences between the two matrices are shown in Table 2. The zone-to-zone trips estimates for 1962 resulted in the lowest overall errors after calibration. The D-statistic, a measure of errors due to distribution, ranged from between 12 to 15%.

The F-factor curves obtained from the BPR calibration for the four years are shown in Fig. 3 in their absolute values. The curves have the typical exponentially decaying shape indicating the disutility associated with increasing travel time. It can be seen that the F-factors in their absolute forms did not remain constant but changed over the years as travel conditions changed. The shape of the curves also indicates that there has been a gradual shift towards relatively higher F-factors for longer trip length. This is particularly true for 1981, where the F-factor curve is above and to the right of the other curves. The three years 1976,

Table 3. Observed trip length frequency distributions.

Year	Average observed trip length (min)	Standard deviation of the distribution (min)	Coefficient of variation
1962	13.4	5.0	0.37
1971	15.5	6.6	0.43
1976	16.3	7.7	0.47
1982	18.0	7.0	0.39

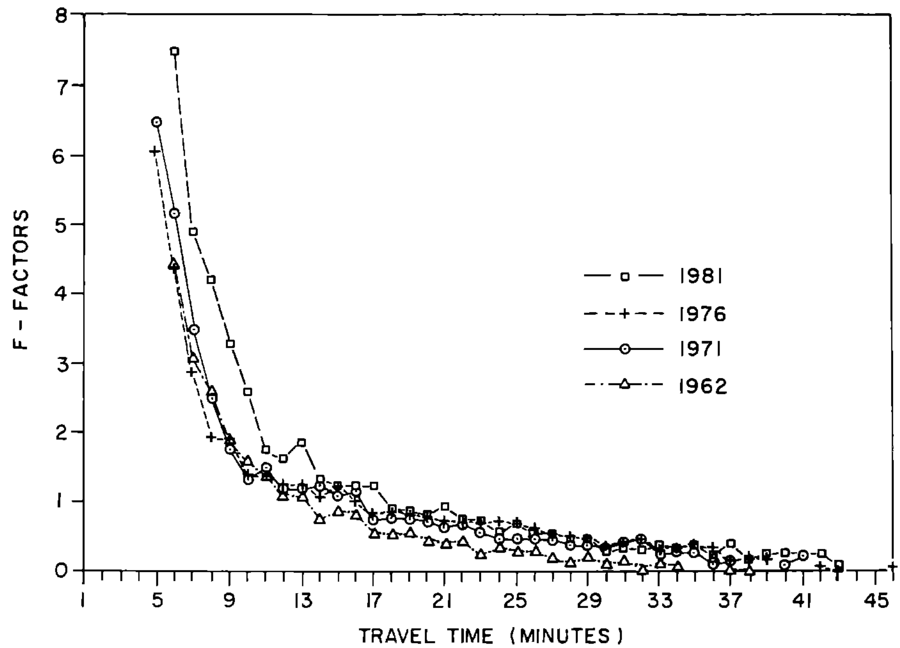

Fig. 3. F-factor curves.

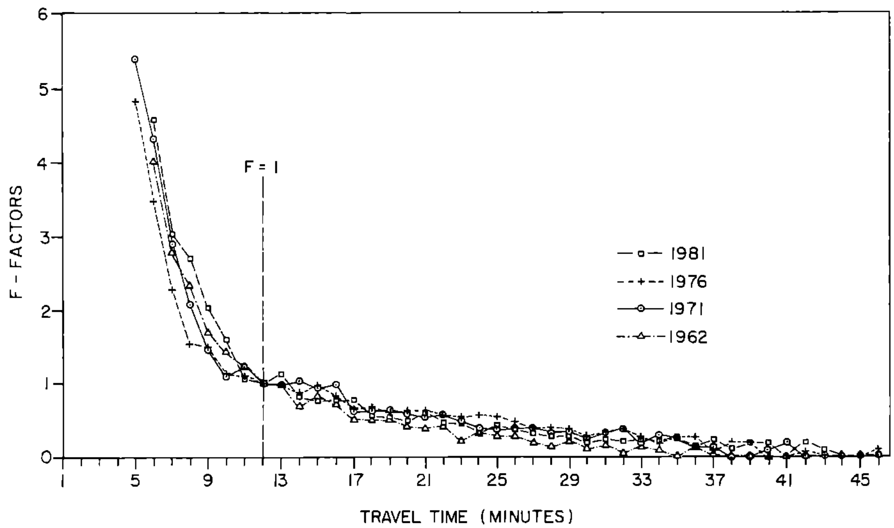

Fig. 4. Normalized F-factor curves.

1971 and 1962 have F-factor curves that are very similar in shape. To make a meaningful comparison of the F-factors these absolute values will have to be normalized. The twelve minute point seems to be the clear break in the F-factor curves.

To compare the F-factors across the years, the curves have been normalized so that, for the twelve minute time interval, the F-factor is 1.0 for all years. The normalized curves are shown in Fig. 4. The average difference between the maximum and minimum values of the set of four curves, for travel times less than 12 min is 0.69 with a standard deviation of 0.37. For travel times greater than 12 min, the average F-factor difference is 0.21 with a standard deviation of 0.06. It can be generalized that the F-factors curves, across the years, are stable within a range of values, for the City of Winnipeg.

The K-factors, calculated for each pair, were analyzed in order to note any patterns. The O-D pairs were studied to determine if any relationship existed between the K-factors over the 20 year period. Only a few of the zone pairs maintained about the same magnitude of K-factors over the period of time. As an example, trips originating from a residential area and destinated to the larger of the two universities in Winnipeg, had K-factors ranging from 1.0 to 1.2, over the 20 year period. Another superzone to the same destination, had K-factors consistently increasing from 0.70 to 1.60 between 1962 to 1981. The O-D pairs destined to the downtown, had K-factors predominantly between 0.80 and 1.20. These observations however, were limited only to these O-D pairs. In general the K-factors for the study area as a whole were not consistent from one year to the next. Using these K-factors for prediction may be inappropriate.

Forecasting with the calibrated model

To assess how well the BPR calibrated model can forecast future trips, 1962 was used as the base year and its F-factors alone were used to forecast to 1971, 1976 and 1981. Following this, 1971 was used as the base year and its F-factors were used to forecast 1976 and 1981 trips. The 1976 base year F-factors were then used to forecast 1981 trips. Similar forecasts were carried out using both F-factors and K-factors. Forecasts were then made using neither F-factor nor K-factor (i.e. using only production and attraction data). This was used to study and compare the errors that would be generated if the model had not been calibrated. The results are shown in Tables 4, 5 and 6.

The model predictions using only F-factors resulted in the forecasts with the better results. They gave lower overall errors, than when no F-factors were used or when the combination of F-factors and K-factors were used. The percent absolute mean errors for the forecasts with no F-factors were on average 13% higher than when only F-factors were used. The percent root-rmse was on average 26% higher, and the percent of the estimated trips that needed to be reassigned (based on D-statistics) were 7% higher.

When forecasts were carried out with the combination of F-factor and K-factor, the percent absolute mean errors were on the average 4% higher than when F-factors alone were used. The percent rmse was on the average 10% higher, and the percent of estimated trips that needed to be reassigned (D-

Table 4. Errors in gravity model forecast no F-factors and K-factors.

(a) Mean absolute error (as % of study area mean trips)

Year	1962	1971	1976	1981
1962	22 (39%)	32 (44%)	35 (45%)	38 (48%)
1971	–	–	–	–
1976	–	–	–	–
1981	–	–	–	–

(b) RMSE (as % of mean study area trips)

Year	1962	1971	1976	1981
1962	65 (116%)	76 (105%)	82 (106%)	83 (104%)
1971	–	–	–	–
1976	–	–	–	–
1981	–	–	–	–

(c) D-statistics (%)

Year	1962	1971	1976	1981
1962	20	22	23	24
1971	–	–	–	–
1976	–	–	–	–
1981	–	–	–	–

statistic) was 1% higher. From this analysis, it is apparent that the use of K-factors are not necessary since they produced trip forecasts with larger errors than when they are not used.

The errors generated by the forecasts using F-factors alone resulted in the lowest overall errors. The reliability of the model predictions was therefore assessed on the basis of these errors. Expressed as a percent of the average zone-to-zone trips, the resulting percent errors were very consistent (i.e. monotonic or constant). Forecasts made from the base year to the most immediate horizon year (e.g. from 1962 to 1971, 1971 to 1976, and 1976 to 1981) had percent absolute mean errors ranging from 31% to 32%, and rmse and D-statistics for the same pair of years were also consistent. The rmse ranged from 65% to 79% and the D-statistics from 16% to 17%. One may thus conclude that future trip distribution predictions for Winnipeg, using the gravity model, will most likely produce about the same magnitude of errors.

The reliability of a model prediction is akin to how consistent its error estimates are. From this analysis, one can thus conclude that the gravity model forecasts, using only F-factors, is generally reliable for predicting trips in Winnipeg. It produced fairly consistent prediction errors. It is difficult to categorically specify an acceptable level for the magnitudes of the errors. This study does not

Table 5. Errors in gravity model forecast with F-factors only.

(a) Mean absolute error (as % of study area mean trips)

Year	1962	1971	1976	1981
1962	13 (23%)	23 (32%)	27 (35%)	27 (34%)
1971	-	20 (28%)	24 (31%)	26 (32%)
1976	-	-	23 (30%)	26 (32%)
1981	-	-	-	24 (30%)

(b) RMSE (as % of mean study area trips)

Year	1962	1971	1976	1981
1962	38 (68%)	55 (76%)	76 (99%)	68 (85%)
1971	-	46 (64%)	61 (79%)	52 (65%)
1976	-	-	55 (72%)	52 (65%)
1981	-	-	-	50 (63%)

(c) D-statistics (%)

Year	1962	1971	1976	1981
1962	12	16	18	17
1971	-	14	16	16
1976	-	-	15	17
1981	-	-	-	15

attempt to do that. The decision to accept any error is subjective. It can depend on the objectives of the planner and/or on the importance of the results.

Conclusion and recommendations

The gravity model calibrated with the Winnipeg data produced F-factors that resulted in consistent overall errors in predictions of future trips. The F-factors found over the four study year period were generally stable within a narrow range of values. The pattern of K-factors was very irregular and varied for the majority of zone pairs, from one year to the next. From this forecasting analysis, it was apparent that K-factors are not necessary since they resulted in predictions with larger errors than without their use.

The results of the forecasting analysis, using F-factors alone, indicated that the prediction errors are rather consistent from one prediction year to the next. This means that the model can be used with greater confidence since the relative magnitudes of the expected errors are known.

A number of recommendations are made following this analysis of the gravity

Table 6. Errors in gravity model for forecast with F-factors and K-factors.

(a) Mean absolute error (as % of study area mean trips)

Year	1962	1971	1976	1981
1962	0	25 (35%)	32 (41%)	33 (41%)
1971	–	0	23 (30%)	26 (32%)
1976	–	–	0	27 (34%)
1981	–	–	–	0

(b) RMSE (as % of mean study area trips)

Year	1962	1971	1976	1981
1962	0	62 (86%)	98 (127%)	98 (123%)
1971	–	0	50 (65%)	49 (61%)
1976	–	–	0	53 (66%)
1981	–	–	–	0

(c) D-statistics (%)

Year	1962	1971	1976	1981
1962	0	17	21	21
1971	–	0	15	16
1976	–	–	0	17
1981	–	–	–	0

model, with respect to the Winnipeg data. First, since the F-factors are stable within a range of values for the 20 year period, present forecasts should be made using the upper and lower ranges that bound the curves, instead of choosing just one of the four curves. Both the upper and lower values of the range can be used for testing future land use scenarios. Second, K-factors are not necessary and should not be used in the forecasting process. Third, further study should be carried out using the cell-by-cell regression method for obtaining F-factors. Despite the poor regression fit, the errors resulting from the comparison of the estimated and observed zone-to-zone base year trips were only slightly higher than the BPR method. Had the regression fit been better, the procedure might well have produced much better calibration and forecasting results than the BPR method. An additional procedure was also suggested that might improve the results. In order to make the O-D specific F-factors more representative of a function that incorporates trip making variables such as travel time, etc., perhaps the factors could be grouped, based on their similarities, by a cluster type analysis. This would result in clusters of F-factors that exhibit similar trip making characteristics.

Finally, while the accuracy of the predictions are of primary concern, it is important to consider the effects of the trip interchange estimates with respect to the next phase of the transport demand modelling; trip assignment. How will errors associated with the predicted O-D trips affect trip assignment results? For example, if the number of trips between an O-D pair is overestimated by 500 trips, should that be considered significant in the assignment phase? Perhaps when assigned over the network, these trips may be evenly distributed over the links of the network and result in neglible errors on an individual link basis. In which case the errors in O-D trip estimate become insignificant. However, if these trips are not evenly distributed over the network, but concentrate on one or two links, then the errors become quite significant. It is in this situation that the reliability of the forecasts becomes critical.

Acknowledgements

The authors gratefully acknowledge Mr Ed Guertin, from the City of Winnipeg, Streets and Transportation Department, who made the data for this study available to them. The authors would also like to thank the referees of this paper for their useful comments and suggestions.

References

The City of Winnipeg Travel and Demographic Trends, 1962–1981. Works and Operation Division, Streets and Transportation Department. City of Winnipeg, 1982

Dickey, J. W. (1983) Metropolitan Transportation Planning. 2nd edn. (p. 210). McGraw-Hill Publisher

Emmerson, P. (1982) Accuracy of Trip Forecast Based on Simple Distribution Models. TRRL 1044

Freidman, H. 61972) Introduction to Statistics. 1st edn. (p. 105). Publisher Random House Inc.

Sosslau, B., Arthur et al. (1978) Quick Response Urban Travel Estimation Technique and Transferable Parameters: User's Guide TRB 187, 1978

Stopher, P. R., A. H. Meyburg (1975) Urban Transportation Modelling and Plannning (p. 156)

Todes, M. A. (1978) The Inapplicability of the Gravity Model in Non-gravity Situations. Traffic Engineering and Control (pp. 1–6)

Wilson, S. R. (1976) Statistical Notes on the Evaluation of Calibrated Models. Transpn. Res. Vol. X. (pp. 343–345)