A Sequential Logit Dynamic Travel Demand Model For Hurricane Evacuation

By

Haoqiang Fu

Ph.D. Candidate
Department of Civil and Environmental Engineering, Louisiana State University
Baton Rouge, LA 70808
Phone: (225) 749-7092, Fax: (225) 7497092, email: cehfu@lsu.edu

Chester G. Wilmot

Associate Professor
Department of Civil and Environmental Engineering
Louisiana Transportation Research Center
Louisiana State University, Baton Rouge, LA 70808
Phone: (225) 578-4697, Fax: (225) 578-8652, email: cewilm@eng.lsu.edu
(Corresponding author)

Paper Prepared for the

TRB 2004 Annual Meeting
Washington, D.C.
January 2004

(Word count: 6608 words plus 3 tables and 4 figures
Abstract length: 159 words)

Submission date: 11/05/2003

ABSTRACT

Little attention has been given to estimating dynamic travel demand in transportation planning in the past. However, when factors which influence travel are changing significantly over time – such as with an approaching hurricane - dynamic demand and the resulting variation in traffic flow on the network become important. In the research reported in this paper, the decision to evacuate in the face of an oncoming hurricane is considered as a series of binary choices over time. A sequential binary logit model is developed to model the probability of a household evacuating at each time period before landfall as a function of the household's socio-economic characteristics, the characteristics of the hurricane, and policy decisions made by authorities as the storm approaches. Data from Southwest Louisiana collected following Hurricane Andrew was used to estimate a model which produces dynamic travel demand estimates of hurricane evacuation. Based on the results, a sequential logit model appears capable of modeling dynamic evacuation demand satisfactorily.

Keywords: trip generation, dynamic, travel demand, time-dependent, hurricane, evacuation, sequential model, logit model, sequential logit.

1. INTRODUCTION

In the past two decades, one of the fastest growing research areas in travel demand modeling has been Dynamic Traffic Assignment (DTA) [1, 2, 3, 4, and 5]. DTA seeks to assign traffic continuously or in very short time intervals, and then keeps track of vehicles both temporally and spatially. The result is a representation of traffic conditions on each link in the network at all times. This is fundamentally different from traditional static traffic assignment where traffic conditions on the network are assumed to remain static during a period of analysis which may extend for several hours or even a day. It is also assumed that trips occupy all the links on the shortest path between their origin and destination during this period. Thus, an estimate of total travel over an extended period is obtained in static assignment based on the assumption that traffic conditions during that period remain unchanged. DTA, on the other hand, models dynamic conditions on the network by making assignment responsive to varying travel demand, congestion, and changing road conditions due to incidents, road closures, or even reverse-laning of facilities. As a result, DTA provides more accurate and realistic prediction of traffic conditions during an evacuation process.

Hurricane evacuation is very different from day-to-day travel modeled in conventional urban transportation planning. It involves long travel times, high levels of extended congestion, the uncertainty of road conditions on the route ahead, and the possibility that destinations may need to be changed due to closed roads or excessive congestion. In urban transportation planning, many trips are discretionary in that they can be postponed from one time to another or, in certain cases, foregone entirely. However, in evacuation, relatively little flexibility on timing is available and evacuation is sometimes mandatory thereby virtually eliminating the discretion of the individual traveler. Evacuees are also generally more willing to follow directions from officials as to which route to use and are less likely to choose the shortest path than urban travelers making regular trips. These differences point to the fact that while travel is generated in both situations, travel behavior is different between them. It is important, therefore, to be careful in applying urban transportation planning procedures in evacuation transportation planning. One area where this is particularly relevant is in traffic assignment. Static traffic assignment requires that trip travel times be shorter than the analysis period so that the assumption that trips traverse all links between origin and destination is validated [6]. In urban transportation planning, analysis periods of several hours are usually sufficient to satisfy this condition but in evacuation travel times are often 10 or more hours in duration. If static traffic assignment is used, this results in estimates of traffic over long periods of time without any knowledge of how speed, volume, density, delay, and travel time vary within the period. Clearly, this seriously compromises the usefulness of the assignment results in evaluating alternative evacuation strategies and management plans.

To perform dynamic traffic assignment, the estimation of time-dependent origin-destination (O-D) demand is required. However, most researchers working in the field of DTA assume such time-dependent O-D tables are available a priori. Little research has been conducted on estimating dynamic travel demand. Ziliaskopoulos and Peeta [5] pointed out in their recent review on DTA: “Probably, the single most challenging obstacle to overcome, before deploying DTA for planning applications, is that of estimating and predicting the time-dependent origin-destination demand… Surprisingly, the problem of estimating the temporal distribution of demand has been addressed by only a few studies.”

There are several computer packages to model evacuation, some of which can be applied to hurricane evacuation. Some of these packages use dynamic assignment, thereby providing more accurate information about traffic conditions. However, they all assume that either a time-dependent O-D table is given or a default response curve is specified to assign total evacuation demand to time intervals.

Current practice in hurricane evacuation travel demand estimation is to conduct the process in two steps: the estimation of total evacuation demand in the first step and the estimation of departure time in the second. Generally, these steps are conducted using simple relationships such as means, rates, and distributions rather than the more sophisticated mathematical relationships observed in urban transportation planning [7]. For example, the most common method of estimating evacuation demand is to use evacuation 'participation rates' of geographic subdivisions of the area in which evacuation behavior is expected to be homogeneous. Participation rates are the proportion of households in these geographic subdivisions (evacuation zones) that evacuate. The proportions are assumed to change depending on the severity of the storm and its flooding potential. Participation rates are established subjectively based on past observed behavior. Some researchers report the use of logistic regression to model evacuation demand [7, 8, 9, and10]. Artificial Neural Networks have also been used [7].

In current hurricane evacuation modeling practice, a response curve is typically used to portray the percentage of trips evacuating in each time interval [11, 12, 13, and 14]. A response curve is the assumed departure time distribution of evacuees. It is also sometimes referred to as a loading or mobilization curve. It is usually presented as the cumulative percentage of evacuees evacuating by time period, and traditionally has been assumed to take on a sigmoid or "S" shape. Depending on how readily an analyst expects evacuees to respond to an evacuation order, loading curves are classified as "quick", "medium", or "slow". The quicker the response, the steeper the curve. In current practice, choice of loading curve is a subjective decision made by the analyst.

In the research reported in this paper, the position has been taken that the decision to evacuate and the decision to depart are made jointly. That is, we assume that the decision of whether to evacuate and when to evacuate is made simultaneously. We also postulate that this joint decision is an issue that is considered repeatedly prior to it being taken. In other words, we suggest each household reviews the conditions surrounding a storm repeatedly as it approaches, each time deciding not to evacuate, until, if a threshold is reached in their evaluation, a decision is made to evacuate at a certain time. To model this process, we propose the use of a sequential logit model.

2. METHODOLOGY

The logit model has been used extensively in transportation for the last several decades to model choices. The choices modeled have typically been nominal choices (i.e. distinguished by name) as, for example, in the choice among travel modes such as auto and transit. Provided these choices have been distinct, the Independence from Irrelevant Alternatives (IIA) property of logit models has not distorted the model's estimates. However, when ordinal choices are modeled (i.e. choices in which order among the alternatives is significant), some dependency among the choices may exist. For example, in the choice of the number of trips to make or the number of vehicles to own, the choices are not entirely independent of each other because the choice of n implies that the choice of n-1 must have preceded it. Models that are explicitly constructed to handle such ordered choices are needed.

Agresti [15] describes three commonly used ordered probability models: the adjacent-category model, which compares the probability of each outcome to the probability of next outcome in the sequence; the continuation-ratio model, which compares the probability of each outcome to the probability of all higher outcomes (or alternatively, to all lower outcomes); and the proportional odds model, which compares the probability of an equal or smaller outcome to the probability of a larger outcome. In all three models, the coefficients in the utility function can be made to remain constant among the alternatives or attain unique values for each alternative *i*. Different models make different assumptions regarding this issue.

Ordered outcomes considered in the past display a subtle difference. In some cases, ordered outcomes have been described as a ranking without any linking or sequence of choices implied among the outcomes. Examples of this kind of ordering are choices among grades of gasoline (regular, super, premium), choice of level of employment (none, part-time, full-time), choice of theater tickets by price, number of days vacation to take, or size of home to buy. The other type of ordered outcome considered is where the choice of an outcome implies that all earlier outcomes in the ordering had to be considered first. This occurs, for example, when a household considers purchasing an additional vehicle or trip generation is seen as a sequence of decisions to make an additional trip. This type of ordered choice is perhaps more aptly termed sequential choice since higher categories of outcome can only be reached by proceeding through each lower category of outcome and the decision maker cannot reverse former choices. Sheffi [16] calls these sequence of choices nested alternatives.

Sequential choice occurs in dynamic travel demand modeling during evacuation. If we discretize time into time intervals, then in time interval *i* a household has the binary choice to evacuate or not to evacuate provided the decision to not evacuate was made in all earlier choices. If the choice in time interval *i* is not to evacuate, then the household faces the same binary choice in time interval *i*+1, and so on until either a decision to evacuate is made or the end of the analysis period is reached with no decision to evacuate being made. Amemiya [17] describes a model that can handle such sequential decisions based on random utility theory. Fahrmeir and Tutz [18] derive a similar model based on latent variable model.

Amemiya's model is based on a series of continuation-ratio models. It can be illustrated using the random utility principle in the context of hurricane evacuation. Let $U^c{}_i$ be the utility of a household choosing not to evacuate in time interval *i* given that the *i*th interval was reached without evacuation, and $U^s{}_i$ the utility of a household choosing to evacuate in time interval *i* given that the *i*th interval was reached without evacuation (superscripts *c* and *s* stand for "continue" and "stop", respectively). If, in keeping with common practice, we assume that each of the random utilities $U^c{}_i$ and $U^s{}_i$ is composed of a systematic component V_i and an error term ε_i, i.e. $U_i = V_i + \varepsilon_i$, and the error terms associated with each utility function in *i* are independently and identically Gumbel distributed, then the $U^c{}_i - U^s{}_i$'s are logistic distributed and the probability that a household will evacuate at time *i* given that it did not evacuate earlier, can be expressed in the form of a regular binary logit model for each time period *i* :

$$P(i)_{s/c} = \frac{e^{V_i^s}}{e^{V_i^s} + e^{V_i^c}} \qquad i = 1, 2, \ldots I, \tag{1}$$

where *I* is the total number of time intervals.

The assumption that the error term *ε* is independent among the alternatives (i.e. whether to evacuate or not) in any one time period is not difficult to justify as the alternatives are distinctly different. However, if the model in equation 1 is not estimated on observations from each time period separately, but on observations from all time intervals collectively, then repeated observations of the same household will occur in the estimation dataset and the potential exists for correlation among the error terms. To the extent that characteristics of the household affect the decision to evacuate or not, then the potential for correlation among observations of the same household exists. However, the greater impact on evacuation decision is expected from characteristics of the storm which change over time and are unrelated to households (e.g. proximity of the storm or wind speed). Sheffi [16] has also showed that when the correlations among the difference in error terms in a sequential binary probit model are assumed to be related to the overlap of utilities of earlier choices, the covariance among utility differences is zero. This suggests that a similar condition may exist among utility differences in a sequential logit model. If we assume the $U^c{}_i - U^s{}_i$'s are independent for different time intervals i, then $P(i)$, the probability of the household evacuating in time interval i is

$$\begin{aligned} P(i) &= \Pr(U^c{}_1 \geq U^s{}_1 \cap U^c{}_2 \geq U^s{}_2 \cap ...U^c{}_{i-1} \geq U^s{}_{i-1} \cap U^s{}_i \geq U^c{}_i) \\ &= P(1)_{c/s} P(2)_{c/s} \cdots P(i-1)_{c/s} P(i)_{s/c} = P(i)_{s/c} \prod_{j=1}^{i-1} [1 - P(j)_{s/c}], \end{aligned} \tag{2}$$

where $P(i)_{s/c}$ is the probability that the utility of a household to evacuate is greater than the utility of the household to not evacuate in time interval i, provided that the household has not already evacuated.

From equations 1 and 2, the probability of a household evacuating in time interval i, $P(Y = i)$, is the product of i independent binary choices, the first i-1 choices being not to evacuate and the ith to evacuate. Because of this special structure, this sequential model can be estimated using existing methods for binary choice models.

One intuitive method is to apply the continuation-ratio logit model concept to estimate the parameters of each individual binary choice model (the conditional probability). Then the unconditional probability of evacuating in each time interval for every household can be calculated based on equation 2. For this, the data must be arranged as follows. For time interval 1, the outcomes of those who evacuate in this interval are coded as 1, all those who do not evacuate in the interval are coded as 0, and the parameters of a binary logit model for time interval 1 are estimated. For time interval 2, data for those who evacuate in the previous time interval are excluded. The outcomes of those who evacuate in time interval 2 are coded as 1, all those who do not evacuate in time interval 2 are coded as 0, and the parameters of a binary logit model for time interval 2 are estimated on this data. This procedure is repeated for every time interval eliminating records of those who evacuate for all previous time intervals. However, there are several drawbacks to this method. First, multiple models have to be estimated, involving more data manipulation and modeling effort. Second, for later time intervals, data used to estimate the parameters will be smaller, resulting in less reliable estimation of the parameters. This is because as households evacuate in the previous time intervals, there are fewer and fewer households remaining. Third, restrictions such as the same parameters for different time intervals, which might be a valid option, cannot be applied. Last, since parameters are related to time intervals, predictions beyond the scope of the observed time intervals cannot be obtained.

There is an alternative method to estimate this model that allows us to consider all binary choices simultaneously and avoid the disadvantages just mentioned. Let $P_n(i)$ denote the probability that household n evacuates in time interval i. Using equation 2, the likelihood function is

$$L = \prod_{n=1}^{N} P_n(i) = \prod_{n=1}^{N} P_n(i)_{s/c} \prod_{j=1}^{i-1} [1 - P_n(j)_{s/c}], \tag{3}$$

where N is the total number of households. This likelihood requires estimation of a binary model with a pooled dataset constructed in the following way. Each individual binary choice made at consecutive time intervals for the same household is treated as an independent observation. If a household evacuates in time interval i, that household will have i rows in the dataset, along with the variable values of that household for each time interval. The outcome variable for the first i-1 rows of each household will be coded as 0 for not evacuating. But the outcome variable for the ith row of the household will be coded as 1 for evacuating. For example, if a household evacuates in time interval 3, then there will be three rows of data, with the outcome variable coded as 0 for the first 2 intervals and 1 for the third interval. After pooling the data, we can use existing software for binary choice models to estimate the parameter vector β and α_i. Finally, the unconditional probability of evacuation at each time interval for each individual household will be calculated using equation 2. One extra benefit of this format is that time-dependent variables can be easily accommodated.

3. DATA DESCRIPTION

Data used in this paper were collected in Southwest Louisiana following the passage of Hurricane Andrew through that region in August 1992. After cleaning, the dataset contains data from 428 households of which 156 evacuated. Data collected include household socio-demographic information, type and location of residence, past hurricane experience, perceived risk, the ability to protect property, whether a hurricane evacuation order was received, and the time of evacuation if the household evacuated.

The time of evacuation for each household was reported in terms of four time intervals per day (12 a.m. to 6 a.m., 6 a.m. to 12 p.m., 12 p.m. to 6 p.m., and 6 p.m. to 12 a.m.). Since evacuation lasted for three days in this case, the total number of time intervals reported in this study was 12.

The original dataset only had static variables and lacked the dynamic information regarding the hurricane itself and policy decisions by the authorities during the onset of the storm. Using supplemental information, the data was enhanced by adding hurricane advisory information (time and location of hurricane watches and hurricane warnings), characteristics of the hurricane (the speed, intensity, category, and location of the storm), and distance from storm to each household at every time interval. Most of the information was obtained from the National Hurricane Center.

The timing and the type of an evacuation order play an important role in the evacuation decision [19]. However, this critical information was not available from many of the local authorities in which the data were collected. The only information that was available was whether a household perceived receiving an evacuation order as reported in the survey. No time was associated with the answer to the question and as a result, evacuation order was treated as a static variable in this study although it would normally be an important dynamic variable.

The data was subdivided for analysis and testing. 85% of the data was randomly selected and used for model estimation, while the remaining 15% was used for model validation.

4. MODEL STRUCTURE AND ESTIMATION

A stepwise forward selection process was conducted to find the covariates and their interactions in the model. The eight variables that had levels of significance greater than 5% are listed in TABLE 1.

Among the selected covariates, *dist, TOD* and *speed* are dynamic variables. Distance is not expected to have a linear impact on evacuation; a change of 100 miles when a hurricane is 1,000 miles away is expected to have a very different impact on the decision to evacuate compared to the situation when the hurricane is, say, only 150 miles away. We used the natural logarithm of distance to represent that effect. However, once the distance of a hurricane to a household is within a minimum distance or reaches a threshold, d_{min}, it will be too dangerous to evacuate. Thus, at close distances, the change of distance should no longer have an impact on evacuation. From analysis of the data we found that an appropriate value for d_{min} to be 94 miles. As a result, we chose to use a transformed value of distance, $dist(t)$, as shown below:

$$dist(t) = \begin{cases} 0 & \textit{if } d(t) < d_{min} + 1 \\ \ln[d(t) - d_{\min}] & \textit{otherwise,} \end{cases} \tag{4}$$

where $d(t)$ is the distance between the storm and a household at time t. For the dynamic variable *TOD*, morning is defined from 6 a.m. to 12 p.m., afternoon is from 12 p.m. to 6 p.m., and night is from 6 p.m. to 6 a.m. The covariate *speed* is the rate of forward motion of the hurricane (in miles/hour) in the past time interval. No interactions among the covariates (both static and dynamic) were found to be significant. Every covariate was treated as a generic variable because none of them vary across the two choices (to evacuate or not to evacuate) in a time interval. As a result, we specify that the covariates only appear in the choice to evacuate, and the choice to not evacuate is used as the reference choice with no variables and alternative-specific constant.

The parameters for *hurtrisk* and *protect* are significant and inclusion of these variables in the models clearly improves the model fit. However, these variables are subjective responses provided by respondents to the survey but are not variables that can be predicted for a household. Therefore, they were dropped from the models.

The model estimated was the conditional probability model (equation 1) although the model used for prediction was the unconditional probability model (equation 2). The slope parameters β's are assumed to be the same across the time intervals. TABLE 1 also gives the estimated parameters and the statistics of the model.

Households that have missing information for any of the above covariates were eliminated. As a result, data from 320 households were used in the model estimation. The total number of observations was 3390. The log likelihood at market shares LL(c) and convergence $LL(\beta)$ were -505.5 and -420.0 respectively, with the log likelihood ratio test well below 0.0001. The likelihood ratio index was 0.169. The *p-values* in the table are the probabilities of the Wald test that the parameters are zero. All coefficients of the model have *p-values* significant at 5% level.

In general, all the coefficients of covariates have the right signs, and their values are reasonable. From TABLE 1, among all the covariates in the model, *TOD* has the largest absolute parameters. It has a strong impact on the evacuation decision. A more detailed discussion on this will be presented later. The covariate *mobile* has the second largest parameter, implying that households living in mobile homes are 5.2 times ($e^{1.6496}$) more likely to evacuate than people not living in mobile homes. The covariate *flood* also plays a significant role. If a household lives in a location that it believes flooding is very likely, then such a household is more than twice ($e^{0.7809}$) as likely to evacuate as a household who does not believe flooding is a risk to them. The impact of perceived evacuation order is next to that of *flood* in order of influence. Ideally, covariate *orderper* should have been treated as a dynamic variable; instead it is treated as a static variable because no dynamic information was available for it. Past studies have shown that the parameter of a covariate can vary greatly depending on whether the covariate is treated as a static variable or a dynamic variable [20]. The covariate *dist* is a dynamic continuous variable in the model and the negative coefficient means that the nearer the storm, the more likely a household would evacuate. From the data set used for model estimation, the values of *dist* ranges from 0 to 7 and the odds ratio between the two extremes of *dist* is 270, making *dist* the most influential covariate in the model.

Goodness of fit (GOF) of a model is often expressed as a function of the difference between the observed and the model fitted values. In the model shown in TABLE 1, the likelihood ratio index is such a measure. Alternatively, the Hosmer-Lemeshow test [21 and 22] provides a convenient way to assess a binary model GOF and is available from most popular statistical software. The Hosmer-Lemeshow test organizes subjects into g groups based on the values of their estimated probabilities [23]. For example if g =10, there will be 10 groups and the grouping cut-points can be based on deciles of their probabilities. There are 2 rows for every group; one for outcome=1 and one for outcome=0. The Hosmer-Lemeshow statistic $\hat{C}$ is obtained by calculating the Pearson chi-square statistic from the $g\times 2$ table of observed and model estimated frequencies and is given by

$$\hat{C} = \sum_{k=1}^{g} \frac{(o_k - n_k \overline{\pi}_k)^2}{n_k \overline{\pi}_k (1 - \overline{\pi}_k)}, \tag{5}$$

where n_k is the total number of subjects in the kth group, o_k is the number of choices that are Y=1 in the kth group, and $\overline{\pi}_k$ is the average estimated probability in the kth group. $\hat{C}$ is chi-square distributed with g-2 degrees of freedom. A large $\hat{C}$ will result in a small *p-value*, meaning an inferior GOF. Usually if the *p-value* is smaller than 0.05, we will reject the null hypothesis that the model fits the data well. The appropriateness of the *p-value* depends on the assumption of m-asymptotics, which means the estimated expected frequencies in each cell have to be large. A rule-of-thumb is that the frequency be no smaller than 5. Furthermore the number of groups g should not be smaller than 6.

The contingency table derived in conducting the Hosmer-Lemeshow test is given in TABLE 2. After combining low frequency groups, there are 7 groups (g=7). The Hosmer and Lemeshow statistic is 4.439 with 5 degrees of freedom. The *p-value* is 0.488. As a result, we do not reject the null hypothesis that the model fits well.

5. ANALYSIS AND DISCUSSION

The GOF measure above is for the conditional probability model (equation 1). However, the real model we are interested in is the unconditional probability model (equation 2). Perhaps the most important criterion to assess the GOF of the unconditional model is to study the model prediction with real data. This is done by comparing the observed and model predicted evacuation for each time interval on a separate dataset in which evacuation decisions are known. This process is also referred to as model validation. The aggregation technique in this study is complete enumeration of all households. Fifteen percent of the data were retained for this purpose. However, after eliminating the observations with missing covariates values, only 57 subjects remained in the data set with 20 evacuations. As a result, there were too few observed cases to compare for each time interval. To solve this problem we assumed that the evacuation pattern of the 15% subjects is the same as that of the 100% subjects and used the observed evacuations in the 100% sample to compare with model-predicted evacuation on the 15% data factored up to represent the same number of observations. The factor is the ratio of the number of households in the 100% dataset over the number of households in the 15% dataset. The probability of evacuation for each household in each time interval in the 15% sample was first calculated. Then the probabilities were factored up and summed by time interval and compared to the observed number of evacuations for each time interval in the 100% sample. FIGURE 1 plots the observed and model predicted evacuations for all time intervals. The model clearly reproduces the observed evacuation pattern. The total predicted evacuation over all time intervals is 128.5. This prediction is very close to the observed value of 124. The relative error is only 3.63%. If evaluated at the time interval level, the root-mean-square error (RMSE) is 3.09, and the percent RMSE is 37.10%. The percent RMSE does not include the errors from time interval 5 because the observed value is zero. Therefore the real percent RMSE is somewhat higher. Some intervals have very high relative errors, especially for intervals 1 and 2, with nearly 100% and 50% relative errors. The rest of the intervals have relative errors between 10% and 35%. The maximum absolute error is smaller than 5 for every time interval.

After studying 26 hurricane evacuations, Baker [19] has identified the five most important variables. These variables are

1. Risk level (hazardousness) of the area
2. Action by public authorities
3. Type of housing
4. Prior perception of personal risk
5. Storm-specific threat factors

Comparing the above variables with those in TABLE 1, we find that although the names of the variables between the two groups are different, it is clear that the variables used in this study are the core variables identified by Baker, especially the first four variables by Baker's definition. The last variable, the storm-specific threat factors mentioned by Baker includes the severity of the storm, its proximity, issuance of hurricane watch or warning, and the probability that the path of the hurricane is toward the subject. Limited by data availability and narrow spectrums of change of some variables, we only succeeded in using *dist* to simulate the impact of storm proximity and *speed* to represent the severity of the storm. Hurricane warning was found not significant in the model estimation. Hurricane watch was found significant, but the parameter estimated was counter-intuitive and was dropped from the model.

The joint effects of covariates are more complicated to analyze. To do that we consider four scenarios S1 through S4. We first define a high risk household as a household that lives in a mobile home (*mobile*=1) and is considered very likely to be flooded (*flood*=1) during a hurricane, a low risk household as a household that does not live in a mobile home (*mobile*=0) and is not considered very likely to be flooded (*flood*=0). S1 is the reference scenario, and it is a low risk household who does not receive an evacuation order (*orderper*=0). S2 is the same low risk household as in S1, but the household receives an evacuation order (*orderper*=1). Scenarios S3 and S4 are a high risk household without and with an evacuation order respectively.

TABLE 3 gives hypothetical values for distance and speed of the hurricane for each time interval. These values, along with the scenarios, were used in the following analysis of joint covariate effects.

Based on the information in TABLES 1 and 3, we applied the estimated sequential model to calculate the probabilities of evacuation for every scenario in each time interval. The results are plotted in FIGURE 2.

For a low risk household, the probability of evacuation is much smaller than a high risk household (S1 and S2 vs. S3 and S4). The presence of an evacuation order increases the probabilities of evacuation for either household. For a low risk household, the probability of evacuation increases day by day as the hurricane approaches. However, the high risk household exhibit different patterns of evacuation. Without an evacuation order, the high risk household has a smaller probability to evacuate and tends to delay evacuation until a later time (S3); while with an evacuation order, the same household has high probability to evacuate and tends to evacuate early (S4). Such findings conform to our understanding of hurricane evacuation. The high risk household tends to live near water or low-lying areas and they are the first to feel threatened and probably have longer evacuation distances to negotiate. As a result, they are the first ones to evacuate once an evacuation order is received. Without an evacuation order, the same households tend to wait and see how the situation evolves. The low risk households tend to wait until the last minute to evacuate. The sum of probabilities for all the time intervals for each household is the probability of that household to evacuate during a hurricane. The difference between the sum of probabilities for the high-risk household with and without an evacuation order (98.0% and 92.1%) is smaller than that for the low risk household (31.7% and 23.7%). This indicates that the impact of evacuation order is more significant for low risk households than for the high-risk households. The high-risk households tend to evacuate with or without evacuation orders under the same conditions.

Another important phenomenon is the time-of-day pattern revealed in FIGURE 2. Both types of households exhibit large differences of evacuation probability along the time line. FIGURE 2 gives the probability of evacuation for each time interval of 6 hours for 3 days. Obviously there are interesting patterns to uncover. On the bottom of FIGURE 3, the time intervals are marked by different line types to denote night, morning and afternoon. The graph shows that people are least likely to evacuate during nighttime, there are more evacuations in the morning, but people are most likely to evacuate in the afternoon. The parameters for the two dummy variables of *TOD* show that *TOD* is the most influential covariate describing the evacuation pattern. Its impact is manifested by changing the probabilities of evacuation for different times of the day. Considering scenarios 2 and 4, which include the high and low risk households with an evacuation order, FIGURE 3 plots the model predicted probabilities of evacuation for the households for each of the time intervals with and without the impact of *TOD*. The plots without *TOD* show an increase of probability of evacuation as the hurricane approaches. However, the plots with *TOD* portray a different picture. S4 is more likely to evacuate early while S2 is more likely to evacuate late. However, both display a low probability of evacuation at nighttime, an increasing probability in the morning, and the highest probability in the afternoon.

In a typical hurricane evacuation study, quick, medium or slow loading curves are assumed to predict the percentage of evacuation trips in each time interval. They typically take on a sigmoid shape. However, our study revealed a quite different shape of the loading curve. FIGURE 4 shows the two loading curves. Notice that the unit of the horizontal axes is different. In FIGURE 4(a) it is hour and in FIGURE 4(b) it is time interval. Typical loading curves in (a) are flat at the two ends of the evacuation, and steep in the middle parts, indicating that the majority of evacuations take place in the middle part of evacuation process. However, FIGURE 4(b) gives the observed and predicted loading curves and portrays a different evacuation pattern. The loading curve is flat at the beginning (time interval 1), then increases as day-time approaches (time intervals 2 and 3), it then becomes flat again when night-time approaches (time intervals 4 and 5). This cycle is repeated for each of the three days of evacuation. It is also interesting to notice that the loading curve is the steepest in the day-time (time interval 10 and 11) of the last day for this study. FIGURE 4(b) also shows a cumulative percent evacuation for observed and the model predicted. They are relatively close and show good model fit from another prospective.

6. CONCLUSIONS

This paper describes the application of a sequential logit model to model dynamic travel demand for hurricane evacuation. Using data collected after Hurricane Andrew in Southwestern Louisiana, the model was estimated on 85% of the data and then tested on the remaining 15% of the data. The estimated model provided satisfactory goodness of fit with a likelihood ratio index based on market shares of 0.169. The variables in the model were similar to the variables found to be most significant in describing evacuation behavior in other studies [19]. When the model was used to predict evacuation behavior on the 15% testing data set, it reproduced observed evacuation behavior with an RMSE of 3.09 evacuations. Based on the above performance, the method appears capable of modeling dynamic evacuation demand satisfactorily.

When the model was applied to hypothetical scenarios involving households of different housing type (i.e. mobile home or other), different flooding potential, and either an evacuation order was issued or not, the results are intuitively what would be expected. Households of higher risk (mobile home dwellings and those vulnerable to flooding) showed higher rates of evacuation and a tendency to evacuate earlier than those of lower risk. Evacuation orders accelerate evacuation so that evacuations occur earlier than when evacuation orders are not given; evacuation orders also result in more households evacuating.

While the model produced plausible results it was based on only one dataset with some serious limitations. For example, it was not known when an evacuation was ordered in each parish in which the survey was conducted; only whether the respondents perceived that an evacuation order was issued or not. In addition, only one storm was observed so the impact of storm intensity, size, and track could not be included in the model. Intuitively, these factors will affect evacuation behavior but more extensive data involving multiple storms would be needed to estimate these effects. A larger dataset might improve the model and provide deeper insights. For example, a larger dataset would enable us to test the validity of the model at zonal level by each time interval, instead of aggregating all zones together. Then the model would be more helpful predicting the impact of different policies on hurricane evacuation and provides better decision support for planning and management to local authorities.

ACKNOWLEDGEMENT

The research on which this paper is based was funded by the Louisiana Transportation Research Center (LTRC) in an ongoing research project on the modeling of hurricane evacuation traffic. The authors wish to acknowledge their initiative in launching and funding this research. The data from southwestern Louisiana used in the study was graciously provided by Dr. Jeanne S. Hurlbert, Louisiana State University. The authors would like to thank Dr. Luis Escobar of Louisiana State University, who provided valuable insights and help during the execution of the research reported in this paper.

REFERENCES

1. Friesz, T., Luque, J., Tobin, R. and Wie, B. Dynamic Network Traffic Assignment Considered as A Continuous Time Optimal Control Problem. Operations Research. No. 34, 1989, pp. 839-901.
2. Mahmassani, H., Peeta, S., Hu, T., and Ziliaskopoulos, A. Algorithm for Dynamic Guidance in Congested Networks with Multiple User Information Availability Groups. Proceedings of the 26th International Symposium on Automotive Technology and Automation, Aachen, Germany, 1993, pp. 273-280.
3. Janson, B. Dynamic Traffic Assignment for Urban Road Networks. Transportation Research, Vol. 25B, Nos. 2/3, 1991, pp. 143-161.
4. Ran, B. and Boyce, D. Modeling Dynamic Transportation Networks: An Intelligent Transportation System Oriented Approach. Second edition, Springer, 1996.
5. Ziliaskopoulos, A. and Peeta, S. Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. 2002 Annual TRB meeting.
6. Sheffi, Y. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall, 1985.
7. Mei, B. Development of Trip Generation Models of Hurricane Evacuation. Master thesis, Louisiana State University, Baton Rouge, Louisiana, 2002.
8. Irwin, M. and Hurlbert, J. A Behavioral Analysis of Hurricane Preparedness and Evacuation in Southwestern Louisiana. Louisiana Population Data Center, September, 1995.
9. Regional Development Service (RDS), Department of Sociology, and Department of Economics. Executive Summary of A Socioeconomic Hurricane Impact Analysis and A Hurricane Evacuation Impact Assessment Tool (Methodology) for Coastal North Carolina: A Case Study of Hurricane Bonnie. East Carolina University, Greenville, NC, July, 1999.
10. Johnson, J. and Zeigler, D. Modeling Evacuation Behavior during the Three Mile Island Reactor Crisis. Socio-Econ. Plan. Sci. Vol. 20, No. 3, 1986, pp. 165-171.
11. Leik, R., Carter, T. and Clark, J. Community Response to Natural Hazard Warnings. University of Minnesota, Minneapolis, MN, 1981.
12. Lewis, D. Transportation Planning for Hurricane Evacuations. ITE Journal, August, 1985, pp. 31-35.
13. Tweedie, S., Rowland, J., Walsh, S. and Rhoten, R. A Methodology for Estimating Emergency Evacuation Times. The Social Science Journal, Vol. 23, No. 2, 1986, pp. 189-204.
14. U.S. Army Corps of Engineers. Alabama Hurricane Evacuation Study Technical Data Report: Behavioral Analysis. Final Report, 2000.
15. Agresti, A. Analysis of Ordinal Categorical Data. John Wiley & Sons, New York, 1984.
16. Sheffi, Y. Estimating Choice Probabilities Among Nested alternatives. Transportation Research, B, Vol. 13B, pp. 189-205.

17. Amemiya, T. Qualitative Response Models. Annals of Economic and Social Measurement, No. 4, 1975, pp. 363-372.
18. Fahrmeir, L. and Tutz, G. Multivariate Statistical Modeling Based on Generalized Linear Models. 2nd edition, Springer, New York, 2001.
19. Baker, J. Hurricane Evacuation Behavior. International Journal of Mass Emergencies and Disasters, Vol. 9, No. 2, 1991, pp. 287-310.
20. Christensen, E., Schlichting, P., Andersen, P., Fauerholdt, L., Schou, G., Pedersen, B., Juhl, E., and Tygstrup, N. Updating Prognosis and Therapeutic Effect Evauation in Cirrhosis with Cox's Multiple Regression Model for Time-Dependent Variablbes. Scand, J. Gastroenterol, No. 21, 1986, pp. 163-174.
21. Hosmer, D. and Lemeshow, S. A Goodness-of-fit Test for the Multiple Logistic Regression Model. Communications in Statistics, A10, 1980, pp. 1043-1069.
22. Lemeshow, S. and Hosmer, D. The Use of Goodness-of-fit Statistics in the Development of Logistic Regression Models. American Journal of Epidemiology, No.115, 1982, pp. 92-106.
23. Hosmer, D. and Lemeshow, S. Applied Logistic Regression. 2ed edition, John Wiley & Sons, New York, 2000.

LIST OF TABLES AND FIGURES

TABLE 1 Covariates and Summary Results of the Model

covariate	definition	logit model		
		β	se(β)	p-value
intercept	Model constant	-2.8238	0.9123	0.002
dist	A function of distance to the storm at time t.	-0.7995	0.1144	0.0001
TOD(1)	Time-of-day. Periods used – night, morning, and afternoon. Two dummy variables used.	1.4512	0.3096	0<0.0001
TOD(2)		2.0244	0.2811	0<0.0001
speed	Forward speed of the hurricane at time *i* (miles/hour).	0.1463	0.0691	0.0342
orderper	1 if perceived receiving an evacuation order, 0 otherwise.	0.5401	0.218	0.0132
flood	1 if the residence is believed very likely to be flooded, 0 otherwise.	0.7809	0.2276	0.0006
mobile	1 if a mobile home, 0 otherwise.	1.6496	0.2293	0<0.0001
hurtrisk	1 if a serious risk of being hurt is perceived, 0 otherwise.	Not used in the model		
protect	1 if staying home was considered necessary to protect property, 0 otherwise.	Not used in the model		

TABLE 2 Contingency Table

group	evacuate=0		evacuate=1		total
	observed	expected	observed	expected	
1	1381	1380	7	8	1388
2	333	333	5	5	338
3	334	331	4	7	338
4	329	329	10	10	339
5	326	325	13	14	339
6	309	317	30	22	339
7	262	259	47	50	309

TABLE 3 Hypothetical Values in Analyzing Covariate Effects

time interval	1	2	3	4	5	6	7	8	9	10	11	12
distance (mile)	1182	1096	1004	911	815	713	607	500	398	305	218	146
dist	7.0	6.9	6.8	6.7	6.6	6.4	6.2	6.0	5.7	5.4	4.8	3.9
speed (mph)	12.0	12.5	13.0	13.5	14.0	14.5	15.0	16.0	17.0	18.0	19.0	20.0

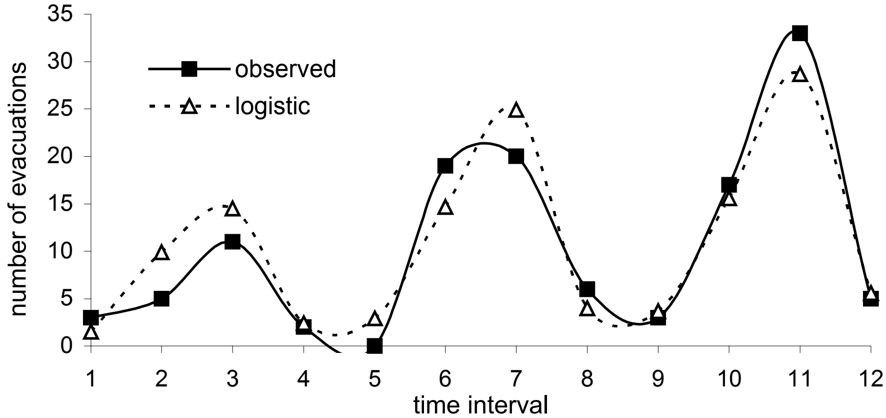


FIGURE 1 Observed vs. model predicted evacuation.

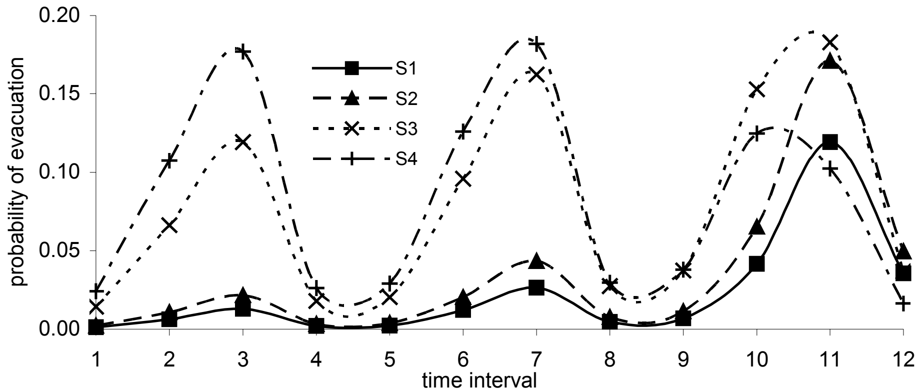


FIGURE 2 Model estimated probability of evacuation for four scenarios.

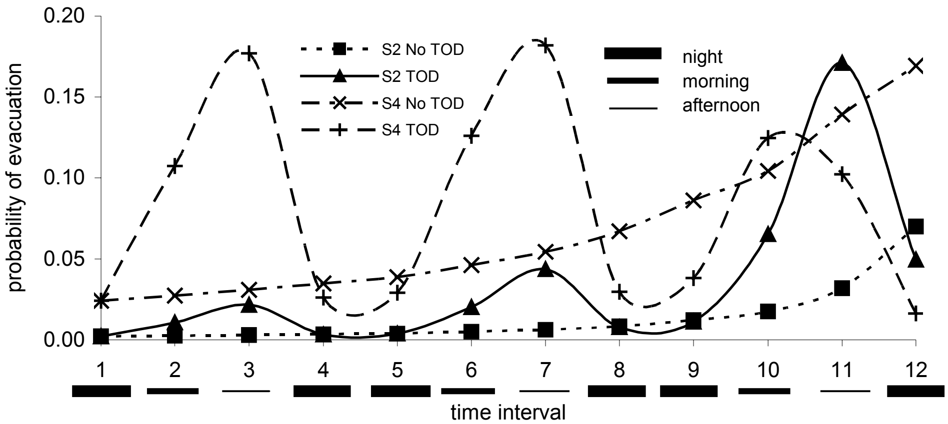
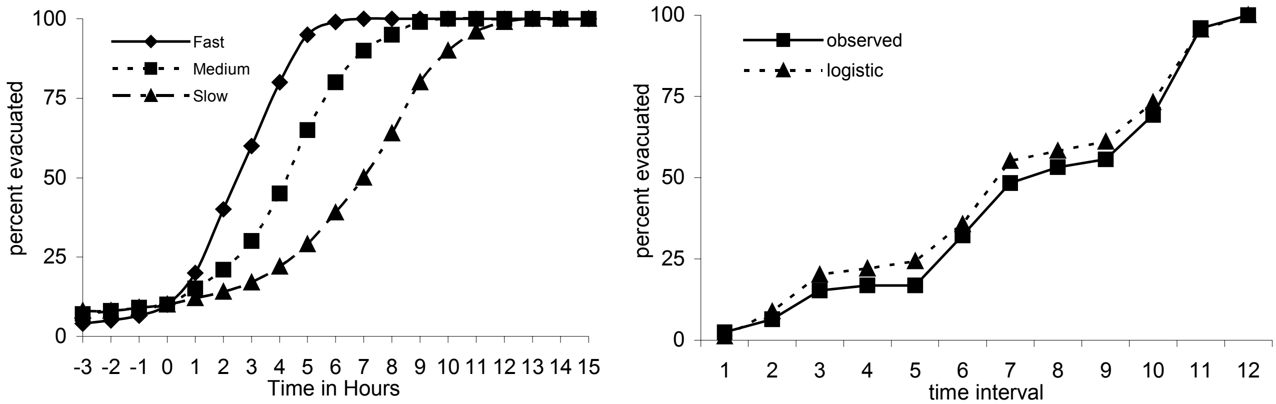


FIGURE 3 Impact of *TOD*.



(a) Typically Assumed [14] (b) Observed and Predicted

FIGURE 4 Comparison of loading curves.