REVIEW

BY GORDON EWING

The six papers in this section discuss what are commonly referred to as trip distribution models. The models range in complexity from a variant of the most simple gravity model to one which is amongst the most generalized in the present literature(1976 literature).

The first three papers by Cheung et al. involve different elaborations of the simple gravity model. In the simplest, "A Method for Predicting Enroute Overnight Park Use", (TN 18), the form of the basic gravity model, $V(i,j)=kP(i)^{\alpha}M(j)^{\beta}/D(i,j)^{\gamma}$ is retained, but with the terms in the numerator redefined. Where P(i) normally refers to the size of the origin it is replaced by T(i,j), the volume of traffic on a section- of highway leading to j, and M(j), destination size, is redefined as the number of campsites in park j. The other two papers by Cheung replace destination mass, M(j), by a measure of j's attractiveness, A(j), which is exogenously defined as a function of several park characteristics. In addition, a term measuring the strength of competition from alternative destinations completes the modified gravity model. As with the attractiveness terms, the "alternative" term is measured exogenously (defined by a mathematical formula using values that do not depend on V(i,j) – this is not to say that development of alternatives and their attractions has not been influenced by V(i,j) observed in the past) and simply included in the model as an extra independent variable. The distinction to notice between the other two papers is that "A Day Use Visitation Model" posits that the independent variables, with the exception of distance, are additive in their effect on V(i,j), whilst "An Application of Mathematical Models to Compare Two Potential Park Sites" (TN 7) assumes the same variables combine multiplicatively to affect V(i,j). As is common in trip distribution modelling studies, there is no discussion of the theory that might give rise to these alternative composition rules. Population growth, change in family composition, population "aging", etc. are not considered in TN 7, although they should be.

Whilst the model in Cesario's TN 4 is a variant of the simple gravity model, the innovative feature is that origin emissivity, E(i), and park attractiveness, A(j), now become parameters to be estimated in calibrating the model. The implications of the innovation are important. In calibrating the traditional gravity model with several origin and destination variables, it is never clear whether poor calibration results reflect the absence of important variables or the incorrect structure of the model or some combination. Removing the need to pre-specify the presumed important variables means that poor calibration is more clearly attributable to an improperly structured model.

It also transpires that the ability to estimate E(i) and A(j) values in what appears to be a simple gravity model, V(i,j) = E(i)A(j) f(D(i,j)), enables these estimates to be used as dependent variables in a second round of parameter estimation yielding in a more complex distribution model. Evidence for this is to be found in the Beaman et al. paper, TN 11. Indeed, although Cesario's concern is not to calibrate a generalized distribution model, the method he describes clearly provides the impetus for the remaining two papers in this section which involve much more general models.

Both "A Work Plan for the Development of a Mathematical Model to Predict and Explain Overnight Use of Parks" by MacDonald, Netherton and Cesario and TN 11 contain terms not found in the other papers. The major difference between the two papers is that the former is concerned only with the formulation of an appropriate mathematical model, whereas Beaman et al. paper considers the mathematical problem of how to estimate parameters in the model. Put simply, the two papers are about the interpretation of emissivity terms as estimated in Cesario's paper. They argue that an origin's emissivity is firstly a function of its size and socio-economic composition. In addition, since E(i) is unique to each origin, they argue that it contains within it an estimate of the spatial supply of facilities which is also unique to any origin. In particular, they see this supply as having two separate and opposite effects. One is referred to as the alternative factor, defined in terms of the sum of all parks' attractivenesses discounted for distance. This is similar to the alternative factor in two of the Cheung et al. papers, except that Cheung had to ignore the effect of alternatives' attractiveness for lack of information. In all cases the size of the alternative factor is seen as having a negative effect on the patronage of any one park, in the sense that these alternatives constitute competition. However, at the same time that all n-1 alternatives provide competition for the nth park, all n parks, and in particular their attractiveness as mediated by distance, can be thought to have a supply-generating-demand effect. That is to say, the more a particular spatially distributed facility is provided, the more use there is likely to be of such facilities. Thus, for example, given a latent demand for cross-country skiing, the more facilities that are provided, the more participation is likely up to a point. TN 11 discusses the problem of separating these two effects embedded in the emissivity term for each origin.

Turning to the individual papers, each has features which merit comment. Cheung's "Day-Use Model", which provided the impetus for many succeeding Technical Notes, is characteristic of earlier interaction models in that the form of the model is simplified to make statistical estimation straightforward. Thus, unlike later work, the attractiveness of destinations is computed exogenously and weights on the distance terms are determined by trial and error. In particular, the unusual procedure of having three different distance exponents depending on a destination's distance is used. The rationale is that the overall R^2 value in the calibrated model is maximized using these three parameters. However, it is of note that a single exponent for all distances of either 1.5 or 2 results in only a 0.01 reduction in R^2. Statistically it is very probable that the increment of 2 parameters in the model accounts for this small increase, rather than the variable distance exponent having any behavioral significance. This variable exponent on a destination's distance contrasts with the treatment of distances to alternatives which are all assigned an exponent of .5 in the alternative factor. No rationale for this inconsistency is suggested.

Turning to the attractivity factor in the model, and in particular to the function used to calculate its value ($T(j) = \Sigma_e S(e) (\Sigma_m R(m,e)Q(m,e))$), its form deserves comment. The relative popularity rating of activity e, S(e), is measured as a direct function of a Canada-wide participation rate in the activity. The importance of facility m (R(m)) in enabling activity e to be enjoyed is calculated from the correlation between a park's attendance and the "amount of" facility m at the park. It would seem that there is an element of "double counting" in the above equation. For example, if horse riding has a low popularity rating (S(e)), it seems likely that horse rental at a park will do little to increase its attendance figures. The latter being the case will produce a low importance rating (R(m)) for horse rental. However R(m) is clearly dependent on S(e) in this case, whereas the implication of the equation is that the terms measure separate and independent factors contributing to park attractiveness.

One other question about the attractivity function is how the "rank" scores (Q(m)) were arrived at for the amounts of certain facilities. For example, why does a 9-hole golf course rate 8.5, and an 18-hole course rate 11, or why 6 showers rate 8 and 8 showers rate 10.5? Why, also, are rank scores then used in calculating T(j) as if they had interval scale properties?

It is noteworthy that the T(j) values calculated contribute one-fifth of one percent to the explained variance of V(i,j), the dependent variable in the model.

As regards the model itself, three points stand out. Firstly, 84% of the 91% of explained variance in V(i,j) is attributable to P(i)/g(D(i,j)), the population of the origin mediated by distance. This dominance of the origin's mass term (population) in interaction models where other non-mass terms are included is a familiar pattern. Indeed it is difficult to cite any published work where origin mass is not the major contributor to the variance explained, along with distance. On the face of it, it might seem that trip distribution modelling is a subject that requires little further investigation, if only 10% or 20% of variance is left unexplained by the simplest gravity model. However, a large value of R^2 in this case depends on there being a large variance of origin mass which produces a large variation in use figures. If an area were studied where origin masses were more uniform, variance attributable to origin mass would be much less. In such a case, R^2 would be greatly reduced if the trip distribution model were not properly specified. In other words, when origin mass has a large variance, just about any interaction model will give a high R^2 provided origin mass is one of the independent variables. But where the variance of origin mass is small, any but the properly specified interaction model will result in a low R^2 value.

The second feature of the model which is common to most empirical calibrations of interaction models is that the objective function used in estimating parameters is the minimization of the residual sum of squares of a linear regression model, or in other words the maximizations of R^2. It has been argued elsewhere (see #Reference 38) that R^2 is often quite insensitive to significant parameter changes. An example of this in the present paper is given in Table 5. There is only a 3% variation in R^2 as the distance exponent varies between 1.5 and 3 and indeed between 1.5 and 2.5 the range in R^2 is only 1%. Yet these exponents represent a sizeable range in terms of the typical values found in a wide variety of gravity model studies. It has been suggested that whilst R^2 is a valuable measure, it can be used in combination with other measures, such as mean predicted distance. This latter has been found to be much more sensitive to parameter changes. Thus an alternative procedure would be to seek parameter estimates which minimized the difference between mean observed and predicted travel distances, subject to a constraint on the degree that R^2 could deviate from its maximum value.

A third feature of Cheung's interaction model which is unfortunately common to most interaction studies is that no test of the homoscedasticity of the independent variables is provided. The effect of a variable violating this assumption can be to artificially increase R^2. Consider origin population since that is the most important variable contributing to R^2. If, for example, there are many small origins and one or a few very large places, with only a few intermediate sized centres, which is a not uncommon urban size distribution, the flows have a distribution that is heteroscedastic (small flows may have high variability compared to their magnitude but large flows will have much larger variances – e.g. consider that variance is proportional to or equal to the expected flow as considered in TN 19) and will undoubtedly result in an appreciably higher R^2 value than if variance was homoscedastic (variance was a constant σ^2 for all flows). Although Cheung does not provide information on the size distribution of the origin areas, Table 8 gives a sample of origin area sizes. Fourteen of the fifteen in that table range from l,000 to 8,000 in population, whilst the fifteenth, Saskatoon, has a population of 117,000. This suggests a number of relatively large flows from Saskatoon and these would have high variability (see TN 19) compared to small flows from small centers thus the homoscedasticity assumption is violated. It might have been interesting to calculate R^2 with and without Saskatoon included in the analysis.

One final question relates to the reliability of data based on a 56% voluntary return of completed questionnaires by visitors leaving parks. Whether this group is representative of the universe or of those given the questionnaires on entry to the park is unknown.

Many of the questions addressed to the first paper can obviously be repeated for "An Application of Mathematical Models to Compare Two Potential Park Sites". A very similar equation is calibrated, but no indication of the value of R^2 is possible. This is a little distressing in that the accuracy of the predictions for the future being made in this paper depend in large part on the accuracy of the model. However, even if R^2 were high, concerns with heteroscedasticity and structural validity of the model should lead to caution in treating modelling results as good approximations to what would occur.

The accuracy of prediction of the future use of parks depends first on the validity of the interaction model and the accurate estimation of its parameters. Secondly, it depends on how accurate estimates of future use levels of parks are. In this paper, the latter are estimated by extrapolating forward figures based on an average of recent years. One obvious question is whether present trends are likely to continue. This depends both on the future rate of population increase remaining steady, particularly amongst the age groups most active in park use, as well as on the future recreational -preferences and constraints of the population not changing. The stability of the latter in particular is questionable, although it is difficult to see how changing preferences or constraints and their effects can be reliably predicted. One possible approach, not mentioned by Cheung, would be to calculate whether the rate of change in park use has itself been changing over recent years. For example, a three-year running mean of the change rates in Table 4 might show whether or not the change rate was steady or not.

One final feature of the paper deserving mention is the treatment of site attractiveness. Although park attractiveness is not part of the model (for reasons explained by Cheung), a method is used to estimate the attractiveness of two existing sites and this is then used as a surrogate for the projected parks, prediction of whose future use is the goal of the paper. However, the grounds on which the parallel is drawn between the two existing and future sites are not indicated. In general, two alternative methods exist for estimating the attractiveness of an as yet non-existent park. One, (possibly rare) procedure would be to do as Cheung did and find an existing park whose characteristics were close to those of the proposed park, both in quantity and quality. The other park's attractiveness relates to a set of measurable park characteristics, so that only a knowledge of the latter quantities in any future park would be necessary to accurately define its attractiveness.

In "A Method for Predicting Enroute Overnight Park Use", a high R^2 is obtained after a logarithmic transformation of the data and is not a measure of goodness of fit between the untransformed predicted and observed data in Table 2. R^2 in the latter is 0.93. Also the percentage error in prediction of the three largest use figures is only about 8% whereas the mean percentage error for the rest of the data is 21%. This indicates that after logarithmic transformation, the three largest values were much more accurately predicted than the average, which suggests that even after logarithmic transformation the homoscedasticity assumption is still not met.

It is interesting to note that origin size, in the form of traffic volume on the nearest arterial highway, and distance to that highway explain relatively little variance in enroute camper volume compared to campsite capacity. It is rare indeed to find either the origin or destination mass term not contributing a large amount to the variance explained. However, the fact that distance contributed so little to the variance explained might provide a clue as to why highway traffic volume also explained very little, vis-à-vis campsite capacity. It is most improbable that distance, per se, has little effect on patronage: so many recreation studies indicate otherwise. However, if the range of campsite-to-highway distances is quite narrow it is possible that, since people are not infinitely sensitive to distance, the differences in distance are not large enough to have any appreciable effect on patronage. Likewise, if the range of highway traffic volumes (or at least camper traffic volumes) is small, they may give the appearance of having little influence on campsite attendance, even though a larger range of values might show them contributing appreciably to R^2. Since the original data are not provided in the paper, the latter must remain speculative. However, if the regression results are taken at face value, and there is assumed to be appreciable variation in highway traffic volumes and distances, the appropriate planning response in developing parks to cater to enroute campers would be to give little attention to arterial traffic volume and distance to highway and consider only park size. In other words parks could be located anywhere and only their size would be a consideration, if the results of the regression model were taken at face value. Such an interpretation is so seriously at odds with common sense that one is inclined at least to accept that either arterial traffic volume or distance or both are important, but that this study simply lacks a sufficient range on these variables to measure their effects. Certainly, the rarity of finding either "origin mass" or distance to be of relatively minor importance in explaining visitation levels is such that the findings should not readily be accepted without further enquiry.

Cesario's paper breaks new ground in trip distribution studies for reasons stated earlier in this review. The paper also should be seen as part of a continuing series of papers by Cesario on this topic (see 1971, 1973, 1974 and 1975). In this regard, it should be noted that his paper in the Journal of Regional Science, January, 1974, describes a solution to the problem of biased parameter estimates which he indicated to be in progress in his recommendation No. 4.

Basically Cesario's analysis of convariance model is used to estimate one set of parameters (E(i), i = 1, 2, . . . , n) which are unique to each origin and another set (A(j), j = 1, 2, m) which are unique to each destination. The major area of debate relates to the interpretation of these parameters. Cesario's opinion in this paper appears to be that "the emissiveness of population centre i, E(i), reflects its relative (to other population centers) propensity to emit trips under identical circumstances -that is, if it were the case that all centres were confronted by the same availability or "supply" of recreation opportunities". He provides no rationalization of this assertion. In contrast, TN 11 interprets E(i) more generally as a composite measure of several. factors unique to origin i, one of which is its unique spatial supply of park facilities. It is, of course, unique in that no other origin has the same spatial relationship to the set of parks under study. This latter, broader interpretation of E(i) would seem to be the more defensible one at the present stage of knowledge, until such time as empirical evidence is provided to show that the spatial supply effect is negligible.

Cesario's interpretation of attractiveness of park j, A(j), is that it "reflects its relative (to other parks) ability to attract trips under identical circumstances -e.g., if it were the case that all parks were equally accessible from each population centre". Since the effect of distance between i and j is separated out in the model, it is certainly acceptable to define "identical circumstances" as all parks being equally accessible from centre i. However, in a vein similar to the argument about emissiveness, the reader may be inclined to suggest that contained within A(j) is the competition which park j faces from the other parks. However, the competition which park j faces varies with the origin i from which it is being evaluated and is a constant for all parks as viewed from origin i.

Therefore the competition effect can properly be assigned to the origin rather than the destination. In other words, E(i) may contain a combined "supply of and competition of alternatives" effect. This appears most clearly in the MacDonald et al. paper (see TN 30) in their model where:

(1) $v(i,j)=-s(i)^{\alpha}P(i)^{\beta}(\Sigma_k (A(k)/C(i,k)^{\gamma}))^{\varphi}(A(j) /C(i,j)^{\gamma})/((\Sigma_k A(k))/C(i,j)^{\gamma})$

(all summations in Equations 1 through 3 are over k),

In this expression $(\Sigma_k (A(k)/C(i,k)^{\gamma}))^{\varphi}$ refers to the supply-generating-demand effect, while $1/((\Sigma_k A(k))/C(i,j)^{\gamma})$refers to j's competition from all parks as measured from centre i. Clearly both expressions are unique to origin i and therefore can be thought of as being contained in E(i) along with $S(i)^{\alpha}$ and $P(i)^{\beta}$, so that the above equation could be rewritten in a form similar to Cesario's (see Equation 2 below in this review) with E(i) defined as in Equation 3 of this review.

The suggestion by Cesario (unpublished) that k in his Equation 5 can be taken to implicitly refer to the "competing opportunity" effect is clearly incorrect, if the level of spatial competition varies with the location of the origin with respect to the competing parks. If the latter is true, then the competition effect varies with origin i, and is presumably contained in the estimate of E(i).

In the second stage of his analysis, Cesario seeks to explain the estimates obtained for all E(i) and A(j) values. A regression analysis shows 77% of the variance in E(i) values to be accounted for by origin population. The fact that origin mass terms are very important in "explaining" trip flows is well documented in many gravity model studies. Cesario suggests that so great is the contribution of population that it is not worth trying to refine the predictive equation by including other variables or using different functional forms. The counterargument is that so long as absolute flows are used as the dependent variable in a trip distribution study, origin mass, be it population, manufactured goods or whatever, will inevitable account for a large proportion of trip flow variance. The relation is almost as inevitable as the results of a regression study showing that the variance in the number of deaths in population centres is largely explained by their population size. Both beg a question. In the latter, the concern should presumably be to explain variance in the mortality rate of certain age groups rather than absolute mortality figures. And in trip distribution studies, the concern should surely be to explain variance in the per capita rate of flow between origins and destinations. Explaining absolute flows and emissiveness would appear in general to be a fairly trivial problem; explaining flow and emissiveness rates almost certainly is not. The dependent variable in this case would be V(i,j)/P(i). The more subtle possible influences of socioeconomic variables which are dwarfed in the present study by the effect of P(i) on V(ij) and E(i) would have a chance to be revealed, if indeed they have an effect on the emissivity rate. Another independent variable which might well be included in an analysis of emissivity rates is:

$(\Sigma_{k}.(A(k)/C(i,k)^{\gamma}))^{\varphi-1}$

which incorporates both the origin specific supply-generating demand and competition effects defined in Equation 1 above.

Despite the above comments, the paper marks a significant methodological advance in the trip distribution literature. It was the first to estimate values of A(i) and E(j) independent of ad hoc assumptions about relevant variables. And from it useful progress in trip distribution analysis has been made, as exemplified by the two remaining papers in this section by MacDonald et al. and Beaman et al., to which we now turn.

"A Work Plan for the Development of a Mathematical Model to Predict and Explain Overnight Use of Parks" by MacDonald et al. is generically similar to the model discussed in the last paper of this section by Beaman et al.. Both are significant extensions of the model discussed in Cesario's paper, although estimation of parameters depends on an initial estimation of attractiveness and emissivity parameters using the analysis of covariance method described in Cesario's paper. Although parameter estimation is not discussed in the paper by MacDonald et al., Equation 9 could be rewritten as:

(2) $V(i,j) = E(i)A(j) / C(i,j)^{\gamma}$

WHERE E(i), i = 1, 2, . . . , n and A(j), j = 1, 2, . . . , m and are parameters to be empirically estimated by the method discussed in Cesario's paper, and E(i) is defined as

(3) $E(i)=\boldsymbol{\kappa}S(i)^{\alpha} P(i)^{\beta} (\Sigma_k (A(k)/C(i,k)^{\gamma}))^{\varphi-1}$

Given estimates of E(i) and A(j), the latter equation can be solved for values of α, β, γ and φ.

Unlike the Cheung models where attractiveness is estimated by a predetermined function of specified site variables, attractiveness in the MacDonald et al, Cesario and Beaman et al. papers, it is empirically estimable from trip distribution data. In addition, the distance exponent is estimated by a least squares minimizing routine, rather than by the trial and error method used by Cheung.

MacDonald et al. include in their model 1 a supply-generating-demand effect defined as a non-linear function of the sum of all parks' attractiveness/distance ratios. In model 2, this term is replaced by one defining the probability of a potential camper from i visiting any of the parks at all. The theoretical underpinnings and derivation of this probability term are much stronger than for the equivalent term in model 1. However, the derivation of this probability relies on the rather strong, and in this case, doubtful assumption that the probability of visiting any particular park from origin i is independent of the probability of visiting any other park. It would seem more likely that the higher the probability of one park being visited, the lower the probability of another being visited, in which case the probabilities would not be independent (but see TN 19). The consequences of not conforming to this assumption deserves to be explored before the “term” is used to define one of the terms of model 2.

With specific regard to the enroute model, which is a special case of the main destination model, the discussion does not make clear what the definition is of alternative parks for an enroute stop. If the definition were all parks as viewed from a particular link on the highway network, C(i,j) would only reasonably be defined as the distance between park j and the nearest point on an arterial highway for parks on that link. For all other parks, C(i,j) would be the arterial distance plus side-road distance between link i and park j. If only parks adjacent to link i are considered as alternatives, then the definition of C(i,j) given in the paper is reasonable. It also bears emphasis that the enroute camper model assumes that on all links the percentage of campers who are looking for an overnight stopping place is a constant.

An issue raised indirectly in the MacDonald et al. paper is that no model exists to predict touring camper choices. It seems probable that many vacation campers have regions rather than sites as destinations, and although the task of modelling spatial choice that involves (first) a region and (second) a route and stopping places may seem formidable, the prediction of park use levels may depend on such a model to an appreciable extent.

Many of the features of the Beaman et al. model have already been commented on with respect to the paper by MacDonald et al. The basic issue involved in the Beaman et al. model is to define a model which is valid for any number of destinations, including a single destination. At the same time the model requires the two features discussed earlier, the supply-generating-demand effect of destinations and their competition effect. Beaman et al. try to satisfy these conditions in Equation 3 which simplifies to Equation 2 in the special case of the single destination, or isolated site as it is described in the paper. The expressions in Equation 3 that may be least familiar to the reader are also those that simplify to the value 1 when there is only one destination. Specifically they are, $PP(o)/\Sigma P(o,d)$ which refers to the "competition from alternatives" effect and $(\Sigma(P(o,d)/PP(o))^{SE})^{SM}$, which refers to the "supply generating demand" effect. The justification for the first expression, which is not completely clear from the text, is that the denominator is a measure of competition defined in terms of the sum of all sites, attractiveness/distance ratios, identical to the denominator of Equation 9 and 10 of MacDonald et al.'s paper. The numerator which is simply a constant for any one origin is required for the isolated site case in which the numerator and denominator become identical and cancel each other. The justification for the second expression is as follows. The exponent SM is required to define the positive non-linear effect of supply on use, but must be less than or equal to 1, to satisfy an earlier condition in the paper that with more than one site, total visitation will exceed visitation to a single one of these sites but will be less than the sum of visitations to each site, if each were the only site.

In passing, the latter part of this condition seems unnecessarily rigid. It seems possible that the sum of visits to, say, two sites might exceed the visitation they would receive if each were an isolated site. In any event, relaxation of the assumption would simply entail setting no limits on the value of SM. SE can be interpreted behaviorally as implying that sites with a smaller attractiveness/ distance ratio contribute even less to the supply-generating-demand effect than their relative size would suggest. Once again this assumption could be readily relaxed without any change to the structure of the expression. The reason for PP(o) in the denominator is to ensure that one of the terms in the summation will always equal l, so that the summation will always be greater than 1 for the case of two or more sites, and equal 1 exactly for the case of the isolated site, no matter what values SE and SM take on. This satisfies the basic assumption that there will be extra demand stimulated by the provision of an extra site, as well as the assumption that the supply generating demand effect is unity in the case of a single site.

Besides the question of solubility (which is not at issue in this paper) the main question that arises is whether or not a more elegant, less complex formulation could achieve the same effect as this model. Certain elements, such as PP(o) in the numerator of the first expression discussed above are disturbing. In the case of PP(o), the effect is to add a different constant to the model for each origin. Without some behavioral interpretation of the role of this term in the general model, its only function seems to be to satisfy conditions for the isolated site model in Equation 2.

However, it is not clear why a simple variant of the more straightforward model, proposed by MacDonald et al. (Equation 9 in their paper) would not more simply accomplish all that TN 11 Equation 3 does. Specifically, using TN 11 notation:

(4) $V(o,d)=C\,Pop(o)G(o)(\Sigma_d P(o,d))^{\varphi}\, A(d)^{[-R(D(o,d))]}) /(\Sigma_d A(d)^{[-R(D(o,d))]})$

will in the case of the isolated site simplify to:

(5) $V(o,d)=C\,Pop(o)G(o)A(d)^{[-R(D(o,d))]}P(o,d)^{\varphi}$

The difference is that in the isolated site case the supply generating demand effect will vary nonlinearly with the value of that site's attractiveness/distance ratio. The two components of Equation 4 can be defined as:

(6) $NN(o)=Pop(o)G(o)(\Sigma P(o,d))^{\varphi}$

WHERE NN(o) = the number of actual park visitors generated by origin o; and

(7) $P(o,d)=A(d)^{[-R(D(o,d))]}/(\Sigma_d A(d)^{[-R(D(o,d))]})$

WHERE P(o,d) = the probability of a park visitor from o choosing to visit d from the set of available alternative parks; and by definition the Σ_o p(o,d) = 1. In the isolated case, p(o,d) = 1, i.e. all park visitors from o visit d. Therefore, in words, the model estimates the number of park patrons generated from o as a function of o's characteristics as well as of its overall access to parks and then assigns each park d a proportion of o's total visitors based on that park's proportion of the total sum of parks' attractiveness/ distance ratios.

In any case, the implication is that estimation problems can be overcome, both the Beaman et al. and MacDonald et al. models constitute important efforts to produce a more general and realistic trip distribution model.