# Cumulative attraction and spatial dependence in a destination choice model for beach recreation

## Abstract

Beach recreation value is an important consideration in a cost-benefit analysis of coastal development or conservation. A destination choice-based travel cost analysis is often used to quantify recreation values but the destination choice only partially reflects the intrinsic characteristics of that site. Visitors are influenced by opportunities available at other sites and can visit multiple sites resulting in spatially correlated errors. For this study about the recreation value of beaches on the Coromandel Peninsula we draw on the theory of cumulative attraction to analyse the compatibility of different beaches for the multiple-destination visitors who comprise more than half our sample. We investigate different random utility model formulations to explain destination choice and find that a cross-nested logit with sites nested by availability of amenities explains the observed patterns of visitation well and is more computationally efficient that non-closed-form models such as mixed logit. We also include inverse distance variables to the nearest amenity of each type and their significance supports the tenet of cumulative attraction that the importance of other spaces is greater when the attributes differ. Overall beach recreation values are maximised when sites are diverse in terms of development level and type.

## 1. Introduction

Destination choices of individual recreationists collectively determine the demand for beach recreation and the potential welfare effect they experience from coastal development or from changes in environmental quality.  A typical approach to modelling recreation demand is to use a random utility model (RUM) which allows the estimation of demand for multiple sites, substitution across sites, and is consistent with utility maximisation theory (Fletcher, Adamowicz, and Graham Tomasi 1990). The problem with traditional destination choice or demand models is that they fail to account for the effects of the spatial distribution of the recreation sites. The first law of geography is that "everything is related to everything else, but near things are more related than distant things" (Tobler, 1970).  Recreation sites do not exist in isolation; tourist flows are enhanced or diminished by attractiveness of neighbouring destinations (Griffith 2007).

Spatial effects can include correlated errors due to unobserved attributes that nearby destinations share. When there are a large number of alternatives destination sites it is likely that tourists use hierarchical processing to simplify their decisions, and evaluate sites in geographic clusters (Pellegrini and Fotheringham 2002). Sites in these clusters likely have common attributes. There can also be spatial knowledge spillovers, where tourists incidentally acquire information about places near their destination of choice. Reduced uncertainty about these sites increases their likelihood of being destinations for future visits (Marrocu and Paci 2013).

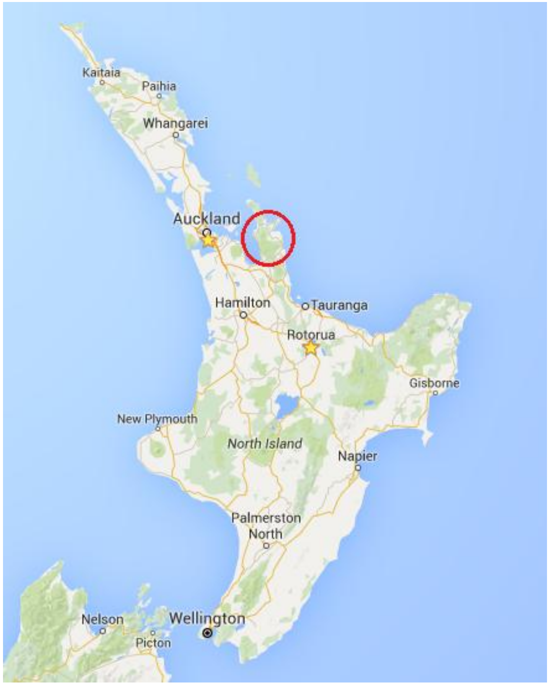
Another reason for spatial correlation is that recreation trips often include multiple destinations in order to satisfy a diverse range of objectives and reduce the risk of unrealised expected benefits (Lue, Crompton, and Stewart 1996). In the context of this study, more than half of sampled visitors to the study area reported visiting more than one beach site per trip. The theory of cumulative attraction (Nelson 1958) recognises this effect, and implies that areas with multiple, differentiated destinations will attract more tourists than single destinations – an agglomeration effect. There may alternatively be competition effects when similar destinations are located near each other. Which effect prevails and for whom remains an empirical question.

Spatially correlated errors violate the assumption of the travel cost method that sites must be substitutes. The value of a recreational site is likely to be overstated if some of the benefits of visiting other sites during a trip are "mistakenly" attributed to the study site (Haspel and Johnson, 1982). When sites share unobserved attributes that influence choice behaviour this also violates the assumption of independence of error terms in the widely-used multinomial logit discrete choice model. Spatial heterogeneity, if ignored, will cause substantial bias in model parameters (Bhat et al. 2015).

In this paper we investigate spatial issues in recreational tourist flows to beaches on the Coromandel peninsula in New Zealand. The Coromandel peninsula is a rather unique context due to the idiosyncrasies of the geography and transport routes, and the fact that there are so many attractive beaches within close proximity. Multiple-destination trips are the norm rather than an inconvenient minority to discard. We review approaches in the literature for dealing with spatially correlated errors and and undertake a model search to compare the results of different specifications. A relatively unique feature of this study is that we use the theory of cumulative attraction to guide the development of spatial variables that better fit the observed beach visitation behaviour.

## 2. Empirical context

The Coromandel is a steep and hilly peninsula that lies across the Hauraki Gulf from Auckland, the largest city in New Zealand. Most of the peninsula interior is forest park and settlements of varying sizes are dotted along the coastline. Coromandel beaches are popular holiday destinations for residents of the nearby urban areas of Auckland and Hamilton, and to a lesser extent, international tourists. There are many beaches with high scenic and recreational appeal. Landscape characteristics and level of human modification vary around the peninsula. Administratively, it comprises five Community Board areas (Figure 2), which are fairly arbitrary political boundaries but do have some distinguishing geographic characteristics. Thames area is named for the town at the southern corner of the gulf and it is the entry point for the majority of visitors who come from Auckland or Hamilton. There is a road going east into Tairua-Pauanui and another winding road heading north along the relatively homogenous shingle-covered West coast. The Coromandel-Colville area has a thriving fishing industry, northern areas with limited accessibility and we further subdivide it by East and West coast. Mercury Bay has the largest population and many exceptionally scenic white sand beaches. Tairua-Pauanui is the gateway to Mercury Bay and provides a wide range of services. Whangamata area is named for a large town and popular surf beach and is the main route for people travelling from the south-eastern coastal city of Tauranga.



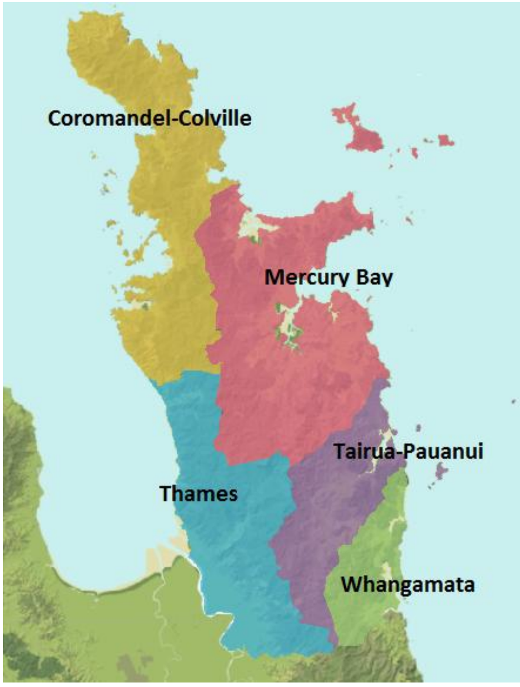




Figure 1 - Coromandel Peninsula (circled)

Figure 2 - Community Board Areas

For this study the coast of peninsula is divided into 109 discrete beach "sites" using a combination of visual inspection and labelled bays and harbours on Google maps. On the East coast a site is either an estuary, cove or bay with an area of sand or shingle flanked by headlands or cliffs. Some longer beaches are divided into two sites, such as Hot Water Beach which has a settlement at the southern end and undeveloped dunes at the northern end and separate access points. The Southwest coast has long stretches of relatively homogenous coast with few named bays. So, some beach sites are defined by the nearest settlement instead.

Our analysis of recreation values is simplified somewhat because the vast majority of visitors travel by car, every urban area is on the coast and the main road forms a loop around the peninsula. It is a simple matter to determine a tourist's probable route to any beach, and which other beaches they would have passed along the way. However, complicating factors include the need to consider not only travel distance and beach characteristics, but also the option for multiple destinations, differences in accessibility and complementary features of other sites on-route or in the area.

# 3. Theoretical framework

## 3.1 Travel cost

The consumption of beach recreation requires the user to incur the costs of travel and access to the site. These costs serve as the implicit price of the trip (Fletcher, Adamowicz, and Graham Tomasi 1990). An individual can only visit one site at a time and is assumed to choose the site that maximises his or her unobserved utility function for recreation benefits.

Multiple-destination trips complicate travel cost analysis because there is the potential for value to be attributed to the wrong site. The most direct solution is to discard multiple-site visitors from the sample if they are rare. However, more than half our sample comprises multiple destination beach trips. One approach is to allocate travel cost by the proportion of time spent at each site, as in Yeh, Haab and Sohngen (2006). Another is to include a dummy variable and price interaction for multiple destination trips (Parsons and Wilson 1997). Yet another approach is to use nested models for additional or "follow on" destinations (Taylor, McKean, and Johnson 2010). Mendelsohn (1992) treats combinations of sites as additional sites, but given the large number of combinations generated by 109 sites this solution is impractical. The most appropriate way to allocate costs largely depends on which travel pattern the individual visitor is using out of en-route, base-camp, regional tour or trip chaining (Lue, Crompton, and Fesenmaier 1993). We find that trip chaining provides the best fit in terms of log-likelihood for the observed data. The travel cost to the first site is therefore divided amongst all sites visited on the trip, while the incremental cost to access the second site is assigned to the second and subsequent sites, and so on. By factoring in the structure of the tour and weighting trips costs we avoid the downward bias from ignoring multi-day trips and the upward bias from attributing all trips costs to a single site.

## 3.2 Random Utility Models and Space

The predominant approach in the literature of travel cost analysis is to model the probability of discrete site choice within a random utility framework which incorporates travel costs, site qualities and allows the estimation of demand and substitution patterns across multiple sites (Phaneuf and Smith 2005). Estimation requires, at a minimum, the specification of a functional form for both the observed, deterministic part of utility and assumptions about the distribution of the unobserved, random component as per Manski's formulation (1977). The destination choices we analyse are conditional on the fact that the respondent has already decided to visit the Coromandel peninsula for the purpose of beach recreation. Hence there is no "stay-at-home" option and no option for other regions about which we have no visitation data.

The destination choice problem differs from types of choice such as consumer products because of the added dimension of space. Early applications of discrete choice models included spatial choices (for example, residential location in McFadden 1978) but the added complexity of spatial dependence was not often recognised (Pellegrini and Fotheringham 2002). Modelling spatial dependencies and interactions is complex due to difficulties in characterizing, defining and measuring them (Sener, Pendyala, and Bhat 2011). True spatial dependence can be multi-directional, but when the dependent variable is latent—as in random utility models—a fully simultaneous model is intractable (Anselin 2002). A more suitable approach is either exogenous spatial lags or spatial error components. There is also a fundamental identification problem since a cross-section of observations does not provide sufficient information to identify the full covariance structure or economic mechanism that leads to spatial dependence (Anselin 2001). Covariance between alternatives follows from the specification of a spatial weights matrix describing the strength of the relationship (contiguity or distance, for example) between each pair of sites. The specification of the weights is somewhat arbitrary and this is a weakness of modelling continuous space as discrete sites. The weights matrix may not be specified directly if a hierarchical approach or direct spatial variables are used but this does not eliminate it (Corrado and Fingleton 2012).

While some spatial effects may be accommodated through the construction of spatial variables, there will inevitably be unobserved effects (Bhat et al. 2015). The challenge is therefore to specify a computationally feasible model that accommodates the important spatial effects and has a firm foundation in economic theory. We now review different random utility models and their advantages and shortcomings for modelling spatial choice.

### 3.2.1.1 Multinomial probit

If the joint density of the random errors is assumed to be multivariate normal, the resulting specification is a multinomial probit (MNP) model. The advantage of the probit is that it allows flexible patterns of substitution because alternatives can have difference variance and correlations with each other (Pellegrini and Fotheringham 2002). The disadvantage is that calculation of a single probability requires integration with as many dimensions as there are alternatives, an analytically intractable task except when the number of alternatives is small (e.g. less than seven). Simulation techniques have been developed such as maximum simulated likelihood or maximum approximate composite marginal likelihood (Bhat 2011), but these methods are not available in standard statistics software and require substantial investment in programming purpose-specific code and are hence of interest only to a small minority of practitioners. We therefore turn our consideration to simpler models.

### 3.2.1.2 Multinomial Logit

The multinomial logit (MNL) model was shown to be consistent with RUM by McFadden (1974) and is the most widely used structure within random utility modelling. The utility that person *n* expects to obtain from site *i* is specified as:

$$U_{ni} = V_{ni} + \varepsilon_{ni} \qquad (1)$$

where $V_{ni}$ is a deterministic, linear-in-parameters component and $\varepsilon_{ni}$ is an unobserved utility component independently and identically distributed (i.i.d) according to a Type I Extreme Value distribution. The MNL approximates a joint multivariate normal distribution but is much simpler to compute. However, the i.i.d. assumption results in the property called Independence from Irrelevant Alternatives (IIA), which is undesirable when patterns of substitution are variable across different alternatives (as in spatial clusters, for example). As McFadden (1978) noted, “there may be a structure of perceived similarities among alternatives” that invalidate the assumption.

A simple way to remove the IIA property is to use a competing destinations model (Pellegrini and Fotheringham 2002) in which the MNL utility function is amended to reflect the probability that an alternative is evaluated. The rationale for this approach is that people do not evaluate every alternative and are more likely to be aware of sites that are large and close. The choice probability is therefore:

$$P_{ni} = \frac{e^{V_{ni}} L_n(i \quad G)}{\sum_{j=1}^{K} \left( e^{V_{nj}} L_n(j \quad G) \right)}$$

Where $L_n(i \quad G)$ is the likelihood that alternative *i* is in the cluster of "awareness" for individual *n*. The probability of evaluation only needs to express relationships not already included in the utility function. We test a competing destinations MNL model with an accessibility measure similar to that suggested by Fotheringham (1983):

$$L_n(i \quad G) = \left[\frac{1}{K-1}\sum_{\substack{j \\ j \neq i}} \frac{1}{d_{ij}}\right]^{\theta}$$

where $K$ is the set of all alternatives and $d_{ij}$ is the distance between alternatives *i* and *j*. Unlike Fotheringham, we do not weight the measure by site population because tourist flows are overwhelmingly inwards and many of our sites have no permanent residents.

### *3.2.1.3 Nested logit*

Another convenient way to allow for correlation across spatially clustered alternatives is to assume a hierarchical choice structure as in the nested logit (NL) model (Daly 1987). The set of alternatives *j* is partitioned into *K* non-overlapping subsets called nests. The assumption is that the individual first chooses a group (spatial cluster) of sites and the choice of site is conditional on this. The unobserved portion of utility, $\mathring{a}_{nj}$, has a generalized extreme value distribution and is correlated within nests. The choice probability for alternative *i* is therefore:

$$P_{ni} = \frac{e^{V_{ni}/\lambda_k}\left(\sum_{j \quad \beta_k} e^{V_{nj}/\lambda_k}\right)^{\lambda_k - 1}}{\sum_{l=1}^{K}\left(\sum_{j \quad \beta_l} e^{V_{nj}/\lambda_l}\right)^{\lambda_l}}$$

where $\lambda_k$ is a measure of the degree of independence in observed utility among alternatives in nest *k*. The nests may be spatial areas or some other grouping that corresponds to similarity of influence (Pooler 1998). The nested logit structure assumes alternatives are equally substitutable within a cluster but not between them. The limitation of the nested logit is that each nest must be exogenously specified, which can be a somewhat arbitrary division of continuous space (Pellegrini and Fotheringham 2002). In our application we test a NL with nests for each community board area.

### *3.2.1.4 GEV models*

Both MNL and NL are special cases of the Generalized Extreme Value (GEV) theorem of (McFadden 1978). The GEV class of models relaxes the i.i.d. assumption of the MNL by allowing the random components of alternatives to be correlated, while maintaining the assumption that they are identically distributed. Several GEV choice models have been developed including the cross-nested logit (Vovsha 1997), paired combinatorial logit (Chu 1989), the generalized nested logit (Wen and Koppelman 2001), the spatial correlated logit (Bhat and Guo 2004) and the network GEV (Daly and

Bierlaire 2006). Multiple-level hierarchies have also been used, for example Bekhor and Prashker (2008). The main difference between these models and NL is that nests overlap.

GEV models are very flexible and maintain closed-form expressions for choice probabilities. But this flexibility can only be realised by estimating a large number of dissimilarity or allocation parameters (Bhat and Guo 2004). Small (1987) noted that likelihood functions can be very flat across large numbers of sites. Paired nests produce a "proliferation of parameters" (Train 2002) which requires the researcher to imposed some sort of exogenous structure; perhaps by including only adjacent pairs. The use of adjacency measures to capture similarity is based on the idea that proximate sites are more likely to share unobserved features and be subject to competition effects, although this ignores the issue of complementarity. True spatial correlation may not be limited to adjacent or nearest neighbour sites.

For this study we test three variations of the general cross-nested-logit specification given by the generator function (Bierlaire 2006):

$$G(y) = \sum_{m=1} \left( \sum_{j \quad K} \left( {\alpha_{jm}}^{1/\mu} y_j \right)^{\mu_m} \right)^{\mu/\mu_m} \tag{2}$$

where $y$ is the deterministic part of the utility function, $j$ refers to an alternative in the set of all sites $K$, $m$ is a nest, $\mu$ is a scale parameter, $\mu_m$ is a nest-specific coefficient and $\acute{a}_{mk}$ are the parameters allocating sites to nests. For model A there are nests for type of site feature (boat ramp, campground, natural dune, estuary, food retail, gravel beach, motel, playground, road access, public toilet, urban area, seawall). Each site which possesses the feature is a member of the nest, weighted by the number of other attributes the site also possesses:

$$\alpha_{jm} = \frac{\omega_{jm}}{\sum_k \omega_{km}} \tag{3}$$

where $\grave{u}_{jm}$ = 1 if the site has the feature and 0 if it does not. For models B and C there is a nest for every site. In model B the allocation parameter is a cliff-distance decay function:

$$\alpha_{jm} = \frac{d_{jm}^{-2}}{\sum_k d_{km}^{-2}} \tag{4}$$

where $d$ is the distance in kilometres between the pair of sites and the cut-off (cliff) is $\acute{a}$ = 0.001 (d=31km). The purpose of the cut-off distance is to reduce the total number of allocation parameters ( 1,761 in this case). The cut-off distance is fairly arbitrary but we find different cut-offs make no difference to the estimation results. In model B we instead define the allocation parameters based on which sites are on-route to the nest site.

$$\alpha_{jm} = \frac{\delta_{jm}}{\sum_k \delta_{km}} \tag{5}$$

where $\partial_{jm}$ = 1 if the route to site *m* involves driving past site *j* and 0 if it does not. The specification is only feasible because of the simple loop nature of roads around the peninsula. This results in 3,104 non-zero allocation parameters. Remote northern sites end up with a lot of southern sites in their nests. ä

#### 3.2.1.5 Mixed Logit

The mixed logit is another generalisation of the MNL that allows for flexible substitution patterns. In fact, it can approximate any random utility model including GEV (McFadden and Train 2000). Complex correlation patterns across alternatives can be captured using additional random parameters or error components (Herriges and Phaneuf 2002). Thiene and Scarpa (2008), for example, used error components to define types of alpine sites which are believed to have a higher degree of substitutability. We test a panel mixed logit model with correlated (across individuals and areas) error components for adjacent community board areas; which are few enough in number to make the model computationally feasible. Mixed logit models have no closed form and must be estimated using simulation methods which is less efficient than GEV models. Utility for this model is specified as:

$$U_{ni} = V_{ni} + \text{ç} z_i + \varepsilon_{ni} \tag{6}$$

where ç is a vector of random normal terms with zero mean and standard deviation $ó_z$, $z_i$ is a vector of dummy variables indicating the site is in the community board area and $å_{ni}$ is still i.i.d. Gumbel. The choice probabilities are derived by integrating over the domain of the parameter distribution:

$$P_{ni} = \int_{\eta} \left( \frac{\exp(V_{ni} + \eta' z_i)}{\sum_{j=1}^{M} \left( \exp(V_{ni} + \eta' z_j) \right)} \right) . g(\text{ç} | ó z) d\text{ç} \tag{7}$$

### 3.2.2 Additional spatial variables

We control for some omitted spatial variables by including dummy variables in every model for the community board areas described above. We also test the inclusion of distance-based variables to express the importance of amenities at other locations in the utility function (since tourists can visit multiple sites). One model is a gravity formulation that includes a sum of "mass" terms for characteristics of other sites in the destination, downscaled by inverse square distance decay (Anselin 2002). However, the notion that distance functions effectively capture spatial dependence has long been challenged (LeSage and Pace 2008). We draw on the theory of cumulative attraction

(Nelson 1958) to develop a novel formulation in travel cost, and more relevant to consumer behaviour.

Nelson developed the theory of cumulative attraction to explain the attractiveness of retail clusters but it has also been applied to tourism (Lue, Crompton, and Stewart 1996; Weidenfeld, Butler, and Williams 2010). The theory implies that multiple attractions in an area will draw more visitors than if the sites were widely scattered. A key component is the principle of compatibility in which total attractiveness depends not only on geographic proximity but how complementary the sites are. Complementary sites must be dissimilar in some way, providing different experiences or services. It seems reasonable to assume that if a nearby site has an amenity that the main destination does not have, this will have additional value to a potential visitor. On the other hand, having additional amenities of the same type further away may provide little (if any) additional value. We analyse site compatibility using the multiple-destination trips in our data to confirm this assumption. We therefore include variables for the inverse travel time to the *nearest* site that possesses each type of amenity included in the utility function[1]. If the destination has the attribute the variable is zero. These variables are a proxy for the distribution and diversity of the sites, a key aspect of cumulative attraction.

## 4. Data Collection

The data was collected via a web-based survey from October 2013 to April 2014 designed to gather information about the revealed and stated preferences of domestic visitors to the Coromandel peninsula for beach recreation. We sourced participants from a pre-recruited panel of New Zealand residents provided by a market research company and a smaller, self-selected sample from online advertisements on Facebook and Google[2]. The use of a pre-recruited panel restricts multiple participations by the same individuals and is an increasingly popular collection mode (Windle & Rolfe, 2011). To qualify for the survey respondents had to live in New Zealand and have visited the Coromandel Peninsula in the past year. The survey included questions about their previous and planned beach visits, environmental attitudes, socio-economic variables and choice experiment questions. In this paper we only report the revealed preference results. Respondents were asked to

[1] Disliked attributes (non-sandy beaches and seawalls) are framed as the distance to the nearest site *without* these features since a negative amenity at another beach neither harms nor benefits a tourist who is under no obligation to visit that site. There is no distance variable for non-estuarine sites because every estuary is adjacent to a coastal beach.

[2] There were demographic and attitudinal differences between the panel and advertisement samples which are discussed in more detail in a forthcoming technical report.

report only trips where beach recreation was the primary purpose of the trip because multipurpose trips are not consistent with the assumptions of the travel cost method (Jeong, Crompton, and Dudensing 2015). They indicated the location of their beach visit(s) using a Google Maps™ API tool which provided the latitude and longitude of each visit. The beach markers were assigned to a beach site based on proximity. We excluded markers that were outside the Coromandel Peninsula, too far off shore or too far inland.

## 4.1 Definition of variables

The value of coastal recreation is highly dependent on the physical appearance of the coastal zone (Coombes, Jones, and Sutherland 2008). So, a large number of variables were calculated for each site including length, wide, surrounding land cover, type of sand/shingle, the presence of a stream, suitability for surfing, length of dune, length of seawalls, headland elevation, boating facilities, public toilets, campgrounds, playgrounds, motels, food retailers, usual population and overall development level. Water quality monitoring data is sparse and limited to a few estuaries where high nutrient levels are suspected, so there are no water quality variables we can include in the model. Many biophysical variables were highly correlated or just not useful explanatory variables. For example, almost all beaches are in close proximity to the forest park that covers the interior of the peninsula. Pohutukawa trees with their iconic crimson flowers also add to the scenic appeal, but again, they are everywhere. The final models include dummy variables for area, natural dune, boat launch facilities, campground, shingle beach, playground, seawall, road access, estuary and development level.

Development level of each site is determined by adjacency to an urban area. Urban areas on the peninsula are classified into three types by the regional authority[3] depending on the size and relative importance to the economy. There are six large towns with usual resident populations ranging from 750 to over 4000. There are eleven medium-size towns with populations in the hundreds and fifteen small settlements which may only have a few dozen permanent residents but provide important services such as food retail. Beaches with no buildings visible from the foreshore are coded as "undeveloped". Beaches with residential development outside any urban area are the base case and are labelled as "rural".

The travel distance and time by car between each origin and destination was calculated using Google Distance Matrix[4]. As discussed above, an assumption of destination-chaining was used to allocate travel cost amongst multiple destinations on the same trip. A standard fuel cost of 20 cents per

[3] http://www.waikatoregion.govt.nz/Environment/Environmental-information/REDI/882842/

[4] https://developers.google.com/maps/documentation/distance-matrix/

kilometre is assumed, based on the assumption of $2 per litre of petrol and 10 kilometres to the litre. For sites with no road access to the foreshore we added the additional walking time. The opportunity cost of travel time is defined as 25 per cent of hourly household income. For the "distance" variables in the gravity and cumulative attraction models we use travel time rather than distance. This is because many stretches of road on the peninsula are narrow, windy or unsealed and travel speed is variable.

## 5. Results

A total of 2,447 trips and 3,946 beach visits by 1,137 unique respondents are in the final data set. The following table shows a selection of descriptive statistics. Women and people with degrees are apparently over-represented in the sample. On-site surveys also found beach visitors were more likely to have a degree than the general population (Thomson 2003).

Table 1 - Descriptive statistics

| Measure | |
|---|---|
| Count of respondents | 1,137 |
| Count of trips | 2,447 |
| Average beaches per trip | 1.61 |
| Average travel time to site (hours) | 2.33 |
| Average age of respondent | 43 |
| Proportion of female respondents | 0.59 |
| Proportion of university-educated respondents | 0.47 |
| Proportion from Waikato region | 0.41 |
| Proportion from Auckland region | 0.38 |
| Proportion from Bay of Plenty region | 0.21 |
| Proportion of visits with an overnight stay | 0.39 |

Figure 3 shows the relative intensity of beach visits around the peninsula with obvious hotspots around large towns and the Mercury Bay area. It also illustrates how close the sites are to each other. Within a 15 minute travel time radius there are an average of 6 other beaches. Almost three quarters of beaches have an urban area within their radius.

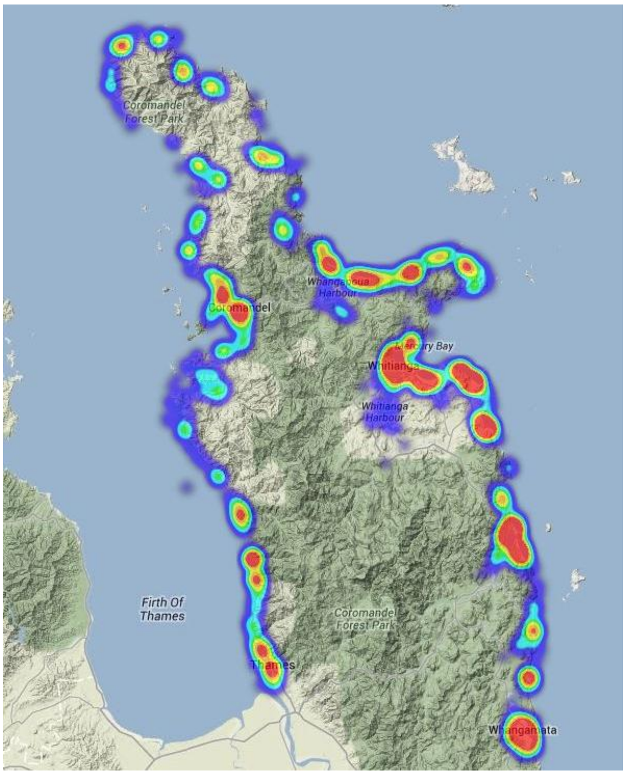



Figure 3 – Heat map of beach visits

## 5.1 Site compatibility

Compatibility is defined as the proportion of visitors to site A that also visit site B. With 109 beach sites there are a large number of potential combinations. We restrict the combinations to pairs of sites because only 11 per cent of people visited more than two beaches and the large number of possible three-site combinations results in miniscule compatibility measures for each trio.

Every site has a most compatible other site, and the median compatibility rating for these pairs is 29 per cent, “highly compatible” according to Nelson (1958). The average travel time between each site and its most compatible site is 18 minutes so compatible sites are not necessarily the closest. Table II shows visitor counts by beach development level cross tabulated with the other type(s) of sites they visited. Visitors to large urban sites are most likely to be on a single destination trips or visit another large urban site. These are typically two beach sites within the same urban area rather than two different towns. Visitors to urban medium sites are most likely to also visit large or small urban sites.

Visitors to urban small sites often also visit urban large sites. Visitors to undeveloped sites are most likely to visit another site of a different type, most often urban large.

Table II - Visitors to other development levels as a proportion of total visitors to development level[5]

| Development level | Unique visitor groups | Development type of other site(s) visited None | Urban large | Urban medium | Urban small | Rural | Undeveloped |
|---|---|---|---|---|---|---|---|
| Urban large | 1293 | 63% | 17% | 12% | 14% | 5% | 10% |
| Urban medium | 538 | 47% | 28% | 14% | 28% | 8% | 17% |
| Urban small | 546 | 38% | 34% | 27% | 18% | 9% | 22% |
| Rural | 389 | 52% | 16% | 12% | 12% | 19% | 23% |
| Undeveloped | 470 | 38% | 28% | 20% | 25% | 19% | 15% |
| All sites | 3236 | 51% | 23% | 16% | 19% | 10% | 16% |

Table III shows the relative compatibility of the different areas based on number of shared visitors. Thames area has the highest proportion of single-destination visitors who visit no other beach on that trip (74 per cent). People who visit a second site generally stay within the same area so most cross-site compatibility ratings are low. Mercury Bay has the highest proportion of people who stay in the area, with only 10 per cent visiting other areas.

Table III – Visitors to other areas as a proportion of total visitors to area

| Area | Unique visitor groups[6] | Area of other site(s) visited None | Coro East | Coro West | Mercury Bay | Tairua-Pauanui | Thames | Whanga mata |
|---|---|---|---|---|---|---|---|---|
| Coro East | 424 | 65% | 24% | 9% | 6% | 2% | 2% | 3% |
| Coro West | 122 | 45% | 32% | 27% | 4% | 1% | 0% | 2% |
| Mercury Bay | 1010 | 55% | 3% | 0% | 41% | 3% | 1% | 3% |
| Tairua-Pauanui | 363 | 71% | 2% | 0% | 7% | 21% | 2% | 4% |
| Thames | 255 | 74% | 4% | 0% | 4% | 3% | 20% | 2% |
| Whangamata | 446 | 73% | 3% | 0% | 8% | 4% | 1% | 18% |
| All sites | 2620 | 63% | 8% | 3% | 20% | 5% | 3% | 6% |

We estimate logistic regression between each pair of sites to see how well the shared visitors can be explained by site characteristics and site differences. Model 1 includes total visitor counts, travel time and site B characteristics. The positive and statistically significant influences on compatibility are a higher number of visitors overall (to site B), being on-route to site A, being in Mercury Bay, Coromandel-Colville or Whangamata areas, having a boat ramp, campground, natural dune, public

[5] Percentages add up to more than 100 because some people visit a second site of the same type as well as a different type

[6] The total is not the same as the previous table because when people visit multiple sites in the same area there are fewer “unique” visitors to the area

road access and public toilet. The negative explanatory variables are total number of site A visitors, site B being estuarine, small or medium urban area, or having a seawall. Model 2 includes additional variables representing differences between sites A and B. The log-likelihood ratio test statistic is significant at $p < 0.001$ so the additional variables improve model fit. If site B is non-estuarine, sandy, has a playground, or is undeveloped when site A is the opposite, this improves compatibility. Being in a different area or a larger or smaller urban area is associated with lower compatibility. These finding generally conform to the theory of cumulative attraction, with the exception of different sized urban areas being less compatible. In the next section we show the destination choice model results.

## 5.2 Model results

Table IV shows the results for all the models discussed above. The basic MNL model has a relatively good model fit to the data, with a McFadden pseudo r-squared of 0.189. The travel cost parameter is negative and significant, and the travel cost times wage interaction variable is positive, which means that high income individuals are willing to travel further. The area dummy variables are all positive which means every other area is preferred to Thames area. Site characteristics associated with a higher probability of visit are boat ramp, campground, motel, playground, public road, public toilet, sandy (as opposed to shingle or pebble) and a large urban area. The negative variables are estuary sites (which tend to be silty and colonised by mangroves), undeveloped, seawalls and food retailers. We expected food would be a positive variable, but it is possibly correlated with other less desirable characteristics or poorly measured. Tourists cannot have motels and playgrounds without the associated urban areas, but after controlling for these amenities small and medium urban areas have a residual negative effect. The parameter for large urban areas is positive and significant in the basic MNL model, yet negative and/or insignificant in models which include distance variables.

The competing destinations model did not offer any improvement in fit as measured by AIC and BIC. Theta is insignificant when area dummy variables are included in the model. The gravity model (with mass terms for amenities at every other site weighted by inverse squared travel time) is a better fit with a pseudo R-squared of 0.205. However, some of the mass variables have negative coefficient estimates (boat ramp, natural dune, playground and undeveloped), which is counter-intuitive. In the cumulative-attraction-inspired model 4, the inverse travel time[7] variables are all with positive coefficient estimates, though not all significant. This supports our assumption that distance to other

[7] We also tested inverse squared but the fit was worse

locations only matters if they have an amenity that is not found in the main visitation site, which is consistent with a complementarity relation.

The panel error components model 5 took several days to converge and fits relatively well with a pseudo R-squared of 0.21; although only one area random parameter is significant. There are significant covariances between Coromandel-Colville East and West, and between Coromandel-Colville East and Mercury Bay areas. It was not possible to estimate random parameters for every attribute; the model was too unstable.

A nested logit model with nests for the different community board areas offered no improvement in fit over the basic MNL with area dummy variables so is not shown below. The cross-nested logit models with allocation parameters based on distance on on-route measures are also relatively poor fits[8]. CNL model 8 with attribute-based nests fits better however, being superior to the ECL and equal to model 4. This implies that beach characteristics are a better indicator of substitutability than distance. Half of the inclusive value (IV) parameters are significant meaning that variance is different across sites with different attributes. For the final model 9 we add the distance variables from model 4 to the CNL model 8 and this improves the fit even further with a pseudo R-squared of 0.218 and the lowest BIC. Hence, based on fit to the data, model 9 is preferred. The log-likelihood ratio test statistic is significant at one per cent when compared with model 8. Not all the inverse-distance variables have significant coefficient estimates at 5 percent—only campground, food, toilet, undeveloped and large urban. Both the dummy and distance variables for large urban area have negative coefficient estimates, which implies a typical visitor prefers not to locate at or near a large urban area all else being equal. However, the large urban areas have accommodation, infrastructure and facilities so a lot of people visit anyway.

[8] We also tried a model with nests for area but these were insignificant with area dummy variables already included in the model

Table IV - Estimated models

| | Variable | 1. MNL | 2. MNL (Competing Destinations) | 3. MNL (Gravity) | 4. MNL (Cumulative attraction) | 5. ECL (Random areas) | 6. CNL (Distance nests) | 7. CNL (Onroute nests) | 8. CNL (Attribute nests) | 9. CNL (Attribute nests +) |
|---|---|---|---|---|---|---|---|---|---|---|
| Model fit | Log-likelihood | -15158 | -15083 | -14712 | -14595 | -14621 | -14927 | -15051 | -14602 | -14479 |
| | Psuedo-r2 | 0.181 | 0.185 | 0.205 | 0.211 | 0.210 | 0.193 | 0.187 | 0.211 | 0.218 |
| | No. parameters | 22 | 23 | 37 | 34 | 31 | 131 | 131 | 37 | 49 |
| | AIC | 30359 | 30211 | 29499 | 29258 | 29304 | 30117 | 30364 | 29158 | 29056 |
| | BIC | 30497 | 30355 | 29731 | 29471 | 29498 | 30939 | 31187 | 29390 | 29363 |
| Site attributes | Travel cost | -0.0775*** | -0.0776*** | -0.0749*** | -0.0774*** | -0.0834*** | -0.0740*** | -0.0633*** | -0.0622*** | -0.0642*** |
| | Travel cost x wage | 0.0007*** | 0.0007*** | 0.0007*** | 0.0007*** | 0.0008*** | 0.0007*** | 0.0006*** | 0.0006 | 0.0006*** |
| | Area CE | 0.913*** | 0.967*** | 1.050*** | 1.130*** | 1.060*** | 0.543*** | 0.924*** | 0.753*** | 0.881*** |
| | Area CW | 1.690*** | 1.670*** | 1.430*** | 1.770*** | 1.900*** | 1.860*** | 2.860*** | 1.380*** | 1.500*** |
| | Area M | 2.200*** | 2.220*** | 1.640*** | 1.980*** | 2.390*** | 2.080*** | 3.420*** | 1.810*** | 1.800*** |
| | Area TP | 0.959*** | 0.923*** | 0.789*** | 1.160*** | 0.987*** | 1.120*** | 1.460*** | 0.880*** | 1.120*** |
| | Area W | 1.070*** | 1.010*** | 0.784*** | 1.240*** | 1.040*** | 1.710*** | 2.820*** | 0.853*** | 1.090*** |
| | Boat ramp | 0.354*** | 0.344*** | 0.311*** | 0.243** | 0.367*** | 0.540*** | 0.561*** | 0.244*** | 0.345*** |
| | Campground | 0.373*** | 0.379*** | 0.260*** | 0.696*** | 0.373*** | 0.624*** | 0.308*** | -0.037 | 0.115* |
| | Natural dune | 0.049 | 0.039 | 0.397*** | 0.064 | 0.050 | -0.154** | -0.168** | -0.064 | 0.153* |
| | Estuary | -1.880*** | -1.870*** | -1.170*** | -1.820*** | -1.900*** | -1.990*** | -1.850*** | -2.760*** | -2.770*** |
| | Food retailer | -0.204*** | -0.192*** | -0.094 | 0.555*** | -0.214*** | -0.010 | -0.535*** | -0.303*** | 0.049 |
| | Motel | 0.224*** | 0.197** | 0.510*** | 0.115 | 0.240*** | 0.405*** | 0.474*** | 0.417*** | 0.429*** |
| | Playground | 0.265*** | 0.268*** | 0.189*** | 0.370*** | 0.268*** | 0.307*** | 0.375*** | -0.011 | -0.045 |
| | Public road | 1.030*** | 1.030*** | 1.070*** | 1.270*** | 1.030*** | 1.180*** | 1.130*** | 0.274*** | 0.456*** |
| | Public toilet | 0.248*** | 0.253*** | 0.246*** | 0.263** | 0.241*** | 0.234*** | 0.084 | -0.266 | 0.163** |
| | Sandy beach | 0.493*** | 0.502*** | 0.291*** | 0.717*** | 0.512*** | 0.557*** | 0.697*** | 0.476*** | 0.637*** |
| | Undeveloped | -0.286*** | -0.267*** | -0.267*** | -0.201** | -0.295*** | -0.018 | -0.442*** | -0.356*** | -0.306*** |
| | Small urban | -0.320*** | -0.309*** | -0.074 | -0.369*** | -0.339*** | -0.447*** | -0.774*** | -0.729*** | -0.823*** |
| | Medium urban | -0.329*** | -0.319*** | -0.134 | -0.474*** | -0.341*** | -0.445*** | -0.552*** | -0.440*** | -0.687*** |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Large urban | 0.468*** | 0.468*** | -0.096 | -0.368** | 0.451*** | -0.063 | -0.237*** | -0.102* | -0.693*** |
| | Seawall | -0.422*** | -0.423*** | -0.454*** | -0.231 | 0.427*** | -0.529*** | -0.513*** | -0.372*** | -0.168** |
| | Theta (accessibility) | | -0.139 | | | | | | | |
| Mass variables (gravity model) / inverse travel time (cumulative attraction model) | Boat ramp | | | -0.851*** | 1.200*** | | | | | 0.371 |
| | Campground | | | 1.290*** | 2.020*** | | | | | 0.898*** |
| | Food retailer | | | 2.450*** | 4.030*** | | | | | 2.680*** |
| | Motel | | | 1.140 | 0.355 | | | | | -0.075 |
| | Natural dune | | | -0.634*** | 0.616* | | | | | -0.058 |
| | Playground | | | -1.630*** | 0.520 | | | | | -0.279 |
| | Public road | | | -0.327* | 0.010 | | | | | -0.012 |
| | Sandy beach | | | 0.181 | 0.312 | | | | | 0.281 |
| | Public toilet | | | 0.045 | 0.420 | | | | | 0.598** |
| | No seawall | | | 0.634*** | 2.830*** | | | | | 1.030* |
| | Undeveloped | | | -2.860*** | 0.292** | | | | | 0.280** |
| | Large urban | | | -0.069 | -5.140*** | | | | | -2.970*** |
| Random parameter standard deviations | Area CE | | | | | 1.120*** | | | | |
| | Area CW | | | | | 0.046 | | | | |
| | Area M | | | | | 0.067 | | | | |
| | Area TP | | | | | -0.057 | | | | |
| | Area W | | | | | 0.447 | | | | |
| Random parameter covariances | Cov(CE, CW) | | | | | 1.090*** | | | | |
| | Cov(CE, M) | | | | | 1.110*** | | | | |
| | Cov(M, TP) | | | | | 0.028 | | | | |
| | Cov(TP, W) | | | | | -0.218 | | | | |

## 5.3 Scenario

We analyse a hypothetical scenario to illustrate the difference in predicted market share between the simple MNL model 1 and preferred model 9. The scenario involves the closure of the popular camping ground at Hahei beach. As coastal property values increase it is not uncommon for camping grounds to be sold and developed with houses or apartments (Collins and Kearns 2010). Table V shows the ten sites with the largest percent change in market share. Model 1 predicts this would result in a 29.6 percent decrease in market share for Hahei, with visitors redistributed all around the peninsula. Model 9, instead, predicts a smaller effect at Hahei (possibly because there is another campground a few kilometres away) and negative impacts at several beaches on the far side of Hahei (Gemstone bay, Stingray Bay, Cathedral Cove and an un-named bay). Model 1 completely ignores the likely impact on visitors of these undeveloped beaches near Hahei, many of whom will want low-cost accommodation nearby.

Table V – 10 most affected sites

| Site Name | Current Market share | % Change in market share | |
|---|---|---|---|
| | | Model 1 | Model 9 |
| Hahei | 4.7% | -29.6% | -10.5% |
| Gemstone bay | 0.5% | 2.2% | -20.8% |
| Stingray bay | 0.3% | 2.2% | -18.4% |
| Cathedral cove | 2.0% | 2.4% | -17.0% |
| Unnamed bay near Hahei | 0.1% | 2.2% | -14.2% |
| Orua Bay | 0.2% | 2.1% | 7.2% |
| Red Bay | 0.2% | 1.4% | 5.4% |
| Humbug bay | 0.0% | 1.4% | 5.1% |
| Waikawau river estury | 0.4% | 1.3% | 4.1% |
| Lonely bay | 0.7% | 2.1% | 3.8% |

# 6. Discussion & Conclusion

There are many several different modelling options for addressing the issue of spatial correlation in destination choice data, but not all of them are feasible or even practical for a study modelling choice between 109 alternative sites. By analysing site compatibility we ascertain that people on multiple-destination trips tend to visit sites that are close together, but different in terms of attributes. We find that a cross-nested logit model with nests defined by attributes fits better than paired distance-based nests. Once included, attribute distance variables improve the fit further. The total improvement in model fit is small but the preferred model appears to generate more intuitive patterns of substitution in response to a change in attributes at one site, making it more useful for policy analysis. Our recognition of the importance of complementary differences between sites is apparently rare in a literature that appears to be more focussed on unobserved similarities.

The expanded model with inverse distance variables also highlights the importance of site diversity in a context where multiple-destination visits are the norm. Being near a food retailer and near a completely undeveloped beach both have value but are mutually exclusive at a single destination. Care should be taken to preserve remaining undeveloped coves near more developed areas, since this is where they will have the highest recreation value. New coastal housing developments, on the other hand, have the potential to reduce the cumulative attraction of the wider area. When considering a change to the coastal landscape or services offered there, decision makers should consider the distance to substitute sites and services of a similar type.

Cumulative attraction implies destination marketing is likely to be most effective when cooperative rather than competitive strategies for tourism are developed in the same area (Hunt and Crompton 2008). The Coromandel regional tourism operator meets this need with an apparently successful marketing strategy for the whole peninsula[9]. Individual tourism operators could also use more cooperative strategies.

One issue we do not address in this paper is visitor heterogeneity. Individuals have different motivations and preferences—some prefer secluded and natural coastal landscapes while others want family-friendly destinations with plenty of services. However, there is no information source about the incidence of these preferences in the population. So, we only include an income interaction variable in the models. We are primarily interested in explaining overall market share of the sites. Future research may investigate whether and to what degree the importance of cumulative attraction varies amongst different types of visitor. For example, some visitors may prefer to go to one destination and stay there, while others want to experience attractions of the wider area.

## 7. References

Anselin, Luc
2001 Spatial Effects in Econometric Practice in Environmental and Resource Economics. American Journal of Agricultural Economics 83(3): 705–710.
2002 Under the Hood Issues in the Specification and Interpretation of Spatial Regression Models. Agricultural Economics 27(3): 247–267.

[9] www.thecoromandel.com

Bekhor, Shlomo, and Joseph N. Prashker
2008 GEV-Based Destination Choice Models That Account for Unobserved Similarities among Alternatives. Transportation Research Part B 42(3): 243–262.

Bhat, Chandra R
2011 The Maximum Approximate Composite Marginal Likelihood (MACML) Estimation of Multinomial Probit-Based Unordered Response Choice Models. Transportation Research Part B: Methodological 45(7): 923–939.

Bhat, Chandra R, Subodh K Dubey, Mohammad Jobair Bin Alam, and Waleed H Khushefati
2015 A New Spatial Multiple Discrete Continuous Modeling Approach to Land Use Change Analysis. Journal of Regional Science 55(5): 801–841.

Bhat, Chandra R., and Jessica Guo
2004 A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling. Transportation Research Part B: Methodological 38(2): 147–168.

Bierlaire, Michel
2006 A Theoretical Analysis of the Cross-Nested Logit Model. Annals of Operations Research 144(1): 287–300.

Chu, Chaushie
1989 A Paired Combinatorial Logit Model for Travel Demand Analysis. *In* Transport Policy, Management & Technology Towards 2001: Selected Proceedings of the Fifth World Conference on Transport Research.

Collins, Damian, and Robin Kearns
2010 "Pulling up the Tent Pegs?" The Significance and Changing Status of Coastal Campgrounds in New Zealand. Tourism Geographies 12(1): 53–76.

Coombes, Emma G, Andy P Jones, and William J Sutherland
2008 The Biodiversity Implications of Changes in Coastal Tourism due to Climate Change. Environmental Conservation 35(4): 319–330.

Corrado, Luisa, and Bernard Fingleton
2012 Where Is the Economics in Spatial Econometrics? Journal of Regional Science 52(2): 210–239.

Daly, Andrew
1987 Estimating "tree" Logit Models. Transportation Research Part B: Methodological 21(4): 251–267.

Daly, Andrew, and Michel Bierlaire
2006 A General and Operational Representation of Generalised Extreme Value Models. Transportation Research Part B: Methodological 40(4): 285–305.

Fletcher, Jerald J., Wiktor L. Adamowicz, and Theodore Graham Tomasi
1990 The Travel Cost Model of Recreation Demand: Theoretical and Empirical Issues. Leisure Sciences 12(1): 119–147.

Fotheringham, Stewart
1983 A New Set of Spatial-Interaction Models.

Herriges, Joseph A., and Daniel J. Phaneuf
2002 Inducing Patterns of Correlation and Substitution in Repeated Logit Models of Recreation Demand. American Journal of Agricultural Economics 84(4): 1076–1090.

Hunt, Michael A., and John L. Crompton
2008 Investigating Attraction Compatibility in an East Texas City. International Journal of Tourism Research 10(3): 237–246.

Jeong, Ji Youn, John L. Crompton, and Rebekka M. Dudensing
2015 The Potential Influence of Researchers' "Hidden" Procedure Decisions on Estimates of Visitor Spending and Economic Impact. Journal of Travel Research. http://jtr.sagepub.com/content/early/2015/09/24/0047287515605932.abstract.

LeSage, James P., and R. Kelley Pace
2008 SPATIAL ECONOMETRIC MODELING OF ORIGIN-DESTINATION FLOWS*. Journal of Regional Science 48(5): 941–967.

Lue, Chi-Chuan, John L Crompton, and Daniel R Fesenmaier
1993 Conceptualization of Multi-Destination Pleasure Trips. Annals of Tourism Research 20(2): 289–301.

Lue, Chi-Chuan, John L. Crompton, and William P. Stewart
1996 Evidence of Cumulative Attraction in Multidestination Recreational Trip Decisions. Journal of Travel Research 35(1): 41–49.

Manski, Charles F
1977 The Structure of Random Utility Models. Theory and Decision 8(3): 229–254.

Marrocu, Emanuela, and Raffaele Paci
2013 Different Tourists to Different Destinations. Evidence from Spatial Interaction Models. Tourism Management 39: 71–83.

McFadden, D.
1974 Conditional Logit Analysis of Qualitative Choice Behavior. *In* Frontiers in Econometrics. P. Zarembka, ed. New York: Academic Press.
1978 Modelling the Choice of Residential Location. Institute of Transportation Studies, University of California.

McFadden, Daniel., and Kenneth. Train
2000 Mixed MNL Models for Discrete Response. Journal of Applied Econometrics 15: 447–470.

Mendelsohn, Robert, John Hof, George Peterson, and Reed Johnson
1992 Measuring Recreation Values with Multiple Destination Trips. American Journal of Agricultural Economics 74(4): 926–933.

Nelson, RL
1958 The Selection of Retail Locations (Dodge, New York).

Parsons, George R, and Aaron J Wilson
1997 Incidental and Joint Consumption in Recreation Demand. Agricultural and Resource Economics Review 26: 1–6.

Pellegrini, Pasquale A, and A Stewart Fotheringham
2002 Modelling Spatial Choice: A Review and Synthesis in a Migration Context. Progress in Human Geography 26(4): 487–510.

Phaneuf, Daniel J., and V. Kerry Smith
2005 Chapter 15 Recreation Demand Models. *In* Handbook of Environmental Economics. Karl-Gran Mler and Jeffrey R. Vincent, ed. Pp. 671–761. Elsevier. http://www.sciencedirect.com/science/article/pii/S1574009905020152.

Pooler, Jim
1998 Competition among Destinations in Spatial Interaction Models: A New Point of View. Chinese Geographical Science 8(3): 212–224.

Sener, Ipek N., Ram M. Pendyala, and Chandra R. Bhat
2011 Accommodating Spatial Correlation across Choice Alternatives in Discrete Choice Models: An Application to Modeling Residential Location Choice Behavior. Journal of Transport Geography 19(2): 294–303.

Small, Kenneth A
1987 A DISCRETE CHOICE MODEL FOR ORDERED ALTERNATIVES. Econometrica (1986-1998) 55(2). 214864977. ProQuest Central: 409.

Taylor, R. Garth, John R. McKean, and Donn Johnson
2010 Measuring the Location Value of a Recreation Site. Journal of Agricultural and Resource Economics 35(1): 87–104.

Thiene, Mara, and Riccardo Scarpa
2008 Hiking in the Alps: Exploring Substitution Patterns of Hiking Destinations. Tourism Economics 14(2): 263–282.

Thomson, Jill
2003 Coastal Values and Beach Use Survey Report. 2003/09. Technical Report. Waikato Regional Council. http://www.waikatoregion.govt.nz/Services/Publications/Technical-reports/2003-technical-reports/.

Train, K.
2002 Discrete Choice Methods with Simulation. Cambridge Univ Pr.

Vovsha, Peter
1997 Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area. Transportation Research Record: Journal of the Transportation Research Board(1607): 6–15.

Weidenfeld, Adi, Richard W Butler, and Allan M Williams
2010 Clustering and Compatibility between Tourism Attractions. International Journal of Tourism Research 12(1): 1–16.

Wen, Chieh-Hua, and Frank S Koppelman
2001 The Generalized Nested Logit Model. Transportation Research Part B: Methodological 35(7): 627–641.

Yeh, Chia-Yu, Timothy C. Haab, and Brent L. Sohngen
2006 Modeling Multiple-Objective Recreation Trips with Choices Over Trip Duration and Alternative Sites. Environmental & Resource Economics 34: 189–209.

# 8. Appendix

Table VI – Logistic regression of the number of visitors to site A who also visit site B

| | Model 1 | Model 2 |
|---|---|---|
| Intercept | -4.6278*** | -4.2584*** |
| Site A total visitors | -0.0014*** | -0.0005** |
| Travel time between sites | -0.0386*** | -0.0380*** |
| *Site B Characteristics* | | |
| Site B total visitors | 0.0057*** | 0.0058*** |
| Site B is on-route to Site A | 0.4846*** | 0.4706*** |
| Mercury Bay area | 1.2012*** | 1.1896*** |
| Tairua-Pauanui area | -0.0698 | -0.1466 |
| Coromandel-Colville East | 0.2228** | 0.1958** |
| Coromandel-Colville West | 1.0736*** | 1.0234*** |
| Whangamata Area | 0.3704*** | 0.2615** |
| Boat ramp | 0.3661*** | 0.3710*** |
| Campground | 0.2527*** | 0.2191*** |
| Natural dune | 0.3331*** | 0.5546*** |
| Estuary | -1.0611*** | -0.9835*** |
| Food retailer | 0.1609*** | 0.1237** |
| Shingle or pebble beach | -0.0404 | 0.0061 |
| Motel | -0.0871 | -0.1553 |
| Playground | 0.0763 | 0.0523 |
| Public road access | 0.9124*** | 0.9287*** |
| Public toilet | 0.1463*** | 0.1833*** |
| Undeveloped | -0.1382* | -0.2835*** |
| Small urban | -0.1460** | -0.1750*** |
| Medium urban | -0.2263*** | -0.3274*** |
| Large urban | 0.0215 | -0.2286*** |
| Seawall | -0.2994*** | -0.2736*** |
| *Differences - characteristics possessed by Site B but not Site A* | | |
| Different area | | -0.2631*** |
| Boat ramp | | -0.0161 |
| Campground | | 0.0605 |
| Natural dune | | 0.4269*** |
| Food retailer | | 0.1023* |
| Not on an estuary | | 0.2883*** |
| Sandy beach | | 0.2349*** |
| Motel | | 0.0868 |
| Playground | | 0.1110** |
| Public road access | | 0.0298 |
| Public toilet | | -0.1885* |
| Undeveloped | | 0.3458*** |
| No seawall | | 0.0224 |
| More developed | | -0.1903*** |

| | | |
|---|---|---|
| Less developed | | -0.6058*** |
| Observations | 11881 | 11881 |
| Null deviance | 21850.9 | 21850.9 |
| Residual deviance | 8175.7 | 7881.3 |

* significant at 10%, ** significant at 5%, *** significant at 1%