Evacuate or Not? A POMDP Model of the Decision Making of Individuals in Hurricane Evacuation Zones

Adithya Raam Sankar
Institute for Artificial Intelligence
University of Georgia
Athens, GA 30602

Prashant Doshi
Department of Computer Science
University of Georgia
Athens, GA 30602
pdoshi@cs.uga.edu

Adam Goodie
Department of Psychology
University of Georgia
Athens, GA 30602

Abstract

Recent hurricanes in the Atlantic region of the southern United States triggered a series of evacuation orders in the coastal cities of Florida, Georgia, and Texas. While some of these urged voluntary evacuations, most were mandatory orders. Despite governments asking people to vacate their homes for their own safety, many do not. We aim to understand the observable and hidden variables involved in the decision-making process and model these in a partially observable Markov decision process, which predicts whether a person will evacuate or not given his or her current situation. We consider the features of the particular hurricane, the dynamic situation that the individual is experiencing, and demographic factors that influence the decision making of individuals. The process model is represented as a dynamic influence diagram and evaluated on data collected via a comprehensive survey of hurricane-impacted individuals.

1 INTRODUCTION

The category 4 hurricane is approaching. Should I follow the official orders and evacuate, or stay in place? Millions of individuals situated in vulnerable areas face this question as imminent disaster threatens. Many choose to leave while several other individuals do not. Those individuals (single persons or heads of households), who choose to stay put have ostensibly made a sub-rational decision. Yet, numerous interviews with such persons clearly convey their conviction in having made the right choice, although some do regret their choice. How can we correctly predict these decisions of grave consequences made by such individuals?

With the 5 costliest hurricanes in U.S. history occurring during the past 13 years, 3 of which were in 2017,[1] there is a sense of urgency to better understand such evacuation decision making. In August and September 2017, two of these five storms - Harvey and Irma - made U.S. landfall within two weeks of each other. Hurricanes have increased in frequency over recent years just as coastal populations have spiked [16, 7]; therefore, evacuating successfully when needed is a critical component to reducing the personal risk associated with hurricanes. Both Harvey and Irma prompted widespread mandatory evacuation orders.

Affected individuals must reason with the uncertainty associated with several decision variables. For example, a hurricane's path cannot be forecast exactly, and even a slight shift in the anticipated path has a deep impact on which geographical areas are affected. Furthermore, the extent of flooding due to the rainfall that accompanies hurricanes is often uncertain and hard to predict. Thus, probabilistic frameworks capable of modeling and reasoning with uncertainty strongly present themselves.

We administered a Qualtrics survey in December 2017, which was completed by 822 unique respondents residing in the evacuation zones of Hurricanes Harvey and Irma. Among these, 330 self-reported as having evacuated while the remaining 492 respondents self-reported as not having evacuated. The survey gathered data on the demographic and experiential variables among others that relate to the evacuation decision making of the respondents. These collected data and other known results from the hurricane modeling literature served, in part, to inform a new computational *process* model of the decision making of individuals in impending disaster areas. We identify the variables that significantly influence the decision making, understand how the variables po-

[1] NOAA Office for Coastal Management, `https://www.coast.noaa.gov/states/fast-facts/hurricane-costs.html`

tentially interact, and present a factored partially observable Markov decision process (POMDP) [8] that predicts an individual's decisions.

We represent the POMDP in the language of dynamic influence diagrams (DID) [4, 17] that naturally support a factored representation in both formulating the problem and solving it, and are easy to comprehend. We evaluated this process model using a 3-fold cross-validation on the collected survey data. The training folds were utilized to learn some of the parameters of the DID while the evacuation decision data in the test folds allowed measuring the model's accuracy. Our model demonstrated an average prediction accuracy of 71.77% with a low standard deviation. A sensitivity analysis reveals that the model responds intuitively to hypotheses about how various factors affect the decisions. A discussion of the cases where the model mispredicts identifies avenues for some improvement, albeit these are challenging to implement. Decision-making models that are reasonably accurate for the classes of persons studied here will have a disproportionate and positive impact on human lives saved and on planning for disaster aid.

2 SURVEY DATA

Hurricane Harvey made landfall 25 August 2017, prompting evacuation orders to be issued on that day and following days. Hurricane Irma first made landfall in the mainland US on 10th September. We administered a survey over the course of one week in December 2017 to participants located in the evacuation zones of these two hurricanes to collect data for understanding and modeling their evacuation decision making.

2.1 Description

Participants were recruited from regions of Texas, Florida, and Georgia where authorities had issued mandatory evacuation orders during Hurricanes Harvey and Irma. In Georgia, evacuation regions were determined by ZIP code and residents in 27 zip codes, comprising the entirety of two coastal counties and coastal portions of four additional counties, were ordered to evacuate. In Florida and Texas, evacuation regions were determined by county. While 23 counties in Florida were asked to evacuate, Texas ordered residents of 12 counties to evacuate on the day of the landfall and following days. These regions are depicted in Fig. 1. All participants included in the analysis reported residing in one of the mandatory evacuation zones at the time that evacuation orders were issued.

Participants were recruited to an online survey through a Qualtrics panel. While 825 participants responded to the survey, 822 participants were included in the final analysis. Three were removed for duplicate participation, identified by IP address. To maximize the validity of this time-sensitive sample, we sought to collect the largest possible sample within a one-week window.

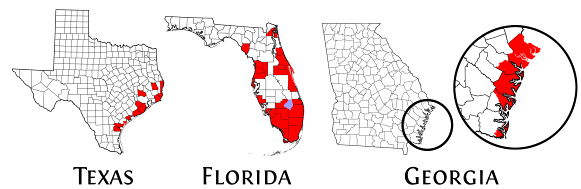

Figure 1: Maps of the target states with county boundaries showing the surveyed regions in red

The survey consisted of 39 questions addressing demographic variables, participants' previous experiences with hurricanes or other similar traumatic events, performance on tests of cognitive bias, and whether the participants evacuated. Twenty questions from the larger survey provided the data for our modeling. The decision to evacuate was self-reported by respondents with an optional reason for their choice. Out of the 822 unique participants, 330 self-reported evacuating during the hurricane, while 492 reported as not having evacuated.

2.2 Statistical Analyses

Relationships were explored among evacuation, demographic variables, and experiential variables. We summarize the statistical findings relevant to the modeling in this subsection. [2] These were computed using ANOVA for nominal predictor variables, and bivariate Pearson correlations were computed between evacuation status and the demographic and experiential variables.

Among the demographic variables, only age and number of children in the home were statistically associated with evacuation, with younger adults and those with children more likely to evacuate. Other demographic variables such as sex, race, relationship status, education level, number of elder dependents in the home, the level on which one lived (1st floor, 2nd floor, etc.), and whether the participant owned or rented his or her home, were all unrelated to evacuation.

Among the experiential variables, we found that the number of times people were asked to evacuate, the perception of neighbors having evacuated, and a personal assessment of the level of risk they experienced, were significant factors in predicting the decision to evacuate. While an experience of prior trauma was a significant

[2] Full details are available in a separate article that is under review. We expect to disseminate the anonymized survey data.

predictor, surprisingly, prior experience with hurricanes did not correlate with the decision to evacuate.

Logistic regression was conducted to predict self-reported evacuation (a binary, yes/no variable), from those variables demonstrating statistically significant correlations with evacuation. The logistic regression was conducted in two blocks: first demographic variables and then experience based variables. The logistic regression predicted evacuation from the significant demographic predictors (age and number of children) in Block 1, and from the significant experience based predictors (times asked to evacuate, assessed risk, preparation level, prior trauma) in Block 2. Note that 301 participants (37.2%) did not respond to at least one item in the regression, and were excluded. The evacuation choices of about three quarters of the remaining participants were predicted correctly using the full model with both blocks of variables.

The most common reasons provided for not evacuating were that respondents did not receive an order to evacuate or did not believe (falsely) that they were in the evacuation zone (118 out of 492); risk had been mitigated through other steps that had been taken (104 of 492); did not believe they were at high risk in general (70 of 492); or did not believe risk was significant in their particular location (65 of 492). No other category of reasons was provided by more than 35 respondents. Among those who did evacuate, top reasons for their decisions included warnings from government and media (70 out of 330); concern that the storm is severe (56 of 330); general references to safety or comfort (51 of 330); and concern for the well-being of family members or other associates (42 of 330). These observations are consistent with the results from our statistical analysis.

3 EMPIRICALLY-INFORMED POMDP

While logistic regression offers good predictive performance, its performance is sensitive to missing predictors. More importantly, we seek a process-oriented and principled computational model with predictions that are consistent with the observed data of the previous subsection. These models differ from statistical curve-fitting (such as using generalized linear models) by providing insights into the decision-making process that possibly led to the observed data, and are better suited for responding to queries that explain its inference and predictions.

Real-world decision making generally involves reasoning with uncertainty, and decision making for evacuation is particularly fraught with it. Several environmental and perception variables that influence the decision are uncertain and evolve in nondeterministic ways. Consequently, the situation "on the ground" is dynamic and may not be perfectly perceived. This motivates modeling an individual's decision making using the well-known POMDP framework. As several distinct variables constitute the decision-making situation, a factored representation of the state is obviously needed. Dynamic Bayesian networks [4, 11] offer a general and popular factored representation for the state and its evolution across time. In contrast to an enumerated representation, these networks allow us to naturally exploit conditional independences between the variables thereby promoting efficiencies. Furthermore, we may combine them with decision and utility nodes, and the resulting dynamic influence diagram (DID) [17] offers a language that is sufficiently general to represent factored POMDPs. A side benefit of using probabilistic graphical models such as DIDs is that their predictions can be robust to missing data – several of our participants did not respond to all questions in the survey.

3.1 State Variables and Priors

Factors that make up the state for the hurricane domain reflect information about the hurricane along with the common effects and precautionary measures that follow. These are modeled using random variables whose values may change over time. We mention the corresponding prior distributions at the first time-step as well.

- *Hurricane level*: This variable represents the strength of the hurricane. Hurricanes are categorized using the standard Saffir-Simpson hurricane wind scale [15]. This five-point scale is determined by the peak one-minute sustained wind at a height of 10 meters over unobstructed exposure. The wind speeds for each category level are shown in Table 1.

Category	**Wind Speed**	**Level of Damage**
1	119-153 km/h	Some
2	154-177 km/h	Extensive
3	178-208 km/h	Devastating
4	209-251 km/h	Catastrophic
5	>252 km/h	Catastrophic

Table 1: Saffir-Simpson hurricane wind scale.

  Values of this random variable are the five categories and a value of 0 that denotes a weakening into a tropical storm or depression. The prior distribution for this variable is the distribution of hurricanes that occurred in 2017 [12].

- *Hurricane path*: The direction a hurricane is headed is usually predicted with sufficient accuracy up to 5 days in advance. An individual located in an impending disaster zone is primarily concerned about whether the hurricane is heading toward or away from her. Therefore, this variable takes one of these two values with a prior distribution of $\langle .33, .66 \rangle$ obtained from the probable track of centers of hurricanes historically (NHC's forecast cone).

- *Evacuation order?*: The presence of an official evacuation order is an important factor in determining the expected risk. The distribution for this variable assigns a probability of .42 to the true value of this Boolean variable. This variable is influenced by the hurricane level and we obtain the probability as a marginal of the prior over various hurricane levels. As we seek to model the decision making of individuals under evacuation orders, this variable is set to true.
- *Rain*: Hurricanes are often accompanied by rain. The amount of rainfall, however, is not influenced by the hurricane's intensity. This random variable can assume a value from this set, {Light, Medium, Heavy}, and the prior over this set is $\langle .5, .4, .1 \rangle$.
- *Preparation to stay*: A factor in not evacuating despite an order is the precautionary measures taken by the individual. This includes stocking up on necessary supplies, boarding up windows and doors, and safeguarding vehicles or other property. This random variable assumes a value {None, Somewhat prepared, Very prepared}, and the prior is $\langle .11, .35, .54 \rangle$ obtained from all the responses in the survey.
- *Neighbors evacuated?*: A key variable in the decision-making situation is whether neighbors facing a similar situation have evacuated or not. Neighbors' actions either add further or mitigate the threat perception of the impending hurricane. This is a Boolean variable whose prior $\langle .4, .6 \rangle$ is obtained by averaging the percentage of other people in the same zip code as the responder who evacuated.
- *Traffic*: The most common mode of transportation for evacuees is by road due to its accessibility and low cost. This may subsequently congest the major highways and block the smaller roads as well. Indeed, multiple individuals cited the worsening traffic situation as a reason for not evacuating in our survey. This random variable can assume a value from {Heavy, Manageable}, and the prior over this set is $\langle .46, .54 \rangle$ inferred from the prior over evacuation order.

In addition to these experiential variables, demographic variables such as age, health condition, prior trauma, number of dependents, availability of transportation, employment status, and the safety of the house also compose the state of the decision-making problem. Age takes one of {$\leq$20, 21-40, 41-60, $\geq$61} values and its prior is $\langle .05, .35, .28, .32 \rangle$ based on the distribution in the survey. Health condition takes one of {Good, Need Support} values and its prior is uniform. Prior trauma is a binary variable and is true with probability .35 based on the survey responses. The dependents variable includes younger and older dependents as well as any pets. It takes a true value with a probability of .26 obtained from our survey data. The availability of transportation is a binary variable which is true 90% of the time. Employment is taken as true with 60% probability based on the 2017 US population statistics. The safety of the house is a binary variable with a uniform distribution.

Notice that the values of these demographic variables do not change with time, and are considered time invariant.

3.2 Individual's Observation Variables

An individual experiences the situation as a hurricane approaches through her observations. Perception is noisy and offers a hazy lens through which the individual learns about the state. As such, the observations are factored as well and are influenced by the corresponding state variables in the same time-slice.

- *Flooding? and Power failure?*: Commonly observed effects of a hurricane are flooding due to excessive rain and power failures due to high wind. We model both these events as Boolean variables. While high winds are a given, an observation of flooding primarily depends on the state variable Rain.
- *Times asked to evacuate*: An order of evacuation is often followed up by repeated announcements on television, radio, emergency alert systems, and door-to-door canvassing. As such, individuals in our survey reported receiving multiple reminders to evacuate. Multiple alerts were seen as indicators of a greater hurricane threat. This observation variable depends on whether an evacuation order is issued and takes one of {0, 1-2, 3-4, $\geq$5} values.
- *Rain and Wind*: These two observation variables capture an individual's perception of whether there is heavy rain currently and high winds are blowing, respectively. Of course, the latter is influenced by the hurricane level.
- *Traffic perception*: Despite traffic not being heavy, individuals may perceive it to be dense based on their prior experiences. This false assumption eventually affects their decision. This observation variable is influenced by the actual state of traffic and takes one of {High, Low} values.
- *Neighbor evacuate?*: It is natural to check whether your neighbor in a similar situation is evacuating or not, and a neighbor's presence early in the timeline may be misjudged as the neighbor staying put. This Boolean observation variable is influenced by the corresponding state variable.

- *Tweets to evacuate?*: Social media allows families and friends to remotely keep track of the disaster situation and any evacuation orders. It allows them to quickly convey their advice and worries about the individual's decision so far. We collected several such tweets issued during the time period hurricanes Harvey and Irma were active. This variable is influenced by evacuation order and takes a Boolean value.
- *Call from Work?*: Respondents in our survey reported receiving calls to report to work despite the presence of evacuation orders.

3.3 Hidden Variables

We introduce meaningful hidden variables each of which models the collective impact of its parents. These nodes help in aggregating related variables and avoiding a combinatorial explosion of the conditional probabilities due to multiple variables affecting a single node.

- *Weather threat perception*: Wind and rain observations together indicate the current status of the weather. This hidden variable may assume a value in the set {Good, Bad} and its conditional probability table (CPT) is shown in Table 2.

O Rain	O Wind	Good	Bad
None	Normal	.95	.05
None	High	.7	.3
None	Extreme	.1	.9
Normal	Normal	.8	.2
Normal	High	.5	.5
Normal	Extreme	.05	.95
Heavy	Normal	.5	.5
Heavy	High	.1	.9
Heavy	Extreme	.01	.99

Table 2: CPT for hidden variable representing *weather threat perception*.

- *Safety threat perception*: Flood and power failure observations, and safety of the house are combined to understand how safe is the individual in his or her geographical area. The conditional probability table of this hidden variable is shown in Table 3.
- *Mobility*: Finally, the demographic factors age, health condition, and transportation availability influence a person's mobility. An extract from the full DID with this node and its several parents is shown in Fig. 2.

While the weather and safety threats may change with time, an individual's mobility remains fixed.

A fourth variable that somewhat falls in this category is the *assessed risk* posed by the approaching hurricane to the individual's own safety and her property. In contrast to the previous hidden variables, our survey specifically asked the participants to assess this risk that they faced on a 7-point Likert scale with 1 being the lowest risk.

House Safety	O Power Failure	O Flooding	Safe	Unsafe
Safe	Not Possible	Not Possible	.99	.01
Safe	Not Possible	Possible	.9	.1
Safe	Not Possible	Flooded	.7	.3
Safe	Possible	Not Possible	.8	.2
Safe	Possible	Possible	.7	.3
Safe	Possible	Flooded	.6	.4
Safe	Failed	Not Possible	.65	.35
Safe	Failed	Possible	.6	.4
Safe	Failed	Flooded	.5	.5
Unsafe	*	*	0	1

Table 3: CPT for the hidden variable representing *safety threat perception*.

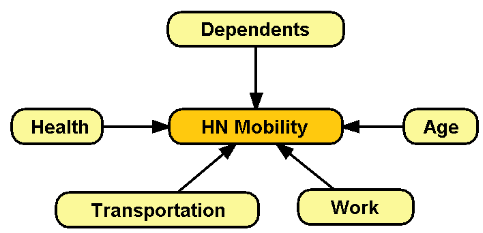

Figure 2: Factors influencing the hidden variable *mobility*. For clarity, we prefix hidden variables with 'HN' and shade them darker. All network extracts visualized in OpenMarkov.

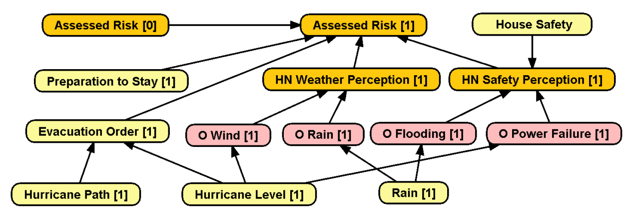

Figure 3: Factors influencing an individual's assessed risk. Variables prefixed with 'O' and shaded in a different color are the observation variables. The number in square parenthesis denotes the index of the time-slice containing that node.

We model the individual's assessed risk as shown in Fig. 3. While hidden variables *weather threat perception* and *safety threat perception* obviously influence the risk, an evacuation order raises the risk considerably. On the other hand, the individual's preparation to stay put and ride out the storm mitigates it.

Participants' responses to our survey question on assessed risk and their responses to its parents were utilized as part of expectation-maximization [9] to learn the CPT of *assessed risk*. During this process, a subset of the survey responses is given as evidence to the network in order to learn the CPT. With each case loaded as evidence, the probability distribution is updated based on the value for the *assessed risk* node and the number of times this combination of states of the parents was encountered earlier (experience). The CPT is initialized with an intuitive Gaussian distribution and the experience values are initialized with a uniform distribution. The assessed risk variable at previous time-step 0 is marginalized using a uniform distribution to obtain the CPT at time-step 1.

3.4 Actions

Individuals in impending disaster areas face the choice of evacuating or not evacuating. They may also choose to make a decision about evacuating after collecting more definite information on the observed variables. We label this third choice as *Get info*. As such, an individual can choose between three actions at each time-step.

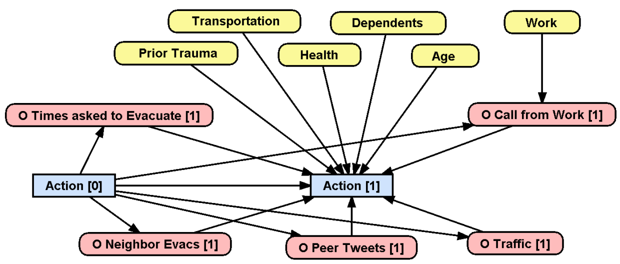

Figure 4: The Action decision node is shaded in blue. Variables that have outgoing edges to the decision node form the information set of action.

Figure 4 is an extract from the full DID that shows the action decision node and its information set. As expected for a POMDP, the observation variables described in Section 3.2 constitute this set (in addition to the previous action) and influence the policy.

3.5 Reward Function

As the DID models the evacuation decision making of an individual, an utility node in this DID models the individual's costs and preferences. Parents of the utility node are those variables that correlated significantly with the binary evacuation variable in our statistical analyses of Section 2.2. Consequently, these include state variables representing prior trauma experienced by the individual, whether neighbors evacuated, age and number of children that influence the hidden variable mobility, the traffic situation, assessed risk, and obviously the individual's decision. We show this subnetwork in Fig. 5. Because the state of the POMDP, which is the ground situation, is dynamic, a utility node is included in each time-slice and the corresponding expected utility may change from one time-slice to the next. The considered utility is the sum of these time-step utilities.

The multiattribute reward function is logically structured

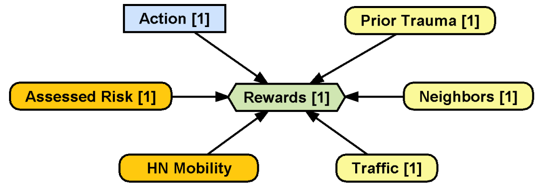

Figure 5: State and hidden variables that influence the reward function (hexagonal utility node) in addition to the action.

as follows. Mobility dominates other variables because an immobile individual is unable to evacuate even if the risk is high. If the individual is mobile, the final reward is a weighted combination of the reward contributions of the variables. After mobility, assessed risk has the highest weight of 3 and the other variables – traffic, neighbor evacuation, and prior trauma – are assigned an equal weight of 1. The *Get info* action is assigned a high cost if the assessed risk is high and a lower cost if the risk is assessed to be low, regardless of the other variables.

3.6 Transition Function

A significant benefit of using a DID to represent the POMDP is that it decomposes the overall transition function of the POMDP into transitions between the state factors due to actions. We show the full two time-slice DID that models the transition function in the supplement.

As we mentioned previously in Section 3.1, variables that pertain to demographics such as age, health, and others do not change with time (see the top portion of the DID). Wooten and Tsokos [19] model the change in hurricane level as a Markovian process. We utilized the transition function of this Markovian model in our DID to model how the hurricane level changes. The path of the hurricane continues to stay the same with a 75% chance. Once an evacuation order has been announced it is rarely revoked unless the strength of the hurricane drops significantly. If an order has not been announced then its announcement depends on the strength as well as the path of the hurricane.

The transition of assessed risk when the individual chooses to not evacuate or to get information is obtained from the expectation-maximization learning method as explained in Section 3.3. If the previous action is to evacuate, then the assessed risk drops to Level 1 regardless of the values of other variables.

Preparation to stay back is significant only when the previous decision is to get information. Otherwise, the distribution is uniform. Also, there is no possibility for the individual to become less prepared than before. A neighbor who has evacuated is generally not expected to return until the hurricane has passed. However, if they had not evacuated then the probability of her evacuating in the next time-step is same as the prior.

If it has started raining, then it continues with the same intensity with a probability of .85 whereas there is a 40% chance for it to start raining and a 10% chance of it being heavy. There is a high chance of traffic increasing when the previous decision is to evacuate. Otherwise, the probability of traffic depends on whether an evacuation order is present or not.

3.7 Observation Functions

The observation functions are designed to reflect the uncertainty in the perception of the environment. Thus, all of these were populated based on the survey responses. For example, we show the CPT for the observation node *times asked to evacuate* in Table 4. Given an evacuation order, we utilized the distribution of answers to our survey question to arrive at the probabilities. Otherwise, an individual most likely did not hear it.

Action (t-1)	**Evacuation Order**	**0**	**1-2**	**3-4**	**≥ 5**
Get info	Announced	.45	.42	.08	.05
Get info	Not Announced	.95	.04	.01	0
*	*	.25	.25	.25	.25

Table 4: CPT for observation variable representing *times asked to evacuate*.

3.8 Time Steps

The two time-slice DID is unrolled into *four* time-steps. Each new time-step signifies a major change in either the intensity or the direction of the hurricane under consideration. The number of time-steps was found to be four using this criterion in the cases of Hurricanes Harvey and Irma. The first time-step occurs when the evacuation order is first announced.

4 SENSITIVITY ANALYSES

Sensitivity analysis allows us to understand the impact of a change in the conditional probabilities of a hypothesis variable on a target random variable. As our model is strongly driven by data, the analysis serves as a tool to partly verify whether the variables and their CPTs are reflecting the influences intuitively. We deploy the sensitivity analysis to answer a series of questions, which clarify and help explain how the model works.

While sensitivity analysis can be performed in multiple ways [6], we rely on the method used by the Hugin Expert system. First, select a cell in the CPT of the hypothesis variable. For various hypothetical probabilities in this cell, we may obtain the corresponding inferred probabilities of the values of the target random variable. This sensitivity function is shown as a line graph – one line for each value of the target variable. A sensitivity value for each state of the target variable is then simply the derivative of the sensitive function (the slope of the line).

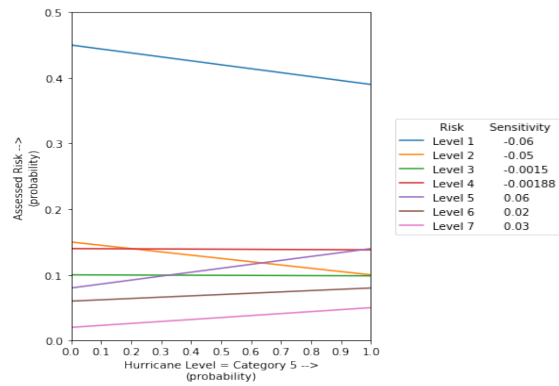

Figure 6: Sensitivity functions showing how the probabilities of various *assessed risk* levels changes with the probability of the hurricane initially being a category 5 at landfall.

Is assessed risk influenced by the strength and path of the hurricane? Figure 6 shows the sensitivity analysis with *assessed risk* as the target variable. Observe that the risk is indeed influenced by the strength of the hurricane: as probability of the hurricane initially making landfall as category 5 increases, risk levels ≥ 5 exhibit a positive slope and for other levels the slope is negative. However, the path of the hurricane did not impact the risk. We believe this is because the network has only seen cases where the hurricane is headed toward the individual and none where the hurricane veers away.

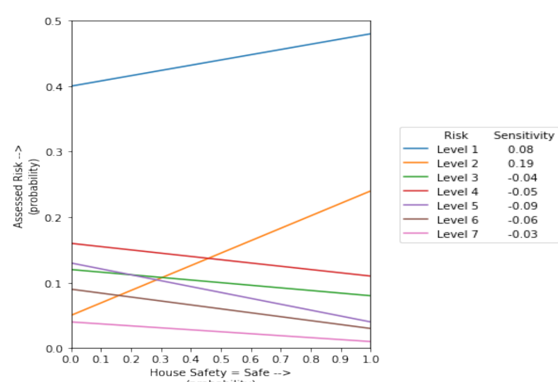

Figure 7: Sensitivity functions showing how the probabilities of assessed risk levels are affected by the changes in probability of individual's house being safe.

Does the safety of the house significantly impact risk as-

sessment? Observe from Fig. 7 that assessed low risk-levels of 1 and 2 exhibit positive slopes as the probability of house being safe increases. Indeed, risk level 2 probability increases steeply. Furthermore, the probabilities of higher risk levels reduce as we may expect.

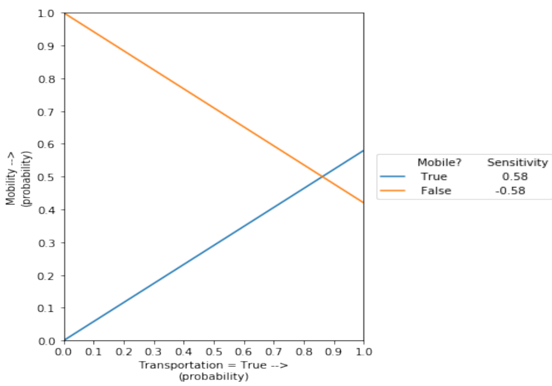

Figure 8: Sensitivity functions showing how the probabilities of mobility values change with the probability of the presence of transportation.

Does the presence of transportation most affect the individual's mobility? What are the other factors influencing mobility? Figure 8 reveals that mobility is highly sensitive to changes in the probability of having transportation (sensitivity value is .58). This was closely followed by the individual's employment situation with a sensitivity value of .32 – employment inhibits the ability to move.

5 PERFORMANCE EVALUATION

We implemented the DID representing the POMDP in Hugin Expert, and utilized Hugin's Java application programming interface (API) to process and test the prediction accuracy of the network.

We evaluated the prediction accuracy of the model using 3-fold cross-validation on data pertaining to the 822 respondents in our survey. A Python program converted the shuffled data from the survey into key-value pairs input to Hugin. During each run of the cross-validation, data in the two training folds were utilized to learn the CPT of the *Assessed Risk* node and others. Recall that Section 3.3 discusses how this CPT is learned.

Each participant record in the test fold (there are 274 records in this fold) was propagated individually through the DID using Hugin API. Demographic state variables for a participant do not change with time and their values were entered as evidence first. This was followed by entering the values of the experiential state variables into the corresponding nodes at the initial time-step. The participant's observations are entered into the nodes prefixed with 'O' in the second time-slice. An example record

Dependents:	"Absent"
Age:	"a61_above"
Prior_Trauma:	"Absent"
Hurricane_Level_0:	"Cat_5"
Evacuation_Order_0:	"Announced"
Preparation_to_Stay_1:	"Very"
O_Flooding_1:	"Possible"
O_Power_Failure_1:	"Failed"
O_Times_asked_to_evacuate_1:	"t1_2"
O_Neighbor_Evacuations_1:	"Not_Evacuated"

Table 5: An example subrecord containing demographic, experiential, and observational data entered as evidence into the DID.

listing evidence values for an individual is shown in Table 5. To allow consideration of these observations, we set the initial decision to be *Get info*. As such, the model predicts a decision to evacuate or not in subsequent time-steps. The evacuation decision with the highest expected utility is chosen as the model's prediction.

(a) *Confusion matrix (NE - Not evacuate, EV - Evacuate). Numbers not in bold are the mispredictions.*

	NE	EV
NE	**105**	23
EV	58	**88**

Accuracy: 70.44%

(b) *Decisions at various time steps (GI - Get Info)*

t	1	2	3
GI	12	1	
NE	**116**	**11**	**1**
EV	**146**		

Table 6: Results for cross validation run 1.

(a) *Confusion matrix*

	NE	EV
NE	**124**	27
EV	43	**80**

Accuracy: 74.45%

(b) *Decisions at various time steps*

t	1	2	3
GI	21		
NE	**130**	**21**	
EV	**123**		

Table 7: Results for cross validation run 2.

(a) *Confusion matrix*

	NE	EV
NE	**105**	21
EV	60	**88**

Accuracy: 70.44%

(b) *Decisions at various time steps*

t	1	2	3
GI	21	3	
NE	**105**	**18**	**3**
EV	**148**		

Table 8: Results for cross validation run 3.

Tables 6, 7, and 8 show the results for each of the three cross-validation runs. We show both the confusion matrix and the predicted decisions at various time-steps for the test fold in each run. Recall that the decision at time-step 0 is set as *Get info* and we do not show it. Once the network recommends evacuate or not, subsequent decisions become irrelevant for that individual. The average accuracy across all three runs is 71.77% with a standard deviation of 1.89% and a high of 74.45%. Our model predicts a decision to evacuate or not at time-step 1 for a majority of the participants. For example, in run 1 it rendered a decision that the participant will evacuate or not in time-step 1 for 95% of the respondents. Indeed, coastal residents in Georgia began evacuating immediately on receiving the evacuation order. The DID takes about 45 seconds on an Intel Core i7, 64GB RAM, Ubuntu PC to render a decision.

Among the mispredictions, our model incorrectly predicts evacuate for more respondents than an incorrect prediction of not evacuate. A deeper analysis of the data reveals that the false predictions in the confusion matrices (numbers not in bold) often reflects human behavior that is hard to understand. We found that 62 respondents had decided to not evacuate despite assessing the risk level to be very high – at 6 or 7 – and they were mobile. Additionally, more than 100 who did not evacuate falsely believed that they were not in an evacuation zone although an evacuation order was issued for their place of residence. Consequently, the number of respondents for whom the model mispredicted evacuate is higher.

6 RELATED WORK

Formal data collection efforts on factors affecting potential evacuees are scarce: we know of just one small survey of affected individuals by NYC's Mental Health and Hygiene Department after Hurricane Sandy [2]. The survey revealed that those who witnessed suffering during the WTC attacks (i.e., prior trauma), those who were staying in lower floors, and those who expected heavy damages to their homes had evacuated more than the others. Many of these findings are consistent with the findings from our survey and serve to reinforce the inclusion of these factors in our model.

Initially, it was believed that the decisions are solely based on the different kinds of warnings. Subsequently, it was identified that different people react differently to the same announcement of disaster warning [10]. Previous research on evacuation decision modeling has explored the use of decision trees composed of binary questions based on personal interviews [5]. Though the model was able to achieve high accuracy of prediction, we observe that it is not probabilistic and requires interviewing the individual. Therefore, there is no room for ambiguity or an option to skip a question and still get predictions.

Hasan et al. [7] developed a statistical mixed-logit composed of numerous factors and variables including state/location, home ownership, prior hurricane experience, education level, dependents, and income. Most of these variables were included in our survey questionnaire. While some of them were found to be significant predictors of evacuation and included in our model, others not found to be influencing the decision were dropped from subsequent analyses. A key point of difference is that the logit model is not a process model.

7 CONCLUDING REMARKS

We presented a POMDP model, represented as a DID with four time-steps, to model the evacuation decision making of an individual in an impending disaster zone. The responses of a directed survey received from the residents of US states of TX, FL, and GA who were asked to evacuate during Hurricanes Harvey and Irma served to inform this model. The interrelations between the variables and the CPTs depict the situation and the differential importance of the factors in the decision-making process. Data from the survey responses were used to enter the evidence for the variables in the model and arrive at the decisions with a reasonably good accuracy. The paper presents a descriptive process model that aims to capture an individual's actual decision making, and not normative recommendations (which was to evacuate for all surveyed individuals). A graphical representation in the form of a DID facilitates comprehension and explainability of its inference and predictions, which promotes its use by agency staff who may not be experts in AI.

Human decision making is known to suffer from cognitive biases and these may affect evacuation decision making as well. Indeed, there is preliminary research [18, 3] on the evacuation behavior of residents in areas frequently affected by hurricanes with a focus on the psychological aspects of decision making (but lacking a computational model). Our administered survey also tested for some cognitive biases in individuals: it addressed three well-established and potentially relevant biases generally evident across populations: gambler's fallacy associated with overconfidence in predictive ability [1]; the illusion of control which occurs when people confuse chance with skill and thus behave as though their actions control the outcome of a random event [14], and the confirmation bias [13]. An ongoing analysis will reveal whether any of these biases were significantly evident in the surveyed population. However, we caution that mathematical models of these biases do not exist, which makes their inclusion in our decision-making model not trivial.

Acknowledgements

This research was supported by NSF through a RAPID grant #1761549.

References

[1] P. Ayton and I. Fischer. The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, 32(8):1369–1378, Dec 2004.

[2] S. Brown, H. Parton, C. Driver, and C. Norman. Evacuation During Hurricane Sandy: Data from a Rapid Community Assessment. In *PLOS Currents: Disasters*, pages 45–50, Jan. 2016.

[3] N. Dash and H. Gladwin. Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, 8(3):69–77, 2007.

[4] T. Dean and K. Kanazawa. A model for reasoning about persistence and causation. *Computational Intelligence*, 5(2):142–150, 1989.

[5] C. Gladwin, H. Gladwin, and W. G. Peacock. Modeling hurricane evacuation decisions with ethnographic methods. *International journal of mass emergencies and disasters*, 19(2):117–143, 2001.

[6] D. M. Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental monitoring and assessment*, 32:135–154, 09 1994.

[7] S. Hasan, S. Ukkusuri, H. Gladwin, and P. Murray-Tuite. Behavioral model to understand household-level hurricane evacuation decision making. *Journal of Transportation Engineering*, 137(5):341—-348, 2011.

[8] L. Kaelbling, M. Littman, and A. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence Journal*, 2, 1998.

[9] S. L. Lauritzen. The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*, 19(2):191 – 201, 1995.

[10] D. S. Mileti and E. Beck. Communication in crisis: Explaining evacuation symbolically. *Communication Research*, 2(1):24–49, 1975.

[11] K. Murphy. *Dynamic Bayesian networks: representation, inference and learning*. PhD thesis, University of British Columbia, 2002.

[12] National Centers for Environmental Information. State of the Climate: Hurricanes and Tropical Storms for Annual 2017. `https://www.ncdc.noaa.gov/sotc/tropical-cyclones/201713/`, 2018. [retrieved March 7, 2019].

[13] R. S. Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2):175–220, 1998.

[14] P. K. Presson and V. A. Benassi. Illusion of control: A meta-analytic review. *Journal of Social Behavior & Personality*, 11(3):493 – 510, 1996.

[15] T. Schott, C. Landsea, G. Hafele, J. Lorens, A. Taylor, H. Thurm, B. Ward, M. Willis, and W. Zaleski. The Saffir-Simpson hurricane wind scale. *NOAA/National Weather Service [Internet]*, pages 1–4, 2012.

[16] S. K. Smith and C. McCarty. Fleeing the storm(s): An examination of evacuation behavior during florida's 2004 hurricane season. *Demography*, 46(1):127—-145, 2009.

[17] J. A. Tatman and R. D. Shachter. Dynamic programming and Influence Diagrams. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2):365–379, 1990.

[18] J. C. Whitehead, B. Edwards, M. V. Willigen, J. R. Maiolo, K. Wilson, and K. T. Smith. Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Global Environmental Change Part B: Environmental Hazards*, 2(4):133–142, 2000.

[19] R. Wooten and C. Tsokos. A Markovian analysis of hurricane transitions. *Neural, Parallel and Scientific Computations*, 16:1–16, 03 2008.