Data-Driven Modeling of Hurricane Evacuee's Individual Decision-Making for Enhanced Hurricane Evacuation Planning: Florida Case Study in the COVID-19 Pandemic

Shijie Chen[1]; Yanshuo Sun[2]; Tingting Zhao[3]; Minna Jia[4]; and Tian Tang[5]

Abstract: Individual evacuation decision making has been studied for multiple decades mainly using theory-based approaches, such as random utility theory. This study aims to bridge the research gap that no studies have adopted data-driven approaches in modeling the compliance of hurricane evacuees with government-issued evacuation orders using survey data. To achieve this, we conducted a survey in two coastal metropolitan regions of Florida (Jacksonville and Tampa) during the 2020 Atlantic hurricane season. After preprocessing survey data, we employed three supervised learning algorithms with different complexities, namely, multinomial logistic regression, random forest, and support vector classifier, to predict evacuation decisions under various hypothetical hurricane threats. We found that the evacuation decision is mainly determined by people's perception of hurricane risk regardless of whether the government issued an order; COVID-19 risk is not a major factor in evacuation decisions but influences the destination type choice if an evacuation decision is made. Additionally, past and future evacuation destination types were found to be highly correlated. After comparing the algorithms for predicting evacuation decisions, we found that random forest can achieve satisfactory classification performance, especially for certain categories or when some categories are merged. Finally, we presented a conceptual optimization model to incorporate the data-driven modeling approach for evacuation behavior into a government-led evacuation planning framework to improve the compliance rate. **DOI: 10.1061/NHREFO.NHENG-1976.** *© 2024 American Society of Civil Engineers.*

Author keywords: Hurricane evacuation planning; Choice behavior modeling; Data-driven approach; Florida; Mixed-mode surveys.

Introduction

This study is motivated by the inadequate modeling of individual choices of evacuees in hurricane evacuation optimization studies. Given the predicted impact of a hurricane, emergency management officials can authorize different types of evacuation orders (i.e., voluntary or mandatory) to protect the lives of residents in designated regions. After a mandatory evacuation order is issued, residents are expected to comply with it and take action to evacuate from risky areas. However, it is known that various factors influence people's evacuation decisions before hurricanes (Baker 1991; Dash and Gladwin 2007; Bowser and Cutter 2015). Important factors include housing type, past experience, availability of transportation means, and information sources (Ploran et al. 2018), among others, that determine how the risk is perceived and subsequently how the evacuation decision is triggered (Huang et al. 2016; Stein et al. 2013). Therefore, mandatory evacuation orders do not guarantee all residents will evacuate, i.e., full compliance, since individual circumstances highly vary (Gladwin et al. 2001). Empirical studies, such as Martín et al. (2017) and Wong et al. (2020), also confirm that residents facing mandatory evacuation orders may not necessarily comply. Nonetheless, very few of the existing hurricane evacuation optimization studies, e.g., Apivatanagul et al. (2012) and Lu et al. (2017), have considered residents' compliance with evacuation orders in the design of government-led evacuation plans. On the contrary, the bulk of the hurricane evacuation planning literature is based on oversimplified assumptions about evacuees' decision making. For instance, Lim et al. (2012) developed new network flow models for the optimization of evacuation paths, flows, and schedules without considering the responses of evacuees to the optimized evacuation decisions. Neglecting evacuees' individual choices or assuming full compliance with evacuees would result in a limited and unrealistic description of their decisions before or during a hurricane evacuation.

The compliance of evacuees with evacuation orders should by no means be overlooked in government-led hurricane evacuation planning because the efficiency and effectiveness of an evacuation plan created by an evacuation management agency largely depend on how residents comply with it (Stein et al. 2013). This issue cannot be addressed by taking coercive measures by governments. Instead, the key lies in understanding how governments can effectively guide and support evacuees' decisions in a human-centric

[1]Ph.D. Candidate, Dept. of Industrial and Manufacturing Engineering, FAMU-FSU College of Engineering, Florida State Univ., 2525 Pottsdamer St., Tallahassee, FL 32310. ORCID: https://orcid.org/0000-0003-4396-1357. Email: sc20hw@fsu.edu

[2]Associate Professor, Dept. of Industrial and Manufacturing Engineering, FAMU-FSU College of Engineering, Florida State Univ., 2525 Pottsdamer St., Tallahassee, FL 32310 (corresponding author). ORCID: https://orcid.org/0000-0003-2943-4323. Email: y.sun@eng.famu.fsu.edu

[3]Associate Professor, Dept. of Geography, Florida State Univ., 113 Collegiate Loop, Tallahassee, FL 32306. ORCID: https://orcid.org/0000-0003-3787-2016. Email: tzhao@fsu.edu

[4]Director, Institute of Science and Public Affairs, Florida State Univ., 296 Champions Way, Tallahassee, FL 32306. Email: minna.jia@fsu.edu

[5]Associate Professor, Askew School of Public Administration, Florida State Univ., 113 Collegiate Loop, Tallahassee, FL 32306. ORCID: https://orcid.org/0000-0003-3588-6828. Email: ttang4@fsu.edu

Note. This manuscript was submitted on July 6, 2023; approved on April 11, 2024; published online on July 13, 2024. Discussion period open until December 13, 2024; separate discussions must be submitted for individual papers. This paper is part of the ***Natural Hazards Review***, © ASCE, ISSN 1527-6988.

manner. Since a lot of factors influence residents' evacuation decision making, from the government's point of view, certain measures can be taken to influence a subset of such factors in order to shift residents' decisions to improve the compliance rate, translating to improved efficiency and effectiveness of an evacuation plan. As illustrated in Fig. 1, a range of factors jointly determine whether an individual evacuates or not. Such a set of factors can be divided into two subsets depending on whether the government is able to influence a factor or not. For example, while the government is unable to change one's age or gender, it can focus on improving transportation services, diversifying information dissemination channels, and establishing financial support mechanisms for evacuation-related expenses for targeted residents. Since the government has a few instruments to influence residents' decisions, an optimization problem can thus be designed with the goal of achieving a higher compliance rate, thus directly enhancing the effectiveness of evacuation plans, subject to resource constraints, e.g., a limited vehicle fleet and a budget limit.

The primary objective of this study is thus to model evacuees' behavioral compliance with evacuation orders with a data-driven approach such that the outcome of this study (i.e., a data-driven model for analyzing or classifying evacuees' choices) can be incorporated into the hurricane evacuation planning framework shown in Fig. 1. Note that incorporating such a data-driven approach in an evacuation optimization problem cannot be realistically done in a single paper focusing on data-driven modeling of evacuation decision making. Therefore, the incorporation of this paper's outcome in evacuation planning optimization is left as an extension of this paper.

Evacuees' responses to evacuation orders involve a range of choices, such as evacuating or staying, using public shelters or hotels, evacuating by car or transit, and evacuating immediately or later. In the literature, there are two major paradigms for choice modeling, namely theory-driven and data-driven, as classified by van Cranenburgh et al. (2022). Specifically, in the case of hurricane evacuations, evacuees' choices are predominantly analyzed using utility-based models. For instance, Wong et al. (2020) employed latent class choice models based on the discrete choice theory to study the evacuation choice (evacuate or not). Theory-driven

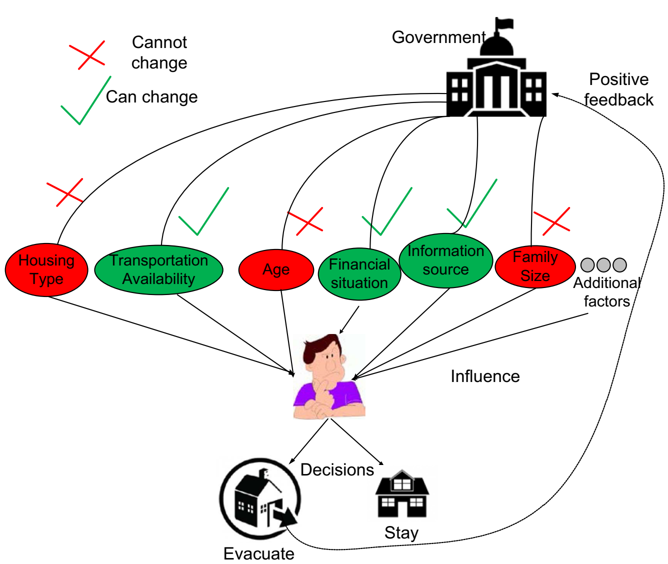

Fig. 1. Government-led hurricane evacuation planning considering evacuees' individual choices.

models, often grounded in random utility theory, offer a structured approach to understanding how individuals might make choices based on utility maximization. They require the careful crafting of theoretical frameworks to incorporate factors like income and housing conditions into the model. While these models are designed around the existing knowledge and hypotheses of researchers, it is important to recognize that they represent approximations or simplifications of complex decision making processes. As such, there may be some divergence from the full spectrum of real-world behaviors. Nevertheless, these theory-driven approaches provide valuable insights and a foundational framework for analyzing choice behaviors.

Another modeling paradigm, namely the data-driven approach, does not rely on any specifications of the relations between independent variables and dependent variables; instead, data-driven models discover those relations by learning from the observed input-output pairs. Various factors that may influence evacuation decisions are modeled as features (or inputs), while evacuation decisions are targets (or outputs). When evacuation decisions are discrete, the choice modeling is essentially a classification task. Thanks to the rapid developments in supervised learning algorithms, a data-driven approach can often achieve higher goodness-of-fit than theory-driven approaches (van Cranenburgh et al. 2022). Therefore, many data-driven models have been developed and compared with theory-driven models in many choice modeling areas, such as high-speed train choice modeling (Sun et al. 2018). Nonetheless, to the best of the authors' knowledge, no studies have adopted data-driven approaches in modeling the compliance of hurricane evacuees using survey data. This paper thus presents the first known effort to analyze the relations between various factors and evacuees' individual choices using a data-driven approach. To the best of the authors' knowledge, this paper is the first to use multinomial logistic regression, random forest, and support vector classifier in a purely data-driven manner for modeling the behaviors of hurricane evacuees using survey data. Although as a variant of the generalized linear models, multinomial logistic regression structurally resembles the multinomial logit model, a variant of discrete choice models rooted in utility maximization theory, the data-driven nature of multinomial logistic regression implies we do not intend to analyze causal relationships, while prioritizing high classification accuracy instead.

In this study, we designed and launched a survey in two coastal metropolitan regions in Florida (namely, Jacksonville and Tampa) between July 2020 and September 2020, coinciding with the COVID-19 pandemic. Based on the survey results, we employed three classification algorithms with different complexities to predict evacuation decisions under various scenarios. Specifically, we considered the following individual evacuation decisions as targets: evacuation likelihood, evacuation order compliance, and destination choice. The predictive performance of each classification algorithm was also evaluated, and some key findings were then summarized.

The remainder of this paper proceeds as follows. We first identify a research gap after reviewing relevant evacuation behavior studies grouped by modeling paradigm and data collection method. Then, we describe the sampling method, questionnaire design, and data processing and show some exploratory analysis results. Next, three classification algorithms are presented along with hyperparameter tuning and model evacuation methods. We further develop conceptual models to illustrate how the data-driven behavior model can be incorporated into a government-led evacuation planning problem. Finally, we present the study findings and discussions, followed by conclusions.

Nat. Hazards Rev., 2024, 25(4): 04024032

Literature Review

We begin the literature review by first examining the assumption of evacuee behaviors in government-led evacuation planning optimization studies. We then show that evacuees are mainly assumed to be fully compliant in such studies, in contrast with the noncompliance assumption in individual hurricane evacuation behavioral studies. Next, we indicate that descriptive analyses and discrete choice models are widely used in modeling the decision making of individual evacuees while data-driven approaches are sparse. Even though some data-driven approaches are applied to evacuation behavior modeling in other contexts, such as cinema or wildfire evacuation, no studies have used data-driven approaches for hurricane evacuation behavioral modeling. Lastly, we note the importance of systematic sampling. If it is absent, it is well known that the resulting sample is biased, and the relevant findings cannot be generalized to the whole population. After those reviews are organized into three parts, research gaps are identified.

Modeling of Individual Decision Making in Evacuation Planning

Given the critical role of hurricane evacuation in mitigating the impact of a hurricane, especially in saving human lives and speeding up disaster response, researchers have been studying people's responses to approaching natural hazards and disasters (Lindell et al. 2018). These include the prediction of people's evacuation decisions, which are discussed in depth in the following section, as well as the modeling of evacuation actions. For example, Lindell and Perry (2012) revised the theoretical protective action model to include the timing of household evacuation. The revised model involved rates of exposure to warning mechanisms (such as radio and TV) over time and mental or logistical preparation required for households to take evacuation actions.

Researchers have developed various optimization models for supporting evacuation planning. In a review article, Galindo and Batta (2013) indicated that evacuation optimization studies involved many decisions, such as network design, shelter location, and emergency vehicle routing and scheduling, which were affected by people's preferences. Galindo and Batta (2013) assessed the appropriateness of common assumptions adopted in evacuation planning and classified each one as unrealistic, limited, or reasonable. In a transit-based evacuation study, Chen and Chou (2009) assumed that the evacuation demand at each bus stop and the total demand for a shelter were known right after the disaster, which was evaluated to be unrealistic by Galindo and Batta (2013). In an emergency supply prepositioning study, Rawls and Turnquist (2010) assumed a limited number of scenarios with the probability of each scenario estimated from historical data. However, the very low frequency of past disasters makes the accurate prediction of such probabilities quite challenging (Galindo and Batta 2013). Another limited or unrealistic assumption, which was not discussed in Galindo and Batta (2013), was the full-compliance assumption. The effect of noncompliant behavior on evacuation planning was not adequately considered, although behavioral researchers were well aware of the noncompliant behavior (Connolly et al. 2020). Therefore, there is a divide between the hurricane evacuation behavioral modeling and evacuation planning optimization communities regarding the treatment of noncompliant behavior.

Evacuation Compliance Modeling Methods

Many research methodologies have been employed to study the individual decision making of evacuees. As classified by Golshani et al. (2019), the first group of studies mainly relied on descriptive analyses and simple statistical tests, especially those early studies. The second group of studies investigated the effect of various factors on evacuation decisions from a behavioral perspective using discrete choice modeling. We next review some representative studies in each category.

As a representative study in the first group, Baker (1991) considered each potential factor and qualitatively examined its impact on evacuation response based on surveys conducted by earlier researchers covering many geographic regions and spanning a few decades. One notable conclusion drawn by Baker (1991) was that many demographic factors, such as age, education, and family status, did not explain the variation in evacuation response well. Instead, Baker (1991) identified risk level, action taken by governments, and type of housing, among others, as the major determinants of a hurricane evacuation. Collins et al. (2021) used similar descriptive statistics to examine people's hurricane evacuation decision making during the time period of COVID-19.

Cross tabulation is a widely used method to show the relations between two categorical variables. For instance, Zhao et al. (2023) studied relationships between hurricanes and COVID-19 risk perceptions and evacuation intention with cross tabulation and assessed the statistical significance of the relationship with a chi-square test. They found that perceived hurricane risks significantly influenced household evacuation decisions regardless of people's concern about getting sick with COVID-19. Brown et al. (2016) used chi-square and t-tests to study hurricane evacuation during Hurricane Sandy. They found a higher evacuation rate for those who witnessed trauma (related to the World Trade Center attacks) in the past. In another representative work, Matyas et al. (2011) used chi-square tests to determine what individual characteristics influenced the likelihood of evacuation based on a survey of Florida tourists. As the chi-square test examines the relationship between two variables, it cannot capture the compound impact of multiple independent variables (e.g., family size, residential type, and education level) on evacuation decisions.

We next review some representative theory-driven models in the second group, which are mostly discrete choice models. Golshani et al. (2019) conducted web-based surveys and employed a multivariate ordered probit model to estimate the likelihood of taking one of the following choices: ignoring the threat, seeking shelter at the same place, and evacuating to a safe place. They found some notable determinants of evacuation. For instance, people who were well-educated or living in multiunit buildings were more likely to evacuate. Ling et al. (2021) studied the effect of information sources and social networks on evacuation behavior based on data from 589 completed postal surveys in Jacksonville, Florida. They developed a mixed logit model and found that larger social networks and durable friendships contributed to shadow evacuation decisions. In a recently published study, Bian et al. (2022) examined the effect of travel delays on household evacuation decisions, including evacuation or not, departure time, and destination choices. Random parameter logit models were calibrated using data from 415 surveys conducted in Virginia coastal areas in 2017. They found that households tended to evacuate when they were informed of significant travel delays. Botzen et al. (2022) employed an ordered probit model to assess how the COVID-19 pandemic impacted evacuation intentions during the 2020 hurricane season among residents of flood-prone areas in Florida. The results revealed that COVID-19 concerns overwhelmed flood risk perceptions, inhibiting evacuation, particularly among older people.

As it is not our intention to conduct a comprehensive survey of all the discrete choice models for evacuation behavior modeling, we direct interested readers to an in-depth review conducted by

Nat. Hazards Rev., 2024, 25(4): 04024032

Murray-Tuite and Wolshon (2013). It is quite clear that theory-driven models are predominantly used in understanding hurricane evacuation behavior, possibly because of their sound behavioral foundations. Many fruitful results have been obtained from those discrete choice analyses, which are very helpful in informing the emergency management practice.

Nevertheless, those discrete choice models are not without shortcomings or weaknesses. The utility functions are handcrafted based on modelers' prior-belief of the effect of certain factors. Therefore, subjectivity in results is unavoidable (Ben-Akiva and Lerman 2018). As utility functions are typically assumed to be linear, any nonlinear relations cannot be captured. In addition, the practice of searching for the most appropriate utility function is labor-intensive (van Cranenburgh et al. 2022). van Cranenburgh et al. (2022) have identified other shortcomings in discrete choice modeling.

Even though data-driven approaches have been successfully used in modeling choices in some fields, such as travel mode choice (Zhao et al. 2020b), applications of data-driven approaches in evacuation behavior modeling and analysis are very rare. One distinction of data-driven approaches is in their capability to learn the fundamental patterns and relations in the raw data without any prior specifications of those relations. In nonhurricane evacuation studies, Zhao et al. (2020a) employed data-driven approaches for understanding preevacuation decision making of building occupants based on data collected from unannounced evacuation drills in a Swedish cinema. Although Zhao et al. (2020a) represent the most relevant study to this paper, notable differences remain. For instance, hurricane evacuation is much more complex than making decisions during preevacuation periods because the latter is less likely dependent on socioeconomics, such as income. In addition, cinema preevacuation behavior is considered one-dimensional by Zhao et al. (2020a), i.e., either the normal stage or the response stage. The normal stage is to continue previous actions, and the response stage is to investigate or evacuate. By contrast, hurricane evacuation decision usually has multiple dimensions, including destination choice, departure time choice, and mode choice.

Roy and Hasan (2021) presented a data-driven model, specifically an input-output hidden Markov model (IO-HMM), to infer individual evacuation behaviors (e.g., evacuation intent and timing) from geocoded tweets during hurricanes. They considered five input variables such as time from landfall and whether a user's home is in a mandatory evacuation zone. Two output variables were considered, namely, the distance of a user's current location from home, and evacuation similarity scores of posted tweets. Their case study demonstrated the great potential of real-time social media data in predicting and understanding individual evacuation dynamics. Nonetheless, Roy and Hasan (2021) noted the limited and uneven penetration of Twitter, now known as X, may cast the generalization of the developed model into question.

Zhao et al. (2021) introduced a conceptual framework that combines artificial intelligence with current wildfire evacuation modeling to enhance the understanding of household evacuation behaviors in such emergencies. Nonetheless, they did not present any case studies.

It should be noted that some data-driven models have been developed for crowd evacuation simulation, such as Yao et al. (2019). As such models focused on pedestrian movements, detailed reviews are not presented here. Interested readers are directed to Dong et al. (2019).

Xu et al. (2023) tested seven different machine learning algorithms using data from the 2019 Kincade Fire. Their findings underscored the superior performance of machine learning models, with the classification and regression tree (CART) model, in particular, offering both high predictive accuracy and clear interpretability. While Xu et al. (2023) provide valuable insights into the potential of machine learning in this domain, it is applied to the study of wildfire evacuation behaviors. Understandably, evacuation behaviors are quite different when the disaster type differs, such as wildfire versus hurricane. Therefore, our study seeks to fill the research gap that data-driven models are barely yet used to understand hurricane evacuation decisions.

Data Collection Methods for Evacuation Compliance Modeling

As high-quality survey data are essential in understanding evacuation behaviors, researchers have employed various data collection methods, including face-to-face interviews, paper-based surveys, telephone surveys, and web-based surveys (email or social media) (Thompson et al. 2017). While all those data collection methods have been widely used, each of them has its advantages and disadvantages. In-person surveys are advantageous because complex interview questions can be asked and follow-ups can be conducted easily. A clear disadvantage of it is its high cost, which usually implies a limited sample size. Paper surveys can be distributed in large volumes and can reach certain targeted respondents. One of its advantages is the significant data compilation efforts. Telephone surveys can be conducted by random digit dialing, while it is widely known that older respondents are much more interested in participating in telephone surveys (Bowser 2013). Web-based surveys require the least amount of money and time, meaning a very large audience can be targeted; however, the representativeness of the collected sample is usually questionable. For instance, younger people rely on social media more than older people, meaning the latter group does not usually participate in surveys launched on social media. Those hard-to-reach individuals are more responsive to traditional surveys based on paper or telephone. More systematic discussions of those survey methods are available in Jones et al. (2013).

In addition, to avoid biased samples, random sampling or stratified random sampling is necessary. Otherwise, it is quite difficult to generalize the findings from analyzing the survey results to the whole population.

Summary

Through the presented reviews, we recognize that while noncompliance of residents with hurricane evacuation orders has been extensively analyzed using descriptive or theory-driven methods, the potential advantages of a data-driven approach have been underexplored. Data-driven approaches offer the benefit of discovering relations between independent and dependent variables from the observed data without relying on prior assumptions or handcrafted specifications. In the context of hurricane evacuations, data-driven approaches allow us to capture potentially complex relations between various influencing factors and evacuees' choices without being bound by preconceived beliefs or theories. In addition, while there are multiple data collection methods for evacuation behavior studies, the use of a sampling approach that effectively captures a representative sample of the population is essential for generalizing the research findings to the broader demographic. This paper thus seeks to advance the literature by investigating residents' partially compliant behaviors with mandatory evacuation orders using supervised learning algorithms based on mixed-mode survey data sampled by Florida's voter registration system.

Data

Sampling Method

Two major metropolitan areas involving three counties in Florida (Duval, Pinellas, and Hillsborough) were selected for the survey, namely, Jacksonville and Tampa, as shown in Fig. 2. During the last several years, all three counties issued mandatory evacuation orders for Hurricane Irma in 2017; Duval County also issued mandatory evacuation orders for Hurricane Dorian in 2019. As only residents in county-designated hurricane evacuation zones (Florida Division of Emergency Management 2022) will be ordered for evacuation in Florida, we selected Zones A to E that are more prone to hurricane threats.

To best represent the targeted population, we randomly selected survey respondents (18 years of age or older) from the latest Florida Voter Registration list, which is a public record in Florida. This voter registration data set contains the voter's name, date of birth, sex, race, party affiliation, and address. The advantage of using voter registration information is that we have background information, especially the mailing addresses, that facilitates our stratified random sampling. More importantly, the physical address from the voter registration database can be used to infer the evacuation zone, while a survey respondent may report a different zone. Since the Florida Emergency Management agencies emphasize "KNOW YOUR ZONE" in their hurricane preparedness education programs, an inconsistent report of evacuation zone by a resident may likely indicate a lack of awareness.

A stratified sampling approach was used. In each metropolitan area, 4,000 residents were randomly selected from the voter registration list with control of the evacuation zone. Half of the residents live in Zone A and Zone B, and the other half in Zone C, Zone D, and Zone E. This sampling approach allowed us to cover high-risk zones (Zones A and B) in both metropolitan areas.

Questionnaire Design

In a hurricane evacuation survey, residents are typically asked about their property type, past evacuation experiences, their evacuation decisions under certain scenarios (actual or hypothetical), and sociodemographic information. To inform our hurricane evacuation survey and ensure it encompassed factors pivotal for understanding compliance with evacuation orders, we extensively reviewed the literature. The insights derived from past studies served as the foundation for our questionnaire design, inspiring us to ask participants about critical aspects such as their property type, past evacuation experiences, evacuation decisions, and sociodemographic information.

In an early study, Baker (1991) concluded that risk perception, actions taken by authorities, and housing conditions were among the key factors when residents made their evacuation decisions, while the role of demographic factors was deemed weak or inconsistent. Then, in a metaanalysis of 38 actual and 11 hypothetical evacuation studies, Huang et al. (2016) evaluated the effect of many factors on household evacuation, such as official warnings, mobile home residence, storm conditions, behaviors of other people, as well as demographic variables. They also found that responses to hypothetical evacuation scenarios were comparable to those to actual hurricane evacuations.

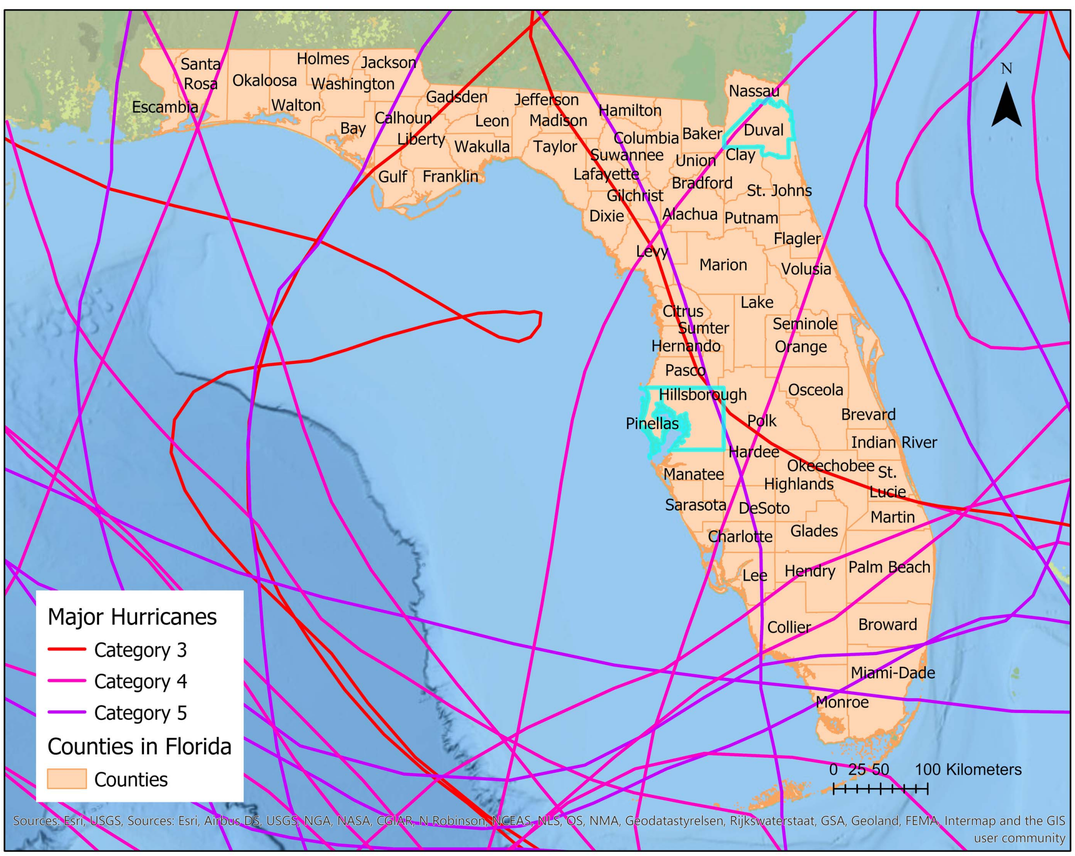

Fig. 2. Two study areas prone to hurricane threats. (Sources: Esri, Aribus, DS, USGS, NGA, N ASA, CGIAR, N Robinson, NCEAS, NLS, OS, NMA, Geodatastyrelsen, Rijkswaterstaat, GSA, Geoland, FEMA, Intermap and the GIS user community.)

Table 1. Overview of survey questions

ID	Survey question (may be shortened or rephrased)	Question type
Q1	Which evacuation zone is your home located?	Multiple choice (single)
Q2	What hurricane mitigation features does your home have?	Multiple choice (multiple)
Q3	If a hurricane of Category X hits your area, would you think it is safe for you to stay in your home?	Matrix
Q4	What was the most recent hurricane that made threats to your area?	Multiple choice (single)
Q5	When this hurricane made threats to your area, did you evacuate or stay?	Yes-or-No
Q6	Where did you go (if applicable, depending on the response to Q5)?	Multiple choice (multiple)
Q7	How did you go there (if applicable)?	Multiple choice (multiple)
Q8	How would you rate your last evacuation experience (if applicable)?	Matrix
Q9	Do you recall the predicted strength of the most recent hurricane?	Multiple choice (single)
Q10	Do you recall hearing evacuation orders, either directly or indirectly?	Yes-or-No
Q11	Do you recall how you heard the evacuation order for this hurricane?	Multiple choice (multiple)
Q12	How much do you rely on the following channels to gain information about emergencies, such as hurricanes and COVID-19?	Matrix
Q13	To what extent do you trust the information about emergencies (such as hurricanes and COVID-19) from the following sources?	Matrix
Q14	If a hurricane threatens your area this year (2020) and the COVID-19 pandemic is still present, how likely would you consider an evacuation?	Matrix
Q15	Under the same conditions as Q14, how likely would you evacuate if the government issues an evacuation order for your area?	Matrix
Q16	If you decide to evacuate this year (2020), where would you plan to go?	Multiple choice (single)
Q17	If you decide to evacuate this year (2020), how do you plan to go there?	Multiple choice (single)
Q18	Would you be more likely to go to a public shelter, if the following measures to contain the spread of COVID-19 are taken at this public shelter?	Matrix
Q19	What would be the major concerns for you not to evacuate if an evacuation order were issued in your area this year?	Multiple choice (multiple)
Q20	Do you have any of the following existing health conditions?	Multiple choice (multiple)
Q21	Has anyone you know in person ever been diagnosed with COVID-19?	Multiple choice (multiple)
Q22	Were you or are you currently sick with COVID-19?	Yes-or-No
Q23	How likely do you think the chance for yourself to be sick with COVID-19 by the end of the year of 2020	Five-level Likert scale
Q24	Which of the following do you practice on a daily base at present to protect yourself and others from COVID-19?	Matrix
Q25	Which year were you born?	Text
Q26	What is your gender?	Multiple choice (single)
Q27	Do you have any children under the age of 18 living in the same household?	Yes-or-No
Q28	Is any family member with special medical needs, disability, or limited mobility living in the same household?	Yes-or-No
Q29	What type of housing do you live in?	Multiple choice (single)
Q30	How long have you lived at your current address?	Multiple choice (single)
Q31	What is the highest grade you completed in school?	Multiple choice (single)
Q32	What was your household income in 2019?	Multiple choice (single)

Through a survey conducted in North Carolina, Whitehead et al. (2000) identified other determinants of evacuation behavior, in addition to those well-recognized ones, such as storm intensity. Whitehead et al. (2000) reported that nonwhite households, pet owners, and those with more education were less likely to evacuate to a shelter or hotel. In a recent study, Wong et al. (2020) launched a 146-question online survey in the aftermath of Hurricane Irma, where respondents were asked to provide detailed demographic information including gender, race, education level, and income.

As the survey was conducted in the Fall of 2020 when Florida led the United States in COVID-19 cases, several questions were asked related to residents' perception of health risks and their underlying health conditions. To ease the spread of COVID-19, residents were supposed to avoid leaving their homes. Those who would like to evacuate in anticipation of a hurricane without a pandemic may change their choice due to concerns about getting sick with COVID-19. The literature review of evacuations under the cooccurrence of a natural disaster and public health crisis conducted by Sakamoto et al. (2020) suggested that the COVID-19 pandemic was supposed to influence residents' evacuation decisions. The risk perception of COVID-19 and its implication for hurricane evacuation were also analyzed by Zhao et al. (2023).

In light of the reviewed studies, our questionnaire consisted of 32 questions, as summarized in Table 1. Note that questions in Table 1 may have been shortened or paraphrased due to space limits. Those questions were of different types, such as multiple choice (single or multiple answers), Yes-or-No, and Likert scale. Certain questions were in the matrix format, which means residents were asked to evaluate multiple choices for each row item. For instance, Q14 asked residents how likely they would evaluate under each hurricane category while considering the impact of COVID-19.

To improve the survey response rate, surveys were distributed by mail, while respondents had the choice to respond online or mail their responses, which is why the survey is called mixed mode (Millar and Dillman 2011). In total, 592 valid survey responses were received between July and September 2020, resulting in a response rate of approximately 8%. This relatively low response rate might be due to the COVID-19 surge experienced by Florida in July and August 2020. Residents at that time avoided unnecessary activities (such as mailing) due to the COVID-19 health risk.

Data Preprocessing

Table 2 presents the three response variables identified from the questionnaire, as follows:

Table 2. Target variables

QID	Variable
Q14	**EvacVolunt** Five-level Likert scale
Q15	**EvacMandt** Five-level Likert scale
Q16	**DestChoice** PubShelt = public shelter HMHSharing = Hotel, motel, or home-sharing housing FamiFrie = Stayed with families, relatives, or friends Other

Note: "QID": Question ID.

- *EvacVolunt*: Represents an individual's likelihood to evacuate for a hurricane under a voluntary evacuation order during the COVID-19 pandemic.
- *EvacMandt*: Evaluates an individual's evacuation likelihood to a mandatory evacuation order during the COVID-19 pandemic.
- *DestChoice*: Depicts an individual's preferred evacuation type (e.g., public shelter versus hotel) during the COVID-19 pandemic.

The *EvacVolunt* and *EvacMandt* represent individuals' inclination to evacuate voluntarily and under a mandatory order, respectively. Hurricane Category 3 is the most challenging one for the government to classify whether people are willing to evacuate, while it is less ambiguous to make decisions under other hurricane categories. Hence, for both *EvacVolunt* and *EvacMandt*, we focus on Category 3 in the subsequent analyses. The last response *DestChoice* reflects the preference for evacuation destination types.

Note that abbreviated names (e.g., *EvacVolunt* and *EvacMandt*) are used, while readers should refer to Q14, Q15, and Q16 for details, as indicated in Table 2.

To handle the three classification problems, feature preprocessing is needed, which consists of missing value processing, nominal variable processing, normalization, and feature selection. The first step is to handle missing value problems. Removing those cases with missing values is a simple way to overcome this issue (Yang et al. 2016) but would introduce substantial biases (Little and Rubin 2019). Instead, we fill those missing values with neutral alternatives that represent no inclination or no opinion, which is an extensively adopted strategy called neutral-value substitution (Phillips et al. 2006). Regarding the categorical variables, when confronted with missing values, our approach is to allocate the absence of a response to the otherwise case for certain questions (e.g., Q3 and Q5) or the No case for others (e.g., Q27). Note that such preprocessing can increase the number of samples, while it may introduce biases to the dataset. If missing values exist in the response variables, the whole sample has to be removed.

Then, dummy encoding is conducted for the nominal variables, e.g., gender and age range. To avoid multicollinearity, a nominal variable of k levels is encoded into $k-1$ dummies by removing the first level because the first level is the reference level. After that, all feature values are normalized by scaling into the range [0, 1].

As the performance of a machine learning algorithm can be improved by removing irrelevant features, we keep only the most relevant ones after feature selection. Two feature selection approaches are adopted in this study, namely, the domain knowledge-based feature selection method and random forest. By referring to the literature and experts' suggestions, some potentially correlated features are first selected. Then, depending on the feature importance score provided by random forest (RF), the most relevant attributes are identified. Higher scores reflect higher correlations between features and the target variable. The important features are kept, while the less correlated ones are pruned. See (Yao and Bekhor 2020) for more details on the RF-based feature selection method. As a result, 29, 33, and 30 most significant features are identified for the three problems *EvacVolunt*, *EvacMandt*, and *DestChoice*, respectively. The associated key variables are presented in Tables 3 and 4. When one row consists of multiple options, such as five-level Likert scales, this row refers to a continuous variable or feature. Otherwise, each row represents a categorical variable. Clearly, some features are not included for classifications since they are not expected to be important, such as those from questions Q7, Q9, and Q10. According to subsequent analyses, they do have little effect on the classification performance.

These preprocessing steps yield 592 valid samples in total. After removing the samples that do not answer the three target responses, 574, 581, and 534 samples are available for the three problems, respectively.

Methodology

Classification Algorithms

As there are many machine learning algorithms for classification, it is not our intent to exhaustively test them. Instead, we seek to compare a few representative ones with various levels of complexity. Specifically, this study involves three classification algorithms: multinomial logistic regression, RF, and support vector machines. Multinomial logistic regression is one of the generalized linear models (GLMs), while RF and support vector machine classifiers can learn nonlinear relations. Each of the three algorithms is introduced next.

We first introduce multinomial logistic regression as a special case of the generalized linear model. Then, we clarify the relationship between multinomial logistic regression and a specific variant of discrete choice models, namely, the multinomial logit model.

Multinomial Logistic Regression

Given predictors $X \in \mathbb{R}^p$ and an outcome Y of the dependent variable, a GLM is defined by three components:

1. A systematic component that relates predictors X to a linear predictor η, i.e., a linear combination of unknown parameters β;
2. A random component that specifies a distribution for Y conditional on X;
3. A link function g that connects the random and systematic components.

The GLM is then written as

$$\mathbb{E}(Y|X) = \mu = g^{-1}(\eta) \tag{1}$$

where $\mathbb{E}(Y|X)$ = expected value of Y conditional on X, also denoted as μ. The linear predictor η is essentially $X\beta$. The link function is denoted as g, which establishes the relation between the linear predictor η and the expected mean $\mathbb{E}(Y|X)$. Different choices of the link function and the distribution of the random component allow GLMs to model a wide range of data types as data-driven models. For instance, when the random component is normally distributed and the link function is identity, the GLM becomes linear regression. More possible choices are presented in Lindsey (2000).

As a special case, when the dependent variable y is a Bernoulli random variable $y \in \{0, 1\}$, i.e., a binary classification task is under consideration, we need to model the probability that y equals 1 given predictors X. In other words, we map the linear predictor $X\beta$ to a probability $p(y=1|X)$. Given that $\eta = X\beta$ can be outside

Table 3. Coding of features (Part 1)

QID	FID	Keyword and options
Q1		**Evacuation zone**
	F1-1	1 = Zone A or B, 0 = otherwise
	F1-2	1 = Zone C, D or E, 0 = otherwise
	F1-3	1 = Others, 0 = otherwise
Q3		**Safety perception**
	F3	1 = Any yes, 0 = otherwise
	F3-1	1 = Category 2 is safe, 0 = otherwise
	F3-2	1 = Category 3 is safe, 0 = otherwise
	F3-3	1 = Category 4 is safe, 0 = otherwise
Q4		**Recent hurricane**
	F4-1	1 = Dorian, 0 = otherwise
	F4-2	1 = Irma, 0 = otherwise
	F4-3	1 = Matthew, 0 = otherwise
	F4-4	1 = Other, 0 = otherwise
Q5		**Previous decision**
	F5	1 = Evacuate, 0 = stay
Q6		**Previous destination**
	F6-1	1 = not evacuated, 0 = otherwise
	F6-2	1 = public shelter, 0 = otherwise
	F6-3	1 = hotel or motel, 0 = otherwise
	F6-4	1 = home-sharing housing, 0 = otherwise
	F6-5	1 = stayed with families or relatives, 0 = otherwise
	F6-6	1 = stayed with friends, 0 = otherwise
	F6-7	1 = other, 0 = otherwise
Q8		**Previous evacuation experience**
	F8-1	Regarding lodging, 1 = very satisfied, 2 = satisfied, 2.5 = no opinion, 3 = dissatisfied, 4 = very dissatisfied
	F8-2	Regarding transportation, 1 = very satisfied, 2 = satisfied, 2.5 = no opinion, 3 = dissatisfied, 4 = very dissatisfied
	F8-3	Regarding overall cost, 1 = very satisfied, 2 = satisfied, 2.5 = no opinion, 3 = dissatisfied, 4 = very dissatisfied
Q11		**Means of hearing evacuation order**
	F11-1	1 = not heard, 0 = otherwise
	F11-2	1 = TV news, 0 = otherwise
	F11-3	1 = Text or voice message by phone, 0 = otherwise
	F11-4	1 = Radio, 0 = otherwise
	F11-5	1 = Family or friends, 0 = otherwise
Q12		**Information channel**
	F12-1	1 = Depend on traditional info (TV, radio, printed newspapers, and digital news websites and apps), 0 = otherwise
	F12-2	1 = Depend on authority info (text by phone, government websites, and government apps), 0 = otherwise
	F12-3	1 = Depend on social media, 0 = otherwise
Q13		**Trust government or family**
	F13	1 = Yes, 0 = no

"FID": Feature ID.

[0, 1], the logistic function is used to ensure the resulting probability is between 0 and 1, as follows:

$$p(y=1|X) = \frac{e^{\eta}}{1+e^{\eta}} = \frac{1}{1+e^{-X\beta}} \tag{2}$$

Since the link function is logit, the method of learning the relationship between predictors and a Bernoulli random variable is called binary logistic regression, described as follows:

$$g(u) = \ln\left(\frac{u}{1-u}\right) = X\beta \tag{3}$$

where $u = p(y=1|X)$. When the dependent variable y has J possible outcomes ($J > 2$), the binary logistic regression can be generalized into the multinomial logistic regression. Given predictors X, the probability of each category j can be modeled. For each category j, there is a separate linear predictor, namely, $\eta_j = \beta_j^T X$. The odds of category j relative to the baseline category J are given by

$$\frac{P(Y=j|X)}{P(Y=J|X)} = e^{\eta_j} = e^{X\beta_j} \tag{4}$$

where $\eta_j = X\beta_j$ for $j = 1, \ldots, J-1$ and $\eta_J = 0$ (baseline category). The link function in the multinomial logistic regression relates the probabilities of the different outcome categories to a set of linear predictors, which is often expressed in terms of odds ratios compared to a baseline category (usually the last category J)

$$g_j(u) = \log\left(\frac{P(Y=j|X)}{P(Y=J|X)}\right) = \eta_j = X\beta_j \tag{5}$$

Therefore, the probability of category j is given by a softmax function

$$P(Y=j|X) = P(Z_j > Z_k, \quad \forall\ k \neq j) = \frac{e^{\eta_j}}{\sum_{k=1}^{J} e^{\eta_k}} \tag{6}$$

The outcome category with the highest associated probability is designated as the predicted class.

Multinomial Logit Model under a Random Utility Maximization Framework

In econometrics, discrete choice models are used to describe and predict choices made by people among multiple discrete alternatives. Discrete choice can be modeled using utility theory (McFadden 2001), positing that rational agents maximize individual utility functions, i.e., utility maximization by making choices. Based on prior beliefs, modelers select the most suitable functional form (relevant attributes only), e.g., coefficient specification (positive versus negative), and error term distribution for each alternative choice (Gumbel versus normal distribution). After being calibrated with data, discrete choice models can be used as a causal model to explain how people make choices and thus are used to explain people's behaviors (Cao et al. 2022).

Suppose agent i faces a choice among J alternatives, the alternative with the highest utility U_{ij} would be chosen, which is formulated as

$$y_i = \text{argmax}_{j\in\{1,\ldots,J\}}\{U_{ij}\} \tag{7}$$

The utility U_{ij} consists of two parts: systematic utility and random utility. Systematic utility V_{ij} is typically modeled as a linear combination of observed attributes. Therefore, V_{ij} can be written as V_j for all agents. The random utility ϵ_{ij} is used to represent the unobserved factors by the modeler and thus this model is called a *random* utility model (McFadden 2012). Facing the utility of alternative j, i.e., $U_{ij} = V_j + \epsilon_{ij}$, it is assumed that a rational agent i prefers alternative l over j, if $V_j + \epsilon_{ij} < V_l + \epsilon_{il}$, $\forall\ j \neq l$.

Numerous random utility models exist, which differ in the underlying assumptions, such as the assumed probability distribution for the agent's random utility. Particularly, when the random utility follows a Gumbel distribution, the random utilities are independently and identically distributed (i.i.d.) across alternatives, and the random utilities are i.i.d. across rational agents, the resulting random utility maximization (RUM) model is called the logit model. If $J > 2$, it is called a multinomial logit (MNL) model. The probability P_l that an agent chooses alternative l among a set of alternatives C can be expressed as

Table 4. Coding of features (Part 2)

QID	FID	Keyword and options
Q18		**Preference for hotels and shelters**
	F18-1	1 = Very unlikely or unlikely to prefer smaller shelter such as government-contracted hotels, 0 = other
	F18-2	1 = Very unlikely or unlikely to prefer the shelter provides people with spacious room, 0 = other
	F18-3	1 = Very unlikely or unlikely to prefer the shelter deep cleans its space before people are admitted, 0 = other
	F18-4	1 = Very unlikely or unlikely to prefer the shelter takes body temperatures before people are admitted, 0 = other
	F18-5	1 = Very unlikely or unlikely to prefer the shelter can provide COVID-19 tests, 0 = other
	F18-6	1 = Very unlikely or unlikely to prefer the shelter requires people to wear a mask all the time except while eating or drinking, 0 = other
	F18-7	1 = Very unlikely or unlikely to prefer the shelter provides hand sanitizers to use, 0 = other
	F18-8	1 = Very unlikely or unlikely to prefer the shelter provides separate areas to isolate people with cold-like symptoms, 0 = other
	F18-9	1 = Very unlikely or unlikely to prefer the shelter recommended by Google Maps and Yelp, 0 = other
Q19		**Evacuation concern**
	F19-1	1 = Don't want to go to public shelters, 0 = otherwise
	F19-2	1 = Concerned about the traffic during evacuation, 0 = otherwise
Q20		**Health issue**
	F20	1 = Have any of specific health issues, 0 = otherwise
Q21		**Know someone diagnosed with COVID-19**
	F21	1 = Yes, 0 = no
Q22		**COVID-19 positive**
	F22	1 = Yes, 0 = no
Q23		**Chance of getting COVID-19**
	F23	1 = very unlikely, 2 = unlikely, 2.5 = no opinion, 3 = likely, 4 = very likely
Q24		**Hygiene practice for COVID-19**
	F24	The number of practices that are never or rarely done
Q25		**Age**
	F25-1	1 = less than 15, 0 = otherwise
	F25-2	1 = 16 ~ 30, 0 = otherwise
	F25-3	1 = 31 ~ 45, 0 = otherwise
	F25-4	1 = 46 ~ 60, 0 = otherwise
	F25-5	1 = 61 ~ 70, 0 = otherwise
	F25-6	1 = 71 ~ 80, 0 = otherwise
	F25-7	1 = over 80, 0 = otherwise
Q26		**Gender**
	F26	1 = Male, 0 = otherwise
Q27		**Have children**
	F27	1 = Yes, 0 = no
Q31		**College or higher degree**
	F31	1 = Yes, 0 = no
Q32		**Income**
	F32-1	1 = Less than $25,000, 0 = otherwise
	F32-2	1 = $25,000 to $49,999, 0 = otherwise
	F32-3	1 = $50,000 to $99,999, 0 = otherwise
	F32-4	1 = $100,000 and above, 0 = otherwise
		Evacuation zone
	F33	1 = Zone A or B, 0 = otherwise
	F33-1	1 = Zone A, 0 = otherwise
	F33-2	1 = Zone B, 0 = otherwise
	F33-3	1 = Zone C, 0 = otherwise
	F33-4	1 = Zone D, 0 = otherwise
		County
	F34	1 = Duval County, 0 = otherwise
		Seaside city
	F35	1 = Yes, 0 = no

$$P_l = \frac{e^{V_l}}{\sum_{j \in C} e^{V_j}} \tag{8}$$

Note that Eq. (8) structurally resembles Eq. (6). The probability formula arises from the key assumption that the random utility components ϵ_{ij} follow a Gumbel distribution, leading to the property that the difference between two Gumbel-distributed variables is logistically distributed. Thus, the probability that the utility of alternative l exceeds that of any other alternative j can be modeled using this closed-form logistic function.

Aside from MNL, there are more various random utility models under different assumptions. For instance, the multinomial probit

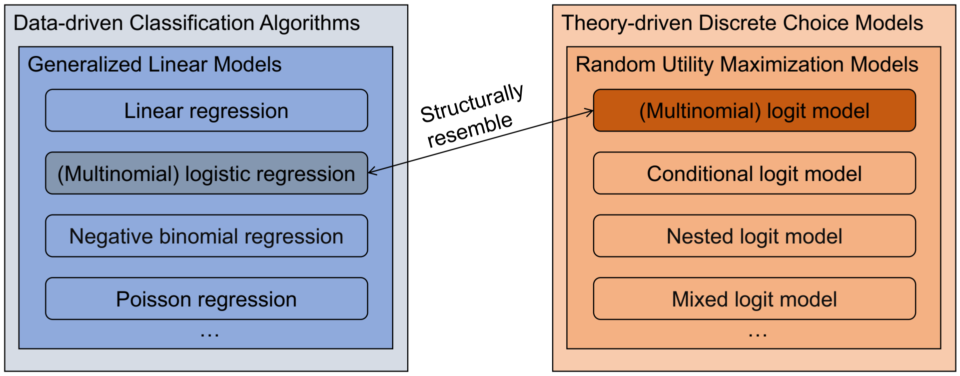

Fig. 3. Relation between data-driven classification algorithms and theory-driven discrete choice models.

model (MNP) has a similar utility function, in which the random utility, however, follows a multivariate normal distribution rather than a Gumbel distribution; the nested logit model further considers the correlation in error terms; the mixed logit model additionally allows random preference variation across individuals. Such discrete choice variants significantly differ from multinomial logistic regression.

Relationship between Multinomial Logistic Regression and Multinomial Logit Model

In the machine learning community, multinomial logistic regression is a common classification algorithm that predicts the probabilities of different outcomes of a categorical dependent variable as a function of the independent variables. For instance, (Arbabzadeh and Jafari 2017) predicted driving safety risk with a regularized multinomial logistic regression using data from the second strategic highway research program (SHRP 2) naturalistic driving study. It can also be used to classify whether a rat is obese or predict the primary food choice for alligators. Multinomial logistic regression is considered a data-driven approach, as it is a variant of GLMs.

Discrete choice models depend on behavioral theory, which is why discrete choice models are considered theory-driven. Under the random utility maximization framework, as demonstrated earlier, when certain assumptions hold, such as the random utility following a Gumbel distribution, a specific variant of discrete choice models structurally resembles the multinomial logistic regression. However, when such an assumption is absent, the resulting discrete choice model variants diverge significantly from multinomial logistic regression.

More importantly, even though a variant of discrete choice models, namely, the multinomial logit model, resembles multinomial logistic regression, we note that multinomial logistic regression is not a utility-based approach. In choice modeling, utility is a pivotal concept representing the perceived value or satisfaction derived from selecting a particular option, reflecting the inherent preferences and decision making processes of individuals. However, multinomial logistic regression operates without reference to utility; it is purely data-driven, focusing solely on discerning relationships between independent variables and the probabilities of various outcomes. This lack of a behavioral foundation means multinomial logistic regression does not assume a utility-based decision making process, thus precluding the attribution of causal relationships. Conversely, due to their behavioral grounding, results from discrete choice analysis offer insights into human decision making processes.

Fig. 3 further illustrates the relationship between data-driven classification algorithms, such as GLMs, and theory-driven discrete choice models. GLMs, designed to capture statistical relationships, stand in contrast to random utility maximization models, which are rooted in economic theory and provide the decision making rationale of rational agents. Although most GLM variants do not align with the majority of random utility maximization variants, it is noteworthy that multinomial logistic regression, as a GLM variant, bears a resemblance to the multinomial logit model, a specific variant of discrete choice models. However, this resemblance does not detract from the data-driven nature of multinomial logistic regression, which remains widely embraced in the machine learning community. Conversely, the multinomial logit model finds extensive use in econometrics, intrinsically linked with utility theory. Table 5 further delineates the distinctions between the logit model and logistic regression in different contexts. For more in-depth comparisons of two choice modeling paradigms, readers are directed to (van Cranenburgh et al. 2022).

In this study, multinomial logistic regression is adopted as a data-driven approach, making no behavioral assumptions and relying solely on data to ascertain relationships between independent and dependent variables. Unlike theory-driven models, we abstain

Table 5. Comparison of logit model and logistic regression

Problem	Logit Model under the RUM framework	Logistic Regression
Evacuation mode choice for coastal residents	Appropriate; supported by behavioral theories; model results are interpretable	Appropriate; relations between evacuation mode choices and independent variables can be learned from the data
Determine whether rats are obese based on heartbeat, blood glucose level, etc.	Not appropriate; no rational agents are involved; utility theory does not apply; no relevant studies can be found	Appropriate; relevant studies can be found
Sentiment analysis of digital texts	Not appropriate; no rational agents are involved; utility theory does not apply; no relevant studies can be found	Appropriate; keywords or other features can be used to determine whether a document has positive sentiment or negative sentiment; relevant studies can be found

from analyzing causal relationships, prioritizing high classification accuracy instead, holding significance in government-led evacuation planning.

Random Forest

Besides, the RF algorithm constructs a user-defined number B of decision trees, each trained independently on a unique bootstrap sample of the original dataset. For each terminal node within a tree, one variable out of m randomly selected ones is chosen to bifurcate the node, based on the principle of impurity minimization, until a predefined stopping criterion is met. The final prediction is made by a majority vote across all trees. RF leverages collective wisdom, enhancing robustness against overfitting and generally outperforming a single decision tree in multiclass classification tasks (Ho 1995).

For multiclass classification problems, impurity minimization, a central tenet of RF, is typically measured using Gini impurity. The Gini impurity for a node with K classes is computed as

$$\text{Gini} = 1 - \sum_{k=1}^{K} (p_k)^2 \tag{9}$$

where p_k = proportion of samples belonging to class k at the node. The algorithm aims to minimize the Gini impurity when selecting the optimal split at each node. The process ensures that the algorithm selects the most informative features to split on, thereby enhancing the predictive capability of the RF model.

Support Vector Machine

Similar to multinomial logistic regression and RF, support vector machine (SVM) is a common machine learning algorithm for classification tasks (Pisner and Schnyer 2020), which aims to find the optimal boundary that separates different classes in the feature space. However, unlike multinomial logistic regression that only considers linear boundaries, SVM is capable of finding both linear and nonlinear boundaries by transforming the feature space using kernel functions.

The fundamental idea of SVM is to construct a hyperplane in the n-dimensional feature space that distinctly classifies the data points into different categories. Mathematically, the equation of the hyperplane is given by

$$\mathbf{w} \cdot \mathbf{x} - b = 0 \tag{10}$$

where $\mathbf{w}$ = weight vector; $\mathbf{x}$ = feature vector; and b = bias term.

The objective of SVM is to maximize the margin between the hyperplane and the nearest data point from each class, which is known as the support vector. Therefore, the optimization problem can be formulated as

$$\underset{\mathbf{w},b}{\text{Max}} \frac{2}{|\mathbf{w}|} \tag{11}$$

$$\text{s.t. } y_i(\mathbf{w} \cdot \mathbf{x}_i - b) \geq 1, \quad \forall\ i = 1, \ldots, N \tag{12}$$

where y_i = label of the i-th data point; and N = number of data points.

When the data are not linearly separable, SVM employs kernel functions to map the feature space to a higher dimensional space where a hyperplane can be used to separate the data points. The most commonly used kernel functions include linear, polynomial, and radial basis function (RBF) kernels.

SVM is known for its effectiveness in high-dimensional spaces and its flexibility to handle different types of data, including text and images. It has been successfully applied in various fields such as bioinformatics, image recognition, and natural language processing.

The implementation of these classification algorithms in this study is facilitated by a Python library, scikit-learn, a popular toolkit known for its efficient algorithms for machine learning and statistical modeling.

Hyperparameter Tuning and Model Evaluation

For learning algorithms, some parameters can be learned from data, such as β_j in Eq. (5), while hyperparameters, which control the learning process, cannot be learned from data and must be tuned or optimized to maximize a learning algorithm's performance.

While multinomial logit model (MLR) does not actually have hyperparameters for tuning, a few settings used in scikit-learn's implementation of MLR are optimized. The *tolerance for stopping criteria* and *maximum number of iterations* are searched on ranges $[1e-8, 1e-2]$ and $[10, 150]$, respectively. The solvers to be searched include "newton-cg," "lbfgs," "sag," and "saga."

A variety of model configurations for RF are considered by varying three hyperparameters *number of estimators*, *minimum number of samples at a leaf node*, and *minimum number of samples to split a node* on ranges $[10, 300]$, $[1, 21]$, and $[1, 21]$, respectively. Two functions evaluating split quality are tried, namely "gini" and "entropy." The proportion of considered features for evaluating splits has a range of $[5\%, 100\%]$.

For the SVM classifier, we have explored a range of hyperparameter values to determine the optimal configuration. The *regularization parameter* is varied over the values $[0.1, 1, 10, 100]$ to control the trade-off between minimizing a training error and a testing error. The *kernel parameter*, with options "linear" and "rbf," determines the type of hyperplane used to separate the data. The *gamma* parameter, with value options $[0.01, 0.1, 1, 10]$, defines the influence of a single training example.

Since it is virtually impossible to test all possible combinations of hyperparameters with grid search, Bayesian optimization, which is designed to optimize black-box functions, is employed for hyperparameter tuning in this study. The iterative process of Bayesian optimization can be stated as follows: given all observations of the objective function (accuracy as a function of hyperparameters) whose structure is unknown, a prior is placed over the random function; after evaluating the function, which yields a new observation, the prior is updated to obtain the posterior distribution over the unknown function; the posterior distribution is then used to construct an acquisition function that finds the next point for evaluation. Compared to evaluating the black-box function, the evaluation of the acquisition function is inexpensive. More details of hyperparameter tuning are available in Yang and Shami (2020).

Five-fold cross-validation is used to evaluate the performance of a classification model. We use accuracy as the evaluation metric, which is computed as the number of correct classifications divided by the total number of predicted samples.

Results

Feature Importance

First, we identify the most significant factors of hurricane evacuation decisions, namely, *EvacVolunt* and *EvacMandt*. Fig. 4 shows the 15 factors/features with the highest importance scores produced by RF. For both decisions, the order of important features is largely the same. In particular, F3-1 Safety perception under Category 3 and F3-2 Safety perception under Category 2 are the two most important features in predicting *EvacVolunt* and

Nat. Hazards Rev., 2024, 25(4): 04024032

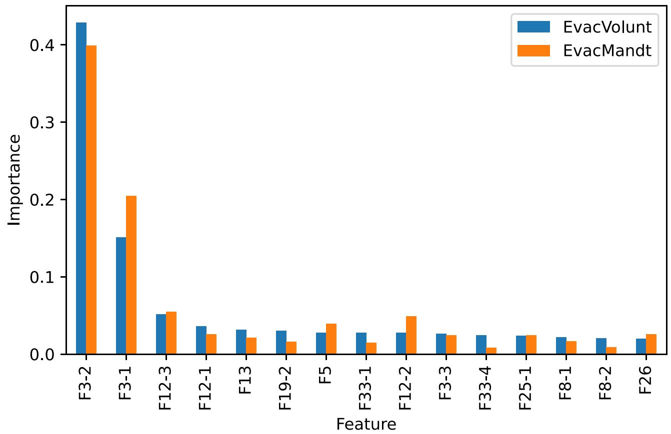

Fig. 4. Normalized feature importance for *EvacVolunt* and *EvacMandt*.

EvacMandt. Note that Tables 3 and 4 provide the coding of each feature. Fig. 4 therefore confirms that the response's perception of a hurricane is the most significant predictor of the evacuation decision, with or without a government-issued evacuation order. This finding is consistent with Whitehead et al. (2000) and Dash and Gladwin (2007).

The strong correlation between hurricane safety perception and voluntary evacuation decision (i.e., *EvacVolunt*) can be further illustrated in Fig. 5. Fig. 5 clearly shows how the evacuation intents vary across two respondent groups: those who perceived it safe to stay at their home if a hypothetical hurricane of Category 3 hit and those who perceived the hurricane risk otherwise. Nearly 90% (30.6% + 57.9%) of the respondents who perceived it safe indicated their evacuation as "unlikely" or "very unlikely." By contrast, about 80% (41.0% + 36.7%) of the respondents who perceived the risk differently indicated their evacuation as "very likely" or "likely." In other words, Fig. 5 effectively illustrates the variation in evacuation intentions between two groups of respondents. Individuals who believe it is safe to remain in their homes during a hypothetical Category 3 hurricane are more inclined to stay, whereas those who perceive a higher risk from the hurricane are more likely to evacuate, even when the evacuation order is voluntary. The correlation between hurricane safety perception and another target variable *EvacMandt* is very similar; therefore, it is not plotted or analyzed further.

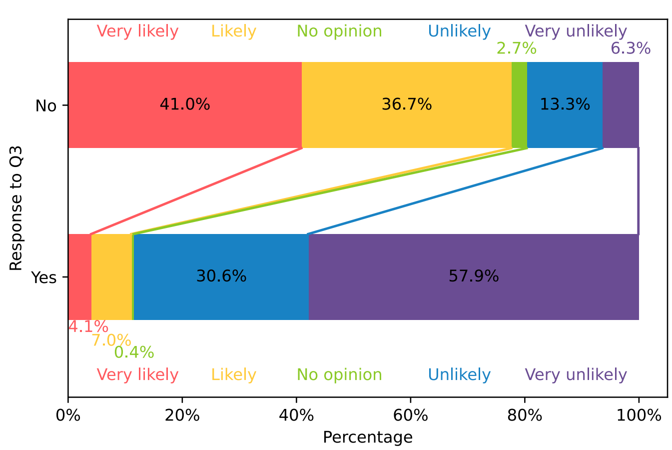

Fig. 5. Hurricane safety perception and evacuation intent.

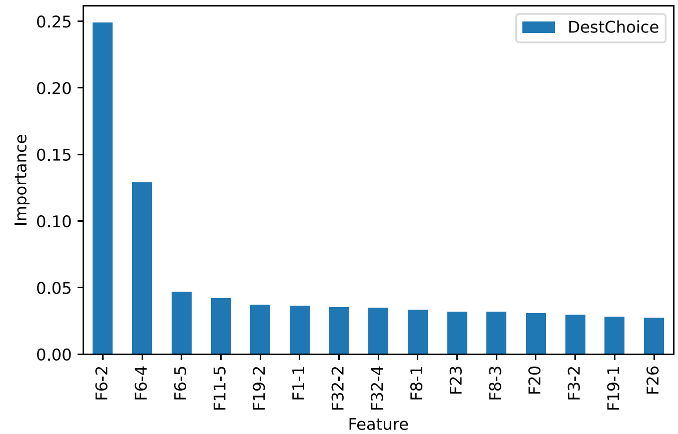

Fig. 6. Normalized feature importance for *DestChoice*.

Fig. 6 shows that the leading features in predicting *DestChoice* are all associated with Question 6, which asked about previous evacuation destination types (e.g., F6-2 Hotel or Model, F6-4 Stayed with families or relatives, and F6-5 Stayed with friends). Note that F23 Self-perceived chance of getting COVID-19 and F20 Existing health conditions appeared in Fig. 6 but not in Fig. 4. In other words, one's self-estimated chance of becoming sick with COVID-19 (i.e., COVID-19 risk perception) and existing health conditions will affect the destination choice, while they are not notable factors for evacuation intentions.

As the choice made during previous evacuations has proven to be a significant predictor of future evacuation destinations (referred to as *DestChoice*), we next examine their relationships. Among the respondents who answered Q5 asking whether they evacuated in the previous hurricanes, only 249 (43.5%) indicated that they evacuated, whose evacuation destination choices were distributed as shown in Fig. 7. Nearly half of the evacuees went to families or relatives' houses, followed by 25.8% evacuees who went to hotels or motels. By contrast, the public shelters provided by governments only accommodated a small fraction of evacuees (3.8%). Similar destination preferences were reported in Whitehead (2003), in which 70.2% stayed with friends or family but only 5.5% went to public shelters. Fig. 8 shows the correlation between past and future evacuation destinations for the 249 respondents. Unsurprisingly, staying with families or friends or at hotels/motels are the three popular choices. The very high correlation explains why

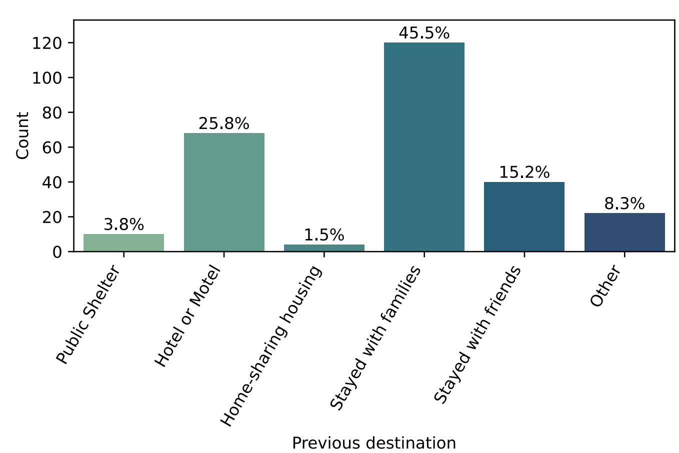

Fig. 7. Distribution of previous evacuation destinations.

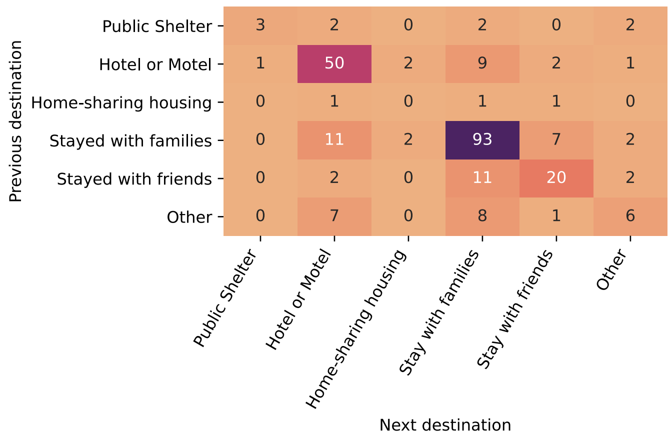

Fig. 8. Correlations of past and future evacuation destinations.

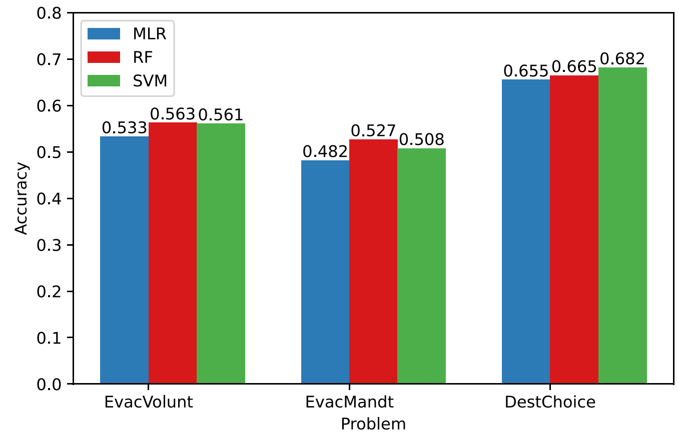

Fig. 10. Classification algorithm comparisons for three target variables.

the past evacuation destination choice is the most important predictor of the future destination choice.

It should be noted that among the 249 evacuees, 222 (89.2%) used personal vehicles as the evacuation transportation mode. Similar findings have also been reported by Lindell et al. (2011) (90%) and Southworth and Chin (1987) (90% in a nighttime evacuation).

As some features about the evaluation of previous evacuation experience are found significant for the classifications in Fig. 6, Fig. 9 presents the evaluation of previous evacuation experience. It can be seen that most respondents were satisfied with their previous evacuation. The experience ratings did not vary significantly over various aspects of evacuations (i.e., lodging, transportation, and overall cost).

Model Evaluation

To achieve the highest classification accuracy, the 29, 33, and 30 features with the highest importance scores are eventually selected for three target variables, i.e., *EvacVolunt*, *EvacMandt*, and *DestChoice*, respectively. As five-fold cross-validation is performed, Fig. 10 examines the comparative accuracies of three classification algorithms across different problems. The findings are as follows: SVM and RF consistently outperform MLR across all categories. Accuracies differ across the classification tasks, with the highest accuracy achieved in *DestChoice* and the lowest in *EvacVolunt*. The mean accuracies of RF for the three problems are 0.56, 0.53, and 0.66, respectively. RF performs the best in the third

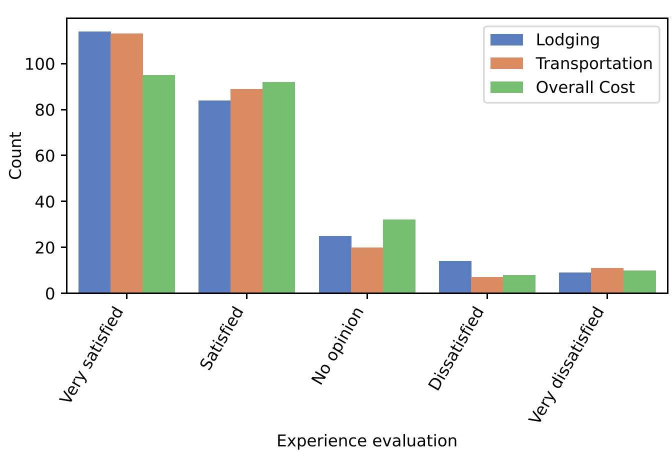

Fig. 9. Evaluations of previous evacuation experience.

classification problem *DestChoice*. RF's mean accuracy is slightly over 0.5. There are two possible reasons for the inferior performance of MLR. First, there might be a multicollinearity issue in the features, such as between F3-1 Safety perception under Category 2 and F3-2 Safety perception under Category 3. Second, MLR, as a generalized linear model, predicts the output depends on the linear combination of inputs, while the linear assumption may not hold in this context. A nonlinear combination of inputs might better characterize the underlying relation.

We next examine the performance of RF by analyzing the confusion matrix. Figs. 11–13 show the confusion matrices for three classification problems. In each of the figures, the left matrix is without normalization and the right one is normalized. The diagonal elements of the confusion matrix represent correct predictions, while off-diagonal elements are those mislabeled by the classifier.

Fig. 11 indicates that when the true evacuation intent is "very likely," the recall is 92%, which is very high. Note that *recall* measures the percentage of relevant samples that are correctly retrieved. Similarly, when the true evacuation intent is "unlikely," the recall is 81%. When the actual intent is "no opinion," the recall is pretty low, as it is frequently mispredicted as "likely" or "unlikely." It should be noted that there is only one sample where the true intent is "very unlikely." The misprediction by RF as "very likely" for this lone sample could be considered an outlier prediction. Therefore, despite the overall accuracy of 56%, the recall achieved by RF for some important classes, such as "very likely" and "unlikely," is substantially high.

Fig. 12 shows similar results for *EvacMandt*: when the true label is "unlikely," the classification recall is as high as 89%; when the true label is "no opinion," the classification recall is quite low. When two classes "very likely" and "likely" are merged as "to evacuate," the classification recall for this class is $60/70 = 86\%$. Similarly, when "unlikely" and "very unlikely" are merged as "to stay," the recall is $25/30 = 83\%$. In addition, when classes are merged for *EvacVolunt*, RF can achieve very high recalls: 95% for "to evacuate" and 79% for "to stay."

A significant improvement in recall is expected, considering that the reduction in the number of classes due to merging inherently simplifies the classification task. Such metrics are comparable to the performance of other binary classification tasks in the literature. For instance, (Zhao et al. 2020a) reported 87.0% as classification accuracy when studying the preevacuation behavior of individuals in a cinema, namely classifying the normal stage (NS) and response stage (RS). It is worth noting that the training data used by

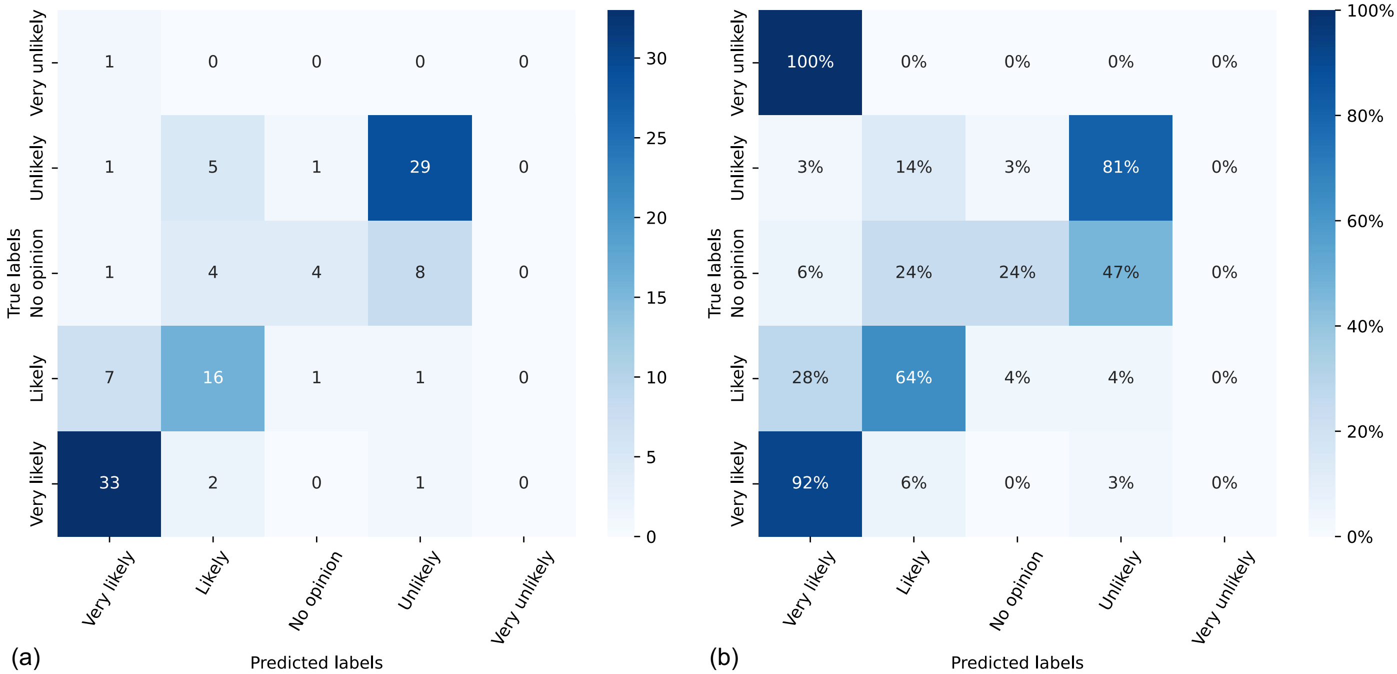

Fig. 11. Confusion matrix for *EvacVolunt*: (a) without normalization; and (b) normalized.

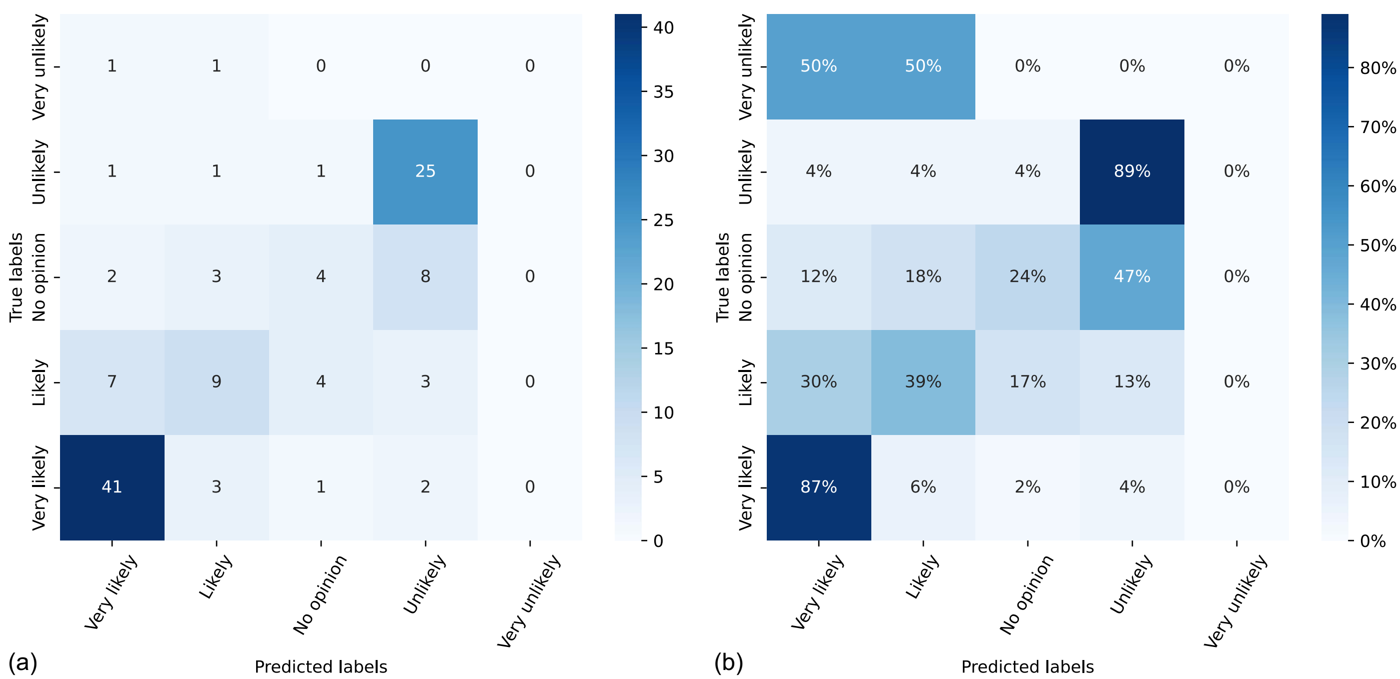

Fig. 12. Confusion matrix for *EvacMandt*: (a) without normalization; and (b) normalized.

Zhao et al. (2020a) is imbalanced: NS has a much higher percentage (i.e., 79.9%) than RS. For this imbalanced classification task, a baseline classifier based on the zero rule, meaning that predicting the majority class all the time can yield an accuracy of 79.9%, which is not very far from 87.0%. In addition, the sample size is over 5,000 in the work by Zhao et al. (2020a), which is much larger than in this study.

As illustrated in Fig. 13, since 64% of respondents selected *FamiFrie* as the destination (which implies imbalanced classes), *FamiFrie* is much easier to predict. 94% of *FamiFrie* respondents are successfully detected. By contrast, RF can correctly predict 55% of respondents choosing *HMHSharing*.

Application of the Proposed Data-Driven Model

The proposed data-driven model for predicting the evacuation decisions of residents can be incorporated into a government-led

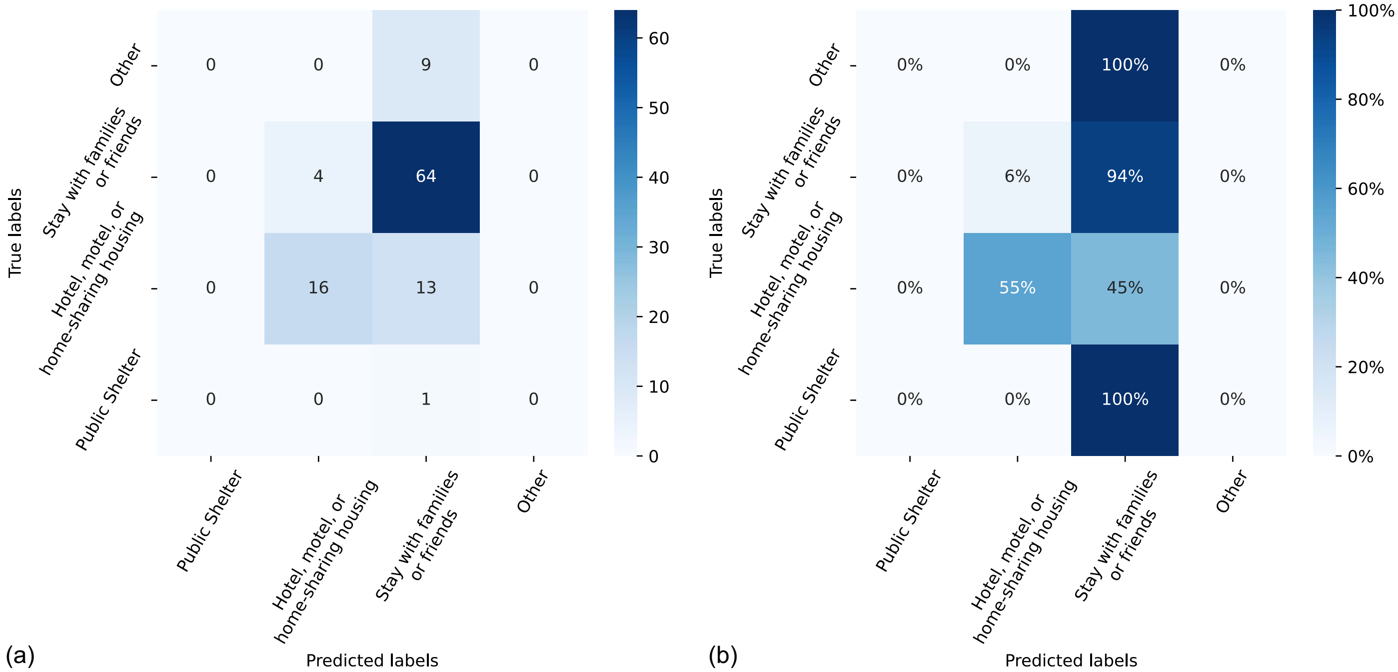

Fig. 13. Confusion matrix for *DestChoice*: (a) without normalization; and (b) normalized.

evacuation planning framework. For resident i, we define α_i and β_i as those factors for evacuation decisions that can and cannot be influenced by the emergency management authority (or interchangeably government), respectively. Since α_i can be affected by a government decision, denoted as x, we also write α_i explicitly as $\alpha_i(x)$. The decision variable y_i denotes the evacuation choice of resident i. For simplicity, y_i can be binary, where $y_i = 1$ indicates the decision to evacuate, and $y_i = 0$ indicates the decision to stay. Alternatively, y_i could be expanded to represent other decisions such as mode of evacuation, time of departure, and route choice. The hurricane evacuation decision making process of resident i can be represented by a conceptual optimization model, as follows:

$$\underset{y_i}{\text{Max}}\ h(y_i, \alpha_i(x), \beta_i) \tag{13}$$

$$\text{s.t. } g(y_i, x) \le 0, \quad \forall\ i \tag{14}$$

In this conceptual decision making problem, resident i pursues some objective, namely, Eq. (13), subject to constraint (14). The government's intervention decision, denoted as x, can impose direct restrictions on the options available to residents. For example, if the government does not provide transit, it restricts residents from using it for evacuation, and similarly, if a roadway is blocked, residents are prevented from using that route. Therefore, in the conceptual optimization model, x appears in constraint (14). Since a government decision x can affect factors $\alpha_i(x)$, it can indirectly influence the evacuation decision made by resident i.

The individual decision making problem, namely, Eqs. (13) and (14), is unfortunately not fully understood by the government. The data-driven model can thus be employed by the government to predict evacuation decisions of resident i for given government decision x, although the predicted decision could deviate from the true decision with a certain probability. Such a data-driven prediction routine can be represented by $y_i = \hat{f}(x)$, $\forall\ i$. Then, the government can solve its evacuation planning optimization problem, conceptually written as

$$\underset{x}{\text{Min}}\ p(x, y_i) \tag{15}$$

$$\text{s.t. } q(x) \le 0 \tag{16}$$

$$y_i = \hat{f}(x), \quad \forall\ i \tag{17}$$

where the government seeks to achieve its goal Eq. (15), such as minimizing evacuation makespan (i.e., the total time required for the entire evacuation process) or maximizing evacuation rate (i.e., percentage of people who evacuate), subject to some resource constraint Eq. (16). Constraint (17) gives the anticipated decision of resident i for a given government decision x. The government's objective Eq. (15) depends on both x and y_i. As the government is unable to predict a resident's decision with 100% accuracy, the government can optimize the expectation of its objective. Detailed formulations for such government-led evacuation planning problems and relevant case studies should be developed in the future.

Conclusions

Unlike most theory-driven hurricane evacuation behavior research in the literature, this study presents a data-driven approach for classifying evacuation decisions of evacuees. A mixed-mode survey was conducted in two metropolitan areas of Florida during the 2020 Atlantic hurricane season. Three widely used supervised learning algorithms, i.e., multinomial logistic regression, RF, and SVM, have been trained on the survey data and evaluated through five-fold cross-validation. We highlight the following findings:

1. Hurricane risk perception is the most significant predictor for both voluntary and mandatory evacuation decisions.
2. Regarding the choice of destination types, respondents tend to follow their prior choices.
3. RF outperforms multinomial logistic regression in classifying evacuation decisions and destination choices.

4. RF achieves very high recall (e.g., over 80%) for certain classes, such as "very likely" and "likely" to evacuate. A high recall in these categories means our model is efficiently capturing a significant portion of individuals genuinely inclined to evacuate. When some related evacuation intents are merged, higher classification recall is achievable.
5. The COVID-19 risk perception and existing health conditions will affect the evacuation destination choice but not the evacuation intent.

The developed data-driven model for hurricane evacuation behavior classification can be incorporated into a government-led evacuation planning framework. For example, the government can estimate the number of evacuees if some factors are changed and maximize the evacuation compliance rate subject to some resource constraints, as illustrated in the section "Application of the Proposed Data-Driven Model." Additional work is needed to connect the evacuation behavior modeling literature to the government-led evacuation planning literature.

The size of the sample used in this study is limited. Gathering a significant number of responses through a survey can be costly, yet it is valuable in order to attain improved accuracy in classification. Considering that evacuation behavioral data are typically very expensive to collect, implying the number of samples is usually limited, transfer learning could be conducted to reuse a pretrained model based on data from nearby or similar jurisdictions to improve the generalization of a new data-driven model being developed by a jurisdiction facing data availability issues.

While a data-driven approach has its comparative advantage, it is not meant to replace theory-driven models, which are essential if interpretable results or causal relations are needed. Instead, given its better predictive capability, the data-driven modeling paradigm serves as an attractive alternative, which can complement the existing theory-driven approaches (van Cranenburgh et al. 2022).

Data Availability Statement

The computer codes for survey data processing and data-driven modeling are available from the corresponding author by request. The raw survey data are protected and confidential in nature.

Acknowledgments

The authors received support from Florida State University under two funding programs, Multidisciplinary Support (MDS) and Collaborative Collision: COVID-19 Pandemic. Constructive comments from three anonymous reviewers and the handling editor are greatly appreciated.

References

Apivatanagul, P., R. A. Davidson, and L. K. Nozick. 2012. "Bi-level optimization for risk-based regional hurricane evacuation planning." *Nat. Hazard.* 60 (Jan): 567–588. https://doi.org/10.1007/s11069-011-0029-9.

Arbabzadeh, N., and M. Jafari. 2017. "A data-driven approach for driving safety risk prediction using driver behavior and roadway information data." *IEEE Trans. Intell. Transp. Syst.* 19 (2): 446–460. https://doi.org/10.1109/TITS.2017.2700869.

Baker, E. J. 1991. "Hurricane evacuation behavior." *Int. J. Mass Emerg. Disasters* 9 (2): 287–310. https://doi.org/10.1177/028072709100900210.

Ben-Akiva, M., and S. R. Lerman. 2018. *Discrete choice analysis: Theory and application to travel demand*. Cambridge, MA: MIT Press.

Bian, R., P. Murray-Tuite, P. Edara, and K. Triantis. 2022. "Household hurricane evacuation plan adaptation in response to estimated travel delay provided prior to departure." *Nat. Hazard. Rev.* 23 (3): 04022010. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000557.

Botzen, W. J. W., J. M. Mol, P. J. Robinson, J. Zhang, and J. Czajkowski. 2022. "Individual hurricane evacuation intentions during the COVID-19 pandemic: Insights for risk communication and emergency management policies." *Nat. Hazard.* 111 (2022): 507–522. https://doi.org/10.1007/s11069-021-05064-2.

Bowser, G. 2013. "Determining the differences in hurricane perception and evacuation behavior in the elderly of South Carolina." Ph.D. thesis, Dept. of Geography, Univ. of South Carolina.

Bowser, G. C., and S. L. Cutter. 2015. "Stay or go? examining decision making and behavior in hurricane evacuations." *Environ.: Sci. Policy Sustainable Dev.* 57 (6): 28–41.

Brown, S., H. Parton, C. Driver, and C. Norman. 2016. "Evacuation during Hurricane Sandy: Data from a rapid community assessment." *PLoS Curr.* 8 (Jan). https://doi.org/10.1371/currents.dis.692664b92af52a3b506483b8550d6368.

Cao, X., C. Ding, and J. Yang. 2022. *Urban transport and land use planning: A synthesis of global knowledge*. Cambridge, MA: Academic Press.

Chen, C.-C., and C.-S. Chou. 2009. "Modeling and performance assessment of a transit-based evacuation plan within a contraflow simulation environment." *Transp. Res. Rec.* 2091 (1): 40–50. https://doi.org/10.3141/2091-05.

Collins, J., A. Polen, K. McSweeney, D. Colón-Burgos, and I. Jernigan. 2021. "Hurricane risk perceptions and evacuation decision-making in the age of COVID-19." *Bull. Am. Meteorol. Soc.* 102 (4): E836–E848. https://doi.org/10.1175/BAMS-D-20-0229.1.

Connolly, J. M., C. Klofstad, and J. Uscinski. 2020. "Leaving home ain't easy: Citizen compliance with local government hurricane evacuation orders." *J. Behav. Public Adm.* 3 (2): 1–13. https://doi.org/10.30636/jbpa.32.99.

Dash, N., and H. Gladwin. 2007. "Evacuation decision making and behavioral responses: Individual and household." *Nat. Hazard. Rev.* 8 (3): 69–77. https://doi.org/10.1061/(ASCE)1527-6988(2007)8:3(69).

Dong, H., M. Zhou, Q. Wang, X. Yang, and F.-Y. Wang. 2019. "State-of-the-art pedestrian and evacuation dynamics." *IEEE Trans. Intell. Transp. Syst.* 21 (5): 1849–1866. https://doi.org/10.1109/TITS.2019.2915014.

Florida Division of Emergency Management. 2022. "Know your zone, know your home." Accessed July 1, 2022. https://www.floridadisaster.org/knowyourzone/.

Galindo, G., and R. Batta. 2013. "Review of recent developments in OR/MS research in disaster operations management." *Eur. J. Oper. Res.* 230 (2): 201–211. https://doi.org/10.1016/j.ejor.2013.01.039.

Gladwin, C. H., H. Gladwin, and W. G. Peacock. 2001. "Modeling hurricane evacuation decisions with ethnographic methods." *Int. J. Mass Emerg. Disasters* 19 (2): 117–143. https://doi.org/10.1177/028072700101900201.

Golshani, N., R. Shabanpour, A. Mohammadian, J. Auld, and H. Ley. 2019. "Evacuation decision behavior for no-notice emergency events." *Transp. Res. Part D Transp. Environ.* 77 (Jun): 364–377. https://doi.org/10.1016/j.trd.2019.01.025.

Ho, T. K. 1995. "Random decision forests." In Vol. 1 of *Proc., 3rd Int. Conf. on Document Analysis and Recognition*, 278–282. New York: IEEE.

Huang, S.-K., M. K. Lindell, and C. S. Prater. 2016. "Who leaves and who stays? a review and statistical meta-analysis of hurricane evacuation studies." *Environ. Behav.* 48 (8): 991–1029. https://doi.org/10.1177/0013916515578485.

Jones, T. L., M. Baxter, and V. Khanduja. 2013. "A quick guide to survey research." *Ann. R. Coll. Surg. Engl.* 95 (1): 5–7. https://doi.org/10.1308/003588413X13511609956372.

Lim, G. J., S. Zangeneh, M. R. Baharnemati, and T. Assavapokee. 2012. "A capacitated network flow optimization approach for short notice evacuation planning." *Eur. J. Oper. Res.* 223 (1): 234–245. https://doi.org/10.1016/j.ejor.2012.06.004.

Lindell, M. K., J. E. Kang, and C. S. Prater. 2011. "The logistics of household hurricane evacuation." *Nat. Hazard.* 58 (3): 1093–1109. https://doi.org/10.1007/s11069-011-9715-x.

Lindell, M. K., P. Murray-Tuite, B. Wolshon, and E. J. Baker. 2018. *Large-scale evacuation: The analysis, modeling, and management of emergency relocation from hazardous areas*. Boca Raton, FL: CRC Press.

Lindell, M. K., and R. W. Perry. 2012. "The protective action decision model: Theoretical modifications and additional evidence." *Risk Anal.* 32 (4): 616–632. https://doi.org/10.1111/j.1539-6924.2011.01647.x.

Lindsey, J. K. 2000. *Applying generalized linear models*. New York: Springer.

Ling, L., P. Murray-Tuite, S. Lee, Y. G. Ge, and S. V. Ukkusuri. 2021. "Role of uncertainty and social networks on shadow evacuation and non-compliance behavior in hurricanes." *Transp. Res. Rec.* 2675 (3): 53–64. https://doi.org/10.1177/0361198120962801.

Little, R. J., and D. B. Rubin. 2019. Vol. 793 of *Statistical analysis with missing data*. New York: Wiley.

Lu, W., L. D. Han, C. Liu, and B. L. Bhaduri. 2017. "Impacts of high resolution data on traveler compliance levels in emergency evacuation simulations." *Int. J. Intell. Transp. Syst. Res.* 15 (2): 108–116. https://doi.org/10.1007/s13177-016-0125-z.

Martín, Y., Z. Li, and S. L. Cutter. 2017. "Leveraging Twitter to gauge evacuation compliance: Spatiotemporal analysis of Hurricane Matthew." *PLoS One* 12 (7): e0181701. https://doi.org/10.1371/journal.pone.0181701.

Matyas, C., S. Srinivasan, I. Cahyanto, B. Thapa, L. Pennington-Gray, and J. Villegas. 2011. "Risk perception and evacuation decisions of Florida tourists under hurricane threats: A stated preference analysis." *Nat. Hazard.* 59 (2): 871–890. https://doi.org/10.1007/s11069-011-9801-0.

McFadden, D. 2001. "Economic choices." *Am. Econ. Rev.* 91 (3): 351–378. https://doi.org/10.1257/aer.91.3.351.

McFadden, D. 2012. "Computing willingness–to–pay in random utility models." In *Trade, theory and econometrics*, 275–296. London: Routledge.

Millar, M. M., and D. A. Dillman. 2011. "Improving response to web and mixed-mode surveys." *Public Opin. Q.* 75 (2): 249–269. https://doi.org/10.1093/poq/nfr003.

Murray-Tuite, P., and B. Wolshon. 2013. "Evacuation transportation modeling: An overview of research, development, and practice." *Transp. Res. Part C Emerging Technol.* 27 (Mar): 25–45. https://doi.org/10.1016/j.trc.2012.11.005.

Phillips, J. G., S. Butt, and A. Blaszczynski. 2006. "Personality and self-reported use of mobile phones for games." *Cyberpsychol. Behav.* 9 (6): 753–758. https://doi.org/10.1089/cpb.2006.9.753.

Pisner, D. A., and D. M. Schnyer. 2020. "Support vector machine." In *Machine Learning*, 101–121. Amsterdam, Netherlands: Elsevier.

Ploran, E. J., M. A. Trasciatti, and E. C. Farmer. 2018. "Efficacy and authority of the message sender during emergency evacuations: A mixed methods study." *J. Appl. Commun. Res.* 46 (3): 291–322. https://doi.org/10.1080/00909882.2018.1464659.

Rawls, C. G., and M. A. Turnquist. 2010. "Pre-positioning of emergency supplies for disaster response." *Transp. Res. Part B Methodol.* 44 (4): 521–534. https://doi.org/10.1016/j.trb.2009.08.003.

Roy, K. C., and S. Hasan. 2021. "Modeling the dynamics of hurricane evacuation decisions from Twitter data: An input output hidden Markov modeling approach." *Transp. Res. Part C Emerging Technol.* 123 (Mar): 102976. https://doi.org/10.1016/j.trc.2021.102976.

Sakamoto, M., D. Sasaki, Y. Ono, Y. Makino, and E. N. Kodama. 2020. "Implementation of evacuation measures during natural disasters under conditions of the novel coronavirus (COVID-19) pandemic based on a review of previous responses to complex disasters in Japan." *Prog. Disaster Sci.* 8 (Dec): 100127. https://doi.org/10.1016/j.pdisas.2020.100127.

Southworth, F., and S. M. Chin. 1987. "Network evacuation modelling for flooding as a result of dam failure." *Environ. Plann. A* 19 (11): 1543–1558. https://doi.org/10.1068/a191543.

Stein, R., B. Buzcu-Guven, L. Dueñas-Osorio, D. Subramanian, and D. Kahle. 2013. "How risk perceptions influence evacuations from hurricanes and compliance with government directives." *Policy Stud. J.* 41 (2): 319–342. https://doi.org/10.1111/psj.12019.

Sun, Y., Z. Jiang, J. Gu, M. Zhou, Y. Li, and L. Zhang. 2018. "Analyzing high speed rail passengers' train choices based on new online booking data in China." *Transp. Res. Part C Emerging Technol.* 97 (Mar): 96–113. https://doi.org/10.1016/j.trc.2018.10.015.

Thompson, R. R., D. R. Garfin, and R. C. Silver. 2017. "Evacuation from natural disasters: A systematic review of the literature." *Risk Anal.* 37 (4): 812–839. https://doi.org/10.1111/risa.12654.

van Cranenburgh, S., S. Wang, A. Vij, F. Pereira, and J. Walker. 2022. "Choice modelling in the age of machine learning-discussion paper." *J. Choice Modell.* 42 (Jun): 100340. https://doi.org/10.1016/j.jocm.2021.100340.

Whitehead, J. C. 2003. "One million dollars per mile? The opportunity costs of hurricane evacuation." *Ocean Coastal Manage.* 46 (11–12): 1069–1083. https://doi.org/10.1016/j.ocecoaman.2003.11.001.

Whitehead, J. C., B. Edwards, M. Van Willigen, J. R. Maiolo, K. Wilson, and K. T. Smith. 2000. "Heading for higher ground: Factors affecting real and hypothetical hurricane evacuation behavior." *Global Environ. Change Part B: Environ. Hazards* 2 (4): 133–142. https://doi.org/10.1016/S1464-2867(01)00013-4.

Wong, S. D., A. J. Pel, S. A. Shaheen, and C. G. Chorus. 2020. "Fleeing from Hurricane Irma: Empirical analysis of evacuation behavior using discrete choice theory." *Transp. Res. Part D Transp. Environ.* 79 (Mar): 102227. https://doi.org/10.1016/j.trd.2020.102227.

Xu, N., R. Lovreglio, E. D. Kuligowski, T. J. Cova, D. Nilsson, and X. Zhao. 2023. "Predicting and assessing wildfire evacuation decision-making using machine learning: Findings from the 2019 Kincade fire." *Fire Technol.* 59 (2): 793–825. https://doi.org/10.1007/s10694-023-01363-1.

Yang, H., E. F. Morgul, K. Ozbay, and K. Xie. 2016. "Modeling evacuation behavior under hurricane conditions." *Transp. Res. Rec.* 2599 (1): 63–69. https://doi.org/10.3141/2599-08.

Yang, L., and A. Shami. 2020. "On hyperparameter optimization of machine learning algorithms: Theory and practice." *Neurocomputing* 415 (Jan): 295–316. https://doi.org/10.1016/j.neucom.2020.07.061.

Yao, R., and S. Bekhor. 2020. "Data-driven choice set generation and estimation of route choice models." *Transp. Res. Part C Emerging Technol.* 121 (Dec): 102832. https://doi.org/10.1016/j.trc.2020.102832.

Yao, Z., G. Zhang, D. Lu, and H. Liu. 2019. "Data-driven crowd evacuation: A reinforcement learning method." *Neurocomputing* 366 (Nov): 314–327. https://doi.org/10.1016/j.neucom.2019.08.021.

Zhao, T., M. Jia, T. Tang, and Y. Sun. 2023. "Perception of hurricane and COVID-19 risks for household evacuation and shelter intentions." *Prof. Geogr.* 75 (3): 396–414. https://doi.org/10.1080/00330124.2022.2103722.

Zhao, X., R. Lovreglio, E. Kuligowski, and D. Nilsson. 2021. "Using artificial intelligence for safe and effective wildfire evacuations." *Fire Technol.* 57 (Mar): 483–485. https://doi.org/10.1007/s10694-020-00979-x.

Zhao, X., R. Lovreglio, and D. Nilsson. 2020a. "Modelling and interpreting pre-evacuation decision-making using machine learning." *Autom. Constr.* 113 (May): 103140. https://doi.org/10.1016/j.autcon.2020.103140.

Zhao, X., X. Yan, A. Yu, and P. Van Hentenryck. 2020b. "Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models." *Travel Behav. Soc.* 20 (Jul): 22–35. https://doi.org/10.1016/j.tbs.2020.02.003.

Nat. Hazards Rev., 2024, 25(4): 04024032