

Predicting hurricane evacuation for local neighborhoods across a metropolitan region

Shakhawat H. Tanim^{a,*}, Steven Reader^b, Yujie Hu^c

^a School of Civil and Environmental Engineering and Earth Sciences, Clemson University, United States

^b School of Geosciences, University of South Florida, United States

^c Department of Geography, University of Florida, United States

ABSTRACT

To mitigate the devastating impacts of hurricanes on people's lives, communities, and societal infrastructures, disaster management would benefit considerably from a detailed understanding of evacuation, including the socio-demographics of the populations that evacuate, or remain, down to disaggregated geographic levels such as local neighborhoods. A detailed household evacuation prediction model for local neighborhoods requires both a robust household evacuation decision model and individual household data for small geographic units. This paper utilizes a recently published statistical meta-analysis for the first requirement and then conducts a rigorous population synthesis procedure for the second. Our model produces predicted non-evacuation rates for all US Census block groups for the Tampa-St. Petersburg-Clearwater Metropolitan Statistical Area for a Hurricane Irma-like storm along with their socio-demographic and hurricane impact risk profiles. Our model predictions indicate that non-evacuation rates are likely to vary considerably, even across neighboring block groups, driven by the variability in evacuation risk profiles. Our results also demonstrate how different predictors may come to the fore in influencing non-evacuation in different block groups, and that predictors which may have an outsize impact on individual household evacuation decisions, such as Race, are not necessarily associated with the greatest differentials in non-evacuation rates when we aggregate households to block group level and above. Our research is intended to provide a framework for the design of hurricane evacuation prediction tools that could be used in disaster management.

1. Background

Hurricane disaster management is critical to reduce fatalities and injuries, restore utilities, manage traffic, and provide humanitarian assistance to those immediately impacted. From 1980 to 2021, hurricanes were responsible for approximately 6700 deaths in the United States. Recently, Hurricane Ian, which impacted Florida in September 2022, caused a high number (119) of deaths [1] with most deaths being older adults [2]. Hurricane Katrina in 2005 resulted in 1833 deaths and a study revealed that older adults and those of Black race were disproportionately the victims [3]. To prepare and implement a comprehensive plan for pre- and post-storm disaster management, disaster managers would benefit from a detailed understanding of evacuation, including the socio-demographics of the populations that evacuate, or remain, and preferably down to disaggregated geographic levels such as local neighborhoods. With such information, disaster managers would be able to design and implement effective policies and plans for such things as hurricane evacuation warning dissemination, traffic management, shelter location and management, and placement of emergency services and utility crews.

In current practice in the USA to predict evacuation, disaster managers typically use a so-called *participation rate approach* to determine the raw numbers of households evacuating from different geographic areas. In Florida, the Florida Division of Emergency Management has developed the Statewide Regional Evacuation Study Program [4]. This program helps bodies such as regional

* Corresponding author. Lowry Hall, Clemson, SC 29634, United States.
E-mail address: shtanim@gmail.com (S.H. Tanim).

planning councils (e.g., Tampa Bay Regional Planning Council) to produce predicted evacuation numbers for their transportation analysis, with the primary aim being to simulate evacuation traffic. The program subdivides each county in Florida into Traffic Evacuation Zones (TEZs) and has household counts for both mobile homes, and non-mobile homes for each TEZ. Hillsborough County in Florida, for example, is subdivided into 505 TEZs. The evacuation rates for each TEZ zone are determined by what official evacuation zone each TEZ is within. Evacuation zones are County-designated geographically defined areas that are used for the issuance of evacuation orders. All TEZs in the same evacuation zone receive the same evacuation rates. The Statewide Regional Evacuation Study Program provides five levels of evacuation rates from which an analyst chooses. Within each level, there is a different rate for mobile homes and for non-mobile homes for each official evacuation zone. Although the large number of TEZs zones may provide high spatial resolution for generating traffic counts of evacuating households, it is unrealistic to assume the same evacuation rates for each TEZ within an evacuation zone. In addition, the participation rate approach is only focused on numbers of households evacuating and provides no information on the socio-demographic profiles of households who evacuate or remain.

Ideally, we need hurricane evacuation prediction models that provide detailed local neighborhood predictions of not just numbers of households evacuating/remaining but also their different socio-demographic profiles. There is little academic research into such detailed evacuation prediction modeling. Only one study is known to the authors that attempted such a prediction. Xu et al. [5] predicted hurricane evacuation for 1000 households per census tract for various locations in North Carolina. More generally, studies explore evacuation predictors to simply better understand the factors that influence evacuation and to what extent (e.g. Ref. [6]). Xu et al. [5] based their model of the household evacuation decision on a random telephone survey of 405 households located in North Carolina. The survey used a hypothetical hurricane situation, asking if the household would evacuate if under either a mandatory or voluntary evacuation order. From the survey they fitted ordered probit models for each type of evacuation order, and their survey identified these predictors with their effect sizes: distance from coast, mobile home occupancy, single family home type, gender, race, education, occupation status, age, presence of children, and income. For the locations they were to predict evacuation for, the authors needed to generate individual household level data for census tracts since such data is not typically available. The US Census is prohibited by law from disclosing individual household data in order to safeguard personal information [7]. To generate individual household data, the authors used *population synthesis*, a process of generating 'synthetic' individual household data with a range of socio-demographic variables, using statistical methods based on available Census data. Given their household evacuation decision models and 1000 synthesized households per census tract, they then predicted evacuation for Beaufort County, NC for Hurricane Irene (2011), the Outer Banks area, NC for Hurricane Isabel (2003), and the surge zone area of SE North Carolina for Hurricane Floyd (1999).

Although the work of [5] is promising for local neighborhood level prediction of evacuation, it also has a number of shortcomings. First, their household evacuation decision model is based on one survey. Hurricane evacuation surveys are known to produce considerable heterogeneity in the range of predictors considered and their effect sizes, so reliance on one survey is problematic. Second, for their population synthesis, they did not use any established method for this procedure but rather, for each household, they simply looked at the distribution of each variable in that census tract and then randomly generated a value from each of these distributions, independent of the values on the others, so that no dependencies between variables were considered. Third, Xu et al. [5], only used 10 predictors in their household evacuation decision model and only one of those, distance to the coast, is related to storm impact rather than socio-demographics, although they did model separately for mandatory and voluntary evacuation orders. They also did not consider storm related risk perceptions and their effects on the evacuation decision [8]. Realistic household evacuation models should consider other storm-related predictors such as surge and flood zones. Finally, the geographic level of analysis chosen for prediction was the census tract, whereas population synthesis methods can be used for even more disaggregated areas such as census block groups.

This paper proposes an approach to local neighborhood prediction of hurricane evacuation that overcomes all these shortcomings. It should be emphasized that this paper is meant to be illustrative of the types of spatially and socio-demographically disaggregated predicted rates of evacuation made possible by the use of a synthetically generated population at the fine geographic scale of block groups. It is illustrative in the sense that an actual real-world implementation of the approach would require an analyst to a) select a household-level evacuation prediction model and to b) evaluate the characteristics and risks of an impending hurricane that is appropriate to their situation. In our narrative below, these aspects are decided somewhat arbitrarily since they are not central to illustrating the approach.

Also, when discussing household-level prediction models the discussion will initially focus, in the Methods section, on evacuation, as opposed to its reverse, non-evacuation. This reflects the fact that research on such prediction models, largely dominated by transport engineers, has tended to focus on evacuation rates and evacuation predictors. However, in our results section, the reader will notice that we adopt the reverse term and discuss non-evacuation rates. This is quite a deliberate decision on our part and based on our conviction that the greatest hardship and loss of life is suffered by those who do not evacuate rather than those that do. Anyone witnessing the scenes in New Orleans after Hurricane Katrina or the scenes in Ft. Myers Beach following Hurricane Ian can likely attest to our conviction. We seek to change the narrative and emphasis in "evacuation" accordingly to provide the greatest assistance to those in disaster management assisting those in greatest need.

2. Methods

There are four connected methodological components for this study:

- i A model to predict the probability of any household evacuating based on the socio-demographics of the household and its geographic location relative to hurricane-related risk factors;

- ii The choice of a hurricane scenario;
- iii The creation of a population of ‘synthetic households’ within small geographic areas across a large metropolitan region;
- iv Prediction, summarization, and mapping of evacuation rates based on the predicted probabilities of households’ evacuating.

We address each of these components, in turn, below.

2.1. A model for household evacuation prediction

Virtually all published studies of household evacuation prediction have been based on surveys of one, or more, hypothetical or real hurricanes and have used logistic regression models to estimate the probabilities of households evacuating. As noted above, a disaster management analyst could select any one, or a combination, of such studies and use their estimated effect sizes in a predictive model, presumably making their selection of studies appropriate to their own situation. Since our paper here is illustrative, we decided to use the estimated effect sizes from a recently published statistical meta-analysis. Tanim et al. [9] conducted a statistical meta-analysis of 33 evacuation decision models to investigate the consistency of hurricane evacuation decision predictors, uncover potential sampling errors contributing to predictor heterogeneity, and extract more information regarding effect sizes from hurricane evacuation models rather than relying solely on binary significance data. This statistical meta-analysis estimated mean effect sizes (expressed as odds-ratios) for 23 predictors associated with hurricane evacuation decisions. Of these 23 predictors, 7 (length of residence, hurricane experience, mandatory work, evacuation plan, window protection, pet ownership, and disabled person in household) reflected survey questions used in contributing studies and they have no correspondent data available in the US Census. Since our method of creating ‘synthetic households’ is based on the use of US Census data, these predictors could not be included in our model. Of the remaining 16 predictors, 13 were socio-demographic in nature and had correspondent data in the US Census, while 3 (mandatory evacuation order issued, perceived risk of flood, and perceived risk of surge) were directly related to any hurricane being modeled. Just to note, the distinction between the perceived risk of flood and the perceived risk of surge reflects how individuals evaluate the likelihood and potential impact of inland flooding, attributed to hydrology and hydraulics, as opposed to coastal inundation resulting from surges due to hurricane. In this paper, the odds-ratios derived from Ref. [9]; illustrated in Fig. 1, were adopted and used to derive household evacuation probabilities for our synthesized population via a predictive logit equation. It should be noted that although the [9] research used studies based on surveys for both real and hypothetical hurricanes, the odds-ratios used here reflect those for all studies based on real hurricanes, since they provide the most robust estimates.

2.2. The hurricane scenario

For the 3 predictors directly related to the hurricane being modeled, we need to use a hurricane scenario for illustration. We could have used a hypothetical hurricane given the illustrative nature of our paper, but to add some realism we decided to use Hurricane Irma, a hurricane that traversed our study area – the Tampa-St. Petersburg-Clearwater Metropolitan Statistical Area in 2017. Hurricane Irma was a major hurricane that threatened the Tampa Bay area in early September 2017. The National Hurricane Center (NHC) forecast for Hurricane Irma shifted westward over time, with the threat and forecast confidence increasing for the Tampa Bay area. Hurricane Irma was a major hurricane upon landfall, and a relatively rare event with most of the state within the cone of uncertainty at some point during its movement. As such, it triggered one of the largest evacuations in US history with fuel shortages a key concern [10]. All of these factors no doubt influenced evacuation decisions in the Tampa Bay area. Fig. 2 shows the location and forecast track of Hurricane Irma at 8am on Saturday, September 9th, 2017, when evacuation orders were issued for the study area. Given this, we should emphasize that we are not advancing this paper as an accurate evacuation model for Hurricane Irma given the unusual circumstances regarding Hurricane Irma noted here, and other limitations we will note later. In that sense, it may be better to refer to our hurricane scenario as Hurricane Irma-like.

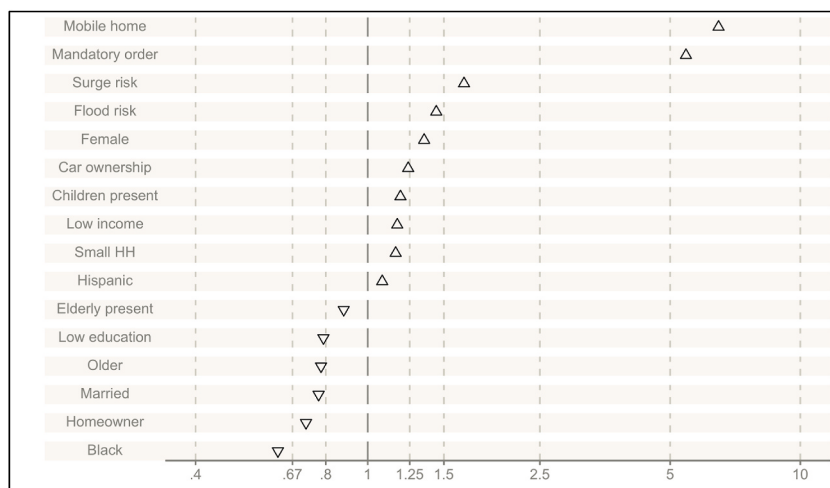


Fig. 1. Mean Effect Sizes (ORs) of Predictors used in our model (from [9]).



Fig. 2. Location and predicted track of Hurricane Irma at 8am on Saturday, September 9th, 2017 [11].

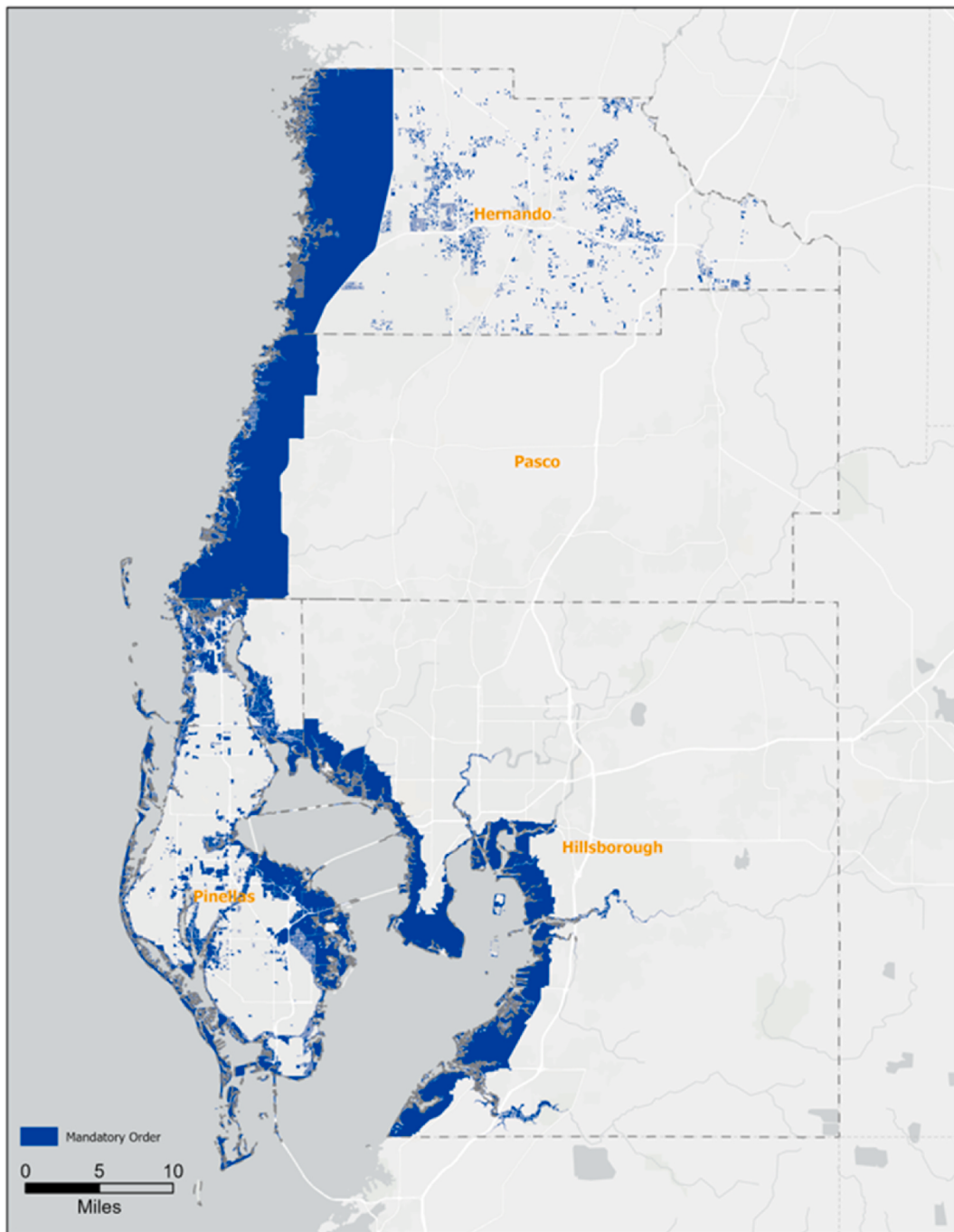


Fig. 3. Mandatory evacuation order locations

To determine the geographic locations under mandatory evacuation orders for our Hurricane Irma-like scenario, we used the different County declarations of such orders as issued on September 9th, 2017 for Hurricane Irma itself (see Appendix, [Figure A1](#)). These locations are illustrated in [Fig. 3](#). The many small inland area locations in both Pinellas and Hernando counties reflect those counties issuance of mandatory evacuation orders for all mobile homes/parks. As such, our model classified any mobile home residing household for these two counties as being under a mandatory evacuation order.

The two predictors of perceived surge risk and perceived flood risk present a challenge in the sense that we have no access to such perceptions in the way that surveys reporting on these measures did. It is also true that risk perception partly assumes a resident is aware of any flood or surge risk zone in which they may be located. Despite these limitations, these predictors are important when it comes to evacuation decisions and while acknowledging the distinction between perceived and actual risk, we can only measure the latter. As such, we determined the location of each block group within high risk surge zones or high risk flood zones, as defined by the Florida Department of Emergency Management [12]. The Florida Department of Emergency Management generates the flood map using FEMA flood maps, which employ hydrology and hydraulics to assess flood risk, while the surge risk map is generated using the

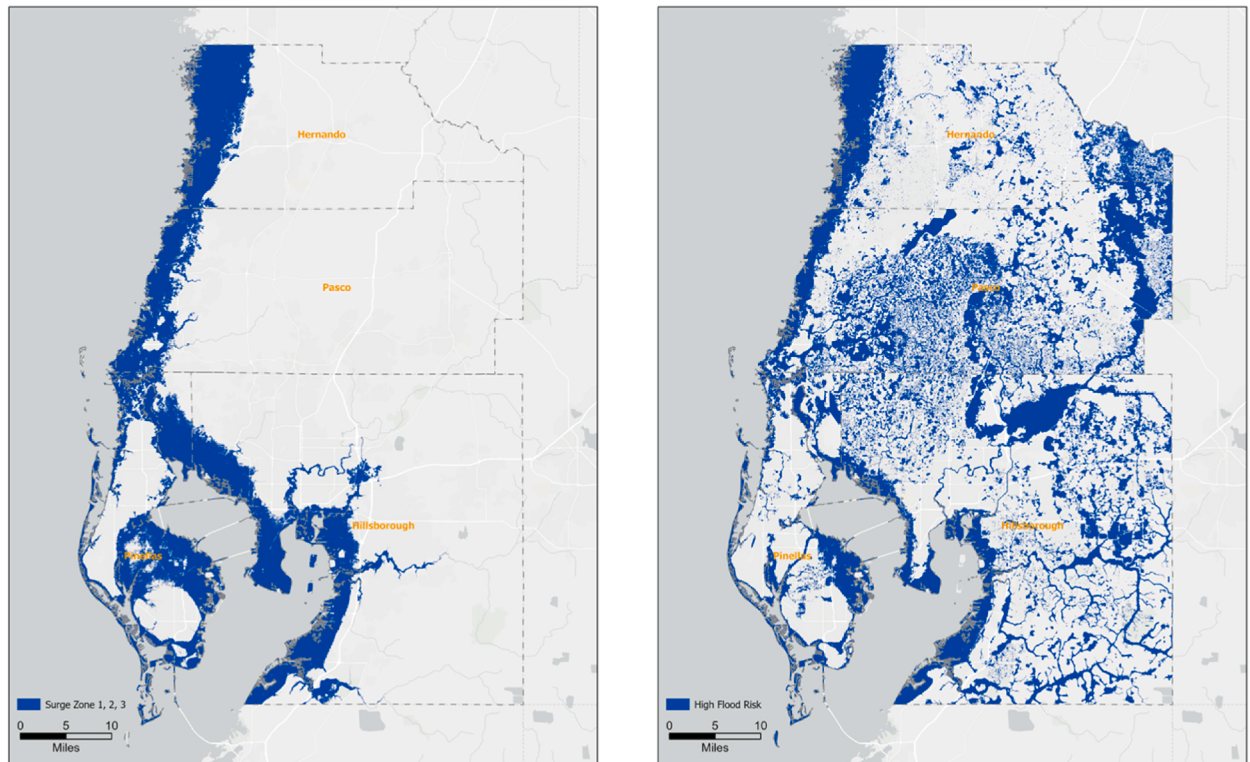


Fig. 4. (a, b): Surge risk and flood risk zones

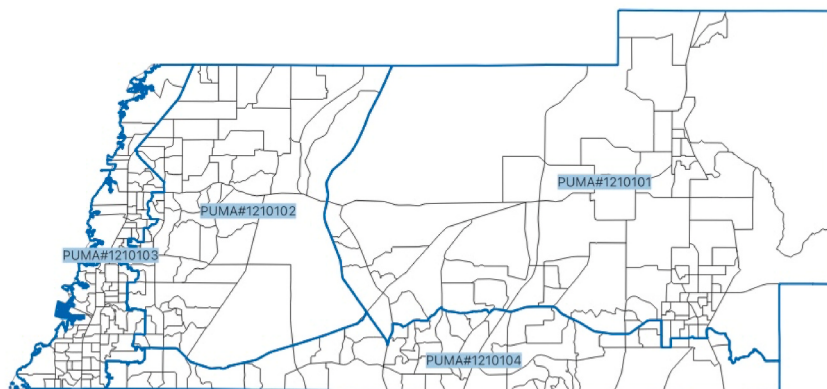


Fig. 5. PUMA's (blue) and block groups (black) of Pasco County, Florida. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article)

Statewide Regional Evacuation Studies of Florida to estimate surge levels corresponding to different hurricane categories. Fig. 4a shows the high risk surge zones for a Category 3 hurricane and Fig. 4b shows the zones deemed to be at high risk of flood.

For all of these hurricane-related predictors (mandatory evacuation order, surge risk, flood risk) we calculated what proportion of the geographic area of a block group was inside any of the zones indicated in Figs. 3–4 (a, b) and randomly assigned that proportion of households in the block group to the relevant designation.

There is another aspect of our predictive logit equation of a household's probability of evacuation that depends on the hurricane being modeled, and it involves the intercept term of the equation. Since all our predictors are either binary or ordinal factors, the intercept term can be interpreted as the estimate of the probability of evacuation for a household in the predictor categories that would yield the lowest evacuation probability. In that sense, it is a base evacuation probability. In using our modeling framework in a real-world situation, this base evacuation probability would be determined based on an analyst's experience and knowledge of evacuations for similar storms in the past. In our demonstration research here, we decided to calibrate this intercept term by setting it at different values, running our equation, and then comparing the predicted evacuation rates that emerged to independent evacuation surveys for Hurricane Irma. In a practical application of our modeling framework a disaster or emergency management analyst could adopt a similar procedure to estimate the base probability – in other words, given chosen effect sizes for chosen predictors, and given an overall estimation of the evacuation rate for the hurricane threat, calibrate the base probability thru iteration.

2.3. Population synthesis for census block groups

Our predictive logit equation to predict hurricane evacuation probabilities is household-based, and our goal is to summarize these probabilities at the fine geographic scale of census block groups. However, to avoid the potential identification of households, the US Census Bureau only releases detailed data on a relatively small sample (5 %) of individual households dispersed across large geographic areas. This data is the Public Use Microdata Sample (PUMS), a subsample of the US Census Bureau American Community Survey, and the geographic areas it is released for are known as PUMA's (Public Use Microdata Areas). Meanwhile, at the census block group level, the Census Bureau releases summary counts of discrete household socio-demographic variables and population counts for other socio-demographic variables. To illustrate the geographic scale relationship of PUMA's to census block groups, Fig. 5 shows the four PUMA's and numerous census block groups designated for Pasco County, FL., one of the 4 counties in our study area.

The process of generating 'synthetic' individual households with a range of socio-demographic variables for small geographic areas, based on using detailed individual household data for larger geographic areas, combined with summary count data for the socio-demographic variables for the small geographic areas, is known as population synthesis. The goal is to generate 'synthetic' households

Step 1: Create multiway table

	Homeownership		Total	BG Count
HH size	Yes	No		
1	3	1	4	30
2	2	6	8	40
3+	2	1	3	30
Total	7	8		
BG Count	60	40		

Step 2: Column adjustment

	Homeownership		Total	BG Count
HH size	Yes	No		
1	$(3*60/7)=25.71$	$(1*40/8)=5.00$	30.71	30
2	$(2*60/7)=17.14$	$(6*40/8)=30.00$	47.14	40
3+	$(2*60/7)=17.14$	$(1*40/8)=5.00$	22.14	30
Total	60.00	40.00		
BG Count	60	40		

Step 3: Row adjustment

	Homeownership		Total	BG Count
HH size	Yes	No		
1	$(25.71*30/30.71)=25.12$	$(5*30/30.71)=4.88$	30.00	30
2	$(17.14*40/47.14)=14.55$	$(30.00*40/47.14)=25.45$	40.00	40
3+	$(17.14*30/22.14)=23.23$	$(5.00*30/22.14)=6.77$	30.00	30
Total	62.89	37.11		
BG Count	60	40		

Step 4: After several iterations

	Homeownership		Total	BG Count
HH size	Yes	No		
1	23.60	6.40	30.00	30
2	15.20	24.80	40.00	40
3+	21.30	8.70	30.00	30
Total	60.00	40.00		
BG Count	60	40		

Fig. 6. Simplified example of the IPF procedure (from [15])

for these small geographic areas that imitate the actual population of households living there.

The method of population synthesis we followed uses weighted bootstrap sampling of the detailed individual household data records at the PUMA level to populate 100 % of the households of the census block groups, but where different sampling weights are used for each block group depending on how their summary count data relate to the PUMA level data. It is the derivation of these weights that is the core aspect of the population synthesis method. The method of population synthesis we followed is actually a sequence of 2 methods: Iterative Proportional Fitting or IPF [13], and then Iterative Proportional Updating or IPU [14].

2.3.1. Iterative Proportional Fitting (IPF)

To illustrate the IPF procedure, we will use a simple example, taken from Ref. [15]; of 100 households for a block group, 15 households for PUMS data, and 2 variables: homeownership (yes/no), household size (1/2/3+). The IPF procedure would then proceed as follows (see Fig. 6):

Step 1. Create the 2-way distribution of household counts for variables from the PUMS data, along with row/column category sums of same, and then include row/column category counts for the block group households.

Step 2. (column adjustment): Each cell is proportionally adjusted so that their column sums equal the column category count of block group households.

Step 3. (row adjustment): Each cell is further proportionally adjusted so that their row sums now equal the row category count of block group households.

Step 4. Steps 2 and 3 are then iteratively repeated until both the column sums and the row sums of the cells equal the respective column and row category counts of block group households (or converge within some specified tolerance).

In this example, and after IPF, each cell represents the number of households with a particular combination of categories from the two variables for that block group. In reality, naturally, households are characterized by many more than two variables, but the above method generalizes to any number of variables, with the adjustments iteratively proceeding variable by variable until all cells in the multiway table sum to the respective row/column variable category counts of the block group households. The multi-way cell counts across all the variables then constitute the synthetic population of households for a block group. Many studies that have used population synthesis have used the IPF procedure exclusively (see Ref. [13]).

However, the IPF procedure has two significant limitations. The IPF procedure can be used on either household count variables or population level count variables separately, but it cannot simultaneously match both the household and population level variable counts [16,17]. Secondly, the IPF procedure suffers from an error when any variable category has a non-zero value in the block group count data but has no representation in the PUMS data – known as the zero-cell problem. For example, block group data may record a small number of mobile home households in the block group, but the 5 % sample of the PUMS data may not include any such households in its PUMA. To overcome these limitations of IPF, Ye et al. [14] proposed the Iterative Proportional Updating (IPU) procedure.

2.3.2. Iterative Proportional Updating (IPU)

Whereas IPF generates synthesized households based on the cell counts in multi-way tables of sets of either household or population level variables, IPU creates synthesized households by direct sampling from the PUMS data itself.

The IPU procedure builds upon two separate IPF procedures, one for household level variables and one for population level variables. In this study we used 7 household level variables with 20 categories across those variables (plus an 8th variable, *Elderly present*, derived post-synthesis based on Age), as well as 5 population level variables with 18 categories across those variables (see Table 1). To

Table 1
Variable definitions

Variable	Definition
Household level variables	
<i>Children present</i>	1 = Having a child, 2 = Not having any child
<i>Mobile home</i>	1 = Lives in a mobile home, 2 = Otherwise
<i>Income</i>	1 = Less than \$10000, 2 = \$10000 to \$29999, 3 = \$30000 to \$59999, 4 = \$60000 to \$99999, 5 = \$100,000 or more
<i>Car ownership</i>	1 = Owned a car, 2 = Otherwise
<i>Marital status</i>	1 = Married family, 2 = Otherwise
<i>Household size</i>	1 = 1 person, 2 = 2 person, 3 = 3 person, 4 = 4 person, 5 = 5 or more
<i>Homeownership</i>	1 = Homeowners, 2 = Not homeowners
<i>Elderly present</i>	1 = Yes, 2 = No
Population level variables	
<i>Age</i>	1 = under 5, 2 = 5 to 17, 3 = 18 to 24, 4 = 25 to 44, 5 = 45 to 64, 6 = 65 above
<i>Female</i>	1 = Male, 2 = Female
<i>Race</i>	1 = Non-Hispanic White, 2 = Non-Hispanic African American,
<i>Ethnicity</i>	1 = Non-Hispanic, 2 = Hispanic,
<i>Education</i>	1 = Below High school, 2 = High school, 3 = College attend and associate degree, 4 = Bachelor degree, 5 = Master or more, 99 = none

Table 2
Simplified example of the IPU procedure (from [14,15])

ID	Ht1	Ht2	Pt1	Pt2	Pt3	Wt0	Wt1	Wt2	Wt4	Wt4	Wt5	Wt100
1	1	0	1	1	1	1	11.67	11.67	9.51	8.05	12.37	1.36
2	1	0	1	0	1	1	11.67	11.67	9.51	9.51	14.61	25.66
3	1	0	2	1	0	1	11.67	11.67	9.51	8.05	8.05	7.98
4	0	1	1	0	2	1	1	13.00	10.59	10.59	16.28	27.79
5	0	1	0	2	1	1	1	13.00	13.00	11.00	16.91	18.45
6	0	1	1	1	0	1	1	13.00	10.59	8.97	8.97	8.64
7	0	1	2	1	2	1	1	13.00	10.59	8.97	13.78	1.47
8	0	1	1	1	0	1	1	13.00	10.59	8.97	8.97	8.64
Sum	3	5	9	7	7		35/	65/	91/	65/	104/	
Constraints	35	65	91	65	104		3	5	111.70	76.80	67.67	
WSum0	3.00	5.00	9.00	7.00	7.00		=	=	=	=	=	
Delta0	0.91	0.92	0.901	0.89	0.93		11.67	13.00	0.81	0.85	1.54	
WSum1	35.00	5.00	51.67	28.33	28.33							
WSum2	35.00	65.00	111.67	88.33	88.33							
WSum3	28.52	55.38	91.00	76.80	74.39							
WSum4	25.60	48.50	80.11	65.00	67.68							
WSum5	35.02	64.90	104.84	85.94	104.00							
Delta5	0.0006	0.0015	0.1520	0.3221	1.37E-16							
W.Sum 100	35	64.99	90.99	64.99	103.99							
Delta100	0	0.002	0.000194	0.000353	0.000353							

Note. ID = Household ID, Ht = Household type, Pt = Population type, Wt = Weight, Constraints = Cell values from IPF, W.Sum = Weighted sum, Delta = |Weighted Sum – Constraints|/Constraints.

undertake IPU, the PUMS data are summarized in a frequency matrix where each row represents a household, and each column represents a particular combination of categories for household type, or a particular combination of categories for population types.

Since we have 7 household level variables with 20 categories, there would be 800 ($=2 \times 2 \times 5 \times 2 \times 2 \times 5 \times 2$) columns, each representing a different household type. Similarly, since we have 5 population level variables with 18 categories, there would be 288 ($=6 \times 2 \times 2 \times 2 \times 6$) total columns, each representing a different population type. A particular PUMS household would be represented by a “1” in the column representing its type, with “0” in every other household column. For the population level columns, the values would indicate how many members of the household matched the different population types. Recall, that an IPF procedure produces cell counts of the number of households, or population numbers, in different combinations of variable categories, and these counts are represented in IPU as “constraints.” To illustrate the IPU procedure we report an example from Ref. [14] for 1 block group, using just 2 household types, 3 population types, and 8 PUMS households (see Table 2). The goal is to produce a weight for each PUMS household (column Wt100) representing its probability of selection when performing a weighted bootstrap sampling of PUMS households to populate a block group. These weights will sum to 100 over all the PUMS households, so dividing each weight by 100 produces each probability.

At initialization, each PUMS household (row) is assigned a weight of 1 (column Wt0), and each value in any household/population type column is multiplied by its corresponding row weight (all equal to 1 initially), and then column-summed to produce the row WSum0. At this point, the values in WSum0 are synonymous with the raw column sums for each household/population type column. The quantity delta represents the deviation between a column’s constraint value (from IPF) and its weighted column sum, and is calculated as:

$$\frac{(\text{constraint} - \text{weighted column sum})}{\text{constraint}}$$

The initial values of delta when all PUMS households have the same initial weight of 1 are shown as the row Delta0 in Table 2.

To update the row weights, the IPU procedure follows a repetitive process column by column. Following the illustration, any non-zero values in the first column (Ht1) have their row weights updated by multiplying the existing row weights (all equal to 1 initially) by the ratio of the current weighted column sum to the column constraint value. In the case of the first column (Ht1) this means the first 3 rows have their weights updated (producing column Wt1) by multiplying their initial weights of 1 by 11.67 (35/3). With these updated row weights in place, new weighted column sums are produced by multiplying each value in any column by its corresponding row

Table 3
Comparison of two survey-based estimates of evacuation rates for urricane Irma versus predicted rates from our model using different base rates

	Mandatory Order	No Mandatory Order
Mason-Dixon	57 %	19 %
Wong	69.50 %	46.40 %
<i>Model</i>		
1 % base rate	48.71 %	9.03 %
3 % base rate	72.26 %	22.46 %
5 % base rate	81.00 %	32.01 %

weight and summing (row WSum1). This same process now repeats column by column, as shown in Table 2. One pass through all the household/population type columns represent an iteration, and an updated value of delta is calculated after each iteration. Table 2 reports values of delta after 1 iteration (as Delta5 since we have 5 columns) and after 20 sets of iterations (Delta100). As the values of delta approach zero we have convergence of the column weighted sums and the constraints. When the values of delta are sufficiently small the IPU procedure terminates (here at 20 iterations), and we have the set of PUMS household weights needed (Wt100) for the bootstrap sampling procedure to take place for 1 block group.

Notwithstanding the simplicity of this illustration, recall that this procedure was actually conducted for 1088 columns, for anywhere from 2579 to 4547 households per PUMA, and for each of 2017 individual block groups. We utilized the software package PopGen [18], an open-source software for population synthesis developed by the Mobility Analytics Research Group at Arizona State University (see Refs. [14,15,19,20]). The block group count data were collected from the American Community Survey (2013–2017) Summary File section of the US Census [21], and the PUMS data were collected from the Public Use Microdata Sample section of the US Census [22].

2.4. Prediction and mapping of evacuation rates

Given a population of ‘synthetic’ households with the socio-demographic predictors used in this study, and with the addition of storm-related variables (evacuation order, flood risk, surge risk) to these households as described above, the prediction of household hurricane evacuation probabilities for our Hurricane Irma-like scenario proceeded using the effects sizes from Ref. [9] in a logit predictive model calibrated with an appropriate base probability (intercept term). The household evacuation probabilities were then summed (by block group) to produce an evacuation rate for each block group, and then mapped. We could also sum the household evacuation probabilities by binary categories of predictor to show differences in evacuation rates by predictor category. Note that for mapping purposes a block group was designated to be in a mandatory evacuation order zone if the majority of its area fell in such a zone.

3. Results

3.1. Model calibration

As outlined above, our logit predictive model required calibration through estimation of its intercept term, or base rate. In a real-world implementation, given a synthetic population of households, a household evacuation probability model, and a predicted overall

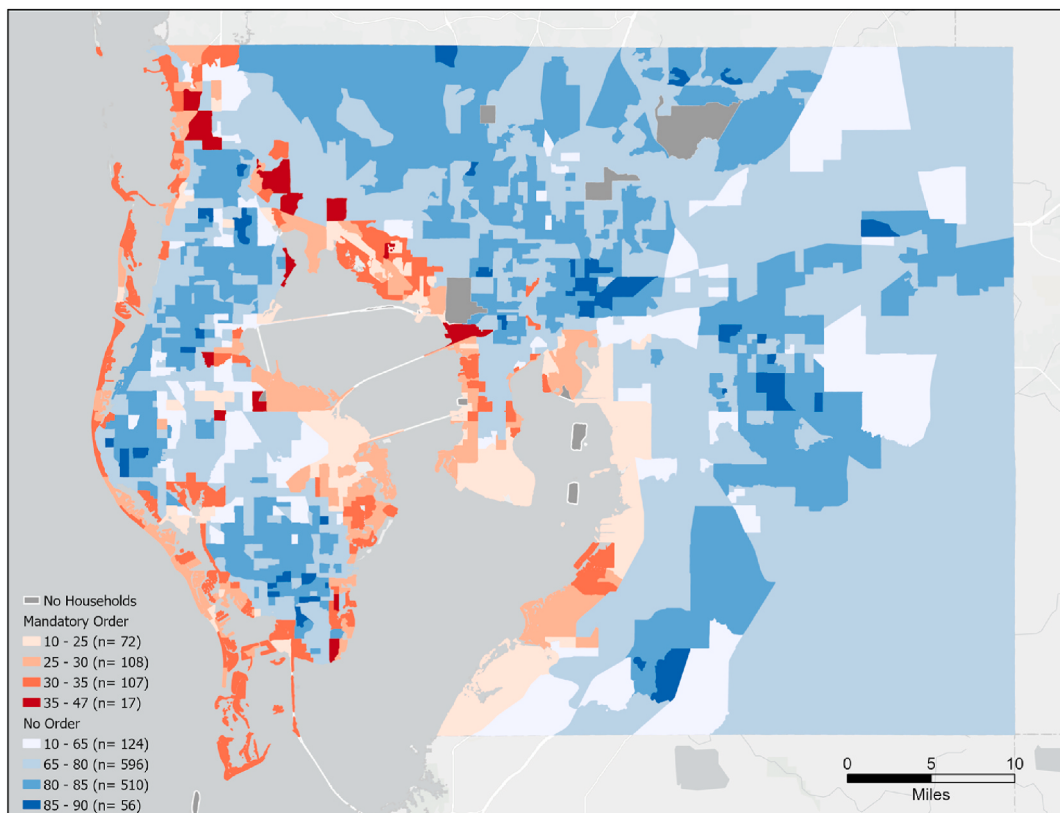


Fig. 7. Percentage of households not-evacuating by block groups for locations both under and not under mandatory evacuation orders: Hillsborough and Pinellas Counties

evacuation rate, estimation of the base probability can proceed in an iterative fashion. Therefore, estimation of the base probability does assume an analyst can estimate a reasonable overall evacuation rate based on their expert knowledge of previous hurricane evacuation scenarios. Again, emphasizing that this paper is meant to be illustrative of the methodological approach, we chose to estimate our base probability in the manner described above by using data from two survey-based estimates conducted in the aftermath of Hurricane Irma. The Mason-Dixon telephone survey [23] had 625 respondents state-wide in Florida. A second online survey by Ref. [24] had 645 responses. Table 3 below reports the evacuation rates in Florida from the two surveys along with our predicted evacuation rates using different base rates of 1 %, 3 %, and 5 %.

In comparing our model predictions to the survey rates in Table 3, we decided to use a base rate of 3 % for the results presented in the remainder of this Results section. The 3 % base rate gives good agreement as regards mandatory evacuation order rates, notably as compared to that of [24]. On the non-mandatory evacuation rate side, our base rate of 3 % gives good agreement with the Mason-Dixon survey [23] but less so with Ref. [24]. However, another hurricane evacuation rate study [4] suggests that the Ref. [24] value may be too high for non-mandatory evacuation locations. Despite the imprecision of our calibration that the foregoing discussion may suggest, any such imprecision is not particularly pertinent to the illustrative nature of this paper. Moreover, even an imprecise estimate of such a base rate in a real-world scenario would not negate the benefits that our approach yields in terms of knowledge of the spatial and socio-demographic variabilities in evacuation.

We attempted to explore existing studies that contained data related to the evacuation patterns during Hurricane Irma. Regrettably, our search yielded no studies providing information on evacuation percentages categorized by order type specific to Hurricane Irma. However, we did come across several noteworthy studies such as [25,26]; and [27]; which offered valuable insights into the overall evacuation dynamics during Hurricane Irma.

In their study [25], utilized Twitter data to determine an overall evacuation rate of 28.4 % for Hillsborough County, a figure strikingly similar to our own finding of 29.09 % for the same region. Similarly [26], employed GPS survey data and reported an overall evacuation rate of 37 % for Florida, which closely aligns with our estimate of 34.72 % for the Tampa Bay Area. In contrast [27], reconstructed evacuation traffic data and found that the Tampa area had a notably higher overall evacuation rate of 52 %, which diverges significantly from our estimations and the data provided by Ref. [24] and the Mason-Dixon survey mentioned earlier. Moreover, it is worth noting that while these studies offer valuable insights into overall evacuation rates, none of them provided data that allows a direct comparison of our results in terms of evacuation order type, finer geographical granularity, or predictors.

3.2. Predicted non-evacuation rates for Hillsborough County and Pinellas County census block groups

The presentation of the modeling results for census block groups (hereafter block groups) switches the focus to *rates of non-evacuation* since, as discussed in our introduction, this is the most important aspect for disaster management in the immediate aftermath of a hurricane. We also confine our discussion for block groups to the two (of four) most populous counties of Hillsborough and Pinellas in the MSA. These results are clearly *after the event* predictions and only applicable to Hurricane Irma, so they really serve

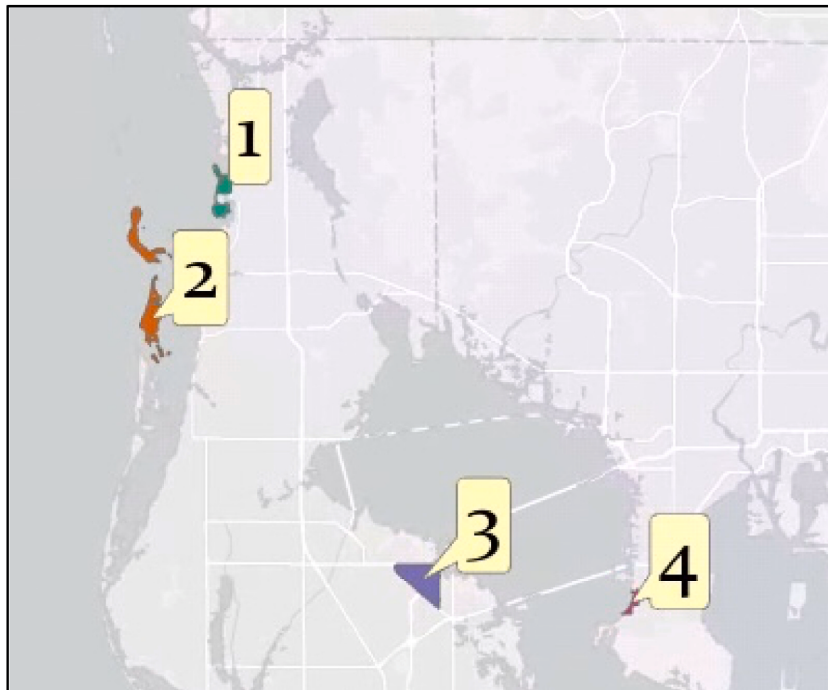


Fig. 8. Locations of the four census block groups summarized in Table 4

Table 4

Rates of non-evacuation and socio-demographic profiles for the four selected block groups shown in Fig. 8.

Block group			1	2	3	4
Non-Evacuation Rate (as %)			34.75	32.02	21.10	18.62
Increases Non-evacuation	Predictors	OR				
	% Black	0.62	0.00	0.00	4.60	10.46
	% Homeowner	0.72	87.60	74.59	16.76	1.11
	% Married	0.77	73.39	47.51	35.73	13.24
	% Older	0.78	39.53	68.51	8.35	1.56
	% Low education	0.79	23.51	25.69	14.43	14.24
Reduces Non-evacuation	% Elderly present	0.88	44.44	70.99	8.93	2.00
	% Hispanic	1.08	0.26	4.14	10.68	10.79
	% Small HH	1.16	75.45	100.00	81.29	86.87
	% Low income	1.17	9.30	37.85	10.74	24.25
	% Children present	1.19	16.02	0.00	17.15	11.23
	% Car ownership	1.24	94.57	95.03	98.77	97.89
	% Female	1.35	30.23	25.97	58.58	49.61

as a demonstration of the variability and granularity of the kinds of predictions that our approach can produce.

Fig. 7 presents a map of our predicted rates of non-evacuation for the block groups of Hillsborough and Pinellas counties. For mandatory evacuation order locations our predictions for non-evacuation range from a low of 10 % to a high of 42.4 %. There is no particular geographic pattern evident for mandatory evacuation order locations, and the block groups with non-evacuation rates above 35 % are located downtown in the city of St. Petersburg, in the west Tampa area around the airport, and then some locations in the hinterland of the coasts, particularly to the north-west of upper Tampa Bay. The block group located under a mandatory evacuation order with the highest rate of non-evacuation (42.4 %) is in south St. Petersburg.

For locations that are not under mandatory evacuation orders our predictions of rates of non-evacuation range from a low of 10.5 % to a high of 88.2 %. There is a more discernible geographic pattern for locations not under mandatory evacuation orders with high rate clusters of non-evacuation evident in the inner-urban, lower socio-economic neighborhoods of Tampa and St. Petersburg. Meanwhile there are regions, such as mid Pinellas County, that have low rates of non-evacuation, and often containing specific block groups that have very low rates of non-evacuation similar to those found in locations under mandatory evacuation order. These regions and specific block groups are typically associated with high levels of mobile home occupancy.

To demonstrate how the socio-demographic hurricane evacuation risk profiles (minus mobile home occupancy) can vary across block groups and impact the predicted rates of non-evacuation at the local block group level, we created a sub-group of block groups with the following characteristics: a) in mandatory evacuation order locations, b) possess highest flood risk, c) possess highest surge risk, and d) have no mobile home occupancy. Within this sub-group we then selected the 2 block groups with the highest predicted rates of non-evacuation (34.75 %, 32.02 %) and the 2 block groups with the lowest predicted rates of non-evacuation (21.1 %, 18.62%). Fig. 8 shows the geographic locations of these 4 block groups and Table 4 summarizes their socio-demographic profiles.

It is noteworthy that the 2 block groups with the highest predicted rates of non-evacuation in Table 4 have 0 % Black households, and yet this is the predictor that most increases non-evacuation at the household decision level. Conversely, the 2 block groups with the lowest rates of non-evacuation in Table 4 have 4.6 % and 10.46 % Black households. More generally, both block groups with the highest predicted rates of non-evacuation have higher levels for all the other factors that increase non-evacuation (homeownership, marriage, older, low education, and presence of the elderly) at the household decision level. For the predictors that might reduce non-evacuation rates, both of the block groups with the highest predicted rates of non-evacuation have somewhat lower levels for 2 of the predictors (Hispanic, female) but the picture is more mixed for the remaining 3 predictors (small households, low income households, presence of children).

Fig. 9 illustrates the distributions of block groups for rates of non-evacuation for binary classifications of households using the 6 socio-demographic predictors that tend to increase the rate of non-evacuation at the household decision level. The binary distributions are most distinct for homeownership and age (younger/older) and fairly distinct for married households and households with an elderly person present. The distributions have more overlap for race (Black/non-Black) and for education.

3.2.1. Predicted non-evacuation rates and race

To investigate the interpretations above further, and to illustrate how our approach can investigate how evacuation rates vary within categories of a predictor, Figs. 10 and 11 present maps of the distribution of our predicted non-evacuation rates by block groups for Black households and non-Black households respectively. In each case, we restrict the block groups used to those with at least 50 households for their category. As expected overall, the rates of non-evacuation for Black households are higher than for non-Black households, as indicated by the greater prominence of the darker hues in Fig. 10 relative to Fig. 11. In fact, in no single block

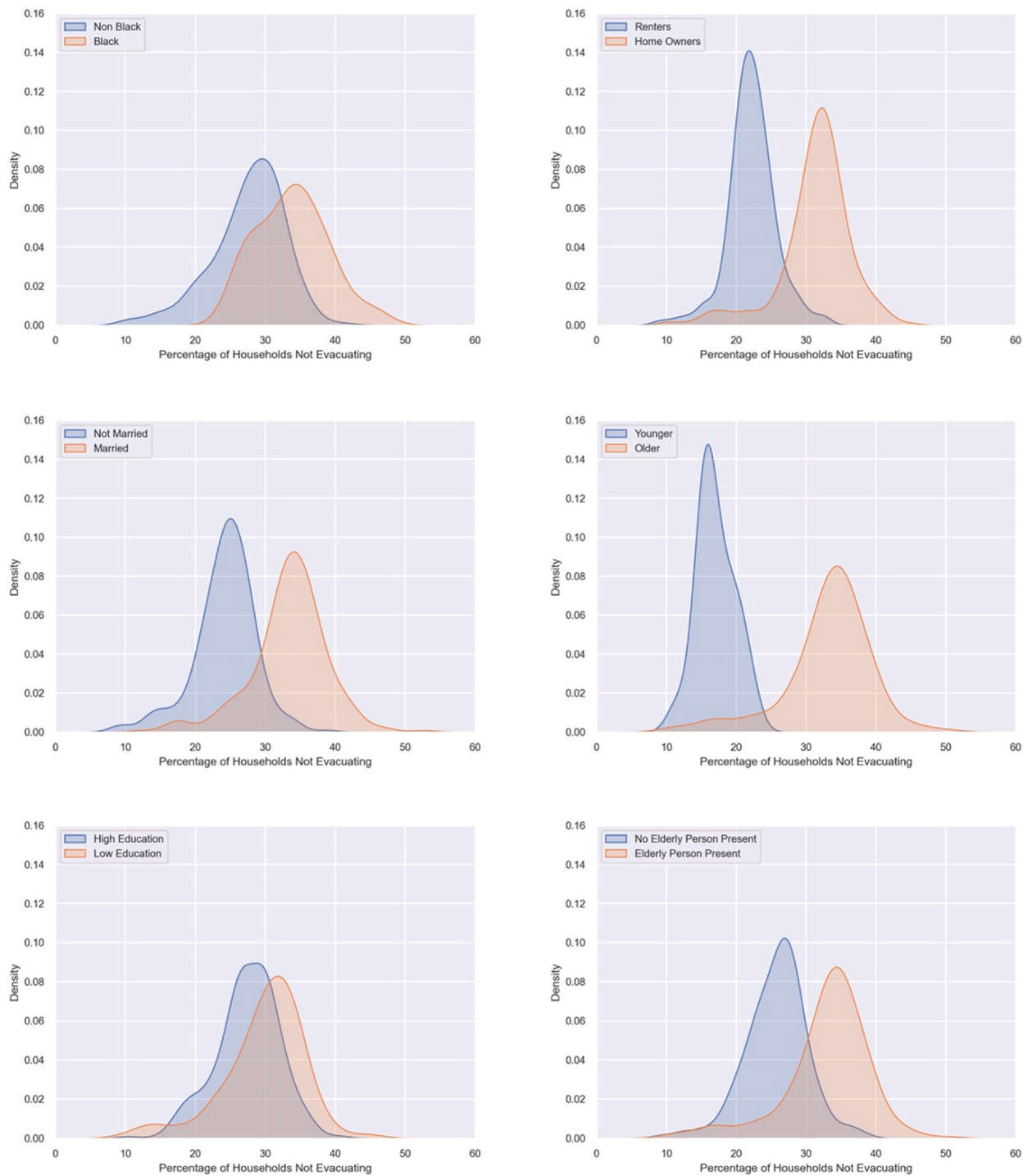


Fig. 9. Distributions of block groups for rates of non-evacuation for binary classifications of households using 6 socio-demographic predictors: Hillsborough and Pinellas Counties

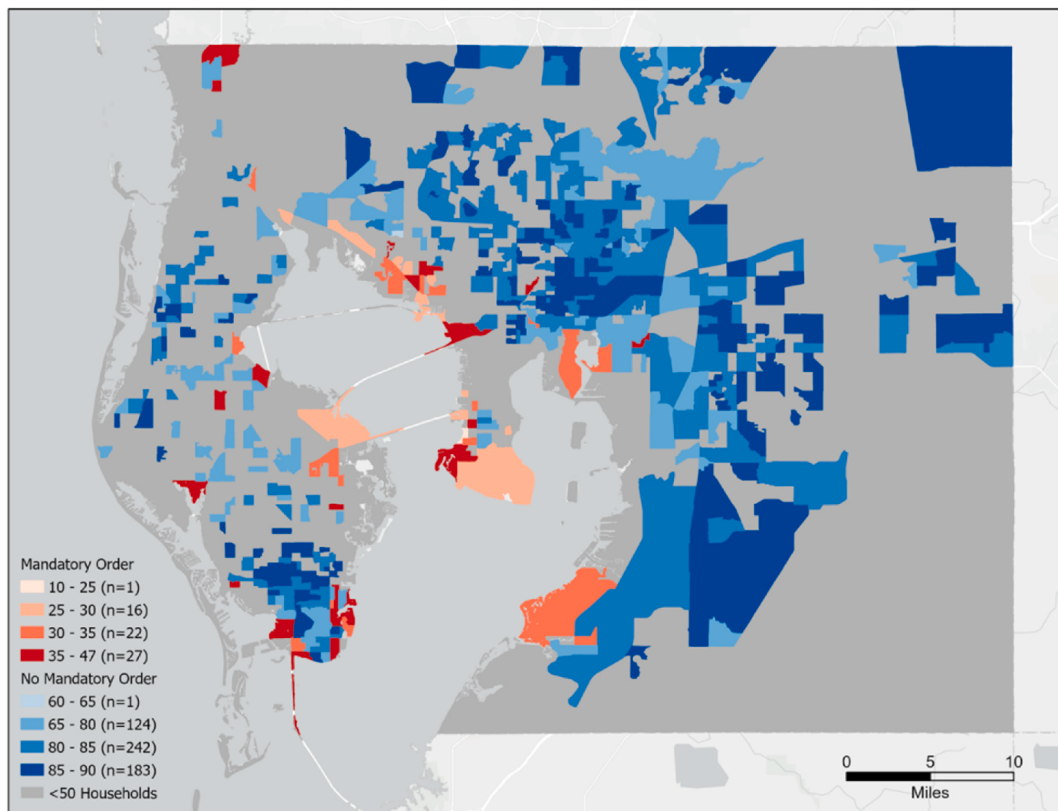


Fig. 10. Percentage of Black households not-evacuating by block groups for locations both under and not under mandatory evacuation orders: Hillsborough and Pinellas Counties

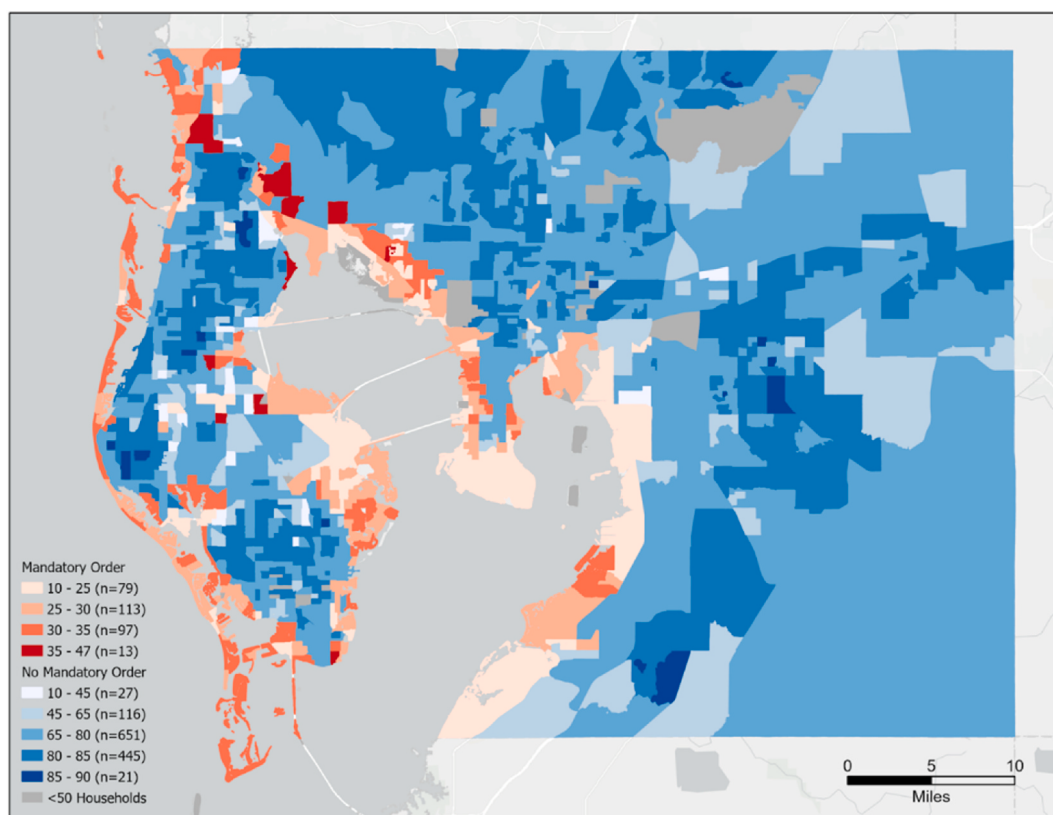


Fig. 11. Percentage of non-Black households not-evacuating by block groups for locations both under and not under mandatory evacuation orders: Hillsborough and Pinellas Counties

group was the rate of non-evacuation for Black households predicted to be lower than that for non-Black households. In terms of spatial patterns of non-evacuation in each case, the map for Black households does seem to indicate some clustering of high rates of non-evacuation (under a mandatory evacuation order or not) in the inner-urban areas of Tampa and St. Petersburg. The two maps also indicate that some specific block groups may exhibit marked differences in non-evacuation rates for their Black and non-Black households. To investigate this further, Fig. 12 presents a map of the difference in non-evacuation rates for Black and non-Black households by block groups.

It is notable that of the 550 block groups mapped in Fig. 12, 16 % (88) have a predicted non-evacuation rate for Black households that is more than 10 % higher than for non-Black households, with 7 % of block groups having more than a 15 % difference. In terms of the spatial distribution of Fig. 12, the greatest differences (>15 %) seem to occur in more suburban locations such as mid-Pinellas County, east of Tampa in the Brandon area, and also in more rural parts of Hillsborough County.

To gain some insight into how the socio-demographic profiles of these block groups are driving the differences in non-evacuation between Black and non-Black households, we selected 2 block groups in locations under mandatory evacuation orders that had similar levels of % Black households but which had very different values when it came to the difference in the non-evacuation rate between Black and non-Black households (see locations A and B in Fig. 13). Similarly, we chose 2 block groups in locations *not* under mandatory evacuation orders and with similar levels of % Black households, but, again, with very different values when it came to the difference in the non-evacuation rate between Black and non-Black households (see locations C and D in Fig. 13).

Table 5 presents the socio-demographic profiles for these 4 block groups. To understand the differences in the Black/non-Black household evacuation rates we need to consider the *relative* percentages of the predictors across Black/non-Black households and whether each predictor is one which tends to increase non-evacuation or reduce non-evacuation. Block group A has a relatively small difference (5.01 %), and this appears to be partly the result of non-Black households having significantly higher rates of homeownership and marriage, both predictors which tend to increase non-evacuation. The small difference is also partly the result of Black households having higher levels of low income, female, and children present households, all predictors which tend to reduce non-

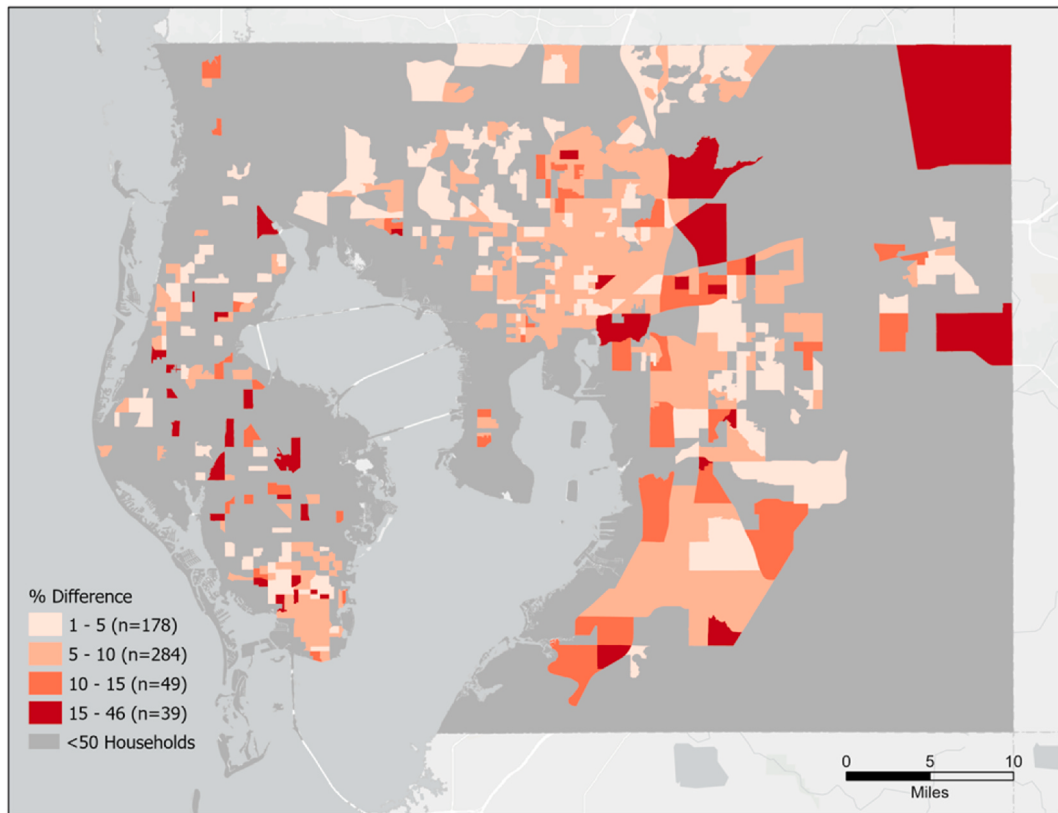


Fig. 12. Difference in the predicted percentages of non-evacuation for Black households and non-Black households by block group: Hillsborough and Pinellas Counties.

evacuation. The only predictors for block group A that would tend to increase the difference would be low education and car ownership.

A somewhat similar story applies to block group C, the other one of our 4 block groups with a small difference between non-evacuation rates for Black/non-Black households (3.31 %). For block group C, non-Black households have significantly higher rates of homeownership, marriage, and, also, older households and households with an elderly person present. These are all predictors that tend to increase rates of non-evacuation. The small difference for block group C also partly results from Black households having higher levels of low income and female households, both predictors which tend to reduce non-evacuation, although the influence of these predictors is probably not as strong as it was for block group A.

Turning to block group B in [Table 5](#), Black and non-Black households have more similar levels of homeownership than we observed for block groups A and C and this tends to increase the difference in non-evacuation rates relative to other block groups where Black homeownership rates are lower. Block group B also has a noticeable difference in % Hispanic with non-Black households having a % Hispanic level that is twice that for Black households. Since % Hispanic reduces non-evacuation, this difference contributes towards the large difference in non-evacuation rates we see for block group B. Block group B also shows a similarity between Black and non-Black in % Low Income, unlike block groups A and C where the rates of % Low Income were much higher for Black households. Since higher levels of % Low Income would reduce non-evacuation this also contributes to the larger difference in non-evacuation rates we see for block group B. Finally, and likely the greatest contributor to the large differential in non-evacuation rates for block group B is the substantial difference in rates of mobile home occupation for Black and non-Black households (60.79 % versus 18.18 %).

Finally, for block group D in [Table 5](#) we have a very large difference in the non-evacuation rates for Black and non-Black households (46.01 %). As with block group B, this is largely driven by the substantial difference in rates of mobile home occupation for Black and non-Black households (63.78 % versus 4.55 %). However, it is also noteworthy that non-Black households also have higher levels of % small households and % low income, both predictors which work to reduce non-evacuation rates.

3.3. Aggregating model results for larger geographic units

Although the main innovation in our approach is to be able to predict evacuation rates for small geographic units, in our case block groups, it is straightforward to then aggregate these small units into larger geographic units as required. For example, [Tables A1](#) and [A2](#) in the appendix provide predictions for non-evacuation at the County level for the four counties of the Tampa-St. Petersburg-Clearwater MSA, along with their predictor risk profiles.

[Fig. 14](#) summarizes how non-evacuation rates vary by different binary classifications of the evacuation predictors for all households

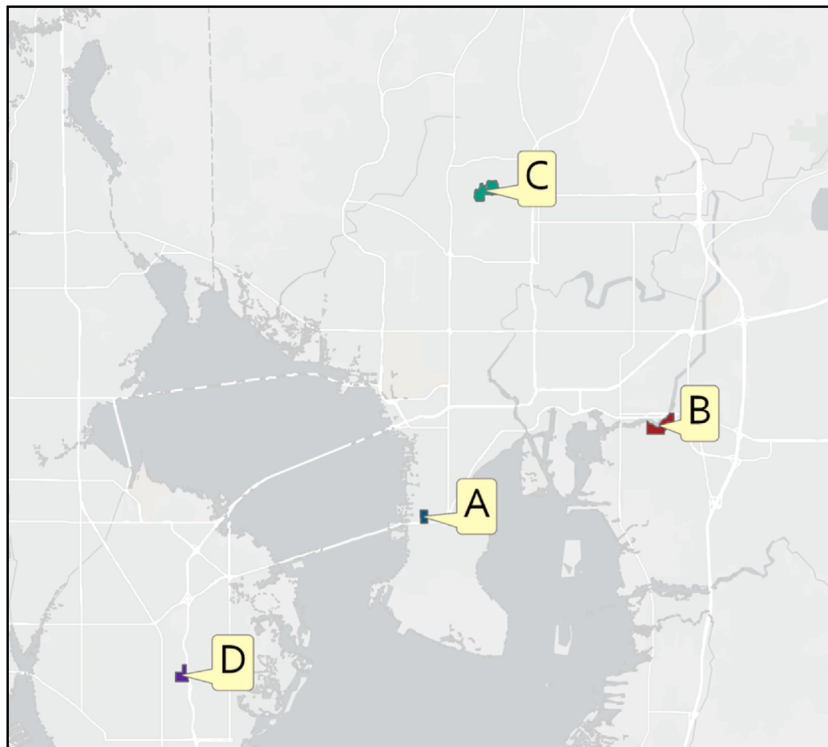


Fig. 13. Locations of the four census block groups summarized in Table 5.

Table 5

Comparative socio-demographic profiles for black and non-black households for the four selected block groups shown in Fig. 13

Block group			A		B		C		D	
% Difference in not evacuating			5.01		19.17		3.31		46.01	
	Predictors	OR	Mandatory Order				No Mandatory Order			
	% Black	0.62	25.16		25.84		20.41		19.64	
Increases			Non Black	Black	Non Black	Black	Non Black	Black	Non Black	Black
Non-evacuation	% Homeowner	0.72	43.98	13.58	38.81	33.52	62.21	37.37	56.89	45.45
	% Married	0.77	28.63	2.47	47.72	34.09	53.44	34.21	23.11	20.91
	% Older	0.78	27.39	27.16	16.24	17.61	28.07	18.95	33.33	26.36
	% Low education	0.79	37.76	56.79	82.18	86.36	29.01	37.37	67.56	57.27
	% Elderly presence	0.88	29.05	27.16	19.60	21.59	32.79	22.63	38.22	27.27
Reduces	% Hispanic	1.08	7.05	7.41	63.76	30.68	26.59	20.53	2.67	0.00
Non-evacuation	% Small HH	1.16	60.17	58.02	65.15	70.45	63.29	66.84	75.33	62.73
	% Low income	1.17	35.27	81.48	63.17	59.09	15.38	35.26	70.89	60.00
	% Children present	1.19	39.83	51.85	25.54	23.86	28.88	30.53	10.44	12.73
	% Car ownership	1.24	92.12	70.37	96.63	91.48	96.36	85.79	75.11	73.64
	% Female	1.35	53.94	70.37	41.58	40.91	47.64	52.63	58.89	77.27
	% Mobile Home	6.47	0.00	0.00	60.79	18.18	0.00	0.00	63.78	4.55

in mandatory evacuation order locations across the whole MSA, along with our predicted mean mandatory non-evacuation rate calculated as the sum of all household non-evacuation probabilities divided by the number of households for our entire study area. The differences in rates for a predictor have *therefore taken into account the influences of all the other predictors (compound socio-demographic profile) across all the households*. The predictors in Fig. 14 are organized into those whose *independent* impact on the household evacuation decision reduces non-evacuation (left side of dashed line) and those whose impact increases non-evacuation (right side of dashed line). They are then ordered from high-to-low for the predictors that reduce non-evacuation and low-to-high for those that increase non-evacuation. This figure again illustrates, now at the MSA level, that although a predictor may have a greater *independent* impact on the household evacuation decision than others, when the overall *composite risk profiles* of households are considered over all households, the net effect may lead to greater differentials for other predictors when summarizing the numbers of households not evacuating. For example, the predictor Black has the largest independent effect on increasing a household's decision to not evacuate, and yet it is not associated with the highest evacuation differential in Fig. 14, where other predictors exhibit greater differentials

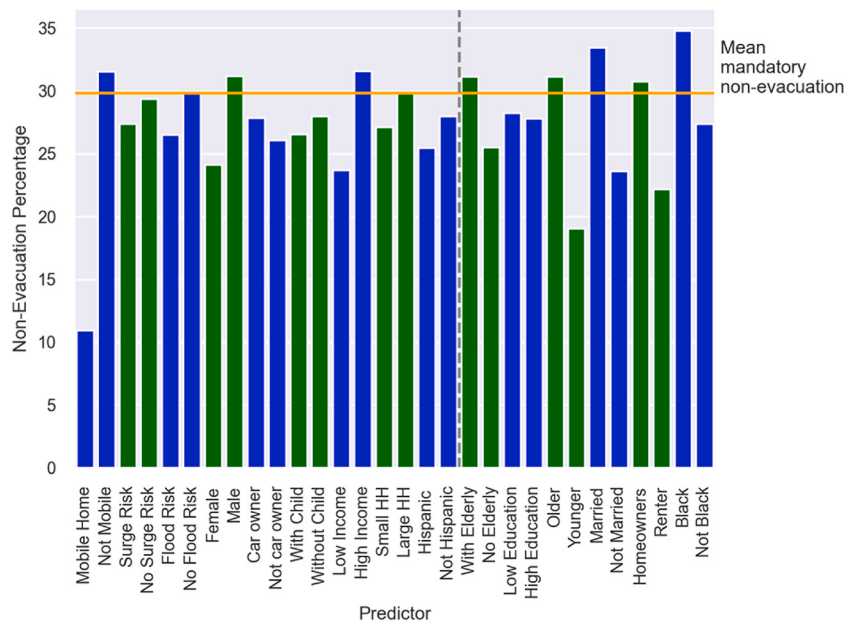


Fig. 14. Non-evacuation rates by different binary classifications of the evacuation predictors for households under a mandatory evacuation order

Table 6

Predictor ranking based on a) independent effect size and b) differential in percentage not evacuating for locations under mandatory evacuation orders: Hillsborough and Pinellas Counties

	Predictors	OR	Rank of OR	% Difference	Rank on model result
Increases Non-evacuation	Black	0.62	1	7.42	4
	Homeowner	0.72	2	8.59	3
	Married	0.77	3	9.85	2
	Older	0.78	4	12.09	1
	Low education	0.79	5	0.47	6
Reduces Non-evacuation	Elderly present	0.88	6	5.66	5
	Hispanic	1.08	9	2.53	6
	Small HH	1.16	8	2.67	5
	Low income	1.17	7	7.85	2
	Children present	1.19	6	1.42	8
	Car ownership	1.24	5	-1.79	9
	Female	1.35	4	7.07	3
	Flood risk	1.44	3	3.28	4
	Surge risk	1.67	2	1.98	7
	Mobile home	6.47	1	20.61	1

(older/younger, homeowner/non-homeowner, married/non-married). A similar figure to Fig. 14 is provided in the appendix for locations not under mandatory evacuation orders (Figure A2).

Table 6 provides greater detail with respect to Fig. 14 and shows how the rank of independent effect of a predictor (based on its model odds-ratio) compares to the rank of the differential in percentages of households not evacuating summarized by each predictor. For predictors that reduce non-evacuation at the household decision level, *mobile home occupancy* clearly has both the greatest independent impact on the household decision to evacuate and the largest differential in the percentages of households evacuating. The next largest differential is associated with % low income even though the independent effect of being *low income* on reducing a household probability of non-evacuation is only ranked 7th (of 9). An interesting predictor is *car ownership* which ranks 5th in terms of independent effect size on reducing a household probability of non-evacuation and yet more car owning households are predicted to remain than non-car owning households. For predictors that increase non-evacuation at the household decision, we have already noted the difference in ranking for the *Black* predictor relative to those of *older*, *homeowner*, and *married*. A similar table to Table 6 is provided in the appendix for locations *not* under mandatory evacuation orders (Table A3).

4. Conclusions

In our opinion, the disaster management and response for hurricanes would benefit from access to spatially localized prediction of evacuation rates by socio-economic-demographic profiles. Such detailed prediction requires an evacuation decision model for

households and a method to generate a synthetic, yet realistic, population of households for local neighborhoods. This paper has addressed these issues by (i) utilizing the results of a statistical meta-analysis of 33 survey-based household evacuation decision models [9], and (ii) utilizing a well proven approach, Iterative Proportional Updating (IPU), to population synthesis. Using our approach, disaster management planners could better plan for, and implement, procedures and policies that could mitigate many of the adverse impacts of hurricanes.

This paper was able to produce detailed predictions of how many, and what types of households, evacuate, or remain, at the disaggregated geographic scale of census block groups for a particular hurricane scenario (Hurricane Irma-like). Our results indicate that it is the composite demographic profiles of the households, acting thru our household evacuation decision model, that largely dictate a very heterogeneous spatial pattern of evacuation. Also, some predictors which may have an outsize *independent* effect on the household evacuation decision model do not always translate to those predictors which may exhibit the largest differentials in rates of non-evacuation. This is due to the composite effect of certain predictors occurring simultaneously in many households. For example, as we have seen above, although the predictor Black has the largest independent effect on increasing non-evacuation, the differentials in non-evacuation rates are higher for the predictors of older, married, and homeownership at the regional scale (Tampa-St. Petersburg-Clearwater MSA). This same phenomenon can also be witnessed in individual block groups, although, as our results indicate, the key aspect is that every block group's risk profile is different, so that in other block groups the non-evacuation differential may indeed be driven by a dominant independent predictor (e.g., Black households). It should also be pointed out that one predictor, household mobile home occupation, appears to be remain dominant throughout, both in terms of independent effect on the household decision model and in evacuation rate differentials.

Since our results demonstrate considerable spatial heterogeneity in evacuation rates at the local neighborhood scale, the key for future disaster management practices will be the ability to apply our framework to particular hurricanes in different locations. A natural outgrowth of this research, therefore, would be to develop a software application that can regularly update the generation of synthetic populations based on latest US Census data, as well as periodically update the statistical meta-analysis results as more hurricane evacuation studies and models are published. Clearly every hurricane is different (track, intensity etc.), as is every locale in terms of unique risks [28] and vulnerabilities [29]. As such, disaster management will always be dependent on the experience and judgement of human managers, but our framework does offer a key component, the base evacuation rate within the logistic regression model, to be able to reflect such judgement calls. It should also be noted that although our approach and framework has been developed in the US context, it could be readily adapted to other national contexts with similar structures as to their national census data, and such other national contexts could draw upon their own set of evacuation studies to determine effect sizes.

Finally, despite extensive research on various aspects of hurricane evacuation, most notably in the social sciences and transportation science, those responsible for planning, and mitigating, the immediate impacts of hurricanes are left wanting in terms of detailed knowledge of how many households, and of what types, are likely evacuating from local neighborhoods. The methods and results of this paper would allow for disaster managers to much better plan and implement strategies and resources that could both enhance evacuation rates, as well as deliver humanitarian and other assistance [29,30] in an efficient and timely manner to those who remain.

4.1. Limitations

A key aspect of any prediction model is validation. This paper has focused on the development of the framework for a predictive model of local neighborhood hurricane evacuation rates but does not present any validation of its results. Although it is certainly possible to imagine a feasible approach to validation, such as post-hurricane evacuation surveys in a range of targeted census block groups, that would require not inconsiderable funding and is left for future work.

In terms of our framework itself, other key limitations must be acknowledged. The household evacuation decision model we used is based on the work of [9] and so the limitations they noted there apply here. Key among them is the dependence on the survey designs and modeling approaches of the numerous studies which contributed to their statistical meta-analysis. Virtually all these studies used logistic regression models for their analysis and only considered linear independent predictors, with no accounting for non-linear relationships or interactions between predictors. A recommendation for future hurricane evacuation survey models would be to investigate, and report, a wider range of modeling approaches, including the use of modern statistical machine learning methods.

More broadly, our approach to generating household level data relies on the use of US Census data and although this allows us to generate data on many of the predictors used in hurricane evacuation models, it cannot be used to generate data on other possibly important predictors. For instance, 7 predictors (length of residence, hurricane experience, mandatory work, evacuation plan, window protection, pet ownership, and disabled person in household) found in Ref. [9] could not be incorporated in the model used here since they are not data collected by the US Census. Just to note, Ref. [9] revealed varied influence of those predictors on evacuation decisions. Longer length of residence, having an evacuation plan, window protections, and disabled household member's presence showed positive relationships with evacuation likelihood. In contrast, mandatory work obligations and pet ownership showed negative effects. The effect of prior hurricane experience was generally negligible, though highly variable. Moreover, some other important predictors such as social networks, information reliability, trust in the source of information, and others were not considered in the Ref. [9] model we used here, partly because the meta-analysis from Ref. [9] established specific criteria for including predictors, requiring that they be present in at least three different models.

As we have indicated, our use of the Ref. [9] model was largely simply illustrative of the use of a household level decision model. We would advocate for emergency management agencies to develop their own household level evacuation decision models that might use different effect sizes and different storm-related variables pertinent to their risk scenario. It is also possible to imagine incorporating estimates of other predictors not found in the US Census into such models by imputing data on other predictors into such

household data (e.g., pet ownership, length of residence) from other data sources. That said, the great advantage of using US Census data in the immediate term is the ability to predict household evacuation rates down to highly localized neighborhoods, and such models, even without the benefit of additional predictors, likely explain a considerable proportion of the spatial variability in such rates. The value of being able to predict the numbers, and types, of households evacuating from local neighborhoods is very considerable, not only in better serving and protecting those households that remain, but also in managing, and servicing, the households who evacuate.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have used data derived from the US census

Appendix

Hernando	All residents west of U.S. 19, as well as mobile homes countywide need to evacuate.
Hillsborough	Hillsborough County, Fla. (Sept. 9, 2017) - Hillsborough County has issued a mandatory evacuation for Evacuation Zone A starting at 8 a.m. today.
Pasco	New Information as of 09/09/17: ☐ A mandatory evacuation has been ordered for residents living west of Little Road, effective 1:00 p.m. today, Saturday, September 9, 2017. ☐ A mandatory evacuation has been ordered for those residents living to the north and west of the following roadway boundary lines: north of 52 and west of Hicks Road; north of Kitten Trail to the intersection of Dennis Rd; west of Dennis Rd, following a northerly line to include properties west of Gopher Hills and west of Thomas Boulevard north to County Line Road, effective 1:00 p.m. today, Saturday, September 9, 2017. ☐ A voluntary evacuation has been ordered for the remainder of all evacuation zones, effective 1:00 p.m. today, Saturday, September 9, 2017. ☐ The hardened shelters of Fivay High School, Schrader Elementary, and the Emergency Operations Center will remain open and are not being evacuated.
Pinellas	Residents in Level B must begin evacuating immediately and secure safe shelter promptly. Residents of Level A and all mobile homes must continue their evacuations and secure safe shelter immediately, as well.

Fig. A.1. Detailed declaration of evacuation orders for Hernando, Hillsborough, Pasco, Pinellas Counties for Hurricane Irma

Table A.1

Non-evacuation rates and evacuation predictor profiles for each of the Tampa-St. Petersburg-Clearwater MSA Counties for locations under a mandatory evacuation order

			Hernando	Pinellas	Hillsborough	Pasco
	Non-Evacuation Rate (as %)		21.53	25.74	26.51	33.53
Increases	Predictors	OR				
Non-evacuation	% Black	0.62	1.46	5.34	8.17	2.55
	% Homeowner	0.72	82.44	66.94	53.87	65.64
	% Married	0.77	46.88	40.93	42.14	42.88
	% Older	0.78	51.25	40.26	19.90	37.56
	% Low education	0.79	57.88	38.03	35.38	59.05
	% Elderly present	0.88	55.79	43.43	23.46	41.10

(continued on next page)

Table A.1 (continued)

			Hernando	Pinellas	Hillsborough	Pasco
Non-Evacuation Rate (as %)			21.53	25.74	26.51	33.53
Reduces Non-evacuation	% Hispanic	1.08	4.34	5.41	23.05	9.06
	% Small households	1.16	82.02	81.70	66.81	72.17
	% Low income	1.17	42.60	31.57	24.22	40.68
	% Children present	1.19	10.86	13.55	24.37	19.84
	% Car ownership	1.24	95.32	92.39	94.50	91.15
	% Female	1.35	48.23	48.91	47.90	50.41
	% Flood risk	1.44	32.36	69.49	81.67	42.47
	% Surge risk	1.67	30.98	86.25	99.33	72.08
	% Mobile homes	6.47	67.25	22.60	5.82	10.87

Table A.2

Non-evacuation rates and evacuation predictor profiles for each of the Tampa-St. Petersburg-Clearwater MSA Counties for locations not under a mandatory evacuation order

			Pasco	Hillsborough	Pinellas	Hernando
Non-Evacuation Rate (as %)			74.6	76.81	78.81	83.25
Increases Non-evacuation	<i>Predictors</i>	<i>OR</i>				
	% Black	0.62	4.98	16.87	10.30	4.79
	% Homeowner	0.72	77.76	58.42	63.35	76.68
	% Married	0.77	55.24	44.12	40.27	51.51
	% Older	0.78	33.23	20.92	29.82	37.48
	% Low education	0.79	46.03	43.65	44.57	55.46
	% Elderly present	0.88	37.48	24.71	33.18	42.49
	% Hispanic	1.08	10.53	22.82	7.62	9.66
	% Small households	1.16	63.83	61.51	71.17	66.22
Reduces Non-evacuation	% Low income	1.17	25.54	29.25	31.65	32.21
	% Children present	1.19	27.16	28.76	20.61	23.72
	% Car ownership	1.24	96.12	92.71	90.85	95.61
	% Female	1.35	47.45	49.73	51.19	49.75
	% Flood risk	1.44	34.16	18.60	14.75	9.93
	% Surge risk	1.67	0.62	16.08	31.59	2.07
	% Mobile homes	6.47	19.36	6.56	0.00	0.00

0 for Mobile homes in Pinellas/Hernando because all mobile homes under mandatory order

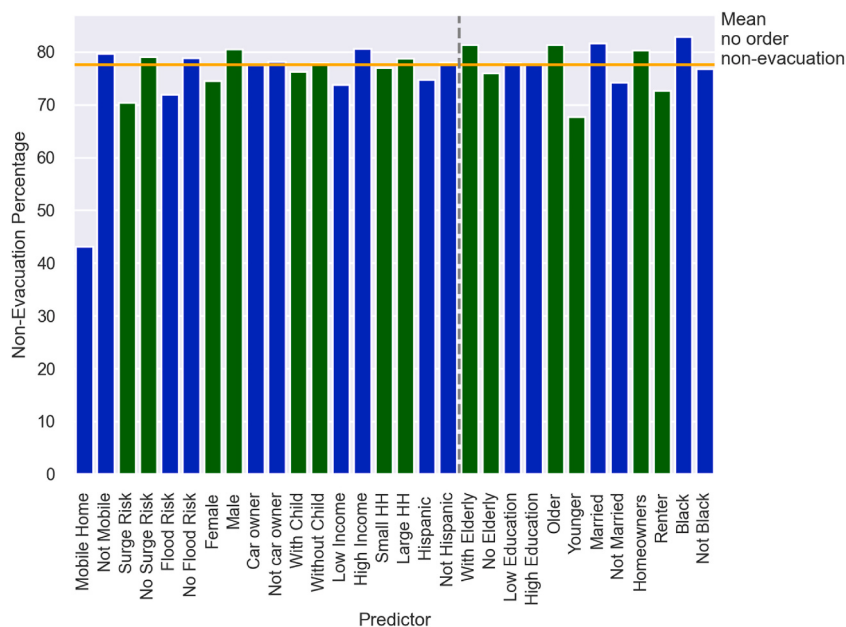


Fig. A.2. Non-evacuation rates by different binary classifications of the evacuation predictors for households not under a mandatory evacuation order

Table A.3

Predictor ranking based on a) independent effect size and b) differential in percentage not evacuating for locations not under a mandatory evacuation orders: Hillsborough and Pinellas Counties

	Predictors	OR	Rank of OR	% Difference	Rank on model result
Increases Non-evacuation	Black	0.62	1	6.09	4
	Homeowner	0.72	2	7.69	2
	Married	0.77	3	7.48	3
	Older	0.78	4	13.62	1
	Low education	0.79	5	0.37	6
	Elderly present	0.88	6	5.38	5
Reduces Non-evacuation	Hispanic	1.08	9	3.35	6
	Small HH	1.16	8	1.76	7
	Low income	1.17	7	6.83	4
	Children present	1.19	6	1.68	8
	Car ownership	1.24	5	0.60	9
	Female	1.35	4	5.97	5
	Flood risk	1.44	3	6.93	3
	Surge risk	1.67	2	8.71	2
	Mobile home	6.47	1	36.50	1

References

- [1] ABC News, Hurricane Ian Death Toll Now More than 100 after Storm Slams into Florida, ABC News, North Carolina, 2022. <https://abcnews.go.com/US/multiple-deaths-reported-hurricane-ian-slams-florida/story?id=90693636>.
- [2] New York Times, Many of Hurricane Ian's Victims Were Older Adults Who Drowned, New York Times, 2022. <https://www.nytimes.com/2022/10/07/us/hurricane-ian-victims-drowned.html>.
- [3] J. Brunkard, G. Namulanda, R. Ratard, Hurricane Katrina deaths, Louisiana, 2005, Disaster Med. Public Health Prep. 2 (4) (2008) 215–223, <https://doi.org/10.1097/DMP.0b013e31818aaf55>.
- [4] Florida Division of Emergency Management, Regional Evacuation Studies, Florida Division of Emergency Management, 2021. <https://www.floridadisaster.org/dem/preparedness/regional-evacuation-studies/>.
- [5] K. Xu, R.A. Davidson, L.K. Nozick, T. Wachtendorf, S.E. DeYoung, Hurricane evacuation demand models with a focus on use for prediction in future events, Transport. Res. Pol. Pract. 87 (2016) 90–101, <https://doi.org/10.1016/j.tra.2016.02.012>.
- [6] A.S. Goodie, A.R. Sankar, P. Doshi, Experience, risk, warnings, and demographics: predictors of evacuation decisions in Hurricanes Harvey and Irma, Int. J. Disaster Risk Reduc. 41 (2019), 101320, <https://doi.org/10.1016/j.ijdrr.2019.101320>.
- [7] U.S. Census Bureau, The 2020 Census and Confidentiality, The Census Bureau, 2020. <https://www.census.gov/content/dam/Census/library/factsheets/2019/comm/2020-confidentiality-factsheet.pdf>.
- [8] E. Adjei, B.C. Benedict, P. Murray-Tuite, S. Lee, S. Ukkusuri, Y. "Gurt" Ge, Effects of risk perception and perceived certainty on evacuate/stay decisions, Int. J. Disaster Risk Reduc. 80 (2022), 103247, <https://doi.org/10.1016/j.ijdrr.2022.103247>.
- [9] S.H. Tanim, B.M. Wiernik, S. Reader, Y. Hu, Predictors of hurricane evacuation decisions: a meta-analysis, J. Environ. Psychol. 79 (2022), 101742.
- [10] FDOT, Hurricane Irma's Effect on Florida's Fuel Distribution System and Recommended Improvements, Florida Department of Transportation, 2018. <https://www.fdot.gov/docs/default-source/info/CO/news/newsreleases/020118-FDOT-Fuel-Report.pdf>.
- [11] NHS, IRMA Graphics Archive: 5-day forecast track, initial wind field and watch/warning graphic. https://www.nhc.noaa.gov/archive/2017/IRMA_graphics.php?product=5day_cone_with_line_and_wind.
- [12] Florida Division of Emergency Management, Florida disaster data, Maps.Floridadisaster (2018).
- [13] R.J. Beckman, K.A. Baggerly, M.D. McKay, Creating synthetic baseline populations, Transport. Res. Pol. Pract. 30 (6) (1996) 415–429.
- [14] X. Ye, K. Konduri, R.M. Pendyala, B. Sana, P. Waddell, A methodology to match distributions of both household and person attributes in the generation of synthetic populations, in: 88th Annual Meeting of the Transportation Research Board, 2009. Washington, DC.
- [15] R. Pendyala, Synthetic population generation for travel demand forecasting, in: *Activity-Based Modeling of Spatial and Temporal Patterns of Human Travel Behavior* [Short Course], Tongji University, Shanghai, China, 2016. https://www.mobilityanalytics.org/uploads/5/0/5/4/5054275/syntheticpopulationgeneration_popgen.pdf.
- [16] P.P. Lim, D. Gargett, Population synthesis for travel demand forecasting, in: Proceedings of the 36th Australasian Transport Research Forum (ATRF), 2013, October, pp. 2–4. Brisbane, Australia.
- [17] D.R. Pritchard, E.J. Miller, Advances in population synthesis: Fitting many attributes per agent and fitting to household and person margins simultaneously, Transportation 39 (3) (2012) 685–704, <https://doi.org/10.1007/s11116-011-9367-4>.
- [18] MARG, PopGen: Synthetic Population Generator [online], Mobility Analytics Research Group, 2016. Available at: <http://www.mobilityanalytics.org/popgen.html>. (Accessed 29 December 2022).
- [19] K.C. Konduri, D. You, V.M. Garikapati, R.M. Pendyala, Enhanced synthetic population generator that accommodates control variables at multiple geographic resolutions, Transport. Res. Rec.: J. Transport. Res. Board 2563 (1) (2016) 40–50, <https://doi.org/10.3141/2563-08>.
- [20] K.C. Konduri, R.M. Pendyala, Updates to the BMC Population Synthesis Model: Incorporating Controls at Multiple Geographic Resolutions, Final Report, 2015.
- [21] U.S. Census Bureau, American Community Survey Summary File Data, The Census Bureau, 2018. <https://www.census.gov/programs-surveys/acs/data/summary-file.html>.
- [22] U.S. Census Bureau, PUMS Data, The Census Bureau, 2018. <https://www.census.gov/programs-surveys/acs/data/pums.html>.
- [23] Mason-Dixon, Scott Receives High Marks for Irma, but Floridians May Hesitate to Evacuate Next Time, 2017. https://media.news4jax.com/document_dev/2017/10/26/Mason-Dixon%20Hurricane%20poll_1509043928726_10861977_ver1.0.pdf.
- [24] S. Wong, S. Shaheen, J. Walker, Understanding Evacuee Behavior: A Case Study of Hurricane Irma, Institute of Transportation Studies, Berkeley, 2018. <https://escholarship.org/uc/item/9370z127>.
- [25] Y. Martín, S.L. Cutter, Z. Li, Bridging twitter and survey data for evacuation assessment of hurricane matthew and hurricane Irma, Nat. Hazards Rev. 21 (2) (2020), [https://doi.org/10.1061/\(ASCE\)NH.1527-6996.0000354](https://doi.org/10.1061/(ASCE)NH.1527-6996.0000354).
- [26] E.F. Long, M.K. Chen, R. Rohla, Political storms: emergent partisan skepticism of hurricane risks, Sci. Adv. 6 (37) (2020) 7906–7917, <https://doi.org/10.1126/SCIADV.ABB7906>.
- [27] K. Feng, N. Lin, Reconstructing and analyzing the traffic flow during evacuation in Hurricane Irma (2017), Transport. Res. Transport Environ. 94 (2021), 102788, <https://doi.org/10.1016/J.TRD.2021.102788>.

- [28] M.L. Hernández, M.L. Carreño, L. Castillo, Methodologies and tools of risk management: hurricane risk index (HRI), *Int. J. Disaster Risk Reduc.* 31 (2018) 926–937, <https://doi.org/10.1016/j.ijdrr.2018.08.006>.
- [29] A.L. Griego, A.B. Flores, T.W. Collins, S.E. Grineski, Social vulnerability, disaster assistance, and recovery: a population-based study of Hurricane Harvey in Greater Houston, Texas, *Int. J. Disaster Risk Reduc.* 51 (2020), 101766, <https://doi.org/10.1016/j.ijdrr.2020.101766>.
- [30] J. Lee, The economic aftermath of Hurricanes Harvey and Irma: the role of federal aid, *Int. J. Disaster Risk Reduc.* 61 (2021), 102301, <https://doi.org/10.1016/j.ijdrr.2021.102301>.