Ratio Between Real And Linear Distances: Estimation of Wiggle Factor of Road Network in Pakistan

Syed Mohammad Farrukh

Co-Chairman, Institute of Business Economics Research & Academia

Abstract

In this research we have calculated the wiggle factor (WF) of Pakistani roads with the help of new features introduced by Google-map, i.e., Straight-line method, which is going to help the other researcher to find out the wiggle factor for their region effortlessly. With the help of this paper the calculated wiggle factor of Pakistani roads will help the logistic managers and transporters to find out the transportation and logistics expenses. This paper is explanatory research and the technique we used in this research for the collection of samples is stratified sampling technique. There are two variables one is straight-line which is independent variable, and the other one variable is road line which is dependent variable. We have collected the distances between eleven major cities of Pakistan for both straight-line and road line variables. Multiple Linear Regression has been used for the analysis of data by the researcher, because we have two variables, to measure the relationship between these two variables Multiple Linear Regression is best statistic tool. In this paper the wiggle of Pakistani roads, calculated by the researcher, is 1.432 which is higher than the traditionally used wiggle factor 1.2 which was calculated by Cooper in 1983. The wiggle factor cannot remain constant; however, it will change after decades because of significant changes occur in the roads network. Therefore, the researchers need to calculate the wiggle factor for their geographic region to find out the near to accurate value of WF for their region.

<u>*Keywords:*</u> *Google Maps, Wiggle Factor, Explanatory Research, Stratified Sampling Technique.*

Introduction

The linear distance between two geographical points is usually different than the actual road distance covered to reach that point. The ratio between the two distances is widely known as wiggle factor (WF). Since different variances are averaged out so it is not the actual figure rather it is an estimation of correction factor which is though widely used by logistics company for estimation of time and cost of the delivery of goods and also for various other usage. Many businesses face the problem in calculating actual time and cost for delivering goods to their customers as mostly it straight line distance which is available but finding the actual distance between two points is a tedious task (Caamaño, 2016). It is further mentioned in Caamaño (2016) research that for estimation of actual distance Cooper (1983) suggested

the use of curvature of the road factor which is now known and called as wiggle factor. This factor now a days is used by academic and business professionals. In view of current China-Pakistan Economic Corridor (CPEC), a long road network is built to connect Gwadar Port in Pakistan to the Southeastern border of China and not only that lot of road transportation of goods is expected on the main CPEC rout but also the peripheral roads between main cities and different economic zones will have lot of transportation activity. The upward pattern in the utilization of road transport has proceeded for a long time, and it appears to be impossible that the significance of road cargo transport will lessen soon. Rail cargo has remained moderately static for quite a while, pipelines are as yet vital for certain particular developments, however, the need to comprehend the relative benefits of, say, intermodal ocean- road cargo is critically important. The purpose of calculating the Wiggle Factor, for Pakistan road network, is to identify the deviation in distances i.e. how much the distance covered through the conventional roads deviates from the straight line between two points. It is generally used to evaluate the route distances for land transport either used as public transport or for the movement of logistics of goods. Though Wiggle Factor is an estimation we understand some degree of accuracy is required, because it is frequently used to calculate not only the fuel costs but also as estimation of the traveled time (which actually largely reflect on both the efficiency of the process and responsiveness).

Review of Literature and Conceptualization

In vehicle scheduling and distribution management problems it is very important to consider the expected distances between two points (the customer and the supplier**).** The parameter used to represent the distance covered by a fleet of trucks is usually assumed to be either the sum of the straight-line distances between the customers and the depot, Or the sum of the weighted distances, the weights being taken as either the average demand of the customer over a period of time or the frequency of service that he requires. This random definition of calculation of the transportation cost has been justified long ago as we had no method that yielded the optimal truck routes and their total associated distance. For the near to optimal routing problem's solution (Dantzig & Ramser, 1959) suggested an algorithm, in which they proposed ideal directing of a fleet of fuel conveyance trucks between a mass terminal and countless stations supplied by the terminal. It is wanted to figure out how to appoint stations to trucks in such a way, to the point that station requests are fulfilled and add up to mileage secured by the fleet is a base. A system in view of a linear programming definition is given for acquiring a close ideal arrangement. (Clarke & Wright, 1964) Used the same algorithm but they modified the method, which tends to lay more accentuation on completely filling trucks than on limiting distances. (Webb, 1968)Investigated the correlation between the measures of distances. Webb's results are to some extent conflicting mainly because he has considered only four particular examples which is lacking of providing enough data. Not exclusively were the issue little limited in number, they were just for the requests requiring conveyance on a solitary day, in these conditions, the purposes of least minute and distances totals were not indistinguishable, the partition could speak to a noteworthy cost contrast, and

the purpose of least distance entirety could be moved huge distances by changes in travel confinements.

The mathematical approaches regarding the calculation of straight line or the shortest distance were discovered by the (Robusto, 1957) using the Cosine-Haversine formula. This method considers the earth as a sphere therefore uses spherical trigonometry to find the shortest distance, so an error is committed. Due to the similar latitude points the error is not significant, in addition (Loves & Morris, 1979) stated that Euclidean distance is more helpful to utilize and seem to assess urban separations all the more precisely. the curved separation display created sensibly exact outcomes, and with its remarkable directional predisposition, might be particularly reasonable where the street arrange isn't exceedingly created., on the other hand (Berens & Korling, 1985) suggested multi parametric estimation instead of single parameter distance function to increase the degree of accuracy. Rather than the outcomes got for the U.S.A by (Loves & Morris, 1979), there were just slight increments in the level of precision of the appraisals when single-parameter separate capacities were supplanted by multi parametric ones. Since, besides, the assurance of a moment or even a third parameter involves a lot of work. it is dubious whether the utilization of a two-parameter capacity can keep on being viewed as a monetarily legitimate technique in Europe, however (Love & Morris, 1988) stated that they agree with the statement of (Berens & Korling, 1985) that "a great deal of extra work has to be done" in order to carry out the calculation with the help of two parameters rather than one, but in any case, there is just a narrow chances of acquiring more exact outcome along these ways (from a two parameter model). Moreover (Loves & Morris, 1979) stated in their previous papers, each geological range is extraordinary. The likelihood of getting significantly more exact outcomes from two parametric models is in this way reliant on the area being demonstrated. Thus, summing up from some other confined informational index is possibly deceptive. Clients should do their own particular examinations utilizing decency of-fit criteria proper to their application.

In addition to this paper (Cooper, 1983) suggested the use of straight-line-distances algorithms concerned with the solution of the vehicle scheduling problems to minimize the transportation cost. On the proof from his experimental investigation, it appears to be obvious that the straight-line separate intermediary for mean transport working expenses on visit joins is alright advocated. The case for receiving another visit measure, specifically the minimization of the likelihood of surpassing a given visit time, appears to be less secure in perspective of input data which is poorer in quality. The data utilized as a part of that paper recommend that connection time difference is less unsurprising than mean connection cost, and this must lessen the unwavering quality of arrangements from any calculation which relies upon the presence of such a relationship. The most certain decide must be that, at whatever point proxies are to be utilized as contribution to rearrange information gathering, at that point an example test is attractive to watch that the proxy holds adequately well. In addition (Stokx, 1991) used a model of (Christofides & Eilon, 1969) for assessing expected course lengths for multi-drop trips (serving a few goals) from radial distances is applicable as an alternate way to stay away from genuine vehicle routing.

Many authors have used straight line distances but (Williams, Schwartz, Newhouse, & Bennett, 1983) were able to obtain road distances, which represent that the actual road distances exceed straight line distance by 20-25%. William study was based on the access of medical facilities to the public, hence they categorize each specialty of medical like General medicine, family medicine, Orthopedic, general surgeon, obstetrics/gynecology and pediatric. William also stated that the greater the availability of specialist, the greater the improvement to access, However (McGuirk & Porell, 1984)stated that the customer are much more sensitive to time differential than to distance differential, since they builds up a spatial request model of hospitals decision to experimentally gauge the effects of distance and time on doctor's facility usage designs. With a cross-item proportion estimation approach, the impacts of physical access are assessed in the wake of controlling for spatial anomalies inferable from the dispersion of healing centers and populace in metropolitan regions. The experimental outcomes propose that separation and time factors unequivocally impact doctor's facility decision, even in metropolitan ranges where options are broadly accessible, and that their belongings change crosswise over administration orders and doctor's facilities, in addition (Phibbs & Luft, 1995) suggested that distances is not accurate proxy for travel time for all applications. Investigations of healing center demand and decision of clinic have utilized straight line remove from a patient's home to doctor's facilities as a measure of geographic access, yet there is the potential for predisposition if straight line separate does not precisely reflect travel time. (Phibbs & Luft, 1995) Underscore that their outcomes don't recommend that distance is an exact intermediary for travel time for all applications. For instance, it is improper to depend just on separate in the outline of a crisis medicinal transport framework.

In the current writing (Witlox, 2007) examined the separation estimation issue utilizing a huge travel overview database in which they discover to what degree can removes as straight-line, most limited way or briefest time be considered as great intermediaries for separations voyaged by means of genuine street systems. In synopsis they presume that the general revealed distances appear to be extremely solid if revised for the effect of exceptions, yet that case subordinate high contrasts happen between the detailed and GIS-registered separation. These deviations are the biggest for stumbles over short distances.

Studies of (Evans, Marrero, & Butler, 1981) introduced an investigation on natural learning and intellectual mapping, one of the goal of this examination was to analyze changes in the exactness of psychological maps with learning, on other hand (Saisa, Garling, Garling, & Lindberg, 2010)expressed that the straight-line separations were judged less precisely for areas, moreover they gave some obvious outcomes concerning the connection between various sorts of cognized separations, straight-line separations, travel separations, and travel times. These outcomes were conceivable to foresee from hypotheses about psychological maps. Past conjecturing about separation discernment appears to have depended much on numerical models of cartographic maps, the consequences of the present examination demonstrate that substantive hypotheses of intellectual mapping may helpfully be applied as a powerful influence for investigate issues in remove comprehension. From the current

writing (Caamaño, 2016) have exhibited a philosophy to figure a Wiggle Factor for any circumstance. This procedure is sufficiently straightforward to be appropriate to other geographic regions with no particular programming device, which makes it extremely valuable in situations when one tries to advance transportation courses rapidly and financially. In addition, the paper shows a refresh of the Wiggle Factor esteem in Spain, recognizing two variations for different context.

Due to frequent changes in the road infrastructure in Pakistan and the rest of the world the wiggle factor for each country need to be re calculated to get the exact or near to accurate result for the respective country or geographic region. Each country optimizes its road line structure in order to minimize the distances between the cities and locations. Pakistan changed its road network by deploying the network of Motorways, bridges and tunnels which helps to reduce the time utilized by traveling the distance and it ultimately reduces the cost of fuel by avoiding the zigzag routes. The recent development in Pakistan road network improved the infrastructure of the transportation routes, specifically road transportation. Likewise, the investment of China in China Pakistan Economic Corridor (CPEC) made Pakistan the gateway to Central Asia which ultimately improved the geographic importance of Pakistan to the central Asian countries. Those countries which are directly or indirectly related to CPEC would also be interested in Pakistan Road network in order to analyze their cost of transportation.

The main purpose of this research is to identify the changes occurs in the circuitry factor (wiggle factor) of any geographic region specifically in Pakistan due the significantly change in the road network. The recent development in the road networks of Pakistan reduced the distances between cities as well as the mega investment of chine in China Pakistan Economic Corridor (CPEC) increased the importance of Pakistan to the other regional countries. Due to the china Pakistan economic corridor (CPEC) the new web of roads and railway line is establishing in Pakistan, which ultimately improves the wiggle factor of Pakistan for the benefits of the host country as well as for those countries which will get benefits from CPEC. By calculating the wiggle factor of Pakistan, the logisticians will be able to calculate the transportation cost, the time duration required for the delivery of their goods, the redesigning and rescheduling of their distribution in order to use their assets efficiently.

Research Questions
The main focus of this study is to identify that how much the wiggle factor of Pakistan has been improved after the recent developments in the road infrastructure of Pakistan. (Cooper, 1983) Suggested that 1.2 wiggle factor for UK roads which was later accepted and widely used by the logisticians and the scientific community for the other regions of the globe as well. Cooper gave that suggestion more than three decades before and that study was limited only for the developed countries such as European countries. Pakistan has improved its road network in the recent years in which the most momentous change occurs after china Pakistan economic corridor. Therefore, it is very important to find out the wiggle factor for Pakistan's

road network.

Research Methodology

Population and Geographical Distribution of the Sample
In this paper we have selected eleven cities in Pakistan which are connected to each other via air routes as well as they are connected through road infrastructure. The distribution of sample in this geographic condition is in such a way that all the major cities are covered and most importantly all the foremost routes are roofed. Moreover, the distribution of sample is in such manner that the most important project "The CPEC" (China Pakistan Economic Corridor) needs to be covered up. All of the four provinces Sind, Punjab, Baluchistan, KPK (Khyber Pakhtunkhwa) including Gilgit-Baltistan are linked, because each province is capable of different land structure, for example the land of Sindh and Punjab are mostly flat and they have fewer mountain which affect the wiggle factor positively, on the other hand the KPK and Gilgit road infrastructure is more like a mountain area due to which the wiggle factor in these two specific regions is higher than the rest of provinces. Therefore, we consider all of the provinces in the geographic distribution of sample to get the average result of the wiggle factor of Pakistan.

Since the research is based on secondary data of distances between the major cities of Pakistan, The inter-city distances of 11 cities of Pakistan has been collected with the help of Google Maps using both straight line distances and actual road distances, which means two pairs of 110 readings used as sample data for analysis. The stratified sampling technique has been selected by the researcher, because of accuracy of availability of data, as the road distances of major Cities of Pakistan has fewer variations as compared to small cities. In this sampling technique we have further used stratum technique because we have one common characteristic in all samples that is airport, which helps us to accurately find the straight line and road line distance to an exact point which is airport.

Procedure of the Study and Data Collection
We have selected eleven major cities of Pakistan as a sample test. Airport is the basic criteria for the selection of each sample city, because the destination point between these cities is from airport to airport. As well as the distance between the cities if we consider road is either by motorway or highway. Previously the researcher used spherical trigonometry or Vincentry formulae to get the accurate distance between two points, but due to the new feature launched by Google map to calculate the straight line between two coordinates is much easier. The road line distances can also be obtained from any GPS navigation software or Google map. As shown in the figure.
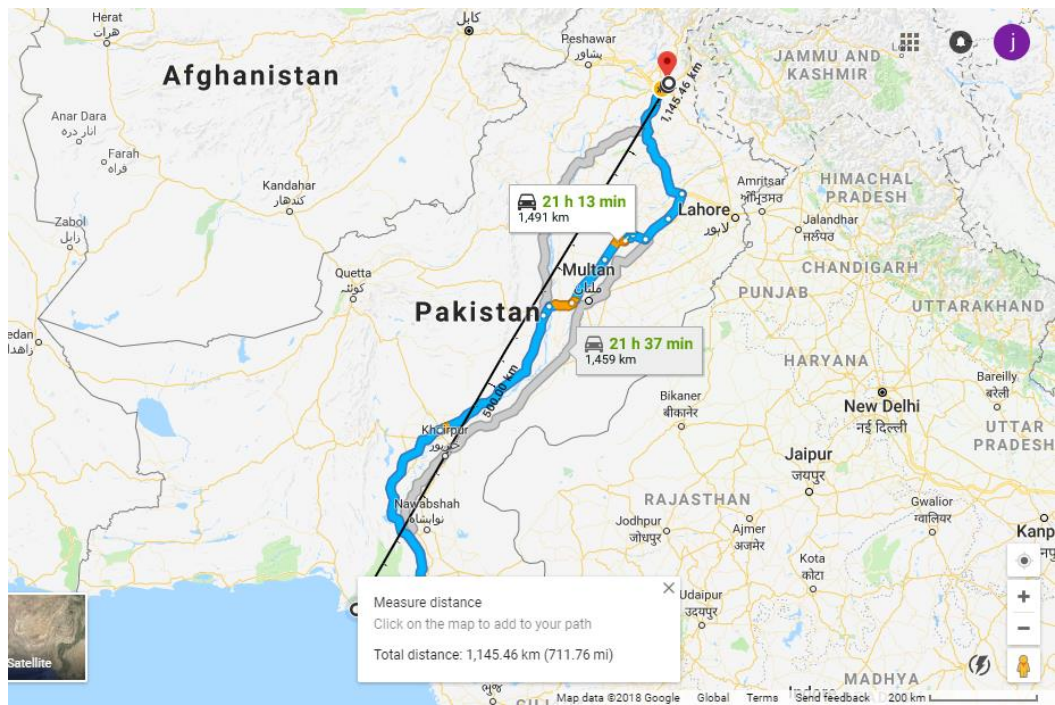
Obtaining the Wiggle Factor and Result Analysis:

Initially we calculated the wiggle factor for each trail by dividing the road distance by the straight-line distance. After analyzing all the 110 routes, a general wiggle factor calculated by averaging each individual wiggle factor.

Reliability and Validity

The data we have collected for this paper is from the widely used geographical map that is Google map. We have not calculated the road distances and straight-line distances from ourselves; therefore, the data is secondary data. The data has been verified several times from both the directions (from start point to end point), although there was a little variation in road line distances, because we have selected airports in the cities as our point of starting and ending journey. Due to one-way route or some deviations in the return route the data is slightly higher which is not inevitable. However, in straight line method the intercity journey has the same distances from both the directions.

Reliability Analysis:

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Straight Line Distance	972.1818	303508.939	.969	.
Road Distance	690.2407	139002.107	.969	.

Since the researcher using only one Independent Variable for Research, so there is no need to exclude any variable from research model, which also shows from Cronbach's Alpha if item Deleted column, as all values are nil.

Road Distance $= \alpha + \beta_1$ Straight Line Distance

The table shows that which variables are entered, and which variables are deleted while all variables are entered and no variable is removed. Under "method," it examines "Enter". This basically means that we are performing a complete entry regression in which all predictors are entered into the model.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.969[a]	.939	.938	137.19186

a. Predictors: (Constant), Straight Line Distance
b. Dependent Variable: Road Distance

R is the multiple correlation coefficients, which measures the strength of linear relationship between Dependent & Independent variables. In this case R is 0.969 i.e. 96.9% which shows the highly correlation between dependent & independent variable.

R square is the proportion of variation in the response variable explained by the regression model. It is also known as *"coefficient of determination"*. In this case 93.9% variation explained by the model which is good enough. Adjusted R square value means that there is no over estimation of R square and also appropriate to report Adjusted R square. The standard error of the estimate is the estimation of standard deviation of residual in the model. Here the value of St. Error of the Estimate is 137.19186.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	31049740.894	1	31049740.894	1649.686	.000[b]
	Residual	2032733.469	108	18821.606		
	Total	33082474.364	109			

a. Dependent Variable: Road Distance
b. Predictors: (Constant), Straight Line Distance

The Regression sum of squares notify us how much variability (not variance yet) is accumulated for with the Regression Model, which is the fitting of the least - squares line in this case it is 31049740.894.

The residual sum of squares enlightens us how much variability (again, not variance yet) is unaccounted for with the Regression Model. The total variability is sum of squares is 2032733.469. DF: Degree of freedom "Regression" is equal to number of independent variables, we have 1 variable (Predictors).

Degree of freedom "Residual" equal to Total number responded (N) and Total number of independent variable (K) whereas N-K-1 is the formula i.e. 109.

The F test for the regression model is calculated from Mean Regression and Mean Residual. It is a test of the null hypothesis that R Square (or just R) is equal to zero. The test of significance discloses that the probability of attaining an F stat as the one we have gained or more. F sharing on 1 and 108 degrees of freedom Sig value is lesser than .05. Therefore, we can't refuse the null hypothesis that R Square in the population is equivalent to zero.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-15.929	27.622		-.577	.565
	Straight Line Distance	1.432	.035	.969	40.616	.000

a. Dependent Variable: Road Distance

Straight Line Distance shows the positive relationship with Road Distance with (1.432) constant value. Show the level of strength of each Independent Variable on that of the Dependent, thus the higher Standardized Coefficients (BETA) of independent variable has the higher affect over dependent variable as Straight-Line Distance shows high impact on Road Distance with 0.969. Sig Value of each Independent Variable provide us information about the rejection and acceptance of Hypothesis, as the research is conducted on 95% Confidence Interval i.e. 5% level of Significance of Straight Line Distance on Road Distance has Sig Value is less than 0.05 i.e. (0.000) which means impact of Independent variable on Dependent is very high.

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	168.9551	2405.8589	972.1818	533.72278	110
Residual	-215.36320	778.21893	.00000	136.56109	110
Std. Predicted Value	-1.505	2.686	.000	1.000	110
Std. Residual	-1.570	5.672	.000	.995	110

a. Dependent Variable: Road Distance

The above table shows that the predicted value Minimum figure is 168.9551; Maximum value is equivalent to 2405. 8589.While the mean of the predicted mean is 972.1818, with a standard deviation (533.72278) and the total number responses are 110 it generally identifies that the statistics has calculated. It is generally difference between Independent and Dependent variable whereas it shows the variability whereas the minimum residual is -215.36320, Maximum residual is 778.21893 and the mean of the residuals is equivalent to .00000. As noted in the residual statistics, the standard deviation of residuals is equivalent to 136.56109. It signifies like for any standard deviation it may signify the average distribution in a statistic set. These STD predicted values mean has equivalent to 0.000 and std deviation is equivalent 1.000.Standardizing is a technique for transforming data so that its mean is 0.000. While STD deviation is falling .995.

Discussion:

Cooper's had suggested the value of 1.2 wiggle factor which is widely used, but (Caamaño, 2016) suggested that the wiggle factor should be calculated for both rural and urban roads separately, because urban roads are more likely straight as compare to rural due to which their wiggle factor will appear less than rural and will impact the overall wiggle factor of the country. In this paper we calculated the wiggle factor of Pakistan which is 1.432 by considering the distances between the eleven major cities of Pakistan. Most of the roads are highways and motorways because the infrastructure of the intra-cities roads is not well established. The consideration of collecting such data is to provide the wiggle factor of Pakistani roads which are widely used by the long haul or the roads which have a high impact on China Pakistan Economic Corridor. Moreover, in this paper we used a slightly advanced methodology of collecting data by taking the advantage of new feature offered by Google map which is easier than all the previous methodology.

Conclusion:

In this paper we have calculated the wiggle factor of Pakistan (specifically motorways and highways) for the first time. A new feature launched by Google-map of calculating straight line distance between two points has been used in this methodology which is going to provide other researchers a lot of alleviate in their works. The wiggle factor for the developed countries like European or American states will remain same, because their road network or railway network has been established well, but the developing countries like Pakistan, India, Iran as well as those European countries which are in developing phase, the wiggle factor might change. This change will occur due to the development of new highways and motorways which reduced the distances between two intended destinations.

Moreover, in this paper the suggested wiggle factor will provide better planning and forecasting and cost determination for the logisticians and transporters who are interested in doing the business in Pakistan.

Findings & Recommendations:

In past the road network of Pakistan was relatively more deviated from the straight line than European countries. The widely used wiggle factor by cooper which is 1.2 would not be suitable for Pakistani roads. The recent development in the road infrastructure of Pakistan which includes the development of highways motorways, tunnels, bridges and bypasses over the cities reduced the distances between the destinations, which ultimately reduced the wiggle factor. The suggested wiggle factor for Pakistan is 1.432 which is still higher than the previously suggested wiggle factor by cooper. We suggest that the wiggle factor for each geographic region needs to be re-calculated after each mega development in their roads, railways infrastructure, because the more the infrastructure developed the lesser will be the wiggle factor which will impact positively on the cost of the logistics.

References

Berens, W., & Korling, F. J. (1985). Estimating Road Distances by Mathematical Functions. *European Journal of Operational Research,21(1985)*, 34-56.

Caamaño, P. D. (2016). An Improved Methodology to Determine The Wiggle Factor: An Application For Spanish Roads. *Brazilian Journal ofoperations and production management,13(1)*, 52-56.

Christofides, N., & Eilon, S. (1969). Expected Distances in Distribution Problems. *Operational Research,20(4)*, 437-443.

Clarke, G., & Wright, J. W. (1964). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research,12(4)*, 568-581.

Cooper, J. C. (1983). The Use of Straight Line Distances in Solutions to the Vehicle Scheduling Problem. *The Journal of the Operational Research Society,34(5)*, 419-424.

Dantzig, G. B., & Ramser, J. H. (1959). The truck disptaching problem. *Informs,6(1)*, 80-91.

Evans, G. W., Marrero, D. G., & Butler, P. A. (1981). Enviromental Learning and Cognitive Mapping. *Enviroment and Behaviour,13(1)*, 83-104.

Love, R. F., & MORRIS, J. G. (1988). On Estimating Road Distances by Mathematical Functions. *European Journal of Operational Research,36(1988), 36*, 251-253.

Loves, R. F., & Morris, J. G. (1979). Mathematical Models of Road Travel Distances. *Management Science,25(2)*, 130-139.

McGuirk, M. A., & Porell, F. W. (1984). Spatial Patterns of Hospital Utilization: The Impact of Distance and Time. *Inquiry,21(1)*, 84-95.

Phibbs, C. S., & Luft, H. S. (1995). Correlation of Travel Times on Roads Versus Straight Line Distances. *Medical care Research and Reveiw,52(4), 52*, 532-542.

Robusto, C. C. (1957). The Cosine-Haversine Formula. *The American Mathematical Monthly,64(1), 64*, 38-40.

Saisa, J., Garling, A. S., Garling, T., & Lindberg, E. (2010). Intraurban Cognitive Distance: The Relationshipbetween Judgments of Straight-LineDistances, Travel Distances, and Travel Times. *Geographical Analysis,18(2)*, 167-174.

Stokx, C. F. (1991). Deriving route lengths from radial distances:Empirical evidence. *European Journal of Operational Research,(1991)*, 22-26.

Webb, H. M. (1968). Cost Functions in the Location of Depots for Multiple-Delivery Journeys. *Operational Research Society,19(3)*, 311-320.

Williams, A. P., Schwartz, W. B., Newhouse, J. P., & Bennett, B. W. (1983). How Many Miles To The Doctor? *The New England Journal Of Medicine,309(16)*, 958-963.

Witlox, F. (2007). Evaluating the reliability of reported distance data in urban travel behaviour analysis. *Journal of Transport Geography,15(2007)*, 172-183.