Transportation Research Part A 87 (2016) 90–101

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Hurricane evacuation demand models with a focus on use for prediction in future events

Kecheng Xu [a,*], Rachel A. Davidson [b], Linda K. Nozick [a], Tricia Wachtendorf [b], Sarah E. DeYoung [b]

[a] Cornell University, Ithaca, NY, United States
[b] University of Delaware, Newark, DE, United States

ARTICLE INFO

Article history:
Received 29 June 2015
Received in revised form 4 December 2015
Accepted 22 February 2016
Available online 26 March 2016

Keywords:
Ordered probit model
Hurricane evacuation prediction

ABSTRACT

Although substantial literature exists on understanding hurricane evacuation behavior, few studies have developed models that can be used for predicting evacuation rates in future events. For this paper, we develop new ordered probit models for evacuation using survey data collected in the hurricane-prone state of North Carolina in 2011 and 2012. Since all covariates in the models are available from the census or based on location, the new models can be applied to predict evacuation rates for any future hurricane. The out-of-sample predictive power of the new models are evaluated at the individual household level using cross validation, and the aggregated level using available data from Hurricane Irene (2011), Hurricane Isabel (2003) and Hurricane Floyd (1999). Model results are also compared with an existing participation rate model, and a logistic regression model available from the literature. Results at the individual household level suggests approximately 70% of households' evacuation behavior will be predicted correctly. Errors are evenly divided between false positives and false negatives, and with accuracy increasing to 100% as the percentage of people who actually evacuate goes to zero or all and decreasing to about 50% when the population is divided and about half of all households actually evacuate. Aggregate results suggest the new models compare favorably to the available ones, with average aggregate evacuation rate errors of five percentage points.

1. Introduction

Empirical and theoretical research on hurricane evacuation behavior has provided a great deal of knowledge about how people make evacuation decisions (Baker, 1991; Dow and Cutter, 2002; Whitehead et al., 2000). Specifically, prior studies have sought to clarify if, when, how, and to where people evacuate under hurricane threat. That research is of great practical importance because an understanding of the anticipated behavior of affected populations is critical to planning effective and efficient evacuations in future events. Nevertheless, few studies have explicitly discussed use of evacuation behavior research for prediction at the regional level in hypothetical or real future hurricanes. A focus on prediction would require two efforts that, with a few exceptions, have not been part of previous research investigations, even those that developed statistical models with probability of evacuation or similarly relevant response variables. They are: (1) ensuring that all

* Corresponding author. Tel.: +1 (607)597 9990.
E-mail addresses: kx46@cornell.edu (K. Xu), rdavidso@udel.edu (R.A. Davidson), LKN3@cornell.edu (L.K. Nozick), twachten@udel.edu (T. Wachtendorf), sedeyoun@udel.edu (S.E. DeYoung).

http://dx.doi.org/10.1016/j.tra.2016.02.012
0965-8564/

covariates in a statistical model are available for a future event and thus the model can be applied in a predictive mode, and (2) evaluating the out-of-sample predictive power, or ability of the model that was fitted with one sample of data to correctly predict behavior for another.

In this study, we aim to address this gap by developing new statistical models of evacuation behavior with a focus on use in a predictive mode. Specifically, ordered probit models are developed to estimate the number of evacuees in each geographic evacuation zone for a specified future hurricane, with the idea that information can be used as input into engineering evacuation planning models. The models are developed and tested using survey data from North Carolina.

Section 2 summarizes the literature on quantitative models of evacuation behavior with specific focus on the extent to which they are suitable for use in prediction. In Sections 3 and 4, we describe the data and analysis approach, respectively. The final fitted models are presented in Section 5, and an evaluation of their predictive power at the individual household and aggregate levels are presented in Sections 6 and 7, respectively.

2. Background

The literature on evacuation behavior is extensive and many useful reviews are available, such as Dash and Gladwin (2007), Yazici and Ozbay (2008) and Murray-Tuite and Wolshon (2013). It is also worth noting that Russo and Chilà (2014) creates a generic framework for demand modeling in evacuation situations that bridge many types of events including hurricanes. We focus in this paper on studies that relate to quantitative models of evacuation behavior, and specifically consider their suitability for use in prediction. As Wilmot and Mei (2004) discuss, two primary methods have been used to investigate evacuation demand (i.e., probability of evacuation or number of evacuees): (1) participation rates, and (2) regression-type statistical models.

The participation rate approach, which is similar to cross-classification used in conventional transportation planning, is the most commonly used approach in practice, including being used in North Carolina currently (Fu, 2004). In this method, participation rates (i.e., percentage of people who will evacuate) are assigned for each geographic region, with different values based on a few key features, such as, hurricane strength and housing type (mobile home or not) (PBS&J, 2002). The product of the population in the affected region and the appropriate participation rate then provides the estimated number of evacuees. Participation rates are relatively coarse, and according to Wilmot and Mei (2004), are typically based on a subjective combination of conclusions from the literature and experience in past storms rather than as output of a particular statistical analysis.

In the second approach, a sample of survey data is used to fit a statistical model in which the household evacuation decision is the response variable and the set of covariates are attributes of the household, hurricane, and/or geographic area. These are typically logit or probit regression models, but include other types as well, and more recently, neural networks. Here we include similar models in which the response variables are related quantities of evacuation timing, or destination type or location. Since our focus is prediction, we consider specifically the extent to which available models may be used in a predictive mode for a future hurricane and/or their predictive power has been evaluated.

Many of the models in the literature were developed with a primary purpose of identifying the variables that are related to the evacuation decision, not being used for prediction. Thus, the final models often include covariates that are not likely to be available for use in prediction (e.g., risk perception), although in some cases, it appears they could be with a relatively small modification (e.g., Riad et al., 1999; Bateman and Edwards, 2002; Whitehead, 2005; Whitehead et al., 2000; Solís et al., 2009, 2010; Lazo et al., 2010; Petrolia and Bhattacharjee, 2010; Huang et al., 2012; Hasan et al., 2012, 2013; Murray-Tuite et al., 2012; Mesa-Arango et al., 2013; Sadri et al., 2013). One of those models did test predictive power with the data they had. Whitehead (2005) included an initial 1999 survey that asked for actual evacuation behavior in Hurricane Bonnie and planned evacuation behavior in future hurricanes, and a follow up 2000 survey of the same respondents that asked if they had actually evacuated during the intervening Hurricanes Dennis and Floyd. The Whitehead study included a computed forecast error as the difference between the predicted probability of evacuation and actual evacuation behavior and found some degree of predictive validity. Since the models include some covariates that would not be available without survey data (e.g., perceived wind risk), however, they still cannot be used directly for prediction in future events. On the other hand, some studies, such as, Cheng et al. (2008), Smith and McCarty (2009), and Reininger et al. (2013) present models that it appears could be used for prediction in a future event, but out-of-sample prediction is not tested or discussed.

Finally, Wilmot and colleagues conducted the only studies we could identify that address prediction directly. Wilmot and Mei (2004) compares five types of evacuation demand models—participation rate, logistic regression, and three types of neural networks—all fitted using 85% of observations in a 1995 household survey in southwestern Louisiana with respect to Hurricane Andrew. The models' predictive performances were assessed and compared at the aggregate and individual household levels. County level evacuation rates were compared to those observed based on the full dataset. For individual households, predicted and observed behavior was compared using the 15% of observations withheld for testing. In Section 7, we compare the Wilmot and Mei (2004) logistic regression to the new models developed in this paper. Fu and Wilmot (2006), Fu et al. (2007), and Gudishala and Wilmot (2012, 2013) discuss different types of models of time-dependent evacuation demand (i.e., probability a household evacuates in time interval t). They all discuss prediction explicitly and after using survey data from one hurricane to fit the models, estimate out-of-sample predictive accuracy by testing with a 15% holdout set, with observed data from a different hurricane, or with a different set of stated preference data. Yazici and Ozbay (2008) compare