

Comparison of Alternative Trip Generation Models for Hurricane Evacuation

Chester G. Wilmot¹ and Bing Mei²

Abstract: The purpose of this study was to compare the relative accuracy of alternative forms of trip generation of evacuation traffic. Conventional participation rate, logistic regression, and various forms of neural network models were estimated and tested using a data set of evacuation behavior collected in southwest Louisiana following Hurricane Andrew. The data set was divided into a 350-household data base on which the logistic regression and neural network models were estimated, and a separate 60-household data base on which all models were tested. Limited and comprehensive model inputs were tested among the neural network models to determine whether more comprehensive specifications enhance the performance of the models. It was found that the limited specification performed almost as well as the more detailed specification. Comparison of the performance of the models considered in this study showed that the logistic regression and neural network models were able to predict evacuation more accurately than the participation rate model.

DOI: 10.1061/(ASCE)1527-6988(2004)5:4(170)

CE Database subject headings: Hurricanes; Evacuation; Traffic management; Comparative studies.

Introduction

The modeling of the evacuation process generated by an approaching hurricane is important for those responsible for the efficient and safe movement of people during evacuation. It can identify bottlenecks in the system before they are experienced in an evacuation, it can be used to determine the impact of road closures due to flooding, the impact of phased evacuation on network loading, and many other possible consequences of an evacuation event. Being able to model alternative evacuation scenarios can lead to the establishment of appropriate evacuation policies, strategies, and contingency plans.

Estimation of evacuation demand is the first step in the modeling of the evacuation process and yet it is the part of the evacuation modeling process that has received the least attention in the past. As noted by Ziliaskopoulos and Peeta (2002) in their review on dynamic traffic assignment (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." The research reported in this paper is an initial attempt at addressing that need by reviewing alternative models of evacuation trip generation and reporting their relative performance.

¹Associate Professor, Louisiana Transportation Research Center and Dept. of Civil and Environmental Engineering, Louisiana State Univ., Baton Rouge, LA 70803. E-mail: cewilm@eng.lsu.edu

²Transportation Planner, Wilbur Smith Associates, Atlanta, GA 30341. E-mail: bmei@wilbursmith.com

Note. Discussion open until April 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on January 27, 2003; approved on June 15, 2004. This paper is part of the *Natural Hazards Review*, Vol. 5, No. 4, November 1, 2004. ©ASCE, ISSN 1527-6988/2004/4-170-178/\$18.00.

Current Practice

Interest in evacuation modeling began in the 1970s with an initial emphasis on hurricane evacuation (Urbanik 1978; COE and SWFRPC 1979). However, after the nuclear accident at Three Mile Island in 1979, the emphasis moved to evacuation from nuclear sites (Urbanik and Desrosler 1981; Sheffi et al. 1982; KLD Associates 1984; HMM Associates 1980). Interest returned to hurricane evacuation in 1990s, primarily because the most expensive hurricane in United States history in terms of damage, Hurricane Andrew, occurred in 1992 and the hurricane that generated the greatest evacuation in United States history, Hurricane Floyd, occurred in 1999. State Departments of Transportation have become more involved in emergency evacuation in recent years and greater attention is being given to the transportation aspects of evacuation than ever before.

Existing evacuation models are typically computerized software packages in which the major emphasis is on traffic assignment. In most of these evacuation models, traffic assignment is conducted using simulation. Sometimes, existing microsimulation packages such as *NETSIM* have been used (Urbanik and Desrosler 1981; HMM Associates 1980). However, in most cases, customized assignment procedures have been developed. For example, *NETVAC*, one of the early macrosimulation evacuation models was developed specifically to model traffic evacuating from nuclear sites (Sheffi et al. 1982). *DYNEV*, developed at approximately the same time (early 1980s), is also a macroscopic simulation model developed to model evacuation from nuclear power plants in the case of an emergency (KLD Associates 1984). *MASSVAC*, using macroscopic traffic flow relationships, was developed to expand consideration to other disasters and was used to model hurricane evacuation of Virginia Beach (Hobeika and Jamei 1985; Hobeika et al. 1985). Another evacuation modeling package, *HURREVAC*, was developed specifically to aid Emergency Managers in making evacuation decisions in the face of an oncoming storm (COE 1994).

A recent addition to the evacuation modeling packages is the Oak Ridge Evacuation Modeling System which can be used to

model evacuation from a variety of disasters (ORNL 1999). Input to the model is time-dependent travel demand with or without specification of the destination of the demand. When origin–destination information is not provided, the model uses a combined distribution–assignment procedure to distribute traffic to destinations. Macrosimulation, based on the *TRAF* family of models, developed over several years for the Federal Highway Administration to ultimately produce *CORFLO*, was adapted for use in this model to simulate travel on the network. The model estimates link flows, speed, and percent evacuated at each time period as well as the total evacuation time. However, like all the other evacuation models mentioned above, time-dependent travel demand must be provided as input.

A model package that does estimate evacuation demand, although it does not estimate it by time period, is the evacuation traffic information system (ETIS) (PBS&J 2000a; Lewis 2001). The ETIS uses default evacuation participation rates by county dependent on the category and speed of the hurricane, tourist occupancy, and type of housing. The model is web-based and operates within a geographic information system environment providing graphical output in the form of maps, diagrams, and tables. The model is proprietary and was developed specifically for the southeastern United States including the states of Florida, Georgia, South Carolina, and North Carolina. In its original form, it considered evacuation flows on interstate routes only but newer versions of the model include other evacuation routes as well. The model is being expanded to include other neighboring states.

Currently, when time-dependent travel demand is estimated, it is estimated in two steps. In the first step, the number of households expected to evacuate are estimated using evacuation participation rates. Participation rates are typically specified by evacuation zone where evacuation zones are geographic areas distinguished by their vulnerability to hurricanes of different strengths. Within each evacuation zone, participation rates typically vary by type and speed of the hurricane, the proportion of the population that are transients, and type of housing (i.e., mobile homes versus other types of structures). Participation rates are highest in those zones that are most vulnerable and these rates are further inflated by stronger and slower hurricanes (because they cause greater flooding), higher proportion of transients, and an increased number of mobile homes. Multiplying the population in each evacuation zone by the appropriate participation rate provides an estimate of the total number of evacuees from an area. In the second step, the time at which the evacuees are expected to begin their evacuation is estimated. This is typically done using a so-called response or mobilization curve that estimates what proportion of the total evacuation demand initiates their evacuation within each time period during evacuation. These curves are based on past evacuation behavior and are often described as mathematical functions to facilitate their application (COE 2000).

The use of participation rates to estimate total travel demand is similar to the use of cross-classification in conventional urban transportation planning, since average trip generation rates are used for categories of users in each case. However, in the preparation of participation rates, statistical analysis of past data has typically not been conducted, as in conventional cross-classification models, but has usually involved the estimation of participation rates using subjective judgment based on the experience of the analyst from past storms.

An alternative form of trip generation for hurricane evacuation that has been used by a few analysts is logistic regression (Irwin and Hurlbert 1995; RDS 1999). Logistic regression is used in place of regular linear regression in evacuation because the de-

pendent variable is a binary variable (evacuate or not) and this introduces violations of some of the assumptions underlying regular linear regression analysis. One study used data from Hurricane Andrew and developed a model including variables such as the respondent's perception of being hurt if staying, the perceived ability of the home to withstand the storm, prior hurricane experience, gender, marital status, education, age, and race of the owner (Irwin and Hurlbert 1995). Another study, using data from Hurricane Bonnie, found that the most influential variables were whether an evacuation order had been issued or not, the risk of flooding, whether the neighbors evacuated or not, and type of structure of the home (RDS 1999). The common practice in the above two models is that some subjective perceptions were included as independent variables because they were found to be statistically significant. However, it must be recognized that while these variables may contribute to explaining evacuation behavior, they are not good variables for forecasting because they cannot be measured and cannot be predicted in hypothetical scenarios in which the analyst may want to apply the model.

Review of the literature reveals that while participation rate models and logistic regression have been used to estimate evacuation demand in the past, no attempt has been to determine how well they perform. The study described in this paper compares the relative performance of several models in an effort to identify superior methods of evacuation travel demand estimation.

Methodology

The methodology employed in this study was to use a portion of a data set of past hurricane behavior to estimate a variety of evacuation travel demand model types, test each model's ability to reproduce observed evacuation decisions on the remaining portion of the data set, and compare their performance. The types of models considered included a participation rate model, a logistic regression model, and three different types of neural network models. The participation rate model was estimated independently by consultants for the region (PBS&J 2000b, 2001). The other models were estimated on data collected following Hurricane Andrew as described in the following section. For the models estimated on the Hurricane Andrew data, the practice often employed in the estimation of neural network models of randomly dividing the data into an estimation (or training) data set and a testing data set, was used. All models were tested on the testing data set.

The participation rate model was estimated based on experience from previous hurricanes and on knowledge of the area (PBS&J 2000b, 2001). Participation rates based on flooding potential of the home, type of housing, and hurricane intensity were developed (PBS&J 2000b, 2001). Flooding potential was described in terms of whether a household would be threatened with flooding based on the category of storm. Thus, a household located in an area (evacuation zone) which could flood with, say, a Category 1 hurricane, would have a higher flooding potential and therefore a higher participation rate than a household that was located in an area that was generally only vulnerable to flooding with a hurricane of Category 2, etc. Type of housing distinguished between single-family dwelling units, multiple-family dwelling units, tourist accommodation, and mobile homes. The model catered for hurricane intensities ranging from Category 1 to 5 on the Saffir–Simpson scale.

The logistic regression model developed in this study was a disaggregate model describing the likelihood that a household

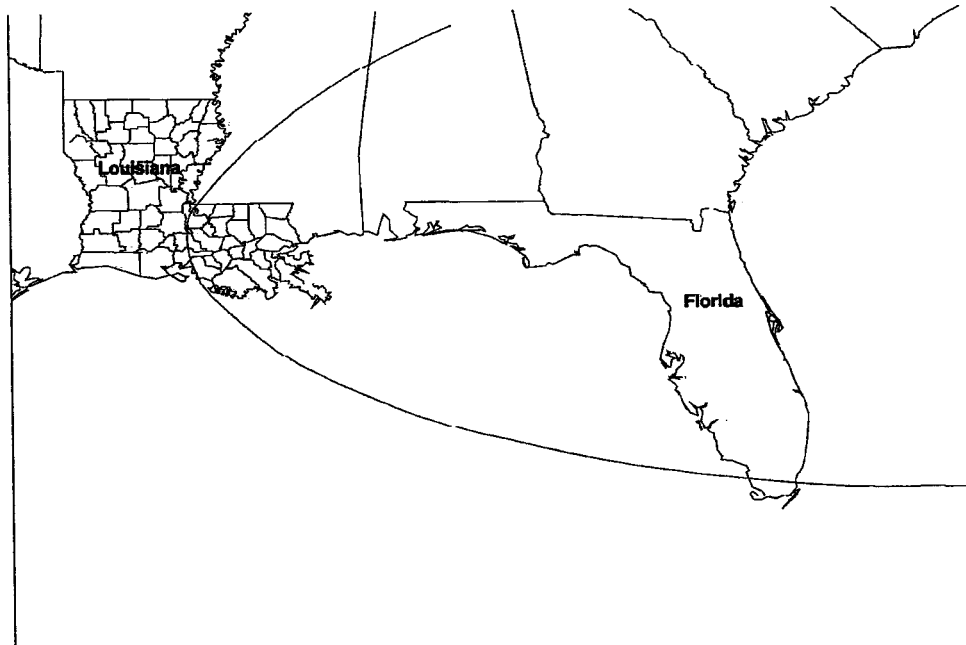


Fig. 1. Track of Hurricane Andrew, 22–27 August 1992

would evacuate based on the characteristics of the household, the location of the home with respect to the closest body of water, and whether an evacuation order was issued or not. The model can be described as

$$y = \frac{e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}} \quad (1)$$

where y = likelihood that a household evacuates; x_1, x_2, \dots, x_n = independent variables; and $\beta_0, \beta_1, \dots, \beta_n$ = parameters.

The ratio of the likelihood that a household evacuates to the likelihood that it does not is the odds $y/(1-y)$. The log of the odds

$$y' = \ln[y/(1-y)] \quad (2)$$

is a linear expression of the independent variables and parameters

$$y' = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (3)$$

in which the parameters can be estimated using maximum likelihood.

The goodness of fit of the logistic regression model was evaluated using the likelihood ratio index, the Hosmer–Lemeshow test, and the receiver operating characteristic (ROC) curve (Hosmer and Lemeshow 2000). The performance of the model was also assessed using the percent correctly predicted (PCP) and root mean square error (RMSE).

Three different kinds of neural network models were estimated and tested in this study. First, a feed-forward model using a back-propagation learning algorithm was employed. This is the most common form of neural network, and is termed a feed-forward neural network (FFNN) in this study. Second, a probabilistic neural network (PNN), which is a pattern classification neural network, was used (Tsoukalis and Uhrig 1997). The PNN uses Bayesian statistics to estimate the probability that a given input vector belongs to a category in a classification scheme. In this case it was used to estimate the probability that a respondent evacuated or not. Third, a learning vector quantizer (LVQ) neural network was applied. It also is a classification-type neural network but

uses a learning procedure adapted from Kohonen's self-organizing mapping network approach (SOM) (Tsoukalis and Uhrig 1997). In a Kohonen SOM, classification is accomplished entirely on the basis of the nature of the data. Cases are assigned to groups based on their similarity to the collective nature of cases in each group. Cases are considered sequentially with each group competing for each case. The group which is the most similar to that case acquires the case and characteristics of the group are updated to be the collective characteristics of the entire membership of that group. Cases can migrate in and out of groups but in the end similar groups are identified. This is the basis for the term self-organizing which is sometimes also referred to as unsupervised learning to distinguish it from the supervised learning that occurs when learning data is available. In the LVQ model, the Kohonen training law is adapted to use knowledge of the true classification of cases to attract those that do belong to the winning group and repel those that do not.

The statistical package for social scientists (SPSS) was used to estimate the logistic regression model. The Neural Network Toolbox in *MATLAB*® was used to estimate the neural network models. In all neural network models the number of neurons in the input layer was the number of input variables. The same variables used in logistic regression were used in all neural network models although a more comprehensive set of variables were also tested to see if they would improve the performance of the models significantly. The performance of all models was assessed using the PCP and RMSE on the testing data set.

Data

The data used to estimate and test the models of evacuation travel demand in this study was a household survey conducted in southwestern Louisiana in 1995 in which information related to Hurricane Andrew was gathered. Hurricane Andrew was a Category 3 hurricane which passed just south of Miami on 23 August 1992, causing some of the greatest damage caused by a hurricane in

Table 1. Characteristics of Survey Data

Variable	Classification	Number	Percent (%)
Evacuated	Yes	194	42
	No	272	58
Housing type	Single-family dwelling	359	77
	Multifamily dwelling	39	8
	Mobile home	68	15
Hurricane experience	Yes	341	73
	No	125	27
Race	White	384	83
	Black and other	78	17
Evacuation order given	Yes	113	26
	No	317	74
Distance to water	≤1.6 km	201	47
	>1.6 km	227	53
Home in flood zone	Yes	55	12
	No	398	88
Children in household	Yes	227	56
	No	176 mi.	44 mi.

United States history. It passed into the Gulf and then made landfall again in Louisiana on 25 August. See Fig. 1 for the track of the hurricane.

The survey was conducted in southwestern Louisiana in September 1995 using computer-assisted telephone interviewing. A total of 651 households were surveyed of which 466 were living in a parish affected by Hurricane Andrew or had evacuated because of it. Some households were not at home at the time of the hurricane or had subsequently moved into the area. Only the 466 households were included in the study. Some characteristics of this sample are shown in Table 1. Missing data result in some classifications not totaling to 466. Variables included in the sample but not shown in the table include income, marital status, education, age of the respondent, household size, and ownership of the residence.

With the exception of household income, age, and household size, all variables in the data set were descriptive variables. The descriptive variables were translated into dummy variables for numerical analysis. This was done by establishing a dummy (or binary) variable for each classification of a descriptive variable in excess of one. That is, for a descriptive variable with two categories, a single dummy variable was created; for a variable with three categories, two dummy variables were created to represent two of the three classifications, with the one left used as a reference, and so on.

Missing data resulted in a total sample of 410 households for which no missing data, with the exception of household income data, were present. Of the 466 households, 121 did not provide information on household income, which would have reduced the number of useful observations considerably if the missing values on household income were allowed to eliminate observations from study sample. However, models estimated using household income found that it was not a significant factor in the presence of the other variables in the model. Subsequently, the 410 household sample, which included some cases with missing data on household income, was adopted for the study. This sample was randomly divided into two subsets: one containing 85% of the sample (350 households) used to estimate or train the models, and the other containing the remaining 15% (60 households) on which the performance of the models were tested.

Analysis

Model Estimation

The 350-household training data set was used to estimate all the models used in the comparison in this study with the exception of the participation rate model which was estimated subjectively by consultants for the area based on past evacuation behavior (PBS&J 2000b, 2001). Initially, correlations among variables were estimated to observe which variables were correlated with the decision to evacuate or not and whether any variables were highly correlated with each other. The former indicated variables that were likely to be good candidates for inclusion in any of the models in the study, while the latter indicated variables that had the potential to replace each other or produce estimation problems if both were included (e.g., produce multicollinearity in the regression equation or make it more difficult to establish stable weights in the neural network models).

In the logistic regression model, individual variables were assessed for inclusion using a forward stepwise variable selection method based on the likelihood ratio statistic. In this approach, the difference in log likelihood with and without an individual variable is used to decide on its inclusion or exclusion. The likelihood ratio statistic method of variable selection is recommended for models with dummy variables (Hosmer and Lemeshow 2000). The results of the variable selection process are displayed in Table 2. Among the dummy variables, the categories of variables that form the reference level for other categories and are not self-evident from the names of the included categories, are multifam-

Table 2. *p* Values for Forward Stepwise Variable Selection in Logistic Regression Model

Variable	<i>df</i>	Step					
		0	1	2	3	4	5
(1) a. Home is a mobile home	2	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
b. Home is a single-family dwelling unit							
(2) Mandatory evacuation order was given	1	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
(3) Age of respondent (years)	1	0.003	0.033	0.033	0.003	0.002	0.001
(4) Distance to nearest body of water ≤1 mi.	1	0.028	<0.001	0.012	0.009	0.009	0.004
(5) a. Respondent is single and never married	2	0.891	0.729	0.849	0.108	0.05	0.05
b. Respondent is married or living with partner							

Table 3. Logistic Regression Model Estimation Results

Variable	β	Standard error	Wald	d.f.	Significance
Constant	1.80	0.77	5.44	1	0.020
Home is a mobile home	2.32	0.66	12.44	1	0.000
Home is a single-family dwelling unit	-1.05	0.46	5.16	1	0.023
Mandatory evacuation order was given	1.44	0.31	22.25	1	0.000
Age of respondent (years)	-0.04	0.01	13.64	1	0.000
Distance to nearest body of water ≤ 1 mil.	0.80	0.28	8.18	1	0.004
Respondent is single and never married	-1.26	0.54	5.47	1	0.019
Respondent is married or living with partner	-0.80	0.39	4.29	1	0.038

ily dwelling units (e.g., apartments) for type of housing, and divorced, separated, or widowed respondents for the marital status variables.

Different classification schemes for descriptive variables with more than two categories were tested and the scheme that produced the best fit of the model adopted. Linearity of the variable age (the only continuous variable in the final model specification) was tested using the Box-Tidwell transformation term, $x \ln x$, in the model specification while observing the significance of the coefficient of the term. The coefficient was found to not be significantly different from zero in this analysis, indicating that the assumption of linearity of the age variable was justified. Interaction among the independent variables was tested using the traditional method of introducing products of variables and observing the significance of their estimated coefficients. No significant interaction was detected among any of the variables.

The best model that resulted from this process is the model shown in Table 3. A likelihood ratio index value of 0.25 was obtained which suggests that a reasonably good fit of the model to the data was obtained. The Hosmer-Lemeshow test failed to reject the hypothesis that the estimated and observed probabilities of evacuation are the same, and the area under the ROC was 0.81, indicating excellent discrimination ability of the model (Hosmer and Lemeshow 2000).

Since the logistic regression model produces probabilities, it was necessary to decide the cutoff point on the probability that distinguishes cases that evacuate from those that do not. This was achieved by applying different cutoff points and observing the performance of the model based on the percent correctly predicted on the model estimation data set. The optimum cutoff point was found to be at a probability of 0.36. That is, model estimates below 0.36 were classified as not evacuating while those equal to or above that value were considered as evacuating. Using this cutoff point, the highest overall percent correctly predicted achieved using the training (estimation) data was 74%.

In logistic regression the coefficients of dummy variables can be conveniently interpreted through a measure called the odds ratio (OR). The odds ratio is defined as the ratio of the odds when the dummy variable equals one to the odds when the dummy variable equals zero. For example, in a logistic regression model with one dummy variable, x , the OR would be

$$OR = \frac{y_{x=1}/(1-y_{x=1})}{y_{x=0}/(1-y_{x=0})} = \frac{\left(\frac{e^{\beta_0+\beta_1x}}{1+e^{\beta_0+\beta_1x}}\right)}{\left(\frac{1}{1+e^{\beta_0+\beta_1x}}\right)} = e^{\beta_1} \quad (4)$$

Thus, the OR of a variable is the exponential value raised to

the power of the coefficient of the variable. For the results in Table 1, this means, for example, that mobile home dwellers are $e^{2.32}$ or 10.1 times more likely to evacuate than those living in multiple-family dwellings (the reference category), or those living in single-family dwelling units are $e^{-1.05}$ or 0.35 times more likely to evacuate than those living in multiple-family dwelling units. Issuing an evacuation order, increases the probability of evacuation 4.2 times over what it is without an evacuation order. Those living within 1.6 km of water are 2.2 times more likely to evacuate than those that live further away. Confidence limits can also be estimated from the standard error of the estimate of the parameter assuming a normal distribution of the estimate.

Evaluating the estimates of the coefficients in Table 3 in terms of sign, magnitude, odds ratio, and past study findings, the estimates seem reasonable. In line with previous studies, type of housing, evacuation orders, and location with respect to water (flooding), are the most significant factors. Age and marital status are less influential. According to the model, the probability of a household evacuating diminishes with an increase in age of the respondent and divorced, separated, and widowed persons are more likely to evacuate than others. While it is intuitively expected that older people would be more reluctant to evacuate and leave their home than younger people, it is more difficult to understand why marital status proves to be significant, and why divorced, separated, and widowed persons are more likely to evacuate than others. One argument may be that, as a group, divorced, separated, and widowed persons are less home-conscious or less home-bound and, therefore, are less concerned about protecting their homes by staying. Another possibility is that they may consider their homes less as a place of refuge than other groups. The marital status variables were retained in the specification in this study because they were significant, but their inclusion is tentative and it may be argued that they should have been omitted in the absence of strong theoretic reasons for their inclusion.

The neural network models were trained (estimated) using two alternative sets of input variables. First, the same variables used in the logistic regression model were used in the neural network models. Second, in an effort to see whether the neural network models could capture explanatory power from the remaining variables, a comprehensive set of 13 variables were used as input. Training and testing were conducted with both sets of input variables for all neural network models but the improvement in model performance was less than 5% on all measures of goodness-of-fit and model performance. For this reason, the more comprehensive specification was abandoned and only the simpler specification was used.

In the FFNN model, the number of neurons in the hidden layer was determined by trial and error based on the percent correctly predicted in the training data set. A tangent sigmoid transfer func-

Table 4. Model Estimation Results

Model	ρ^2	Hosmer–Lemeshow test (<i>p</i> value)	Likelihood-ratio test (<i>p</i> value)	The area under ROC curve	Percent correctly predicted		
					Evacuated (%)	Did not evacuate (%)	Overall (%)
Logistic regression	0.25	0.049	<0.001	0.81	74	74	74
FFNN	—	—	—	0.83	77	80	79
PNN	—	—	—	—	93	100	97
LVQ	—	—	—	—	67	83	76

tion was used in the hidden layer because many of the input variables were binary values and the tangent sigmoid is the preferred function under such conditions (Stornetta and Huberman 1987). The log sigmoid transfer function was used in the output layer so that the output, the probability of evacuation in the single output neuron, would be in the range zero to one. The output layer of the FFNN contained a single neuron in which the probability of evacuation was produced. The optimum cutoff point on the probability, distinguishing between evacuation or not, was identified by running the model repeatedly with alternative cutoff points and observing the number of correctly predicted cases at each level. The optimum cutoff point was found to be at a probability of 0.37. The area under the ROC curve was 0.83 indicating an excellent fit to the data. All the results are shown in Table 4.

In the PNN model, the number of neurons in the pattern layer was equal to the number of cases in the training data set. An exponential activation function was used in the pattern layer. The summation layer had one neuron for each category of the dependent variable (i.e., 2), and the output layer consisted of one output neuron in which the case was classified as either evacuating or not. The PNN model is not able to produce an ROC curve since it is a classification-type model. Estimation results are shown in Table 4.

In the LVQ model, the number of neurons in the Kohonen layer were determined by trial and error based on the percent correctly predicted. The output consists of a classification into either evacuating or not evacuating. The LVQ model is not able to produce an ROC curve since it is a classification-type model. The percent correctly predicted was considerably lower among those that evacuated than among those that did not (see Table 4). While this is a tendency in all the models tested in this study, it was more pronounced in the LVQ model.

Model Comparison

To compare the performance of the models, two alternative approaches were used. First, the ability of the models to reproduce the observed number of evacuations from each parish was assessed. That is, the aggregate number of households estimated by the models to evacuate from each parish were compared with observed aggregate evacuation from each parish. This measures model performance at the level at which evacuation volumes are often estimated. In this approach, model performance was measured in terms of RMSE, a general measure of the average error among a set of observations. Second, the ability of the models to accurately predict individual household behavior was assessed. That is, for each household in the testing data set, the prediction of the model was compared with the observed values. For this test, performance was assessed in terms of RMSE and PCP.

The results of the parish-level evacuation demand estimates

are shown in Table 5. Ideally, these estimates would have been based on the testing data set alone in order that assessment of the models could be made on data other than that on which they were calibrated. However, there were only 60 households in the testing data set (of which only 22 evacuated), and 19 parishes are involved. This would have resulted in a very sparse distribution of households. For this reason, both the testing and training data sets were combined to test the parish-level evacuations reported in Table 5.

The results show that the models reproduce evacuation demand from the different parishes with reasonable accuracy, although the performance varies among the models and some models tend to overpredict while others underpredict. The correlation shown in the table is the Pearson coefficient of correlation between the estimated and observed parish evacuations. The square of this value describes the proportion of the variation in the observed values, explained by the model. Subsequently, the participation rate model describes 76% of the variation at the aggregate level (0.87^2), the logistic regression model 86%, the FFNN model 92%, the PNN model 96%, and the LVQ model 90%. The average error, as expressed by the RMSE, varies from 9.9 evacuations per parish to between approximately two and four evacuations per parish for the logistic regression and neural network models.

In the second test of the performance of the models, the ability of the models to accurately reproduce individual household's evacuation decision, the manner in which the participation rate model is normally applied, had to be altered to permit individual household predictions. In their normal application, participation rate models are aggregate models that estimate the proportion of households in an evacuation zone, facing a particular storm, and living in a particular type of house that are expected to participate in the evacuation process. No attempt is made to predict the evacuation behavior of individual households. However, if households of the same type in the same evacuation zone are considered to be uniform in their evacuation behavior, then the participation rate could be considered the probability that each household in that group will evacuate. This allows households to be ascribed probabilities of evacuation based on their characteristics and the evacuation zone in which they are located. In this study, since Hurricane Andrew was a Category 3 storm at the time it made landfall in Louisiana, participation rates for a Category 3 hurricane were used in the participation rate model. Participation rates of 3, 5, and 100% were predicted by the participation rate model for the respondents in the sample. All respondents with probabilities of 3 and 5% were designated as not evacuating, while all those with 100% probabilities were designated as evacuees.

The results of the disaggregate assessment of the models' performance are shown in Table 6. Overall, the logistic regression

Table 5. Comparative Performance of Aggregate Predictions

County	Observed	Part. Rate	Logistic regression	Feed forward neural network	Probabilistic neural network	Learning vector quantizer
Acadia	9	14	5	7	8	5
Ascension	8	6	10	9	7	7
Assumption	4	9	5	5	4	5
Calcasieu	25	55	21	19	18	16
Cameron	4	4	4	4	4	3
East Baton Rouge	0	0	0	0	0	1
Evangeline	0	0	0	0	0	0
Iberia	22	37	22	23	22	18
Iberville	6	7	7	6	5	5
Jefferson Davis	1	3	1	1	1	1
Lafayette	19	12	32	27	22	22
LaFourche	0	5	3	1	0	2
St. James	1	1	1	1	1	1
St. Landry	3	1	2	5	3	2
St. Martin	11	7	10	10	10	6
St. Mary	28	31	31	29	28	27
Terrebonne	18	42	27	25	18	22
Vermillion	9	8	6	9	6	7
West Baton Rouge	1	0	1	1	1	1
Total	169	229	188	182	158	151
RMSE		9.9	4.1	2.9	1.9	3.1
Correlation		0.87	0.93	0.96	0.98	0.95

and neural network models perform marginally better than the participation rate model. The logistic regression and neural network models estimate the individual evacuation behavior correctly for approximately two out of every three households, while the participation rate model predicts the evacuation decision correctly for only approximately six out of ten households. Since a binary choice is being modeled, a totally random choice would produce the correct evacuation decision (i.e., to evacuate or to not evacuate), on average, 50% of the time. Thus, the models tested do perform better than a model with no discriminatory power (i.e., a model that makes totally random predictions) although their level of discrimination is limited. To quantify the performance, if 50% correct disaggregate prediction is considered “no predictive performance” and 100% correct disaggregate prediction is “perfect predictive performance,” the participation rate model displays approximately 20% of perfect performance, and the logistic regression and neural network models approximately 33% of perfect prediction at the disaggregate level.

Table 6. Comparative Performance of Disaggregate Predictions

Model	Classification	Percent correctly predicted (%)	Overall percent correctly predicted (%)
Participation rate	Evacuated	68	60
	Did not evacuate	55	
Logistic regression	Evacuated	64	67
	Did not evacuate	68	
FFNN	Evacuated	68	65
	Did not evacuate	63	
PNN	Evacuated	55	68
	Did not evacuate	76	
LVQ	Evacuated	46	65
	Did not evacuate	76	

Summary and Conclusion

The purpose of this study was to test different techniques of estimating trip generation of evacuation traffic in order to identify which procedures produce more accurate results. Participation rate, logistic regression, and various forms of neural network models were estimated and tested using data of evacuation behavior observed with Hurricane Andrew. Two model specifications were tested among the neural network models in this study. The first involved five variables found to be the most influential in estimating evacuation in the logistic regression model. The other included a more comprehensive set of 13 variables that included the original five variables. All three neural network models were estimated and tested using both model specifications to determine the impact of the added variables. It was found that the more comprehensive specification improved the performance of the neural network models from that achieved with the simpler specification only marginally (less than 5%). It was therefore concluded that the same specification among the logistic regression and the neural network models could be used to compare the performance of the models.

The five variables found to be the most influential in determining evacuation demand were type of dwelling, whether an evacuation order was issued or not, distance to the nearest body of water, age of the respondent, and marital status. If distance to the nearest body of water is taken as a surrogate for risk of flooding, the first three variables are consistent with the factors identified as the most influential in evacuation behavior in other research (Baker 1991; Irwin and Hurlbert 1995). In the logistic regression model, households living in mobile homes were estimated as 10.1 times more likely to evacuate than those in multiple-dwelling homes (e.g., apartments). An evacuation order was estimated as increasing the likelihood of evacuation by 4.2 times, and those living near water were 2.2 times more likely to evacuate than those living elsewhere. Information such as this is useful to offi-

cial in estimating the impact of alternative policies on evacuation demand. For example, authorities may consider changing residential zoning in their area, imposing higher flood limits, and issuing evacuation orders in a specific phased pattern to allow a more uniform loading of the road network.

When reviewing the accuracy of the models at the aggregate and individual level, the results of this study suggest that logistic regression and artificial neural network models estimate evacuation demand more accurately than conventional participation rate models. From a sample of 410 households, the logistic regression and neural network models estimated the aggregate number evacuating from each of 19 parishes with an average error of between two and four households, while the average error in aggregate prediction with the participation rate model was approximately 10 incorrectly assigned households. Between 87 and 96% of the variation in the aggregate evacuation from the different zones (which varied between zero and 28 per parish), was captured by the logistic regression and neural network models. In contrast, only 76% of the aggregate variation in parish evacuations were captured with the participation rate model. Thus, with this data set, more accurate aggregate estimates were obtained using either logistic regression or neural network models than a participation rate model. However, it must be noted that the logistic regression and neural network models predicted on essentially the same data they were estimated on (the only difference being the 60 households in the testing data set), while the participation rate model was estimated subjectively from past experience in general. The participation rate model used in this study was developed specifically for the region in which it was applied, but its parameters were not estimated from the data set on which its performance is being measured. Thus, while the participation rate model does not enjoy the same advantage of being tested against essentially the same data as were used to estimate the logistic regression and neural network models, it was estimated on data from the same area.

At the disaggregate (i.e., individual household) level, the logistic regression and neural network models produced more accurate estimates than the participation rate model, even though the level of accuracy achieved by all models was relatively low. Specifically, the participation rate model correctly predicted approximately six out of ten household's evacuation decisions correctly, while the logistic regression and neural network models correctly predicted approximately two out of three household's evacuation decisions. Considering that a totally random estimate of the binary decision to evacuate or not would typically produce a correct estimate half of the time, the above statistics of disaggregate performance are not very encouraging. The participation rate model achieves approximately 20% of perfect disaggregate prediction $((6-5)/5)$, while the logistic regression and neural network models achieve approximately 33% $((6.7-5)/5)$ of the best disaggregate prediction possible. However, disaggregate evacuation behavior is not of prime importance in evacuation demand estimation because it is total flows on the network that are ultimately of significance. If these can be estimated with aggregate estimates of travel demand spatially distributed throughout the network, then knowledge of the individual households that make up these flows, are unnecessary.

One of the interesting results of this study is the level of over- and underprediction achieved by the models tested. The participation rate model overestimated evacuation by 36%, and the logistic regression and FFNN models overestimated evacuation by 11 and 8%, respectively. The PNN and LVQ models underestimated evacuation demand by 7 and 11%, respectively. The results

suggest that the current state-of-the-art in hurricane evacuation demand may not be capable of estimating evacuation demand accurately.

The objective of this study was to assess the relative performance of alternative trip generation models of hurricane evacuation. The results at both the aggregate and disaggregate level suggest that better estimation can be achieved using logistic regression or neural network models rather than participation rate models. There is not a clear preference between logistic regression and any of the various neural network models tested in this application, regarding accuracy in demand estimation. However, in general terms, the neural network models appear to perform marginally better than the logistic model, and among the neural network models, the PNN model appears to perform the best although the difference is marginal.

The findings of this study are based on the analysis of a single set of data of evacuation behavior from a specific storm in a particular area. Subsequently, there is no reason to believe the findings are generalizable. Whether the same findings would emerge using data from a different area under different conditions is unknown. Also, whether the models estimated in this study could represent evacuation behavior in other areas is unknown. Evacuation behavior is notoriously different from hurricane to hurricane (Baker 1991) suggesting that only well-specified models will be capable of capturing evacuation decisions in a variety of settings. Further research is needed to establish models that are transferable. The decision to evacuate or not is also closely coupled with the time at which evacuation takes place so that time-dependent trip generation modeling is an important area of future research.

Acknowledgments

The research reported in this paper was funded by the Louisiana Transportation Research Center. The writers gratefully acknowledge their assistance and support. The data used in this analysis were collected for the Louisiana Office of Emergency Preparedness by the Louisiana Population Data Center Survey Research Lab, Louisiana State University. Gratitude is extended to Dr. Jeanne Hurlbert for making the data available for this study.

References

- Baker, E. J. (1991). "Hurricane evacuation behavior." *Int. J. Mass Emerg. Disasters*, 9(2), 287–310.
- Corps of Engineers (COE) and Southwest Florida Region Planning Council (SWFRPC). (1979). "Lee County Florida flood emergency evacuation plan." Fla.
- Corps of Engineers (COE). (1994). *A hurricane evacuation computer model for southeast Louisiana, HURREVAC Version 6.0, Documentation and user's guide*, Prepared for the Louisiana Dept. of Military Affairs Office of Emergency Preparedness, Baton Rouge, La.
- Corps of Engineers (COE). (2000). "Alabama hurricane evacuation study technical data report: Behavioral analysis." *Final Rep.*
- HMM Associates. (2000). "Evacuation time estimates for area near Pilgrim Station." *Rep., prepared for Boston Edison Company*, Boston.
- Hobeika, A. G., and Jamei, B. (1985). "MASSVAC: A model for calculating evacuation times under natural disaster." *Proc., Conf. Computer Simulation in Emergency Planning*, Vol. 15, Society of Computer Simulation, La Jolla, Calif., 5–15.
- Hobeika, A. G., Radwan, A. E., and Jamei, B. (1985). "Transportation actions to reduce evacuation times under hurricane/flood conditions:

- A case study of Virginia Beach city." *Paper Presented at the 74th Annual Meeting of the Transportation Research Board*, Washington, D.C.
- Hosmer, D. W., Jr., and Lemeshow, S. (2000). *Applied logistic regression*, Wiley, New York.
- Irwin, M. D., and Hurlbert, J. S. (1995). "A behavioral analysis of hurricane preparedness and evacuation in southwestern Louisiana." Louisiana Population Data Center, Louisiana State Univ., Baton Rouge, La.
- KLD Associates. (1984). "Formulations of the DYNEV and I-DYNEV traffic simulation models used in ESF." Federal Emergency Management Agency, Washington, D.C.
- Lewis, D.C. (2001). "September's great escape: New information system helps manage hurricane evacuations." *Roads and Bridges*, 39(9), 40–42.
- Oak Ridge National Laboratory (ORNL). (1999). *Oak Ridge evacuation modeling system (OREMS) user's guide*, Center for Transportation Analysis, ORNL, Oak Ridge, Tenn.
- Post, Buckley, Schuh & Jernigan, Inc. (PBS&J). (2000a). "Southeast United States hurricane evacuation traffic study: Evacuation travel demand forecasting system, technical Memorandum 2." *Final Rep.*, Tallahassee, Fla.
- Post, Buckley, Schuh & Jernigan, Inc. (PBS&J). (2000b). "Southwest Louisiana hurricane evacuation study: Transportation model support document." Tallahassee, Fla.
- Post, Buckley, Schuh & Jernigan, Inc. (PBS&J). (2001). "Southeast Louisiana hurricane evacuation study: Transportation model support document." Tallahassee, Fla.
- Regional Development Service (RDS). (1999). "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." Dept. of Sociology and Department of Economics, East Carolina Univ., Greenville, N.C.
- Sheffi, Y., Mahmassani, H., and Powell, W. B. (1982). "A transportation network evacuation model." *Transp. Res., Part A*, 16A(3), 209–218.
- Stornetta, W., and Huberman, B. (1987). "An improved three-layer back-propagation algorithm." *Proc., IEEE 1st Int. Conf. on Neural Networks*, Vol. II, San Diego.
- Tsoukalis, L. H., and Uhrig, R. E. (1997). *Fuzzy and neural approaches in engineering*, Wiley, New York.
- Urbanik, T. (1978). "Texas hurricane evacuation study." Texas Transportation Institute, College Station, Tex.
- Urbanik, T., and Desrosler, A. E. (1981). "An analysis of evacuation time estimates around 52 nuclear power plant sites: An evacuation." *Rep. No. NUREG/CR 7856*, Vol. 1, U.S. Nuclear Regulatory Commission, Rockville, Md.
- Ziliaskopoulos, A., and Peeta, S. (2002). "Foundations of dynamic traffic assignment: The past, the present and the future." *Paper Presented at the 81st Annual Meeting of the Transportation Research Board*, Washington, D.C.