CHAPTER VII-TN 35: STATISTICAL CONSIDERATIONS USING GRAVITY TYPE MODELS TO EXPLAIN VISITOR FLOWS

By M.F. Goodchild

ABSTRACT

This paper reports on an inquiry into the problems of fitting aggregate spatial interaction models to empirical data. The concern is with flows of visitors to recreation sites from a variety of origins, and with the class of models for those flows, normally referred to as gravity models. The paper reviews the conventional approach to spatial interaction analysis, using standard measures of success. Then four basic problems are discussed in the context of a small data set. These relate to: non-linearity of models; integral values of visitor flows; volume of flow observed and goodness of fit; and the weak theoretical basis for the models used. Finally, the paper examines other problems which arise in more complex situations/ and makes general recommendations.

The early sections on problems are statistical and will be of more interest to technical readers; later sections are more general, and the statistical results are restated non-technically in the conclusions.

INTRODUCTION

This paper reports on an inquiry into the problems of fitting aggregate spatial interaction models to empirical data. More specifically, it is concerned with flows of visitors to recreation sites from a variety of origins, and with the class of models for those flows, normally referred to as gravity models from a rather tenuous analogy to the Newtonian inverse square law of gravitational attraction.

The most general form of the gravity model is as follows:

$I(_{i,j}) = f_1()P_i f_2()A_j f_3()D_{i,j}$

$I_{i,j}$ is the flow from origin i to destination j,

P_i is a measure of the potential supply of visitors from origin 1,

A_j is a measure of the attractiveness of place j,

$D_{i,j}$ is a measure of the trip from i to j, and

$f_1()$, $f_2()$, and $f_3()$ are functions calibrated for a given activity and set of origins and destinations.

Since the early sections of the paper consider the problems of fitting models to a single destination, f_2 can be ignored. Assume that P_i is the population of the i^{th} origin, that f_l is linear and that f_3 takes one of the following forms: $e^{bD}{}_{i,j}$ or $D^b{}_{i,j}$. Then there are two possible expressions of the general form presented above:

(1) $I_{i,j}= P_i e^{bD}{}_{i,j}$

(2) $I_{i,j}= P_i\ A_j D^b{}_{i,j}$

While Equation 2 is, the historic gravity model, recent studies have shown increasing interest in Equation B, both on empirical and *a priori* theoretical grounds (Cesario 1974, "More …"; Wilson 1970). Furthermore, the algebraic difference gives rise to rather different methodological problems in each case.

There are several such problems. First, the non-linear form of both models means that if standard linear regression techniques are to be used in calibration, there must be a transformation of the variables, so that the results of the analysis appear in units which are often misleading, and which make comparison with the original data difficult. Second, visitor flows are composed of integral numbers of people. If flows are sampled over some limited time period and used as estimates of long-term interaction, then the sampling process will be quite different to that normally assumed in regression analysis. Third, the success of the analysis, or the degree to which the model fits the data, will depend on the length of the sampled period, so that the greater the total visitor flow sampled, the better the fit. Finally, gravity models have a rather weak theoretical basis. Such explanations as do exist (McConnell & Duff 1976; Schneider 1959; Wilson 1970) contain strong assumptions which are easily broken in the real world, so that the model may suffer from structural inadequacies in specific situations.

The early sections on basic problems are statistical and will be of more interest to technical readers; later sections are more general, and the statistical results are restated non-technically in the conclusions.

THE CONVENTIONAL APPROACH

The data set used for illustration and simulation in this study is the one discussed by Beaman, Knetsch and Cheung (see TN 19). It gives the visitor flows to Rowan's Ravine Provincial Park in Saskatchewan, from seventeen origin areas (see TN 1). In the sampled period, the year 1969, there were 9,828 visitor vehicles, including 5,862 from the observation unit containing the city of Regina, and none from four of the areas. Following the earlier work, each observed flow has been arbitrarily increased by 1. In calibrating logarithmic models one is added to avoid there being an attempt to take the logarithm of zero.

TABLE 1 ANALYSIS OF EQUATION 1

*****	a	b	R2	PE***	RMS****
OLS	-.488	-.0547	.755(.685)*	138.4	217.3
NLLS	-.812	-.0381	.755	75.0	108.7
WNLLS	-.365	-.0585	.296	191.8	337.7(104.1)**
MIN PE	-.380	-.0449	.898	66.6	84.7
MIN RMS	.368	-.0495	.486	67.8	71.9

* R^2 given first for $\ln(I_{i,j}/P_i)$, in parentheses for $I_{i,j}$

* Figure in parentheses computed using weighted observations.

*** Mean absolute % error in flow.

**** Root mean square error in flow. (

*****OLS=Ordinary Least Squares; NLLS=Non-Linear Least Squares; WNLLS=Weighted Non-Linear Least Squares; MIN PE=Minimum Mean Absolute Percent Error; MIN RMS=Minimum Root Mean Square Error.

The conventional approach to calibration is to transform the equations to linearity by taking logarithms. Equation 1 can be transformed as follows:

$$\ln(I_{i,j}/P_i)= a + bD_{i,j}$$

and Equation 2:

$$\ln(I_{i,j}/P_i)=\ln(a)+b \ln(D_{i,j})$$

Both equations are of the form y = c + dx, and involve two constants to be determined by fitting to data. In the first, the negative exponential case, $\ln(I_{i,j}/P_i)$ is regressed against $D_{i,j}$ to estimate values for a and b, while in the second $\ln(I_{i,j}/P_i)$ is regressed against $\ln(D_{i,j})$ to estimate ln(a) and b.

The results of these OLS, Ordinary Least Squares, regressions for the Saskatchewan data can be found in Figures 1 and 2. The R^2 values, describing the degree to which $\ln(I_{i,j}/P_i)$ can be predicted from $D_{i,j}$ and $\ln D_{i,j}$, were .755 and .767 respectively, and the respective equation coefficients a and b can be found in Tables 1 and 2. Statistically, both analyses were highly significant and led to rejection with 99.9% level of certainty that of the null hypothesis could be rejected. Still, virtually the same R^2 values were obtained so R^2 is not showing that one or another model is better.

TABLE 2: ANALYSIS OF EQUATION 2

	a	b	R2	PE	RMS
OLS	10965	-3.40	.767 (.344)*	111.7	141.5
NLLS	163	-2.01	.903	70.8	75.9
WNLLS	8770	-3.50	-.023	236.3	308.6 (93.0)**
MIN PE/MIN RMS	30000	-3.42	-.020	46.7	59.4

* See footnotes to TABLE 1.

THE TRANSFORMATION PROBLEM

The largest residual observed in the negative exponential regression (Equation B) was -3.477; in the power law regression (Equation 2) -3.670. In terms of the ordinate, $\ln(I_{i,j}/P_i)$, which ranged from -12 to 0, these residuals were not large and were visually acceptable (see Figures 1 and 2). But in terms of $I_{i,j}/P_i$, in which the range is from .00001 to 1, the residuals represented an overprediction by a factor of roughly 40. Specifically, the observed flow $I_{i,j}$ for the point with the largest residual was l; the first OLS model predicted a flow of 32.3, the second 39.3. Residuals were then recalculated by taking the difference between the observed flow, and the flow predicted values of $\ln(I_{i,j}/P_i)$. Unlike the earlier values, these residuals will not sum to zero for either model. Equivalents to the earlier R^2 were calculated from the ratio of the sum of squared residuals to the visitor flow sum of squares for each model, to give .685 for the negative exponential and .344 for the power law. Clearly these models are much less successful if measured in terms of their ability to predict $I_{i,j}$ rather than $\ln(I_{i,j}/P_i)$.

Computing the difference between observed and predicted flows inevitably gives most weight to origins which contribute a high flow, and very little to smaller, more distant origins. Yet the planner who is interested in the proportionate difference between the observed and the predicted would give as much weight to a 10 percent error in a flow of 50, as to the same percentage error in a flow of 5000. Various authors have suggested that a percentage error in observed flows is a more useful measurement of residual variation in the recreation context (Elsner 1971; Ellis & VanDoren 1966) and have computed the root mean square percentage error (RMS) as a substitute for the conventional R^2 (see TN 19). The respective values of RMS error for the two models were 217.3 and 141.5, and the means of the absolute percentage error (PE) 138.4 and 111.7. Despite the high R^2 values, the predicted power of the models is rather weak when expressed in these terms.

A more direct approach to the fitting of the general model would be to avoid the logarithmic transformations altogether, and fit Equation 1 and 2 by a least squares procedure, directly minimizing the sum of squared differences between observed and predicted flows. This was done using a combination of Gaussian and Steepest Descent methods to minimize the objective and the relevant statistics are shown in Tables 1 and 2 for comparison with OLS. This procedure will be referred to as NLLS, for Non-Linear Least Squares. Cesario (1974, “more on …”) has discussed a similar approach.

There is a marked reduction shown in both RMS and mean absolute error (PE) and although the R^2 values should not be compared to those derived from the OLS models, they can reasonably be compared to the recomputed values based on flows which are shown in parentheses. It is clear that the use of logarithmic transformations can seriously reduce the validity of gravity-type models, when validity is defined by RMS or PE error measures.

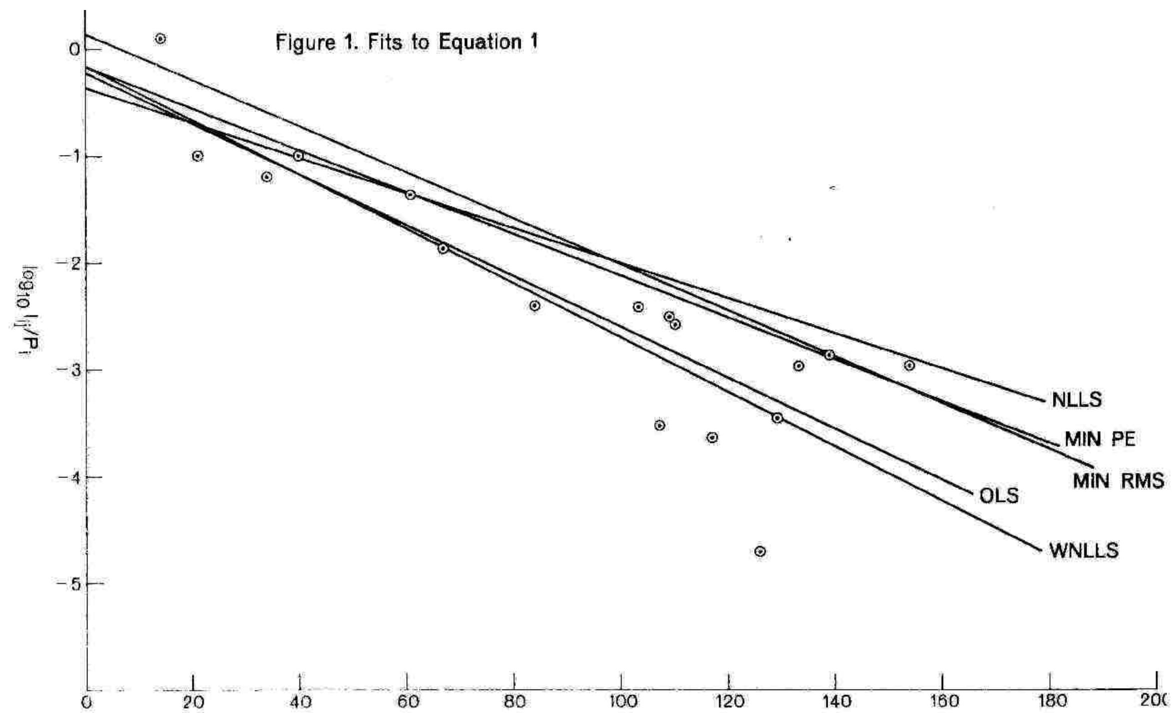

Figure 1. Fits to Equation 1

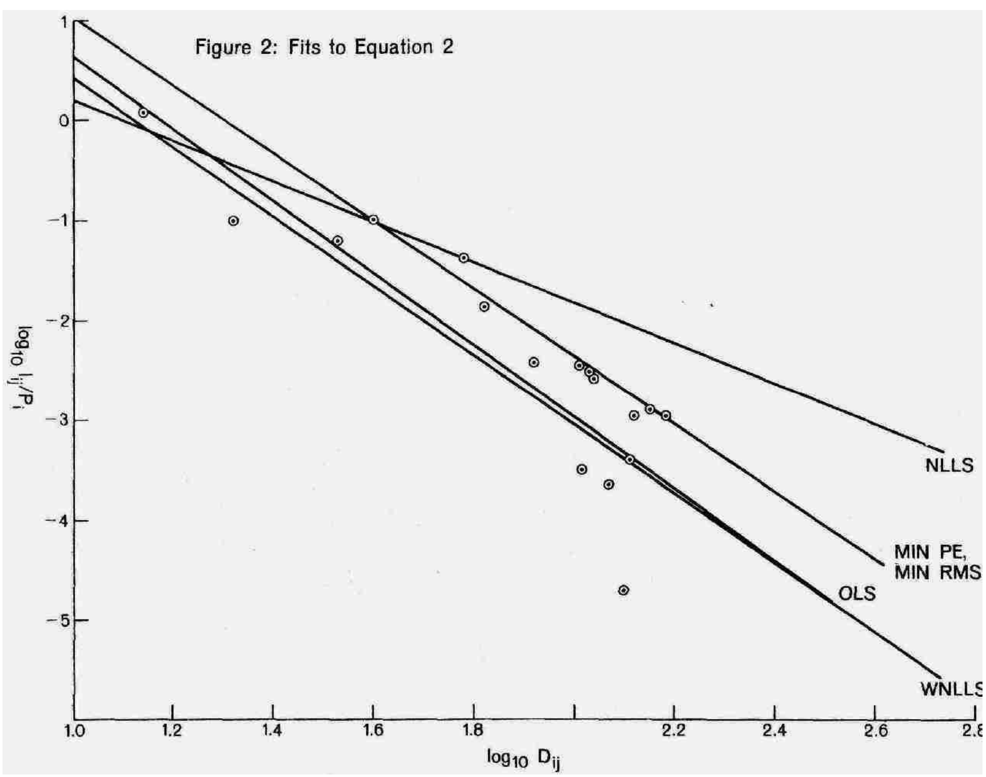

Figure 2: Fits to Equation 2

If it is argued that the RMS or PE measures are the most effective in the planning context, then ultimately the most satisfactory way of fitting or calibrating a flow model must be by direct minimization of these criteria, rather than the conventional sum of squared residuals. In fact for this particular data set the degree of improvement is not great. The minimum mean absolute

percentage error resulted when predictions were made using the equations: $I_{i,j}=P_i e^{-0.380-0.0449\ D_{i,j}}$
$I_{i,j}=P_i\ 0.30000\ D_{i,j}^{-3.42}$ and while the second equation also minimized the RMS error for the power law, the negative exponential was calibrated with a=0.368 and b=-0.0495 on this criterion. The accompanying statistics are shown in Tables 1 and 2 as the MIN PE and MIN RMS procedures. The approach used was an interactive procedure in which the user watched the performance of a general objective function as each constant was incremented over ranges prescribed by the user. The routine was written by the author.

For comparison, the above equations are plotted in Figures 1 and 2 with the results of the earlier procedures. They do not have the same visual impact as the OLS equations, which pass neatly through the points, since their objective functions and residuals are based on $I_{i,j}$ and thus only indirectly related to the ordinate $\ln(I_{i,j}/P_i)$.

THE HETEROSCEDASTICITY PROBLEM

In using the OLS linear regression model, one makes the assumption that residuals are independently distributed with a constant variance (homoscedasticity assumption). In terms of the interaction model transformed by taking the logarithm of $I_{i,j}/P_i$, each observation is assumed to be subject to the same residual error distribution, or in terms of $I_{i,j}/Pi$ to the same proportionate error. If $\ln(I_{i,j}/P_i)$ is assumed to have normally distributed errors with standard deviations of σ, so that errors in $I_{i,j}/Pi$ will be lognormal, or normal when expressed as proportions one has a model for usual (OLS) linear regression.

In reality, this model is quite inappropriate to the expected distribution of $I_{i,j}$. If the observed visitor flow is the result of a very large number of samplings of the origin population under a very small probability, then $I_{i,j}$ can be expected to follow a Poisson distribution (see TN 19). In other words, its distribution is discrete and non-normal, with a variance which is equal to its expected value, and thus varies from observation to observation. The variance in $\ln(I_{i,j}/P_i)$ is compounded by the distribution of P_i, and clearly violates the homoscedasticity assumption of OLS.

Beaman et al. (see TN 19) have calculated the variance in $\ln(I_{i,j}/P_i)$ as approximately proportional to $l/I_{i,j}$ by expanding the logarithm function in a Taylor series. (The error variance of $\ln(I_{i,j}/P_i)$ clearly depends on the value of P_i. Beaman et al. made the implicit assumption that P_i is constant in their Equation 9.) Thus the high-flow observations are subject to a much smaller error variance and merit much greater weight in the regression. To fit the logarithmic transformation of Equation B, Beaman et al. used a generalized least squares procedure (GLS) with flows weighted by the inverse of the dependent variable variance estimates.

A similar approach can be taken in the direct, nonlinear regression. If we assume a Poisson distribution, the error variance of $I_{i,j}$ in Equations 1 and 2 can be taken as $I_{i,j}$ itself. Then the appropriate weighted least squares criteria are:

Minimize Σ (I( )-P( ) exp (a+bD j)) $)^2/I_{i,j}$ and

Minimize Σ (I($_{i,j}$)-P( ) aD $_{i,j}$)**b)2/$I_{i,j}$

These criteria give the greatest weight to the smaller flow values, since in terms of $I_{i,j}$ these are the most reliable, whereas for the logarithmic regressions it was the high flow values that had the least error variance and the greatest weight. The criteria are quite similar to those for the direct minimization of RMS error:

Minimize $(1/n\ \Sigma\ ((I_{i,j}-P_i\ e^{a+bDi,j})^2/\ (P_i\ e^{a+bDi,j})^2))^{1/2}$ and

Minimize $(1/n\ \Sigma((I_{i,j}-P_i a\ D_{i,j}{}^b)^2/(P_i\ aD_{i,j}{}^b)^2))^{1/2}$

WHERE n is the number of points except that the denominator is formed from the predicted rather than the observed flow.

For comparison with the earlier procedures, regression results are shown in Tables 1 and 2 and Figures 1 and 2. Error measures were computed without observation weights. In both cases there is a considerable deterioration in the degree of fit reflected in the error measures. The corresponding weighted RMS figures are 104.1 and 93.0 respectively, and give a fairer representation of the degree of fit. But as would be expected, in neither case do the figures approach those for a direct minimization of RMS error.

It is intuitively reasonable to weight observations in inverse proportion to their error variance. Since the error variances are known from theoretical arguments, it is possible to make corrections for heteroscedasticity by weighting observations. But in reality the observed error variance is composed both of a Poisson-distributed part and a component due to structural inadequacies in the model, with an unknown distribution. Further, the data also violates several other assumptions. Errors are not normally distributed: when expressed as small visitor flows they are Poisson, with a highly skewed discrete distribution, and when expressed as log $I_{i,j}/P(1)$ they are compounded by the distribution of P_i. Further, because the values of the variables themselves are not sampled from a bivariate normal distribution in any of the models, questions of statistical inference and parameter estimation are discussed in the next sections. Discussion is based on the results of Monte Carlo simulations.

SIMULATION TECHNIQUES

The models under examination involve three "sampled" variables, $I_{i,j}$, P_i and $D_{i,j}$. But of these we may assume that $D_{i,j}$ and P_i are known to considerable accuracy, and that all error is concentrated in $I_{i,j}$. Thus error will be of two types: a statistical component due to the use of a flow value sampled over a limited time period as an estimate of a long-term average flow; and a systematic component due to the structural insufficiency of the model in explaining visitor flows.

In the next two sections, the results of simulations of the statistical error component in the Saskatchewan data set are discussed with two objectives. First, it is possible to simulate the sampling distributions of the important regression parameters and so gain a more quantitative notion of the relative importance of the statistical and structural components. Second, simulation can establish the importance of sample size in parameter estimation. Because of the complexity of the error distribution, the results will be specific to the data set and only qualitative generalities can be expected.

The distribution of $I_{i,j}$ is assumed to be Poisson for a given observation. $I_{i,j}$ is an unbiased, maximum likelihood estimate of the Poisson density parameter m such that:

$P(r) = m^r e^{-m}/r!$

WHERE P(r) is the probability that a Poisson process of density m will generate exactly r events in a trial.

To simulate the effect of statistical error, the observed visitor flows were used as estimates of m, and for each simulation run, each flow was independently replaced by a random Poisson deviate of that density. Thus the range of parameters found for models fitted to this simulated data will represent the range of uncertainty attributable to statistical error.

Two methods were used to generate the deviates. Descriptions of these is not provide in this revision of the paper since in 2006 Poission observation generation programs are readily available.

DISTRIBUTION PROPERTIES

Two types of simulations were made, one using real visitor flows as the basis of simulated flows, and the other using predicted flows from the OLS negative exponential model as the basis. The results of the first series of experiments are shown in Table 3. Twenty-four runs were made, and analyzed using the three models based on the negative exponential; the Ordinary Least Squares procedure using $\ln(I_{i,j}/P_i)$ as the dependent variable (OLS), the direct least squares calibration of Equation 1 (NLLS), and the latter using weights estimated from the predicted error variances (WNLLS).

The results show that for this data set, statistical error is very much less important than structural error. For the OLS case, the standard error of the coefficient in the original calibration of the model was estimated to be .00803. Yet successive simulations of statistical errors produced a range of coefficients with a measured standard deviation of only .0005. Similarly, the variation in R^2 as a result of simulation suggests that the major source of unexplained variance is structural. Again, the regression standard error in the coefficient is .808, compared to a standard error in simulation runs of .046.

Results are very similar for the non-linear regression. The skewness in the EMS and PE distributions is largely removed, and the consequent bias in single estimates is reduced. Again, the estimates of a and R^2 are remarkably precise, although there is a slight increase in the standard error for the b coefficients. This is somewhat offset by a proportionate reduction in the standard error of a. On the other hand, the weighted regression results revert to the pattern shown by OLS, giving slightly better estimates of b and biased estimates of RMS.

TABLE 3 FIRST SIMULATION EXPERIMENTS

	a	b	R2	PE	RMS
OLS					
Actual	-.488	-.0547	.755	138.4	217.3
Simulation	-.489	-.0547	.754	140.2	227.7
Std Error	.046	.0005	.010	5.3	16.6
NLLS					
Actual	-.812	-.0381	.905	75.0	108.7
Simulation	-.814	-.0382	.902	75.1	109.2
Std Error	.047	.0011	.010	2.2	3.4
WNLLS		.			
Actual	-.365	-.0585	.296	191.8	337.7
Simulation	-.378	-.0585	.289	196.4	347.6
Std Error	.054	.0005	.024	15.2	38.6

TABLE 4 SECOND SIMULATION EXPERIMENTS

	a	b	R2	PE	RMS
OLS					
Actual	-.488	-.0547	1.000	0.0	0.0
Simulation	-.528	-.0546	.984	20.6	29.3
Std Error	.089	.0019	.010	5.2	7.2
NLLS					
Actual	-.488	-.0547	1.00	0.0	0.0
Simulation	-.536	-.0538	.999	20.0	29.1
Std Error	.049	.0010	.001	5.5	6.8
WNLLS					
Actual	-.488	-.0547	1.000	0.0	0.0
Simulation	-.505	-.0548	.984	21.1	30.7
Std Error	.072	.0015	.015	7.3	9.0

The second set of simulations show the amount of error introduced into a perfectly predictable set of data by adding Poisson-distributed errors to visitor flows (Table 4). They show that the statistical error component has relatively little effect on error statistics by itself so that the vast majority of the error observed in the original data must be structural.

PE and RMS errors are remarkably independent of the-method of calibration when structural error is absent, while R^2 is more sensitive to the statistical error component in OLS and weighted non-linear (WNLLS) regressions. Standard errors of the coefficients tend to be rather higher than they were in the first set of simulations with real data, except in the case of NLLS.

The above simulations, particularly the first sets, provide a rapid method of evaluating the relative importance of statistical and structural error in the calibration of a spatial interaction model. ***Because of the dependence of the result on the specific configuration of the data set, and the immense difficulties of any analytic approach, it is suggested that such Poisson simulations be incorporated into any calibration of this type of model, particularly in cases where samples are small, and statistical error correspondingly high.***

AGGREGATION EFFECTS

Sample size will affect the calibration of a spatial interaction model in various ways. The statistical error component has known properties, if we can continue to assume that its distribution is Poisson, so that the error variance in observed flows will increase in direct proportion to the length of time sampled. The standard deviation will he proportional to the square root of sample size, so that it will decrease in proportion to the observed flow as sample size increases.

While statistical error may be well understood, there is no similar basis for assumptions about structural error. In one particular case, structural error might appear as a normally distributed residuals. Normally distributed residuals can be argued to arise as a result of a variety of “other” external effects on visitor flows combing to yield what appears to be a normal variable. In another case, structural error might be the results of a misspecification of the functional form of the model, perhaps error in specifying the effect of distance. It would seem reasonable on dimensional grounds to assume that the standard deviation of the structural error term will be in constant ratio to the observed flow, and thus rise in direct proportion to sample size. Thus

$I_{i,j}= \{I\}\ (i,,j\ )+c_1 I_{i,j}^{1/2} + c_2 I_{i,j}$

WHERE $\{I\}_{i,j}$ is the expected flow in the sampled interval;

$c_1 \rightarrow I_{i,j}^{1/2}$ is the statistical error component;

$c_2 \rightarrow I_{i,j}$ is the structural component.

STRUCTURAL ERROR

The gravity model is a purely inductive device, open to charges of curve-fitting which can only be countered by relating the excellent degrees of fit invariably obtained. The first part of this paper has concentrated on two problems encountered in calibrating the model: variations in performance between different procedures; and effective measurement of the goodness of fit. But while it is possible to evaluate the relative importance of statistical and structural error, and to make allowances for the latter, it is clear that the existence of the structural component represents an important deficiency in the model.

Any approach to the reduction of structural error must begin at the theoretical level with a discussion of the basis for the model's excellent fit to reality. The literature offers two major classes of explanation, first through hypotheses made at the aggregate level, and second through hypotheses about the behaviour of individuals in the system.

AGGREGATE THEORIES

Wilson (1970) has shown that under certain assumptions, trips made in a completely random way between origins and destinations will aggregate so that the total interaction between origin i and destination j is given by Equation 3, which is algebraically similar to Equation 1.

(3) $I_{i,j} = k_o\, k_d\, O_i\, D_j e^{-bc(i,j)}$

WHERE O_i is the total flow from

D_j is the total flow to j,

$c_{i,j}$ is the cost of one trip, and

k_o, k_d and b are constants,

The equation is derived based on an entropy maximizing strategy using an analogy to Boltzmann's statistical thermodynamics. Suppose the $I_{i,j}$ are known so that a table can be constructed to show the flows between each origin and each destination. It is possible to calculate the number of ways in which specific individuals can be assigned to specific flows, given the table of totals along each path. For example, if all individuals are in origin 1 and travel to destination B, then only one arrangement of individuals is possible, while if trips occur between multiple origins and destinations there are many ways in which the individuals can be rearranged without changing the totals. Not all patterns of flows are possible. The total flow from origin i must equal 0_i, and arrivals at j must equal D_j. Wilson introduces a further constraint that the total cost of all trips made in the system be C.

In a purely random world, in which no behavioural rules exist other than the constraints, it is reasonable to assume that all possible arrangements of trips are equally likely. This means that it is possible to calculate the probability of each pattern of flows from the numbers of alternative arrangements of individuals associated with each pattern, and to find the most likely pattern. The result is Equation 3. Provided the number of individuals in the system is large, this most likely pattern proves to be so much more likely than any other that it is possible to predict with some confidence that it will be the one found in reality.

There are three major ways in which structural inadequacies can arise in Wilson's analysis, and be reflected in structural errors in the model. First, the model invokes what is sometimes known as the Principle of Insufficient Reason (Harvey 1969) and assumes that all arrangements of individuals are equally likely. This will be untrue if, for example, certain kinds of individuals show a preference for certain types of trips, or if a set of individuals is associated with a particular origin and may only make trips from that origin. Second, the constraint that the total cost be C may not reflect behavioural reality. It is more likely that total cost fluctuates as the sum of a number of fluctuating individual expenditures. Third, behavioural traits may take the form of further constraints which limit the set of possible arrangements, and may lead to different conclusions.

BEHAVIOURAL THEORIES

In principle, a theoretical base could be established for the gravity model if it could be shown that a large number of individual actions, each made according to some general rule, would combine to produce aggregate flows which were in agreement with the model. Ewing (1974) discusses some of the relevant problems. In essence, in order to aggregate behaviour in which distance appears as at least a partial criterion, one must assume some kind of geometrical arrangement of origins and destinations, and so the aggregate flows will be unique to the geometry. It is possible for behaviour according to a general principle to combine with a specific geometry to produce a specific aggregate pattern, and it is conceivable that a specific behavioural principle might combine with a specific geometry to produce a general aggregate pattern, fitting the gravity model. But there is a logical impossibility in conceiving of a general behavioural rule which would give a general aggregate pattern independent of the specific geometry. There is no general principle of human geometry.

Niedercorn and Bechdoldt's analysis (1969) has received a great deal of attention. They consider a single individual in origin and the set of trips that he or she makes through time to various destinations j. Each trip will have a certain utility to the individual, which is assumed to depend on the number of purposes which are satisfied by the trip, which in turn is assumed to depend on some attribute of the destination, depending on the context. Each trip has an associated cost, and the individual is assumed to be constrained by a travel budget. Under these conditions, the pattern of trips that maximizes the total utility to the individual is shown to be of the form of Equation 2, under specific assumptions about the utility function.

When aggregated, Niedercorn and Bechdoldt's analysis is capable of predicting the relative magnitudes of flows from a single origin. In a system of many origins and destinations, it predicts that the flows from a single origin will be proportioned in a way which is compatible with the gravity model, but the relative magnitudes of flows from different origins will depend on the geometrical arrangement of destinations around each origin and so can only be in accordance with the model in specific geometrical situations.

DYNAMIC ATTRACTIVITY

One characteristic of the model which has not so far been treated as important is that the independent, predictor variables are usually assumed exogenous with constant values determined from physically measurable parameters. The analysis in the earlier sections was in the framework of a single destination. The model differentiates multiple destinations both on a geometric basis through $D_{i,j}$ and through the parameter A_j, the factor which explains differential flows when distances are constant. To some extent A_j will be exogenous, related in the recreation context to the physical characteristics of a site and its immediate surroundings. But in most cases A_j will also be related to the use that site j is actually experiencing, and thus be at least partly endogenous. A feedback loop will operate to make the site less attractive when the use is heavy, and perhaps also when it is excessively light.

The existence of dynamic determinants of attractiveness in an interaction system is not necessarily a source of structural error. Provided A_j is regarded not as an exogenous variable but as a parameter that can be expected to change whenever the set of $I_{i,j}$ changes, then no contribution to structural error will result. But as such, the model will have no predictive power. Structural errors will occur under sets of hypothetical flows unless the A_j are themselves modelled in terms of the determining flows and exogenous physical parameters. Cesario (1974, “more …”) refers to this as Stage II of a spatial interaction analysis.

MINIMIZING STRUCTURAL ERRORS

Both of the above sources of structural error can be minimized by an appropriate modelling procedure. Suppose that it is possible to assume that any set of individuals, when presented with a set of parks of given attractiveness at given distances, will proportion themselves in the same way, regardless of how many other alternative destinations exist. In effect, this is the Luce choice axiom (1959). Attractiveness may itself be dynamically related to use, and the actual flows from the origin to any site will depend on the alternatives available: the hypothesis concerns the relative proportions only.

The hypothesis can be written as follows:

$$\frac{P(\rightarrow i|j \rightarrow)}{P(\rightarrow k|i \rightarrow)} = \frac{f(a_j, D_{i,j})}{f(a_k, D_{k,i})}$$

OR

$P(\rightarrow j|i \rightarrow) = f(A_j,D_{i,j}))/(\sum f(A_k,D_{i,k}))$

WHERE $P(\rightarrow j|i \rightarrow)$ is the probability that an individual, having left origin i, will go to destination j, and (k indicates the sum over all destinations available from origin i.

In other words, the relative proportions visiting j and k are determined by the relative values of some function of attractiveness and distance for each destination.

Now the probability that an individual resident at i will make the trip to j is:

$P(i \rightarrow j) = P(i \rightarrow) P(\rightarrow j|i\rightarrow)$

WHERE $P(i\rightarrow)$ is the probability of leaving i to go anywhere. Thus

$I_{i,j} = P_i \; P(i \rightarrow j)$

$= P_i \; P(i \rightarrow) \; f(Aj, D_{i,j}))/ \Sigma \; f(A_k, D_{i,k})$

Then the flow from the origin to the j^{th} site will be of the form:

$I_{i,j} = E_i \; f(A_j,D_{i,j})$

WHERE parameter E_i depends on the alternatives available from origin i as well as on p_i.

Provided E_i and Aj are regarded as parameters unique to the given flows and systems geometry, then, and not equated with P_i or with observable properties of the site, neither source of structural error need exist. If the model remains inadequate, it must be because incorrect assumptions are made about the function f, or because the Luce choice axiom does not hold.

PERIPHERAL PROBLEMS

The total use of Rowan's Ravine Provincial Park in 1969 amounted to 9,828 vehicles, or more properly their occupants. None of the modelling procedures recognized this as a constraint. In linear regression, it is true that the sum of predicted values equals the sum of observed values, but not in non-linear regression or in cases where a linear model is fitted to transformed data. Boyet and Tolley (1966) recognized this problem, but it is difficult to deal with in the standard methodological practice.

In Wilson's analysis (1970) there are constraints both on the total arriving at a destination, and on the total leaving an origin, since k_o and k_d are not fitted parameters, but are solved for in the constraint equations. Iterative methods have been developed for fitting the model so that predicted and observed flows will obey the same constraints.

SUMMARY

The points raised in the paper can be divided into two groups. First, there are statistical issues involved in the fitting of models of the form represented by Equations 1 and 2 to spatial interaction data. The conventional method of least squares on transformed variables does not give the "best" fit to the model as that term would be understood by a recreation planner; the method does not minimize discrepancies between flows observed and those predicted by the model. The conventional measure of success is based on transformed variables, and shows quite different results when recomputed for the fit between observed and predicted flows.

Second, several points were made about the structural validity of spatial interaction models, and the generality of model parameters. The paper outlined the conditions under which the model could be consistent with underlying general principles of human spatial behaviour, and thus regarded as a general model.