# Parameter Calibration in Competing Destination Model
# –Using Generalized Additive Model with Golden Section Search

Final Project Paper of GEOG 788P

## 1. Goal Statement

Competing destination (CD) model is an extension of the spatial interaction (SI) model. By introducing the measure of accessibility of destinations, the CD model could alleviate the parameter misspecification in SI models due to the effect of destinations' spatial structure. A lot of methods have been studied to calibrate the CD models and this paper proposed a method combining the generalized additive model with golden section search.

So far, there are no available modules in python for the parameter calibration in the CD model and the main goal in this paper is to try to fill the gap in the python library and use a small dataset to validate the methods.

## 2. Background

### 2.1. Competing Destination Model

The spatial interaction (SI) model is a common modeling method to analyze and simulate various types of flow data. Among them, the gravity model is the most representative, and it is widely used in the study of flow data such as population migration flow(Karemera, Oguledo & Davis, 2000), passenger flow(Tobias, Franz & Armin, 2007), trade flow(Celine, 2006), and information flow(Farmer & Oshan, 2017). The flow data reflect the intensity of interactions between cities or regions, and the analysis of flow data can explore the interrelationships between cities or regions(Roy, 2004), which further strengthens the application value of spatial interaction models.

However, the traditional gravity model does not consider the influence of spatial structure of the destinations. The independence of irrelevant choices (IIA) assumption is implicit in the conventional SI models, that is, that people take into account only the characteristics and spatial resistance of the destination itself in their trip decisions, and do not consider the interactions between destinations or spatial flows. Studies (Fotheringham, 1983; Hu & Poolor, 2002 ) have shown that spatial structure has a significant impact on spatial flows and traditional SI models ignoring this issue are bound to suffer from bias due to omission of variables, resulting in misspecification of model parameters and significant spatial autocorrelation of parameters.

A number of studies have modified the conventional SI models to alleviate or eliminate the effects of spatial structure, and the competing destination (CD) model is popular (Poolor, 1998). This model introduces the effect of competition among destinations to the spatial interaction model, proposed by Fotheringham in the 1980s. It suggests that instead of considering all potential destinations as in traditional SI models, people follow a hierarchical destination choice process in which one simplifies the vast destination information by firstly selecting a large area as the destinations range and then choosing a specific destination from the selected large region.(Fotheringham, 1983)

In Fotheringham's hierarchical destination selection mechanism, a two-stage destination selection process is constructed [16], in which a set *c* containing multiple destinations is selected in the first stage with utility $U_c$ , and a specific destination *j* is selected from this set in the second stage with utility $U_j$. The relationship between the utilities of the two stages can be expressed as

$$U_c = \left( \sum_{j \in c} U_j \right)^{\delta} \quad (1)$$

In Equation (1), $U_c$ denotes the utility obtained by selecting destination set *c* in the first stage; $U_j$ denotes the utility obtained by selecting destination j from set c in the second stage; δ is a parameter indicating the existence of destinations' competition effect. When δ = 1, the utility of selecting destination set c is equal to the direct sum of the utilities of the destinations in c, i.e. there is no competition effect, which is the case of the conventional SI model. When $\delta < 1$, the utility of the set of selected destinations c is less than the sum of the utilities of the destinations in the selected set, i.e., being in a larger set of destinations is relatively disadvantageous to the destinations, so there is a competition effect among the destinations. When $\delta > 1$, the utility of the selected set of destinations c is greater than the sum of the utilities of the destinations in the selected set, i.e., being in the larger set of destinations is relatively advantageous for the destinations, so there is an agglomeration effect among destinations.(Fotheringham, 1983)

The CD model adds to the traditional origin-specific SI model an indicator of spatial accessibility or proximity of each destination to other destinations as a representation of spatial structure. The mathematical form is

$$T_{ij} = Z_i O_i M_j^{\alpha} d_{ij}^{\beta} A_{ij}^{\delta} \tag{2}$$

In Equation (2), $T_{ij}$ denotes the flow from origin *i* to destination *j*; $O_i$ denotes the total flow from *i*; $M_j$ is the attractiveness of destination *j*, usually measured by demographic, economic variables; $d_{ij}$ is the distance from *i* to *j* which could be a straight-line distance, miles traveled, or travel time; $A_{ij}$ denotes the accessibility of destination *j* with respect to origin *i* by comparing to alternative destinations *k*, which is expressed as

$$A_{ij} = \sum_{k \neq j}^{n} \frac{M_k}{d_{kj}^{\sigma}} \tag{3}$$

In Equation (3), $M_k$ denotes the quality of destination *k*; $\sigma$ measures the relevance of distance in capturing people's perception of accessibility and *n* is the total number of origins. Considering that the population size of a city is a comprehensive reflection of its multiple functions and competitiveness, the total population of the city is usually used as the quality indicator; $d_{kj}$ denotes the distance from *k* to *j*; αis the parameter representing the effect of the attributes of *j* on flows. βis the distance-decay parameter representing the effect of movement costs on flows and often is negative since the distance is usually regarded as a deterrent to interaction. $Z_i$ is a balancing factor to ensure that the sum of the flows $T_{ij}$ from each origin is equal to $O_i$, and can be expressed as

$$Z_i = \left( \sum_j^n M_j^{\alpha} d_{ij}^{\beta} A_{ij}^{\delta} \right)^{-1}, \tag{4}$$

where *n* is the number of destinations, all other notations are the same as those in Equation (2) and (3).

**2.2. Parameter Calibration**

The linear regression method and linear programming method are the conventional algorithms for solving gravity models, and the basic idea is to transform a gravity model into an optimization problem of a linear system. Specifically, by taking the natural logarithm of both sides of Equation (2), it is possible to obtain the so-called log-linear CD model, which could be expressed as,

$$ln\left(T_{ij}\right) = k + ln\left(O_i\right) + \alpha ln\left(M_j\right) - \beta ln\left(d_{ij}\right) - \delta ln\left(A_{ij}\right)$$

(5)

In Equation(5), the intercept *k* has the same function as the balancing factor $Z_i$in Equation(2). And the remaining notations are the same as in Equation(2) and (3). Note that the signs before the components $\beta ln\left(d_{ij}\right)$ and $\delta ln\left(A_{ij}\right)$are intended to be negative in order to reflect the underlying hypothesis that spatial interactions decrease with higher costs or distances and competitive effect is usually dominant in relationships between destinations.

However, some limits of the log-normal gravity model are discovered, which would create difficulties with parameter calibrations. These limitations are summarized further below (Oshan, 2016).

1. flows are often counts of people or objects and should be modeled as discrete entities;
2. flows are often not normally distributed;
3. logarithm estimates of actual flows leads to downward skewed flow predictions;
4. zero flows are problematic since the logarithm of zero is undefined.

Flowerdew and Aitkin (1982) have claimed that many of the difficulties listed above are caused by an improper specification of the model itself, and that poisson regression would be a more acceptable model to use in these situations. Specifically, a poisson distributed process with mean $\lambda_{ij}$ could be used to estimate the number of agents travelling from origin *i* to destination *j* when recognizing that

a) the number of agents moving from *i* to *j* must be a non-negative integer;

b) the population in *i* is relatively large;

c) the movement of agents is independent; and

d) assuming that there is a constant probability of an agent in *i* moving to *j*.

$\lambda_{ij}$ is unknown and can be estimated via a SI model, where it is now assumed that $\lambda_{ij}$ is logarithmically linked to a linear combination of the logarithmical independent variables in Equation (5), which is shown below:

$$ln(\lambda_{ij}) = k + ln(O_i) + \alpha ln(M_j) - \beta ln(d_{ij}) - \delta ln(A_{ij}), \tag{6}$$

and by exponentiating both sides of the equation it becomes the poisson log-linear CD model,

$$T_{ij} = exp\left(k + ln\left(O_i\right) + \alpha ln\left(M_j\right) - \beta ln\left(d_{ij}\right) - \delta ln\left(A_{ij}\right)\right),$$

(7)

where $k$ is the estimated intercept and must be included to ensure the total number of flows is conserved.

Calibration of poisson regression can be carried out within a generalized additive modeling framework (GAM), which will be further discussed in the 3.1 section.

As for the value of parameter σ in $A_{ij}$, in most empirical studies, it is generally believed that the fit is better when σ takes a value of 1. However, ideally, σ should be estimated iteratively with the parameters $\alpha$, $\beta$, $\delta$, and $k$. Based on Fotheringham (1983), for the estimation of σ, σ is initially set equal to some arbitrary value and the set of $A_{ij}(\sigma)$ is calculated. A set of parameters ($\alpha$, $\beta$, $\delta$, and $k$) is obtained from the calibration of the CD model and this set is used to define a new set of destination accessibility which in turn is used to recalibrate the model. The process continues until the parameters converge. This project will follow the framework above to calibrate σ, which will be further discussed in the 3.2 and 3.3 section.

## 3. Method

### 3.1. Generalized Additive Model

#### 3.1.1. What is GAM

Generalized additive models (GAM) were proposed by Trevor Hastie and Robert Tibshirani in 1986. A simple and attractive mental model forms the basis for the GAM framework's success, that is, individual predictor-dependent variable relationships follow smooth patterns that could be linear or nonlinear. By simply adding these smooth functions up, we can estimate them simultaneously and then predict $g(E(Y))$.

In general the model has a structure like:

$$g(E(Y)) = \alpha + s_1(x_1) + \ldots + s_p(x_p),$$

(8)

where $Y$ is the dependent variable, $E(Y)$ denotes the expected value, and $g(Y)$ denotes the link function that links the expected value to the predictor variables $x_1, \ldots, x_p$.

The terms $s_1(x_1), \ldots, (x_p)$ denote smooth, nonparametric functions. In contrast to parametric functions, which are defined by a restricted number of parameters, nonparametric predictor functions are totally determined by the data. This allows for a more flexible estimate of the underlying predictive patterns without prior knowledge. Note that GAMs can also contain parametric terms as well as two-dimensional smoothers. Moreover, like generalized linear models (GLM), GAM supports multiple link functions such as normal, binomial, gamma, poisson, and inverse link, etc.

### 3.1.2. Local Scoring Algorithm

Hastie and Tibshirani (1986) introduced the local scoring algorithm as a way for estimating GAM. In the local scoring algorithm, the estimation technique for GAM consists of two loops. A weighted backfitting algorithm (inner loop) is utilized within each step of the local scoring procedure (outer loop) until convergence or the residual sum of squares (RSS) fails to decrease. After that, a new set of weights is created based on the estimates from the weighted backfitting process, and the scoring system's next iteration begins. When the convergence requirement is met or the deviation of the estimates stops falling, the scoring system comes to a halt.

GAM is a generalization of GLM in the same way that additive models are a generalization of linear regression models. Since the algorithm for additive models is the basis for fitting generalized additive models, the algorithm for additive models is discussed first. Numerous approaches to the formulation and estimation of additive models exist. Backfitting is a generic approach for fitting an additive model using any regression-type fitting technique. The following is the unweighted version of the backfitting method:

1. Initialization: $s_0 = E(Y),\ s_1^{(1)} = s_2^{(1)} = \ldots = s_p^{(1)}, m = 0$

2. Iterate: $m = m + 1;\ for\ j = 1\ to\ p\ do$:

$$R_j = Y - s_0 - \sum_{k=1}^{j-1} s_k^{(m)}\left(X_k\right) - \sum_{k=j+1}^{p} s_k^{(m-1)}\left(X_k\right)$$

$$s_j^{(m)} = E\left(R_j | X_j\right)$$

3. Until: $\frac{\sum_{i=1}^{n}\sum_{j=1}^{p}\left(s_j^{(m-1)}\left(x_{ij}\right) - s_j^{(m)}\left(x_{ij}\right)\right)^2}{1+\sum_{i=1}^{n}\sum_{j=1}^{p}\left(s_j^{(m-1)}\left(x_{ij}\right)\right)^2} \leq \epsilon\ (where\ \epsilon\ is\ 10^{-8}\ by\ default).$

In the preceding notation, $s_j^{(m)}(\ )$ denotes the estimate of $s_j(\ )$ at the *m*-th iteration; p is the total number of the predictors; *Y* is the response variables, $E(Y)$ is the expected value of *Y.*

Since the CD model is calibrated by poisson regression, for the backfitting loop in the local scoring algorithm, a weighted backfitting algorithm will be applied, which has the same form as for the unweighted case above, except that the smoothers are weighted. Important components of a weighted smoother for poisson distribution in local scoring algorithm is shown below:

| Distribution | Link Function | Adjusted Dependent (*z*) | Weight (*w*) |
|---|---|---|---|
| Poisson | $log(\mu)$ | $\eta + (y - \mu)/\mu$ | $\mu$ |

$\mu\ is\ the\ mean\ value,\ \eta\ is\ the\ predictor$

Table 1. Link functions, Weights and Adjusted dependents for Poisson Regression

As previously stated, the local score algorithm consists of two loops: an inner loop (weighted backfitting method) and an outer loop, which is the local scoring procedure. We have reviewed the steps involved in the backfitting algorithm so far; now it is time to present the process of local scoring, which is shown below:

1. Initialization: $s_i = g(E(Y)),\ s_1^{\ 0} = s_2^{\ 0} = \ldots = s_p^{\ 0}, m = 0$

2. Iterate: $m + m + 1$;For a poisson regression, form μ and *w* and *z* based on the equations in Table 1 with their corresponding values from the previous iteration:

$$\eta_i^{(m-1)} = s_0 + \sum_{j=1}^{p} s_j^{(m-1)}\left(x_{ij}\right)$$

$$\mu_i^{(m-1)} = g^{-1}\left(\eta_i^{(m-1)}\right) = the\ mean\ of\ \eta_i^{(m-1)}$$

$$w_i = \mu_i^{(m-1)}$$

$$z_i = \eta_i^{(m-1)} + \left(y_i - \mu_i^{(m-1)}\right) \cdot \left(\frac{\partial\eta}{\partial\mu}\right)_i^{(m-1)} = \eta_i^{(m-1)} + \left(y_i - \mu_i^{(m-1)}\right)/\mu_i^{(m-1)}$$

Then, fit a additive model to $z$by using the backfitting algorithm we mentioned before with weights $w$to obtain the estimated smooth functions $s_j^{(m)}(\ ),\ j = 1, \ldots, p$;

3. Until: $\frac{\sum_{i=1}^{n} w_i \sum_{j=1}^{p}\left(s_j^{(m-1)}\left(x_{ij}\right) - s_j^{(m)}\left(x_{ij}\right)\right)^2}{\sum_{i=1}^{n} w_i\left(1 + \sum_{j=1}^{p}\left(s_j^{(m-1)}\left(x_{ij}\right)\right)^2\right)} \le \epsilon^s\ (where\ \epsilon^s\ is\ 10^{-8}\ by\ default)$

### 3.1.3. Golden Section Search and Smooth Function for Spatial Accessibility

This paper will estimate σin Equation (3) iteratively with the parameters α, β, δ, and *k* in the GAM framework. In this case, we need to find an optimal value of σ in each iteration of the backfitting loop in the local scoring algorithm. In this case, the ideal value of σshould come with the minimized Akaike Information Criterion(AIC) of the model.

The AIC (Akaike, 1974) takes into account the different number of degrees of freedom in different models so that their relative performances can be compared more accurately. A model with a lower AIC than another is held to be a 'better' model. And we will use golden section search to find the optimal value of σwith the lowest value of the AIC. The process of golden section search is shown below (Fothringham, 2002):

1. Initialization: $a < b < c,\ and\ calculate\ f(a),\ f(b),\ and\ f(c).\ f()$denotes the function to calculate AIC value in this case.

2. Iterate: choose a new value $d$ between $a$ and $b$ or between $b$ and $c$

    calculate$f(d)$

    if $f(b) < f(d),\ a = a, b = b, c = d$

    if $f(b) > f(d),\ a = b, b = d, c = c$

3.Until: two successive values of $f(d)$are almost the same.

To sum up, in order to fit the GAM model and estimate the parameters (α, β, δ, and $k$) of Equation (6), Hastie and Tibshirani's local scoring algorithm will be implemented, where the optimal value of parameter (σ) in Equation (3) will be obtained by the golden section search method in each iteration in the backfitting loop.

**3.2. Validation**

The CD model can be derived from either the unconstrained model of the spatial interaction model or the production-constrained (PC) model (Fotheringham, 1983), and the latter usually performs better (Hu & Poolor, 2002). Therefore, most of the empirical studies on the CD model compare the latter and the models that the CD model derived from( Ishikawa, 1987; Fotheringham, 2001; Poolor, 2002). Therefore, this paper will compare the model fit indicators of the CD model with those of the unconstrained SI model and the PC model. Two measures will be used and further discussed below.

**Percent Misallocated (PM)**

The mathematical represent of PM is shown below:

$$PM = \frac{50}{T}\sum_i\sum_j\left|\widehat{T_{ij}} - T_{ij}\right|, \tag{9}$$

where $T_{ij}$ is the number of actual flows from origin *i* to destination $j$; $\widehat{T_{ij}}$ is the number of estimated flows from origin *i* to destination $j$. $T$is the total number of all actual flows. The smaller PM value means a better fit of the model.(Oshan, 2016)

**Sorensen Similarity Index (SSI)**

The mathematical represent of SSI is shown below:

$$SSI = \frac{1}{(nm)}\sum_i\sum_j\frac{2min\left(T_{ij}, \widehat{T_{ij}}\right)}{T_{ij}+\widehat{T_{ij}}}, \tag{10}$$

where *n* and *m* is the total number of origins and destinations; $min()$is the function to compare and remain the smaller value between $T_{ij}$ and $\widehat{T_{ij}}$. All other notations are the same as those in Equation (9). The reason why choosing SSI as an indicator of model fit is that it is more and more popular in recent SI model studies, especially in the non-parametric models. The SSI value is confined between zero and one, with values closer to one suggesting a better fit to the model.(Oshan, 2016)

**4. Data Source**

The flow data used here is a small toy dataset, which is previously used to demonstrate spatial interaction modeling in the R programming language (Dennett 2012). The data are migration flows between Austrian NUTS level 2 regions in 2006, which has 81 flows total and includes the name of origins and destinations, the total number of spatial flows, the attractiveness of destinations and the distance between the origin and destination in each flow.

At first, this project planed to use a county-to-county migration flows dataset during the period 2017-2019 to do an empirical study by using the proposed model, however, it took a long time for the writer to figure out the principles behind CD model, GAM and how to combine them. Thus, the remaining time for the empirical study is insufficient since it is time-consuming to use a new un-tidied dataset for the model evaluation and it is computer-inefficient to use a large dataset. Additionally, it is a methodology paper focusing on building models, therefore, this paper used this example dataset for testing the proposed model instead of the county migration dataset, which will be applied in the future.

Figure 1 is the boundary of the toy dataset.

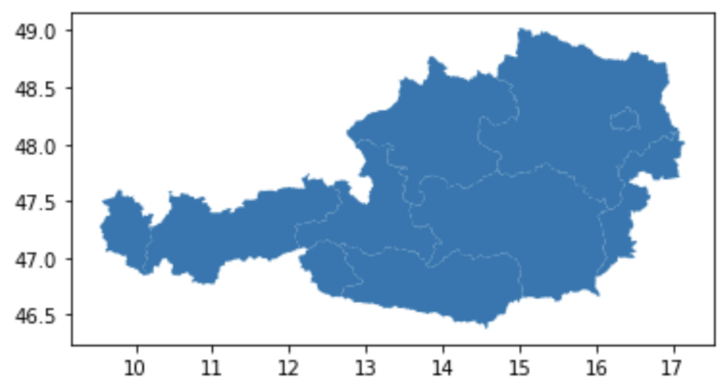


Figure 1. The boundary of Austria

**5. Results**

First, we compared the estimated parameters of the unconstrained model, the PC model and the CD model, which is shown in Table 2. From table 2, we could see that the estimates of the distance decay parameter βbecome less negative from -1.155 to -1.011when comparing the PC model with the CD Model with the intercept. The similar result was found when comparing the unconstrained model with the CD Model with the intercept (from -1.059 to -1.011). This change indicates that, in the case of this study, when the accessibility variable is included into the PC model, the effect of distance impedance on migration appears to be decreasing.

Comparing the CD model with and without the intercept, we could find that the σ value dramatically changed, from 2.69 to -2.69, which leads to two totally different indication towards the relationship between destinations because the parameter δ also changed dramatically (from 0.240 to -0.129). As we discussed before, δ is negative when there is a strong competition effect while an agglomeration effect is dominant as δis positive.

| Parameters | Unconstrained Model | PC Model | CD Model with an intercept | CD Model without an intercept |
|---|---|---|---|---|
| $k$ | -0.858 | 5.031 | -8.364 | — |
| α | 0.738 | 0.737 | 0.821 | 0.775 |
| β | -1.059 | -1.155 | -1.011 | -1.055 |
| δ | — | — | 0.240 | -0.129 |
| σ | — | — | 2.697 | -2.697 |

Table 2. The estimates of parameters of the four models

Then, as indicated in Table 3, we compared the PM and SSI values. According to its PM and SSI values, the PC model has the greatest goodness-of-fit. The unconstrained model performs the poorest, with the greatest PM value and the lowest SSI value. When compared to the traditional unconstrained model, the PC and CD models do reduce estimate bias. However, after including accessibility into the CD model, it seemed as if the parameter misspecification remained in this case study.

| Model Fit Indicator | Unconstrained Model | PC Model | CD Model with an intercept | CD Model without an intercept |
|---|---|---|---|---|
| PM | 13.594 | 11.455 | 13.536 | 13.016 |
| SSI | 0.751 | 0.779 | 0.761 | 0.767 |

Table 3. The model fit statistics of of the four models

## 6. Discussion

Obviously, the proposed approach for estimating the CD model requires modification, since the findings of the estimation are inverse to the implication in CD model theory. For instance, although the CD model should have a higher model fit than the PC model, the PC model has the best prediction performance in this research. Furthermore, the estimates of σand δ are different between the CD model with and without an intercept. Given the importance of the intercept in the CD model, this might be one explanation for the discrepancy in the findings. However, this might be due to the suggested approach itself.

The guess for the reason why the issues above happened is because the codes for the proposed method (the local scoring algorithm with golden section search) are not exactly correct. To validate that, taking the estimate of parameters of the CD model without the intercept as the control, this paper first use the estimate σvalue (-2.679) to calculator the accessibility $A_{ij}$ in Equation (3) and then fit a possession regression by using other available functions in python and obtain a set of estimate of the parameters $\alpha,\ \beta,\ and\ \delta$. The results is shown below:

| Parameters | CD Model fitted by the proposed method in this paper | CD Model fitted by the available function in python |
|---|---|---|
| α | 0.775 | 0.775 |
| β | -1.055 | -1.055 |
| δ | -0.129 | -0.129 |

Table 4. The estimate of CD Model fitted by different modules

Table 4 shows the same estimates of parameters, which indicated the codes for the local scoring algorithm of the Poisson GAM work. However, when setting σvalue as another value, let's say -1, using the set of parameters in Table 4 to create a new set of response variables and fitting the new response variables and the same predictors into the proposed model, the new estimates of parameters are not the same. In fact, no matter which value is set to be the σ, the proposed method would give the almost same set of parameters as the set in Table 4 with different estimates of σ. The result is shown below:

| Parameters | CD Model fitted by the proposed method in this paper | $true\,\sigma = -1$ | $true\,\sigma = -2$ | $true\,\sigma = -3$ |
|---|---|---|---|---|
| α | 0.775 | 0.740 | 0.762 | 0.651 |
| β | -1.055 | -1.132 | -1.089 | -1.209 |
| δ | -0.129 | -0.010 | -0.087 | 0.067 |
| σ | — | 4 | -2.697 | 2.697 |

Table 5. The estimate of CD Model setting different σvalues and response variables for validation

Therefore, further modifications to the code are required to implement the local scoring algorithms using golden section search.

## References

Akaike H (1974) A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19[6]: 716–723

Carrere, Celine, (2006), Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model, European Economic Review, 50, issue 2, p. 223-247.

David Karemera & Victor Iwuagwu Oguledo & Bobby Davis, 2000. "A gravity model analysis of international migration to North America," Applied Economics, Taylor & Francis Journals, vol. 32(13), pages 1745-1755.

Dennett A (2012) Estimating flows between geographical locations:‘get me started in spatial interaction modelling. Working Paper 184, Citeseer, UCL

Farmer, C. and Oshan, T. (2017). Spatial interaction. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed.).

Flowerdew, R., & Aitkin, M. (1982). A method of fitting the gravity model based on the Poisson distribution. Journal of Regional Science, 22(2),191-202.

Fotheringham, AS, Brunsdon, C, Charlton, M (2002): Geographically Weighted Regression: The Analysis of Spatially Varying Relationships, Wiley, ISBN: 978-0-471-49616-8

Fotheringham, A. S. (1981). Spatial structure and distance-decay parameters. Annals of the Association of American Geographers, 71(3), 425-436

Fotheringham AS (1983) A new set of spatial-interaction models: the theory of competing destinations. Environment and Planning A 15[1]: 15–36.

Fotheringham AS. Some Theoretical Aspects of Destination Choice and Their Relevance to Production-Constrained Gravity Models. Environment and Planning A: Economy and Space. 1983;15(8):1121-1132.

Fotheringham A S, Nakaya T, Yano K, et al. Hierarchical destination choice and spatial interaction modelling:a simulation experiment[J]. Environment and Planning A, 2001, 33(5): 901-920.

Fotheringham AS. Spatial Flows and Spatial Patterns. Environment and Planning A: Economy and Space. 1984;16(4):529-543. doi:10.1068/a160529

Grosche, Tobias & Rothlauf, Franz & Heinzl, Armin, 2007. "Gravity models for airline passenger volume estimation," Journal of Air Transport Management, Elsevier, vol. 13(4), pages 175-183.

Hastie, Trevor and Tibshirani, Robert. (1990), Generalized Additive Models, New York: Chapman and Hall.

Hastie, Trevor and Tibshirani, Robert. (1986), Generalized Additive Models, Statistical Science, Vol. 1, No 3, 297-318.

Hu P, Pooler J. An empirical test of the competing destinations model[J]. Journal of Geographical Systems, 2002, 4(3): 301-323

Ishikawa Y. An empirical study of the competing destinations model using Japanese interaction data[J]. Environment and Planning A, 1987, 19(10): 1359-1373.

Pooler J. Competition among destinations in spatial interaction models:a new point of view[J]. Chinese Geographical Science, 1998, 8(3): 212-224

John R. Roy, 2004. "Spatial Interaction Modelling," Advances in Spatial Science, Springer, number 978-3-540-24807-1, July.

McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed), Frontiers in Econometrics. Academic Press, New York, 105–142

Oshan, Taylor M, (2016), A primer for working with the Spatial Interaction modeling (SpInt) module in the python spatial analysis library (PySAL), REGION, 3, issue , p. R11-R23.