May 2022

An Agent-Based Exploration of the Hurricane Forecast-Evacuation System Dynamics

Austin Reed Harris
University of Wisconsin-Milwaukee

AN AGENT-BASED EXPLORATION OF THE HURRICANE FORECAST-EVACUATION SYSTEM DYNAMICS

by

Austin R Harris

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Atmospheric Science

at

The University of Wisconsin-Milwaukee

May 2022

ABSTRACT

AN AGENT-BASED EXPLORATION OF THE HURRICANE FORECAST-EVACUATION SYSTEM DYNAMICS

by

Austin R Harris

The University of Wisconsin-Milwaukee, 2022
Under the Supervision of Professor Dr. Paul Roebber

In the mainland US, the hurricane-forecast-evacuation system is uncertain, dynamic, and complex. As a result, it is difficult to know whether to issue warnings, implement evacuation management strategies, or how to make forecasts more useful for evacuations. This dissertation helps address these needs, by holistically exploring the system's complex dynamics from a new perspective. Specifically, by developing – and using – an empirically informed, agent-based modeling framework called FLEE (Forecasting Laboratory for Exploring the Evacuation-system). The framework represents the key, interwoven elements to hurricane evacuations: the natural hazard (hurricane), the human system (information flow, evacuation decisions), the built environment (road infrastructure), and connections between systems (forecasts and warning information, traffic). The dissertation's first article describes FLEE's conceptualization, implementation, and validation, and presents proof-of-concept experiments illustrating its behaviors when key parameters are modified. In the second article, sensitivity analyses are conducted on FLEE to assess how evacuations change with evacuation management strategies and policies (public transportation, contraflow, evacuation order timing), evolving population characteristics (population growth,

urbanization), and real and synthetic forecast scenarios impacting the Florida peninsula (Irma, Dorian, rapid-onset version of Irma). The third article begins to explore how forecast elements (e.g., track and intensity) contribute to evacuation success, and whether improved forecast accuracy over time translates to improved evacuations outcomes. In doing so, we demonstrate how coupled natural-human models – including agent-based models –can be a societally-relevant alternative to traditional metrics of forecast accuracy. Lastly, the fourth article contains a brief literature review of inequities in transportation access and their implication on evacuation modeling. Together, the articles demonstrate how modeling frameworks like FLEE are powerful tools capable of studying the hurricane-forecast-evacuation system across many real and hypothetical forecast-population-infrastructure scenarios. The research compliments, and builds-upon empirical work, and supports researchers, practitioners, and policy-makers in hazard risk management, meteorology, and related disciplines, thereby offering the promise of direct applications to mitigate hurricane losses.

TABLE OF CONTENTS

LIST OF FIGURES

.

LIST OF TABLES

LIST OF ABBREVIATIONS

ABM	Agent-Based Model
ASP	Advanced Study Program
EM	Emergency Management
EO	Evacuation Order
FLEE	Forecasting Laboratory for Exploring the Evacuation-system
LF	Landfalling
NCAR	National Center for Atmospheric Research
NHC	National Hurricane Center
NSF	National Science Foundation
NWS	National Weather Service
PADM	Protective Action Decision Model
RI	Rapid Intensity
RO	Rapid Onset
US	United States
UWM	University of Wisconsin-Milwaukee
WDTD	Warning Decision Training Division

ACKNOWLEDGEMENTS

This dissertation, the completion of the doctoral degree, and my growth as an interdisciplinarian research scientist is made possible through the support of many people.

Mentors Paul Roebber, for among many things, showing me how to innovate, be efficient and creative, and for taking the training wheels off; Jon Kahl, for showing me how to research and teach; Rebecca Morss, for helping me find my community in meteorology; Clark Evans, for always being supportive; Mike Westendorf, for showing me how to leave a legacy; and Susan Postawko, for believing in me.

To Rebecca Morss, Julie Demuth, Paul Roebber, Heather Lazrus, Olga Wilhelmi, Gina Eosco, Jen Henderson, Kim Klockow and the many other creative, interdisciplinary scientists whose scholarship inspires me — thanks for forging unique paths. To my ASP Colloquium family, thanks for walking them with me.

My actual family: Brian Harris, Janet Reid, Kevin Lee, Maddie Harris, Almeda and Harvey Harris, Bob Reid and the late Norma Reid — for grounding me on my trips home. To Harvey, for always knowing my estimated time of arrival.

Brothers from other mothers, Alex Strout, Seth Strout, Cedric Griffin, Ryan Kloesel and Jacob Tsotigh III — for challenging me intellectually throughout our twenty years of friendship.

My Wisconsin family: Caitlin Crossett, Kyle O'Connell, Amanda Quesnell, and Zachary Freudenberg — for being there during many stressful times. The Burlingame's, my

chosen older siblings, for the life advice and meals — I admire you greatly. And Zoe and Eddy Zibton, for reminding me what's important.

Friends and colleagues at the University of Wisconsin-Milwaukee (UWM), in particular: Kevin Prince, Andrew Westgate, Dillon Blount, Victoria Lang, and Alex Moxon. Thanks for helping make UWM special.

My new Boulder family: Hailey Winterton, Max Grover, Tom and Taylor Gowan, Zach Hiris, Sam Lillo, Manda Chasteen, and many others — thanks for keeping me sane during the home stretch. Special thanks to Russell Danielson, for housing the homeless, and for the excellent trips; to Austin Coleman, the community builder, for always seeing the positives; and to my first meteorology friend, Dylan Reif — look how far we've come since our days at Jackson Elementary.

This research was funded by the National Science Foundation (NSF) under Grant No. 2100801; UWM's Distinguished Dissertator Fellowship; and the National Center for Atmospheric Research's (NCAR) Advanced Study Program's Graduate Student Fellowship.

Lastly — but most importantly — to my parents, Janet Reid and Brian Harris, for recognizing this dream, and for doing everything in your power to make it happen. This dissertation is yours, too.

PREFACE

This dissertation marks a major moment in my career. I want to take a minute to acknowledge – and reflect on – the journey to becoming a Doctor in Atmospheric Science.

On 5/3/99, the “weather seeds” were planted on my brain. That F-5 tornado in Moore, OK passed 1-mile from our home. I wrote afterwards, “I will go to the Oklahoma school of meteorology. There I’ll work as a storm chaser. Eventually, I’ll work in the severe storms lab issuing warnings etc.” Basically, I had the unique privilege of recognizing a dream early in my life, and with the support of two wonderful parents, actually achieved it.

I graduated from the University of Oklahoma in 2013 with a B.S. in Meteorology. After applying – and being denied – entrance to 15+ graduate programs, UWM was the first and only graduate program to offer me a scholarship. I took it, and after two years at UWM, left with an M.S. in Atmospheric Science, forecasting experience at Innovative Weather, and an offer for my first “big boy” job at the Warning Decision Training Division (WDTD) of the National Weather Service (NWS). There, I joined a team of instructors teaching forecasters how to issue thunderstorm, flash flood, and tornado warnings. It was my childhood dream, actualized.

Though I was immensely proud of my new role in meteorology, it was not my meteorological home.

After Katrina (2005) and evacuations from the El Reno tornado (5/31/13), I felt that “accurate” forecasts were not necessarily “good” forecasts and that the weather community needs research focusing on societal impacts of forecasts (e.g., like the interdisciplinary work of Drs. Rebecca Morss, Julie Demuth, Kim Klockow, and many

others). Knowing this belief and interest, and seeing an opportunity to simulate evacuations using agent-based models, Dr. Paul Roebber asked if I would return to graduate school to study this topic. After many conversations, I agreed, knowing Paul was the right advisor, and that this was a golden opportunity to get into an area of research I felt strongly about.

Early in the Ph.D., Paul connected me with Dr. Rebecca Morss at NCAR. Rebecca kindly agreed to serve on my committee and help me navigate the social science literature needed to model evacuations. Through Rebecca and Paul's support, the project has become something I'm extremely proud of. It would not be possible without them. In addition, Rebecca connected me to other interdisciplinary scientists through NCAR's ASP colloquium and the Graduate Visitor Program. In this community, I've found my meteorological home.

This dissertation marks my entrance into this scholarly community. From here, I intend to support the weather enterprise for many years to come.

INTRODUCTION

Across the US, the hurricane-forecast-evacuation system is uncertain, dynamic, and complex. Take Hurricanes Irma (2017) and Rita (2005), cases where accurate but uncertain forecasts triggered evacuations — and severe traffic jams — in Florida and Texas, respectively. As the forecasts shifted and traffic worsened, some evacuees became even more exposed than had they remained in-place (Cangialosi et al. 2018; Wong et al. 2018; Zhang et al. 2007; Knabb et al. 2006). Despite forecast information being as useful as one can reasonably expect given current forecast skill, the cases illustrate the complexities of people using inevitably imperfect information to make evacuation decisions before the storm, and they demonstrate how evacuations involve many physical-social parts and uncertainties that change as the storm approaches the coast (e.g., Morss et al. 2017; Barton 2014; Miller and Page 2007; Watts et al. 2019). Because of these complex dynamics, safe and efficient hurricane evacuations can be difficult to achieve.

Empirical studies explain important aspects of evacuations, such as how forecasts, warnings, and other factors influence evacuation decisions (e.g., Huang et al. 2016; Lindell and Perry 2012; Baker 1991). However, it's impossible to empirically study all aspects of evacuation across multiple cases. Computational models, on the other hand, provide a complementary tool where empirical knowledge can be codified and used to run virtual experiments for many scenarios, real or synthetic (e.g., Morss et al. 2017; Blanton et al. 2018). The entire forecast-evacuation system has not been represented in one computational framework, however, with models focusing on specific aspects only (e.g., evacuation traffic only, or evacuation decision-making only).

This dissertation explores the hurricane-forecast-warning system dynamics from an agent-based, computational perspective which compliments, and builds upon, empirical work. The model includes representations of the forecast, evacuation decision-making, and evacuation traffic together in one framework for the first time. By performing experiments on the unique model system, the study becomes the first to examine the hurricane evacuation system holistically i.e., to establish the relative importance of factors, key interactions between systems, and non-evident emergent patterns.

The work takes the form of a series of manuscripts which have been – or are in the process of being – published by academic journals. Detailed literature reviews are provided in each article and tailored to their specific purposes. Drs. Paul Roebber and Rebecca Morss are co-authors on the manuscripts, and thus, are co-authors for the dissertation (note: I will reference “we” throughout, as they greatly contributed to the articles).

The first article details and describes the agent-based modeling framework i.e., its conceptualization, implementation, and the empirical data which it is based. Though intended for a broad, international audience, the work is most useful for evacuation modelers, as it contains detailed information about the model’s structure and reasoning behind its development. The research has been peer-reviewed and published in the *International Journal of Disaster Risk Reduction* (Harris et al. 2021).

The second article provides a first-order look at the systems dynamics, by using the model to explore how evacuations change with evacuation management strategies and policies (public transportation, contraflow, evacuation order timing), evolving population

characteristics (population growth, urbanization), and real and synthetic forecast scenarios impacting the Florida peninsula (Irma, Dorian, rapid-onset version of Irma). The primary intended audience are researchers, practitioners, and policy-makers in hazard risk management.

The third article begins to explore how individual forecast elements, such as forecast track and intensity – and their improved accuracy over time – translates to evacuation success. As such, the research benefits meteorology by demonstrating how coupled natural-human models provide a societally-relevant alternative to traditional metrics of forecast "accuracy" e.g., by measuring the impact of forecasts elements on how people make evacuation decisions and physically evacuate.

The fourth article is a brief literature review on carless households in the US and whether inequities exist in hurricane evacuations as a result (they do). Through the literature review, the idea is to ensure future evacuation modeling studies do not contribute to inequities, and instead, are used to identify and reduce them across many forecast-population-infrastructure scenarios.

To my colleagues in the weather and hazards communities, I offer the following articles to help advance new methods for studying hurricane evacuations, to better support practitioners/policy-makers aiming to improve evacuations, to change how forecast verification is done across the weather enterprise, and to help cultivate a shared understanding across disciplines of the entire hurricane-forecast-warning system. This last point is especially important for meteorology, as the community often struggles to understand why improved forecast accuracy does not always translate into increased public safety.

ARTICLE 1

AN AGENT-BASED MODELING FRAMEWORK FOR EXAMINING THE DYNAMICS OF THE HURRICANE-FORECAST-EVACUATION SYSTEM

1. Introduction

Hurricanes Irma (2017) and Rita (2005) demonstrate how, in the mainland US, the forecast-evacuation system is uncertain, dynamic, and complex. For example, Irma's 3-10-day forecasts indicated the storm was likely to make landfall as a major hurricane somewhere in Florida, with the most likely track near Miami, triggering the largest evacuation in US history (FDEM 2017). However, the forecast track shifted slightly westward as the storm approached, with eventual landfall near Tampa Bay–St. Petersburg, a common evacuation destination in the event, while leaving Miami largely unscathed (Cangialosi et al. 2018; Wong et al. 2018). Similarly, uncertainties in Hurricane Rita's track and intensity forecasts, combined with the aftermath of Hurricane Katrina, led to mass evacuations and severe traffic jams in Houston–Galveston. The worst of the storm missed the area, but had Rita struck Houston–Galveston directly, the consequences could have been severe, as many evacuees were stranded on area roads (Zhang et al. 2007; Knabb et al. 2006).

The events are relevant since the forecasts were fairly accurate, with the westward shift of Irma's track falling within the National Hurricane Center's cone of uncertainty (Cangialosi et al. 2018), and Rita's forecast track being less erroneous than most (Knabb et al. 2006). However, forecasts were less successful in providing useful guidance for many affected by the events, despite being as useful as one can expect given current

forecast skill. These cases illustrate the complexities of people using inevitably imperfect forecasts to make evacuation decisions well before the storm arrives, and they demonstrate how evacuations involve many interacting physical-social parts and uncertainties which evolve over time (e.g., Morss et al. 2017; Barton 2014; Trainor et al. 2012; Miller and Page 2007). Because of these complex dynamics, safe and efficient evacuations can be a formidable challenge.

Empirical studies provide insight to different aspects of hurricane evacuations, such as how forecasts, warnings, and other factors influence evacuation decisions (e.g., Huang et al. 2016; Lindell and Perry 2012; Baker 1991). However, it is difficult to empirically study all aspects of evacuations across multiple cases. Computational models, on the other hand, provide a complementary tool where empirical knowledge can be codified and used to run virtual experiments for many different hurricane scenarios, real and synthetic (e.g., Morss et al. 2017, Watts et al. 2019). Recent research demonstrates the potential of modeling the hurricane evacuation system together in one framework (e.g., Watts et al. 2019; Blanton et al. 2018). With that we ask: can a modeling framework be designed to *holistically* investigate the complex dynamics of the hurricane-forecast-warning system i.e., to determine which factors are important and how they interact across a range of scenarios?

To answer this question, we introduce a new modeling framework, FLEE (Forecasting Laboratory for Exploring the Evacuation-system). FLEE includes several empirically-informed models representing key, interwoven aspects of real-world hurricane evacuations: the natural hazard (hurricane), the human system (information flow, evacuation decisions), the built environment (road infrastructure), and connections

between systems (forecasts and warning information, traffic, impact zones). The hurricane and forecast information are represented using data and products from the National Hurricane Center (NHC), a component of the U.S. National Weather Service (NWS) which is the leading authority for real-time hurricane forecasting. Two agent-based models (ABMs) replicate 1) the flow of information and evacuee decision-making, and 2) evacuation infrastructure, routing, and traffic. These models are conceptually and numerically interconnected as shown in Figure 1.

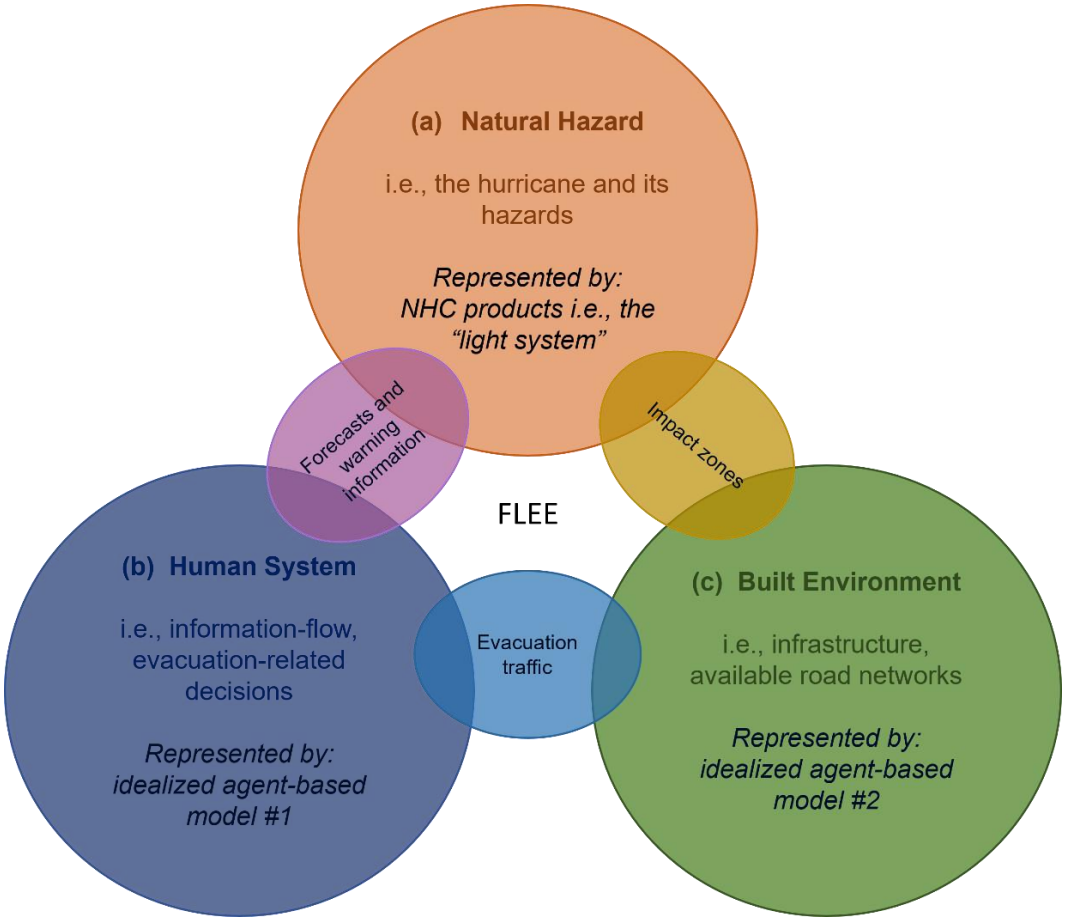

Figure 1: A conceptual overview of FLEE which includes models of the three interconnected systems of hurricane evacuations: (a) the natural hazard (b) the human system, and (c) the built environment, represented by NHC forecast products and two ABMs, respectively (italics). Forecast and warning information (purple), evacuation traffic (light blue), and impact zones (gold) serve as conceptual links between systems. Coupling the individual models (a-c) via these links makes FLEE a hybrid agent-based and system dynamics model (Martin and Schlüter 2015) uniquely positioned to perform experiments impossible to conduct in the real-world.

This article has two primary objectives. First, to overview the conceptualization and implementation of FLEE. This includes describing the model components, which are designed to represent key aspects of real-world hurricane evacuations, while remaining sufficiently idealized to build fundamental and practical knowledge (e.g., see Watts et al. 2019; Sun et al. 2016, discussion in Section 2). The article's second aim is to show results from experiments demonstrating how FLEE is uniquely positioned to examine the hurricane-forecast-warning system dynamics. That is, how it can explore the effects of altering different factors, interactions among system components, and to show how large-scale patterns of evacuation can emerge from individual decisions of many heterogeneous agents interacting with each other and with their physical-informational environments.

Preliminary experiments are performed on a simplified representation of the Florida peninsula – a place frequently visited by tropical systems (Keim et al. 2007) – and for Hurricane's Irma and Dorian, which affected these areas in 2017 and 2019. FLEE was designed to be flexible, however, and thus the modeling framework can be modified to study other regions, hurricane scenarios, and multi-hazards e.g. hurricanes followed by flooding or cascading failures such as loss of power networks, damage to roads etc.

This research builds on previous work which models the hurricane evacuation system by expanding the components of the full system represented within the same modeling framework. For example, one body of work uses ABMs to study evacuation planning (e.g., Madireddy et al. 2011; Zhang et al 2009; Chen 2008, 2012; Zhan and Chen 2008; Chen and Zhan 2004, 2006). Such work focuses on evacuation traffic while using highly

idealized representations of the forecast and warning information and evacuation decision-making. Meanwhile another body of work uses ABMs and other models to study information flow and evacuation decision making but does not include representations of evacuation routing and traffic (see, e.g., Dixon et al. 2017; Yin et al. 2014; Widener et al. 2013; Watts et al. 2019; Morss et al. 2017; Czajkowski 2011; Hasan et al. 2013). Arguably the most comprehensive model of the hurricane evacuation system is Blanton et al. (2018) and Davidson et al. (2018), as they integrate the forecast, evacuation decisions, and evacuation traffic into one system. However, its representation of information flow and evacuation decision making were fairly simplistic as these models were designed for operational use.

Modeling frameworks like FLEE, which represent the entire hurricane-forecast-warning system, can support researchers, practitioners, and policy-makers in a variety of disciplines. This includes hazard risk management, which would benefit from increased knowledge of the relative effectiveness of evacuation management strategies. The evacuation modeling community would benefit from improved understanding of evacuation, which provides better rationale for variable selection in future models. In meteorology, modeling frameworks like FLEE can provide a societally-relevant alternative to traditional measures of forecast accuracy, by showing how forecasts influence evacuation success. Lastly, by looking at the system holistically, these modeling frameworks can cultivate shared understanding across these disciplines, a need emphasized by Bostrom et al. (2016).

2. Modeling framework and implementation

This section describes FLEE's components and design (Grimm et al. 2020). The modeling framework was developed using Fortran due to familiarity with the language but could be developed using existing agent-based software. For further details, the commented code, a model description, and input files are available for download at the CoMSES model library (https://www.comses.net/codebase-release/4cd05855-f387-48bd-8899-9d62375518cb/).

FLEE can run on multiple operating systems, including MacOS, Linux, and Windows, and on computers with average memory and cores (e.g., we used computers with 2 cores and 4 GB memory). Simulations typically require 3-5 days of real-time. Though it cannot run in quasi-real time on a desktop computer, the paper's goal is proof of concept – improving run time is a key next step for more practical use.

The modeling framework includes a spatially explicit virtual world representing a geographical area of interest (described in section 2.1); a dynamic hurricane – and forecast information about it – that passes through that world (section 2.2); a multi-agent model where information is interpreted by millions of heterogenous agents and used to make evacuation decisions (section 2.3); and a traffic model where agents move across the virtual world as the hurricane approaches (section 2.4).

ABMs were chosen to represent the human system (Figure 1b) and the built environment (Figure 1c) as the models capture individual's decision-making processes and interactions between agents, making them excellent tools for investigating complex system dynamics (e.g., see Miller and Page 2007; Barton 2014; Hammond 2015, Rand

and Rust, 2011). Another reason is that ABMs have proven capable of simulating evacuation decisions in hurricanes (see, e.g., Dixon et al. 2017; Yin et al. 2014; Widener et al. 2013; Zhang et al. 2009; Watts et al. 2019; Morss et al. 2017) and hurricane traffic dynamics (e.g., Gehlot et al. 2019; Ukkusuri et al. 2017; Liang et al. 2015; Chen and Zhan 2008). One drawback of ABMs is their high computational expense, which makes them less suitable for operational use (e.g., pros/cons in Bazghandi 2012).

To design and implement FLEE, we integrated across multiple relevant areas of expertise, including agent-based modeling, meteorology, emergency management, protective decision making, risk communication, social vulnerabilities, and traffic modeling. As in any modeling effort, aspects of FLEE are simplified and some real-world processes are not represented. Decisions about what to include were based on our research goals (e.g., to explore the broad system dynamics), review of relevant literature, and discussions among our research team. These decisions are discussed throughout Sections 2.1–2.4.

2.1. The virtual world

FLEE's virtual world is a 10 x 4 cellular representation of the north-south axis of Florida, an area susceptible to hurricanes (Keim et al. 2007) and which has experienced mass evacuations such as Irma (2017). The grid spacing is coarse by design (40 grid spaces of 69-km x 69-km each) as the project's goal is to explore the broader system dynamics, and to provide a starting point for more complex experiments. Census data informs the spatial distribution of agent households on the abstracted grid as well as household characteristics (which then influence evacuation decisions as discussed in section 2.3).

For the built infrastructure, virtual highways and interstates designed to simulate key aspects of Florida's road network are overlaid on the model grid (section 2.4). These roads allow agents to move between grid cells for evacuation. Details regarding the construction of each model system (i.e., the natural hazard, the human system, and the built environment) and the key connections between them is provided in the next three subsections.

2.2.The natural hazard (hurricane, forecasts, and warning information)

FLEE includes a hurricane that approaches and can move through the model domain (Figure 1a). The storm and its forecasts can be real or synthetic; here we simulate real, historical storms using archived NHC forecast products which were issued in real-time. The products include information about the observed storm characteristics (Table 1) and official forecast information (Table 2), both of which update every 6-hours (both in FLEE and in the real-world). When taken together, the products capture the critical storm information and its evolution as the storm approaches. We chose to use NHC products in this implementation rather than meteorological model ensembles (as used in Blanton et al. 2018; Davidson et al. 2018) because they more closely resemble forecasts seen by the public (Demuth et al. 2012), and can be systematically perturbed to assess the evacuation's sensitivities to the forecast. Note, the NHC products are a starting point, but FLEE can be extended to include additional or more complex information about the storm and forecasts and warnings, if desired. In this article, NHC forecast products are obtained for Hurricanes Irma (2017) and Dorian (2019), which represent forecast scenarios with different tracks, speeds, forecast errors, and subsequently, different evacuation behaviors (e.g., Wong et al. 2018, Mongold et al. 2020; Long et al. 2020).

Observed storm characteristics	Archived NHC products where the data are located
Observed wind radii (i.e., 64, 50, and 34 knot wind speeds in each of 4 quadrants)	Advisory Wind Field
Observed maximum sustained winds (i.e., current storm category)	Advisory Forecast Track
Observed forward speed	Advisory Forecast Track

Table 1: Observed storm characteristics used in FLEE and the NHC products from which the data are located. Storm characteristics includes the storm's observed location, size, intensity, and forward speed as it moves across the virtual world (left). This information was taken from archived NHC forecast products (right) which were issued in real time (available at https://www.nhc.noaa.gov/gis/). Consistent with the wind speeds in the NHC data, winds are discussed here in the unit knots (nautical miles per hour, equivalent to approximately 1.15 mph or 1.85 km/h).

Forecast information	Archived NHC products where the data are located
Forecast track	Advisory Forecast Track
Forecast maximum sustained winds (i.e., forecast storm category)	Advisory Wind Field
Forecast wind radii (i.e., 64, 50, and 34 knot wind speeds in each of 4 quadrants)	Forecast Wind Radii
Uncertainty in forecast track	Cone of Uncertainty
Expected arrival time	Arrival time of tropical storm force winds

Table 2: Forecast information used in FLEE and the NHC products from which the data are located**.** The forecast information (left) includes the storm's expected track, category, size, the amount of uncertainty associated with the forecast track (i.e., the cone of uncertainty), and the expected arrival time of the storm. Specific archived NHC products where the data was taken is shown (right) and is available for download at https://www.nhc.noaa.gov/gis/. Note: the cone of uncertainty represents the probable track of the center of a tropical cyclone, and is formed by enclosing the area swept out by a set of circles along the forecast track (at 12, 24, 36 hours, etc.). The size of each circle is set so that two-thirds of historical official forecast errors over a 5-year sample fall within the circle (see full explanation at https://www.nhc.noaa.gov/aboutcone.shtml).

Each time a new forecast is entered into the model, information from the NHC products is synthesized into a "light system" forecast of the three major hazards known to drive hurricane evacuation decisions: wind, storm surge[1], and rain. The approach resembles the Meteoalarm web platform (http://www.meteoalarm.eu) where hazard risk are displayed in traffic-light color-coding (green, yellow, orange, red). Reds are reserved for severe and rare events, while also capturing some degree of immanency (i.e., reds are warnings, yellows are watches) (Alfieri et al. 2012). We chose to use this type of light system in the modeling system because it (1) represents a synthesis of the forecast for

[1] Storm surge is defined by the National Oceanic and Atmospheric Association (NOAA) as the abnormal rise in seawater level during a storm, measured as the height of the water above the normal predicted astronomical tide. The surge is caused primarily by a storm's winds pushing water onshore.

public consumption like TV personnel do (Demuth et al. 2009), and (2) provides means to connect forecast products with the model grid where evacuation decisions are made (Figure 1b).

Light system forecasts are created with ArcGIS by overlaying products onto the 10 x 4 model grid. Then, at each grid cell, forecast products are combined and weighted to estimate risk for wind, surge, and rain. Weights are based on current knowledge of the contributions of different factors to these types of hazards (e.g., Rezapour and Baldock 2014; NOAA 2021; team expertise in meteorology and risk perception), combined with an empirical validation that the progression of hazard risks for Irma and Dorian is reasonable. Sensitivity tests on the light system weighting (not shown) indicated that shifts in the weightings of the different factors did not have a significant effect on evacuations. The exact process of combining and weighting information to create light system forecasts is provided in Tables A1–A3.

Figure 2 presents the light system forecasts for Hurricane Irma (2017) at 24 hour intervals. The early NHC forecasts depict the most likely scenario as a landfalling major hurricane near Miami. However, the forecasts shifted westward as the storm approached Florida, with the storm eventually making one mainland U.S. landfall in the Florida Keys and a second in southwest Florida near Naples. The light system captures the gradual westward shift in threats. Moreover, as the storm approaches Florida and track uncertainty decreases (confidence increases), the light system estimates increased risk focused on areas inside the narrowing cone of uncertainty. Because of these features, the light system appears to be a reasonable way of representing the risks associated with hurricane hazards and is good enough to proceed. As a result, FLEE becomes the first

to use synthesized NHC products with ABMs, and alongside Watts et al. (2019) and Morss et al. (2017), contains one of the most sophisticated representations of hurricane forecast information in models of the hurricane evacuation system to date.

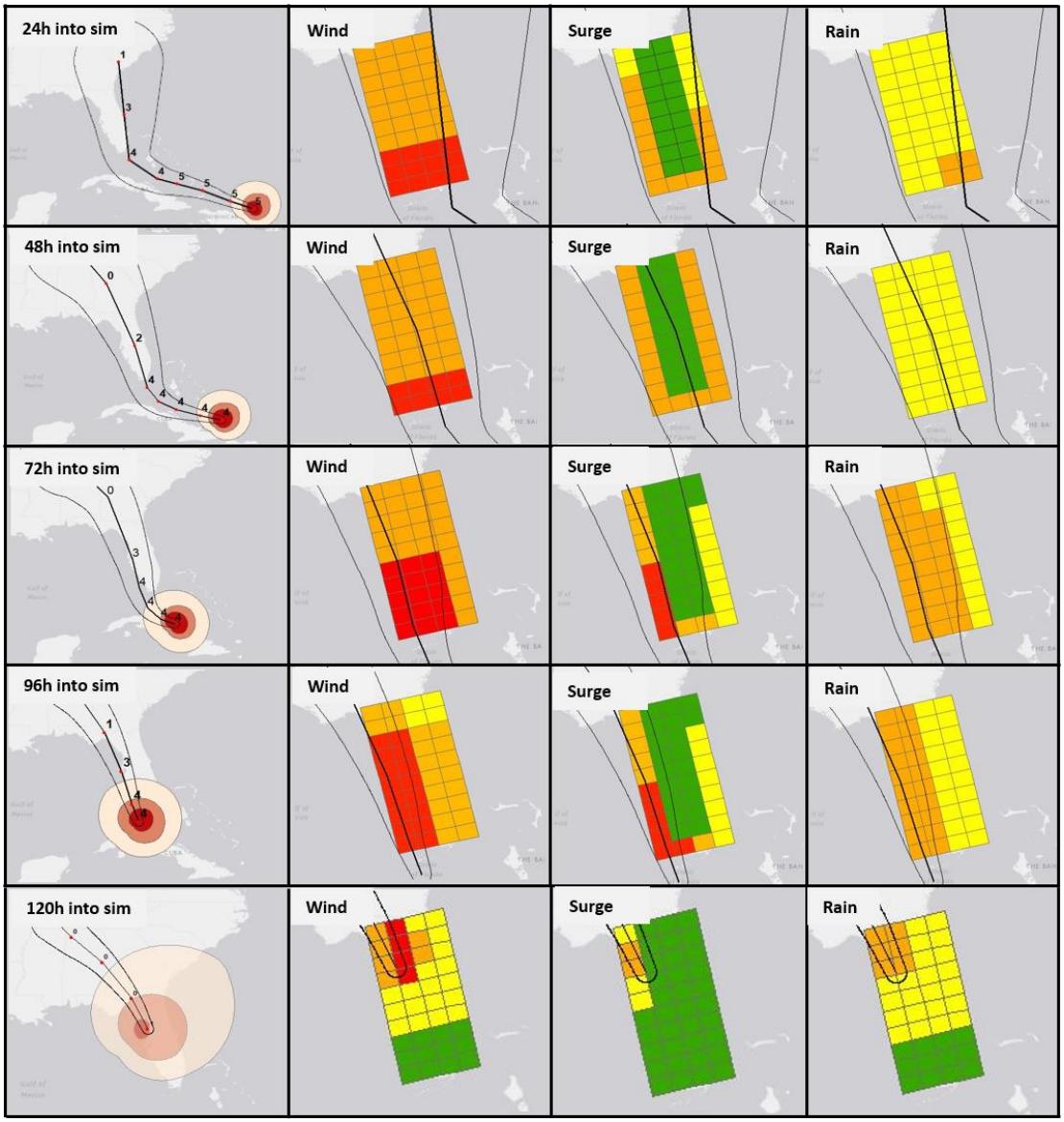

Figure 2: Light system forecasts for Hurricane Irma (2017) as the storm approaches and travels through the Florida-like, model grid. Forecasts are shown at 24 hour intervals, but update every 6 hours in the model simulations (not shown). Left column: Evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. Right three columns: The light-system threats for wind, surge, and rain are shown for equivalent times in the simulation, with the forecast track (center black line) and cone of uncertainty (outer black lines) included for reference. Note: threats are highest when near the center of the forecast cone and when hazards are most imminent, among other factors.

2.3.The human system (information flow, evacuation-related decisions)

With the synthesized light system forecasts as inputs, an ABM simulates the "human system" i.e., information flow and evacuation-related decisions (Figure 1b). This system includes two types of agents: emergency management agents who issue evacuation orders, and household agents (i.e., the public) who collect information, assess risks, and make protective decisions. An overview of the agents and their decision-making algorithms, which run every 30 minutes in FLEE, is described in this section.

As the hurricane approaches the coastline, emergency management agents (EMs) decide whether to issue evacuation orders for each grid cell. The decision-making process is represented schematically in Figure 3 and is based on research by Demuth et al. (2012), Dye et al. (2014), and Bostrom et al. (2016), as well as the analysis in Cutter (2019). Clearance times are subjectively assigned to FLEE's grid cells using data from the Florida Statewide Regional Evacuation Study Program (2019) which accounts for available road networks and the number expected to evacuate per county (based on population density and forecast intensity). For example, high clearance times (40–60 hours) are located in Miami and Tampa Bay for intense (red) surge forecasts; low clearance times (5–20 hours) occur in rural areas upstate with less intense (yellow) surge forecasts. Since surge is not expected inland, only coastal EMs issue evacuation orders in FLEE.

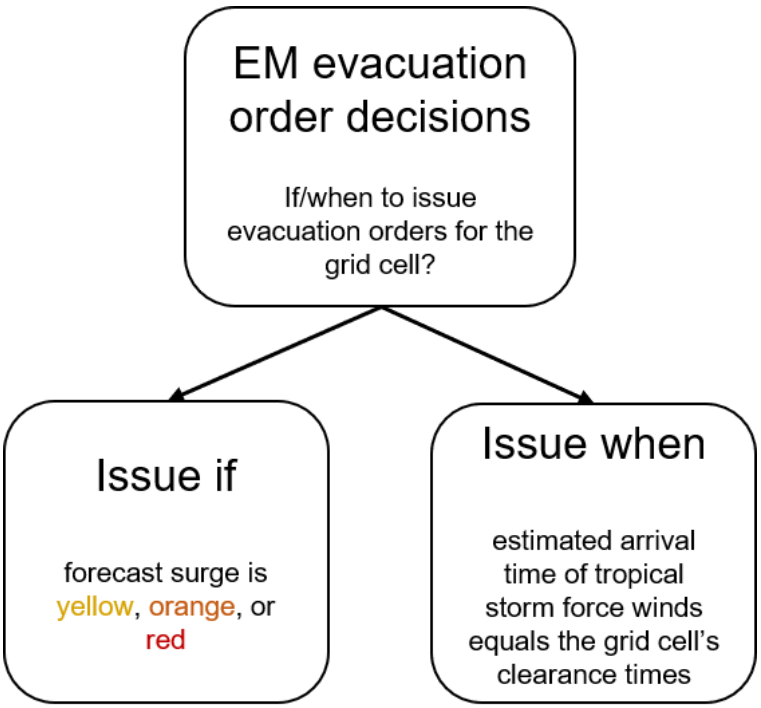

Figure 3: The evacuation order decision-making algorithm prescribed to emergency management agents (EMs) along the coastline. Information used by EMs include the surge light system forecasts (Section 2.2), estimated arrival time of the storm, and clearance times for each grid cell.

The second type of agent, household agents, represent groups of 4 individuals, bringing the number of estimated households in FLEE to 4.1 million (note: the literature suggests people generally make household-based evacuation decisions e.g., summary in Murray-Tuite et al. 2019). This is a simplification to reduce model run-time, as the average household size in Florida is estimated at 2.7. Since the paper's goals are to describe FLEE and demonstrate its capabilities, we believe this assumption is okay, for now. Future experiments building fundamental knowledge of the system dynamics should accurately reflect household size.

Household agents collect information about the hurricane, assess risk posed by the storm, and decide whether the risk warrants evacuation. The design of the evacuation decision-making algorithms prescribed to these agents was adapted from conceptual

models of protective decision-making for hazards, such as the Protective Action Decision Model (PADM; Lindell and Perry, 2012; see hurricane applications in Lazo et al. 2015; Huang et al. 2017; Watts et al. 2019), and findings from empirical research on decision-making for hurricanes (e.g., Baker, 1991; Dow and Cutter, 2000; Dash and Gladwin, 2007; Morss and Hayden, 2010; Bowser and Cutter 2015; Huang et al., 2016, Morss et al., 2016, Demuth et al., 2016; Cuite et al., 2017; Bostrom et al., 2018; Demuth et al., 2018). As noted in Watts et al (2019), a major challenge is to synthesize the conceptual PADM model and information from empirical analyses into simple yet sufficiently specific instructions for agents. For the purposes of our model, we are not seeking a fully realistic algorithm, but one that captures the main processes underlying public evacuation decisions in the context of the modeling system so we can examine the broader evacuation dynamics holistically.

To develop the household decision algorithm, we synthesized the relevant literature which suggests that people generally evacuate when they believe that the hurricane poses a risk to themselves or their family, and that different people perceive risk differently and have different evacuation barriers (e.g., Baker, 1991; Dash and Gladwin, 2007; Lazo et al., 2015). This literature also finds that factors with the strongest, most consistent influence on evacuation decisions include the risks indicated by forecast information and evacuation orders, as well as household characteristics associated with risk perceptions and evacuation barriers (Huang et al. 2016). Thus, we construct the decision-making algorithms by combining time-varying information about the evolving risk (from light system forecasts and EM's evacuation orders) and household characteristics related to perceived and actual risk (age, mobile home residence) to form a risk assessment. This

risk assessment is then compared with evacuation barriers (socioeconomic status, car ownership) which vary across the agent population and the model grid. Undecided agents seek information and update decisions every 30 minutes, making agents active participants in the evacuation decision making process (Watts et al. 2019; Morss et al. 2017; Mileti and Sorensen 1990, Sadri et al. 2017). A high-level schematic of the decision-making algorithm is presented in Figure 4; details regarding the algorithm's variables and formulation is provided in Table A4.

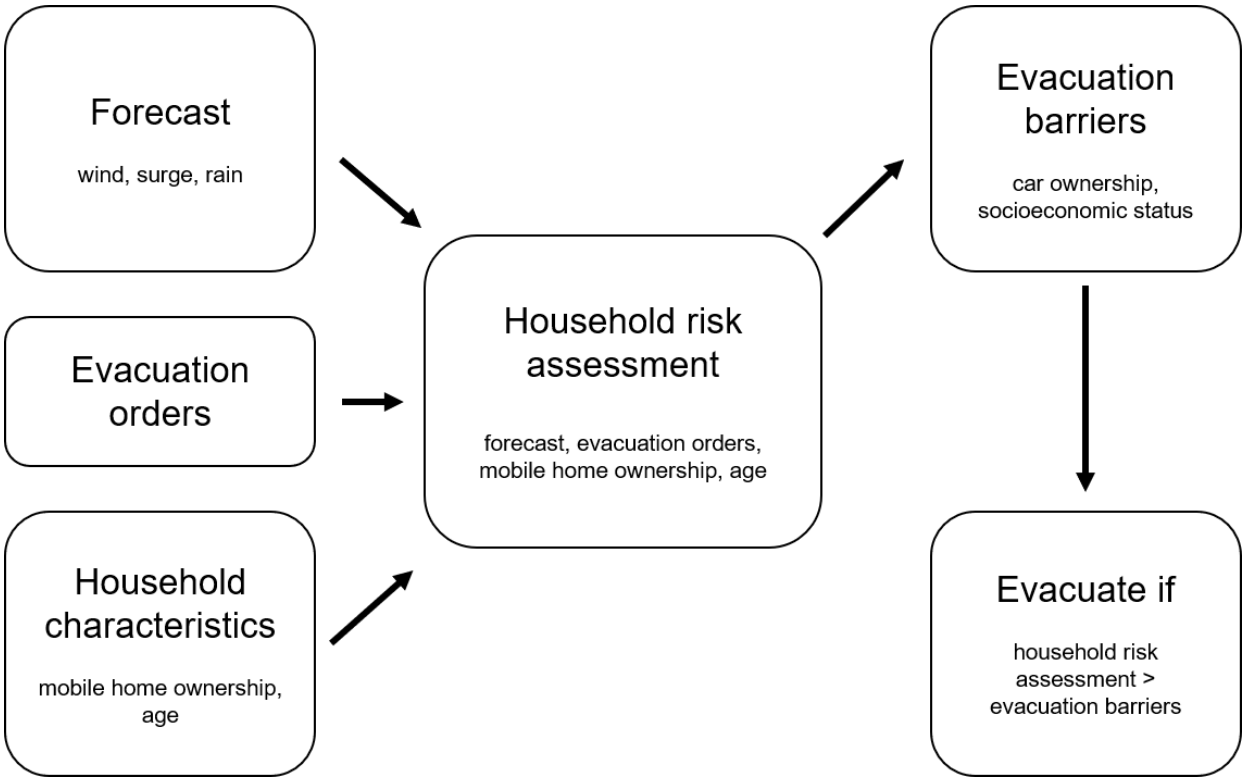

Figure 4: The household evacuation decision-making algorithm in FLEE. Based on the PADM of Lindell and Perry (2012), the process begins when agents combine information obtained from multiple sources (e.g., forecast information, evacuation orders, and household characteristics) into a household risk assessment, which is then compared with evacuation barriers (i.e., socioeconomic barriers, car ownership) that vary across the agent population. A household will evacuate if the household's risk assessment is greater than the household's evacuation barriers.

Agent's household characteristics are prescribed by subjectively projecting county-level census and social vulnerability data regarding mobile home ownership, age, car ownership, and socioeconomic status (which includes poverty rates, unemployment, and income) onto FLEE's model grid (Figure A1; Flanagan et al. 2011). Once the geographical distribution of variables is sorted between cells, specific characteristics are stochastically assigned to individual households (Table A5). The idea is to not perfectly represent the real-world characteristics, but to generally capture its geographical distribution, and have an appropriately wide range of household characteristics within grid cells. This results in many heterogeneous agents with unique preferences and characteristics.

To account for complexities in how people process and value different information, factors influencing a household's risk assessment are weighed differently between households (Table A6). For example, some agents are concerned about evacuation orders while others are not; some are concerned about their mobile home's durability while others are not, and so on. Varying the weights captures these differences. In addition, varying the weights indirectly represents other factors such as culture and worldviews which are sometimes important (Lazrus et al. 2020; Morss et al. 2020). Weight distributions are stochastically generated for each household with specified ranges informed by the literature (e.g., Senkbeil et al. 2019; Bostrom et al. 2018; Petrolia et al. 2011; Meyer et al. 2014; Morss and Hayden 2010; Brommer and Senkbeil 2010; Peacock et al. 2005). The idea is to reflect the relative importance of each factor (e.g., evacuation orders, forecast information, mobile home ownership, and age, in that order) as established in Huang et al. (2016).

One noteworthy simplification of the decision-making algorithm is that households do not share forecast information with other agents. In other words, everyone has the exact same forecast and evacuation order information i.e., it is a world with perfect, instantaneous communication of updated forecast information. Another is that they do not consider social cues, such as seeing other people evacuate, which can increase one's risk perception. We also do not consider previous experience of disasters, social-media influence, or the structural integrity of buildings, which can influence people's risk assessments and behaviors (e.g., Dash and Gladwin, 2007; Lindell and Perry, 2012; Demuth et al., 2018). Again, the idea is to capture the main processes underlying public evacuation decisions so we can examine the hurricane-forecast-evacuation system dynamics holistically. Such features could be added in future model versions, depending on the intended research goals.

2.4 The built environment (infrastructure, evacuation routing, and traffic)

If a household decides to evacuate, they enter another ABM – this time representing evacuation traffic – which moves the household across an idealized road network toward a (presumably) safer location (Figure 1c). An overview of this traffic model, its vehicle agents, and the idealized road infrastructure is described in this section.

FLEE's idealized road network, and its relationship with the 10 x 4 model grid, is depicted in Figure 5. The built environment consists of two five-lane interstates (blue arrows) situated on the edges of the model grid. These interstates, representing Florida's I-75 and I-95, transport evacuees northward along FLEE's "coasts." Additionally, two east-west running, three-lane interstates (purple arrows), representing Florida's I-75 and I-4, allow

residents to move horizontally across the grid. For example, these interstates let households move from Miami (yellow star) towards Tampa Bay (blue star) or inland towards Orlando (orange star). Lastly, eight, two-lane highways (red arrows) allow inland residents access to the interstates where they can flee northward/inland to safety. Though idealized, FLEE's built infrastructure is designed to capture the main elements of Florida's real world road network that influence large-scale evacuation dynamics. However, future models could add complex road structures, such as including local and intra-city road networks, if desired.

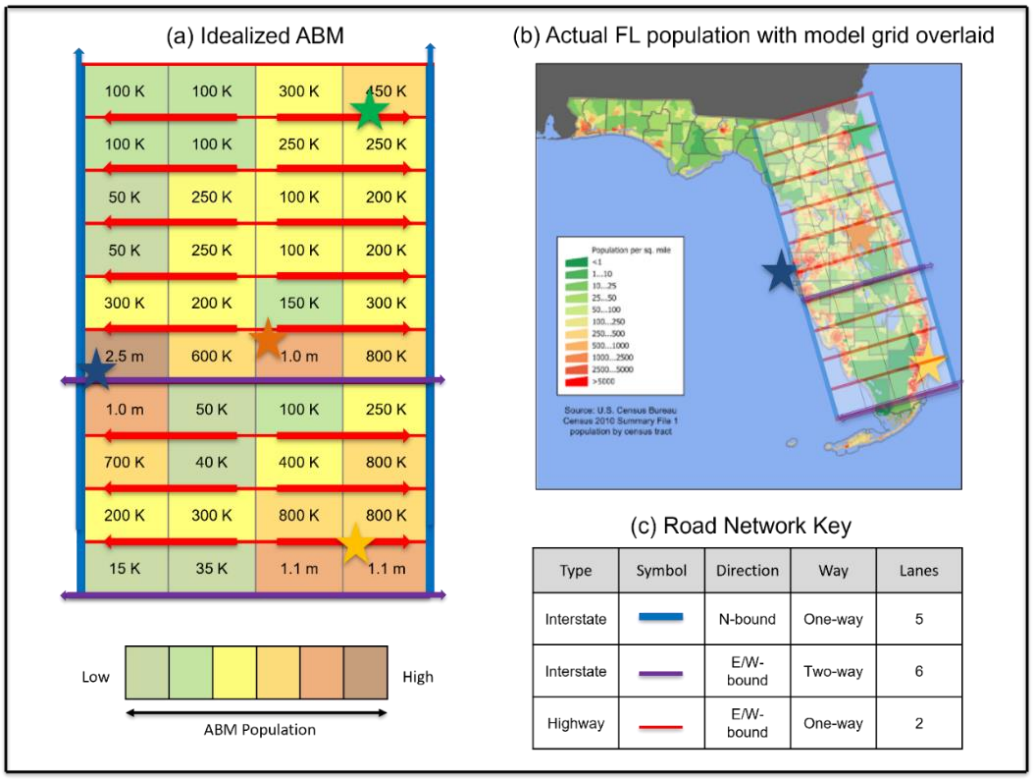

Figure 5: The idealized road network and population distribution on the model grid. Agents inside the idealized grid (a) are subjectively populated and characterized based on 2019 census data (color filled cells). Note there are 16,390,000 agents total, which equates to 4,097,500 households/vehicles, and that grid cell dimensions are 69 x 69 km each. Major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). The grid cell corresponding to Tampa Bay (blue star) contains the most evacuees of any grid cell at 2.5 million. Cities are also depicted in Florida's actual population map (b), with the semi-transparent, 10 x 4 model grid overlaid, for reference. The available road network (e.g., road type, direction, number of lanes) is shown (left) with supporting table (c). Agents are generally instructed to flee onto the primary interstates (blue) and then northward (arrows) to areas of lower risk.

Evacuating households are instructed to depart within twelve hours of the evacuation decision (Huang et al. 2012, Lindell et al. 2005, Murray-Tuite et al. 2019). Departure times are generated stochastically within this twelve hour timeframe. When it's time to depart, households are assigned a vehicle and look for spots on the nearest highway (Figure 5; red and purple lines). Specifically, households search for any unoccupied spot along the 69 km stretch of highway corresponding to their home grid cell. If an open spot exists, they are immediately placed in this spot. If spots are unavailable due to traffic for a period of time, evacuees can lose patience, abandon the evacuation and shelter in-place instead (this process is detailed in Table A7). In this way, the amount of evacuation traffic influence evacuation decision-making for households.

In regard to destinations, nearly half of the evacuees are randomly selected to evacuate out-of-state (e.g., based on Wong et al. 2018; Murray-Tuite et al. 2019). For the remaining in-state evacuees, evacuation destinations are chosen based on where the forecast hazard risk is lower (e.g., from red to green) and where accommodations are available, which is typically in more populated areas (Murray-Tuite et al. 2019). In the case of Hurricane Irma, in-state evacuees typically moved upstate (e.g., towards Tampa Bay, Jacksonville) and inland (e.g., towards Orlando). Carless households move to local shelters, meaning they do enter the road networks and influence traffic (Wong et al. 2018). Regarding route selection, we simplify the complex process by assigning agents the shortest route (Sadri et al. 2014). Once assigned, evacuee routes do not change. The amount of time required to reach destinations is not considered, though this could be added in future models.

For those who enter the road, rules governing vehicle movement are simple: drivers accelerate when they can, slow down if they must, and do not accelerate at the speed limit (70 mph on interstates, 50 mph on roads) or behind another car. Lane switching is not permitted but could be added in future models. Some drivers exhibit erratic behaviors by randomly braking, potentially leading to traffic jams. Accidents are stochastically generated, with a frequency based on Robinson et al. (2009). Default settings for these parameters are described in Table A7.

An example of FLEE's evacuation traffic is shown in Figure 6. The traffic model, which has a 1.2 second timestep, captures interactions between vehicles at micro-scales, e.g., over-reactive and/or erratic drivers cause other drivers to slow down, triggering realistic-looking traffic jams (Figure 6; blue streaks). These interactions are important for investigating complex system dynamics such as traffic (Miller and Page 2007; Barton 2014). Congestion and slowdowns – similar to what is shown in Figure 6 – occur at intersections, in densely populated regions, surrounding accidents, or when vehicles run out of gas. In Section 4.1, we show that, before Hurricane Irma, severe traffic occurs along I-75 and I-95 northbound due to Miami and Tampa Bay being in the storm's path. During Irma's actual evacuation, severe traffic was also observed in these areas (e.g., Zhu et al., 2020; Cava 2018; Wong et al. 2018). Because the traffic model captures important vehicle interactions at microscales, and generates reasonable traffic phenomena at regional scales, we believe FLEE's built environment represents evacuation traffic sufficiently well to examine the hurricane-forecast-evacuation system dynamics holistically.

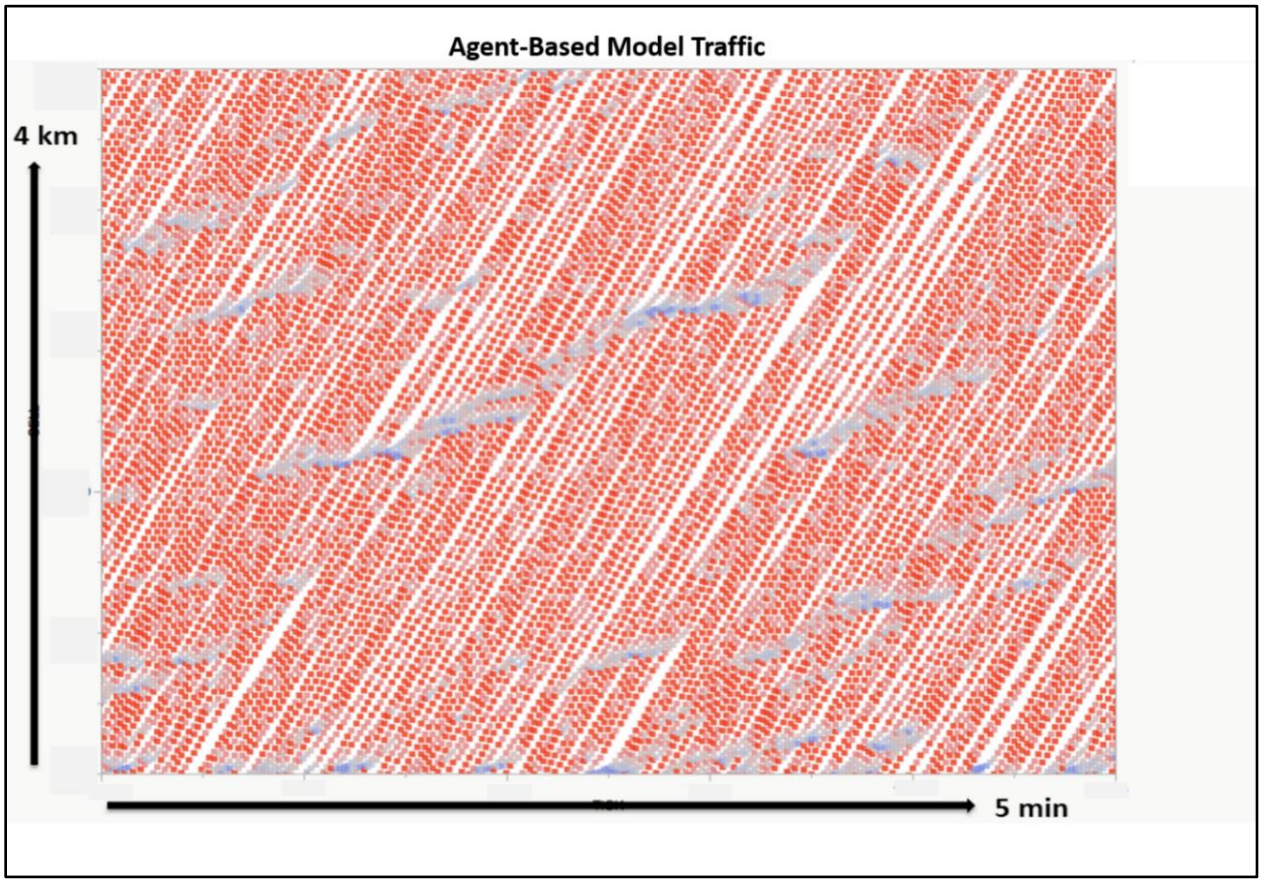

Figure 6: Sample of evacuation traffic generated by FLEE's built environment during Hurricane Irma (2017). Specifically, we show one lane of a zoomed-in, 4-km segment of I-95 (y-axis). Vehicles (dots) move along the interstate segment over a 5-minute period (x-axis) (i.e., vehicles move from bottom-left to top-right). Colors depict vehicle speed – full speed traffic moves unobstructed (red dots), while erratic drivers cause vehicles to slow down (blue dots) or stop altogether (dark blue).

3. Experimental methods and data analysis

3.1 Model validation

There are no governing equations to model human behavior. Therefore a thorough understanding of the FLEE's behavior – and a validation the behavior is realistic as possible – must be achieved. This was accomplished in several ways. First, the modeling framework was tested throughout implementation to ensure the model code is error-free. This includes conducting sensitivity analyses on FLEE i.e., components were perturbed,

one-by-one, to check if it behaves reasonably (e.g., sensitivity tests on light system weights described in Section 2.2). Second, the model framework was calibrated against existing observational data, namely for Hurricane Irma (e.g., Wong et al. 2018; FDEM 2017; Long et al. 2020; Feng and Lin 2021). These empirical studies provide an overview of Irma's evacuation behaviors, including the total number of evacuees, how Irma's evacuation rates change with time and vary spatially, and when/where significant traffic occurred. Throughout Section 4, we compare FLEE's default evacuation behaviors to these observations in an effort to validate the model framework, and in turn, demonstrate that FLEE portrays key aspects of real-world evacuation dynamics sufficiently well to be suitable for experimentation.

3.2 Experimental Design

Table 3 provides an overview of the different experiments reported in this article. The first experiment (Table 3a) uses the default model parameters described in Section 2.1–2.4 for Hurricane Irma. It provides a baseline of evacuation behaviors which are compared to existing observational data for validation. Based on this default simulation, we then systematically modify model parameters one-by-one, while holding other variables constant, to explore FLEE's behaviors and sensitivities. These experiments include varying the evacuation order timing (Table 3b), implementing contraflow (Table 3c), and changing the storm to Hurricane Dorian (Table 3d). Additional experiments changing the evacuation decision-making inputs (Table 3e) and the population density (Table 3f) are included in Appendix A. Together, these proof-of-concept experiments are intended to demonstrate how FLEE can serve as a virtual laboratory uniquely positioned to advance our understanding of the hurricane-forecast-evacuation system.

Experiment	Storm	Goal	Run details
a) Default	Irma	Establish a baseline of evacuation behaviors (section 4.1) for comparison with experiments b-f and observational data for validation.	1: Inputs described in Sections 2.1-2.4.
b) Varying evacuation order timing	Irma	Examine the influence of changing evacuation order timing by adjusting clearance times at each grid cell (section 4.2)	2: Evacuation orders 10 h earlier 3: Evacuation orders 10 h later 4: Clearance times equal 5: Clearance times equal and reduced by 10 hours
c) Implementing contraflow	Irma	Examine the influence of contraflow on evacuations (section 4.3) by adjusting the number lanes on various highways	6: +1 lane on I-95 7: +1 lane on I-75 8: +1 lane on both I-95/I-75
d) Default	Dorian	To examine how the default parameter values carry over to a new storm scenario (section 4.4)	9: Default inputs (Sections 2.1-2.4) but with Dorian's light system forecasts
e) Evacuation decision-making inputs	Irma	Determine the relative influence of each decision-making input by turning them off, one-by-one (Appendix A)	10: Forecast weight = 0 11: Evacuation order weight = 0 12: Age weight = 0 13: Mobile home weight = 0
f) Varying population density	Irma	Adjust population distribution to examine the influence of population density on evacuations (Appendix A)	14: Uniform population distribution

Table 3: Description of experiments. The main goals are to establish the broader spatial and temporal patterns of evacuation behaviors for Hurricane Irma (2017), then intentionally perturb FLEE's key parameters to assess the relative importance and general response of the factors (b-f). In doing so, we demonstrate how FLEE can investigate the dynamics of the hurricane-forecast-evacuation system. Note: experiments e and f were included in Appendix A.

3.3 Data Analysis

To compare evacuation patterns and behaviors quantitatively across simulations, FLEE tracks evacuation statistics for all grid cells. The primary model output analyzed here are the percent of households that successfully evacuated (i.e., evacuation rates), and the percent who intended to evacuate but "gave up" due to traffic. The latter statistic provides insight to where the excessive traffic may be preventing successful evacuations. In addition to displaying data by grid cells, values are broken down into multiple impact zones, designed as first-order approximations of areas likely to experience different levels of impacts based on the actual meteorological conditions produced by the storm. Here, we use four impact zones, defined by whether the grid cells: a) are coastal or inland, and b) primarily experiences winds that are greater than 64 knots (hurricane-force) or less than 64 knots. Using the impact zones, we can determine who evacuated from locations that did not end up experiencing hazardous conditions. In addition, we examine compliance rates (i.e., the percentage of residents under evacuation orders who evacuated) and shadow evacuation rates (i.e., the percentage of residents who evacuated from areas not under evacuation orders; McGhee and Grimes 2006; Murray-Tuite et al. 2019). Note: evacuation orders are issued for entire grid cells i.e., everyone in that grid cell either gets an evacuation order or not.

In looking at the results, we compare multiple metrics that might indicate successful outcomes in different ways. For example, high compliance rates may not be "good" if the storm ends up not having much impact in those areas, and shadow evacuation rates may not matter if those at highest risk can get out safely.

Because FLEE includes stochastic elements, it can exhibit some run-to-run variability. For example, in a series of tests where simulations were repeated five times, evacuation rates ranged from 0–2% within grid cells. This run-to-run variability is smaller than other agent-based evacuation simulations (e.g., Watts et al. 2019, Chen and Zhan 2008), likely because there are many more agents in this model (nearly 4.1 million households/vehicles). Nevertheless, when interpreting results, changes less than this 0–2% variability within grid cells are considered insignificant.

4. Results

4.1 Spatial and temporal patterns of evacuation

First, we examine results from a simulation with the default FLEE configuration for Hurricane Irma (Table 3a). By comparing these results with observations of Irma's actual evacuation (e.g., in Wong et al. 2018; FDEM 2017; Long et al. 2020; Feng and Lin 2021), they provide a first-order assessment that agents in the model are behaving reasonably based on the processes implemented. They also illustrate key aspects of FLEE's behavior, including the spatial and temporal patterns of evacuation, which provide a baseline for interpreting results from subsequent experiments (sections 4.2–4.4; Appendix A, supplementary results 1–2).

Based on the default model settings for Irma, EM agents issue evacuation orders in a similar pattern to what was observed (Figure 7; red cells). Evacuation orders were first issued around Miami-Ft. Lauderdale 36–48 hours into the simulation (Figure 7b), and spread northward along both coastlines over the next several days (Figure 7c–e). The last evacuation orders were issued in Jacksonville 120 hours into the simulation, which

coincides with the time Irma makes landfall along the southwest Florida coast (Figure 7e). By the end of the simulation, Irma's hurricane-force winds (Figure 7f–g; dotted cells) impacted the western two-thirds of the model – particularly the southwest and western coastlines – while leaving the east-coast generally unscathed. This general progression of evacuation orders being issued from south-to-north along both coasts matches what occurred with Irma (e.g., see Page 14-15 and Figure 2 of Wong et al. 2018 for evacuation orders by county). This increases our confidence that the EM decision-making algorithm – and the storm surge forecasts on which its based – behaves reasonably and realistically.

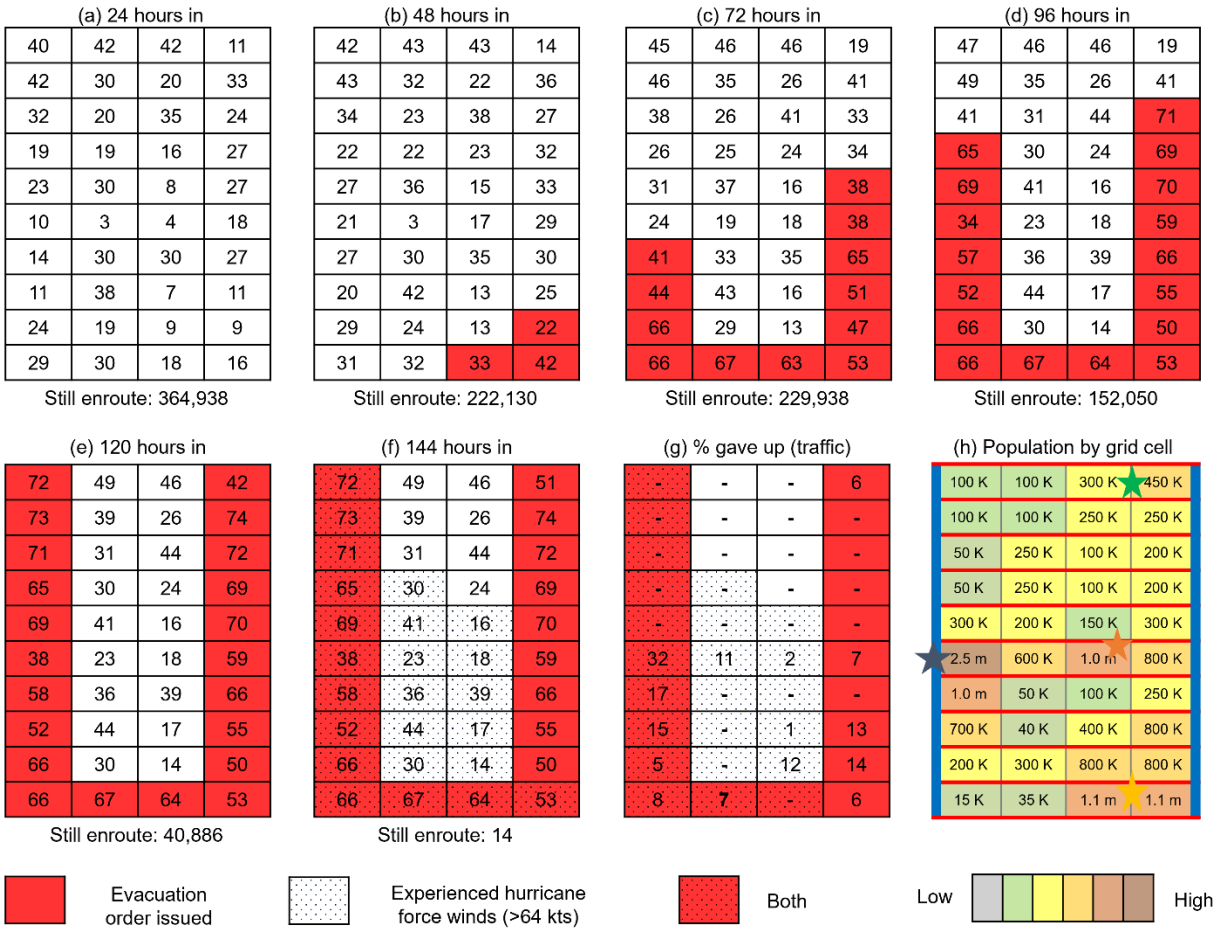

Figure 7: Evacuation rates in Irma's default model run. Rates are presented every 24 hours throughout the 144 hour simulation (a-f) for each grid cell. The percentage which intended to evacuate but could not due to excessive traffic is also expressed (g), as are the spatial and temporal patterns of evacuation orders (red cells) and the swath of hurricane force winds experienced (dotted cells). Also shown are the number of evacuees still enroute (bottom of panels a-f) and the population by grid cell (h). Major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). These provide a frame of reference for the evacuation rates in a-f.

The percentage of households who evacuate is shown at 24 hour intervals for each grid cell (Figure 7 a–f). The results depict spatial and temporal patterns that are similar to real hurricane evacuation behaviors. First, evacuation rates increase after evacuation orders are issued, showing its importance to decision-making (Huang et al. 2016). Secondly and relatedly, evacuation rates are higher along the coasts than inland (Baker 1991). Thirdly, evacuation rates are still high for most areas. This arises because the forecasts in this simulation were dire everywhere, especially before the storm's track shifted westward

(Figure 2). The dire forecasts prompted EMs to issue evacuation orders along both coasts, and as a result, many agents evacuated areas which did not experience hurricane force winds.

FLEE's simulated evacuation rates generally match existing observational data for Irma, which suggest evacuation rates vary from 40–60% along Florida's east coast, to 60-80% across the south and west coasts, and around 5-40% inland (e.g., see breakdown of evacuation rates by region in Figure 4 of Wong et al. 2018; breakdown by voting precinct in Figure 1c of Long et al. 2020). One area for improvement is that FLEE produces evacuation rates higher than realistic early in the simulation, especially in the northern part of Florida (Long et al. 2020).

Table 4 depicts evacuation rates in different impact zones. In total, 45.1% of households on the model grid evacuate, which equals 7.38 million people. Note, estimates from the Florida Department of Emergency Management (2017) suggest actual evacuation numbers totaled 6.9 million. For a given level of wind impact, evacuation rates are higher along the coasts than inland (52.3% coastal vs. 22.2% inland for >64 knots, 58.1% coastal vs. 36.7% inland for <64 knots). Interestingly, areas experiencing hurricane force winds had lower evacuation rates than areas less affected by the storm. This could be due to the potentially higher than realistic evacuation rates early in the simulation. They may also partially result from the east coast receiving evacuation orders, albeit unnecessarily, which increased evacuation rates in these areas, combined with excessive traffic along the west coast. For example, 17 to 32% of the populated Tampa Bay-St. Petersburg gave up evacuating due to excessive traffic (Figure 7g). The severe congestion, which did occur with Irma's actual evacuation, also reduced evacuation rates along the southwest

and southeast coasts (e.g., see traffic information in Page 15 of Wong et al. 2018; FDEM 2017; Feng and Lin 2021).

	% Successfully evacuated							
Experiment	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 knot zone	Inland < 64 knot zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5

Table 4: Evacuation rates by impact zones for Irma's default run. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

A second pattern illustrated by the evacuation rates is the variability in evacuation decisions among households i.e., some households decide to leave, but many do not, despite seeing similar information and having similar characteristics. This is consistent with real-world hurricane evacuations, and more generally with the heterogeneity exhibited by US households in the real-world (e.g., Hasan et al., 2011; Dixon et al., 2017). In the model, the variability arises from household's different weighting of information as well as their different characteristics and barriers, which create differences in household risk perception.

Figure 8 illustrates the temporal evacuation patterns. Despite not receiving evacuation orders, many households (black dotted line) evacuate in the first 0–36 hours. Evacuation rates increase linearly between 36–108 hours as evacuation orders expand along the coasts. Just before the storm moves ashore around 126 hours, evacuation rates

decrease, while the number of households giving up due to excessive traffic (black dashed line) increase. The latter occurs as household agents' patience is influenced by the forecast arrival time of the storm. In other words, agents see the impending landfall, then decide to abandon the evacuation and stay home. These temporal patterns of evacuations, as with the spatial patterns, generally match existing empirical data, which suggests that evacuation rates increased semi-linearly throughout this period (e.g., see Figure 6 of Wong et al. 2018; Figure 2c of Long et al. 2020). As a result, we believe FLEE's simulated evacuations provide a realistic baseline for interpreting results from subsequent experiments (sections 4.2–4.4).

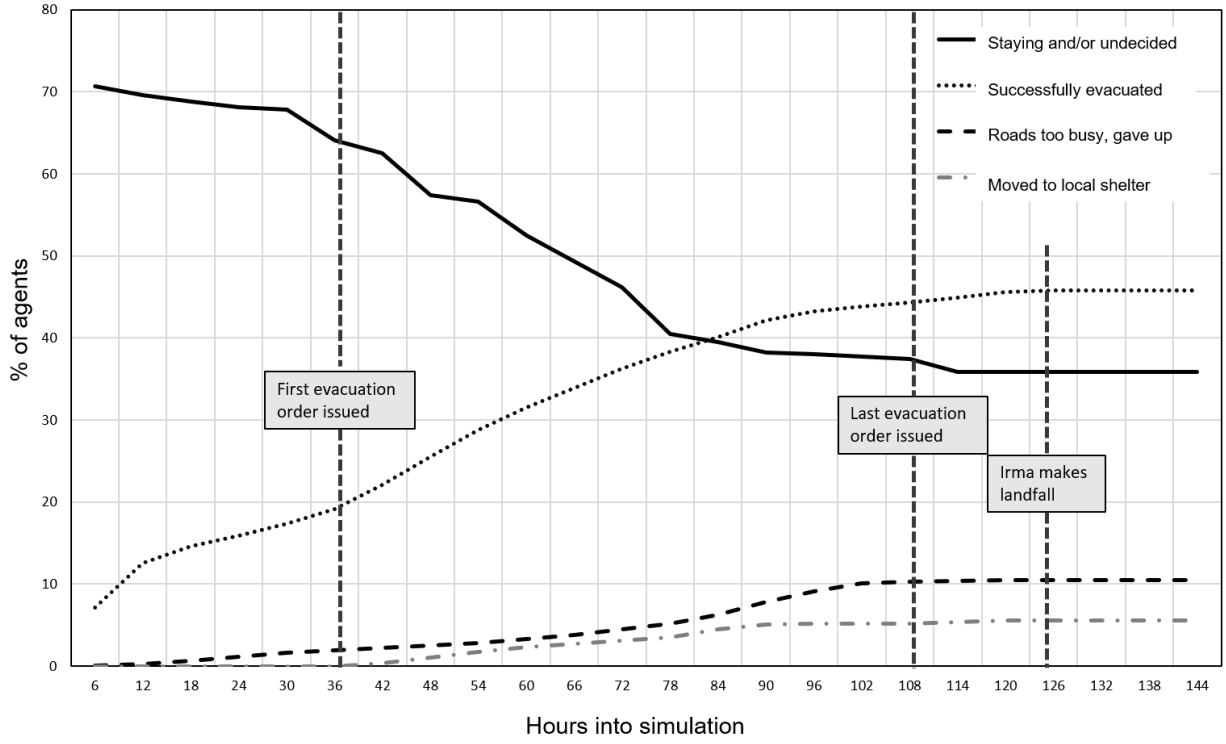

Figure 8: The temporal patterns in evacuation for Irma's default simulation. Successful evacuation rates are shown (black dotted line), averaged across all grid cells, as are the percentage of households giving up due to traffic (dashed line), the percent staying and/or undecided (solid black line), and the percentage of households moving to a local shelter (grey dot dashed line). The latter do not officially enter the road network. Key times in the evacuation simulation, such as evacuation order issuance and storm's landfall, are indicated by the vertical dotted lines. The results illustrate key aspects of the model's behavior and provide a starting point for interpreting results from subsequent experiments (sections 4.2–4.4).

4.2 Varying timing of evacuation orders

Now we investigate the effects of changing the evacuation order timing in FLEE (Table 3b). Specifically, we conduct four experiments: 1) shifting evacuation orders 10 hours earlier, 2) shifting evacuation orders 10 hours later, 3) equalizing the clearance times for all grid cells, making the storm's forecasted arrival time the only factor influencing differences in evacuation order timing across grid cells, and 4) shifting evacuation orders 10 hours earlier than in experiment 3. These experiments build on the results examined in section 4.1, and begin to explore interactions among the evolving forecasts, evacuation orders, and household evacuation behaviors.

Evacuation rates, broken down by impact zones (Table 5), indicate that changing evacuation order timing in the four experiments reduces the overall evacuation rates from 45.1% in Irma's default simulation (top row) to 43.3 – 44.6%, which is 295,200 – 82,000 less evacuees. Similarly, rates of evacuees giving up to traffic increases from 10.5% in the default simulation to 10.9 – 13.0%, which is 65,600 and 410,00 more people. This is surprising, as one might expect evacuation rates to increase if evacuation orders are issued earlier, as this creates more time to evacuate.

Experiment	% Successfully evacuated							
	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 knot zone	Inland < 64 knot zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5
EO +10h	44.6	51.5	20.9	56.2	36.9	53.6	24.7	11.8
EO -10h	43.9	51.7	23.7	51.5	36.8	49.9	30.5	10.9
CTs equal	43.6	51.7	22.4	51.6	37.0	51.6	25.9	12.8
CTs equal, reduced 10h	43.3	50.4	22.8	51.7	36.8	51.0	26.1	13.0

Table 5: Experiments varying evacuation order EO timing. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic. Note: EO is short for evacuation orders; CT is short for clearance times. Irma's default model run is included for reference.

When examining the results for every grid cell (Figure 9), results indicate that, despite only affecting evacuation rates by 1–2% overall, changing the evacuation order timing has significant and sometimes opposite effects between neighboring areas. For example, shifting evacuation orders 10 hours earlier (Figure 9a) increases evacuation rates (and decreases traffic) in Tampa Bay–St. Petersburg by 4%, while decreasing evacuation rates (and increase traffic) from 2% to 16% in neighboring cells to the south. This points to the importance of coordination amongst EMs for issuing evacuation orders within a region and a need for follow-up experiments to unpack these complex processes.

Shifting evacuation orders 10 hours later (Figure 9b) across all grid cells results in evacuation orders not being issued in the Jacksonville metropolitan area. This is because,

during the additional 10 hours where EMs are deciding whether to issue evacuation orders, the forecast shifted westward and away from Jacksonville (Figure 2), thus prompting EMs to decide against issuing evacuation orders for the area. These results demonstrate how the model captures the real-world tradeoffs between issuing evacuation orders earlier (when the uncertainty is greater) versus waiting until closer to the storm's arrival (when the forecast uncertainty is reduced).

Figure 9c and Figure 9d show results from experiments where clearance times are equalized. Recall that clearance times is meant to account for differences in available road networks and the number expected to evacuate e.g., clearance times are highest in populated metropolitan areas and in south Florida where people travel longer distances to evacuate. Thus, equalizing the clearance times, which makes the storm's arrival time the only influence on evacuation order timing, is meant to demonstrate the importance of clearance times in EM decisions. The experiments produce a slight increase in evacuation rates for Tampa Bay–St. Petersburg (1-4%) but with a general decrease in evacuation rates everywhere else. This is especially true in Miami, where evacuation rates drop by 10 to 18%. In this experiment overall, removing the default clearance times worsened hurricane evacuations by 1–2% in total, which is a decrease of 164,000–328,000 evacuees (Table 5). This demonstrates how evacuations can be made more successful by accounting for clearance times in EM's evacuation order decision-making.

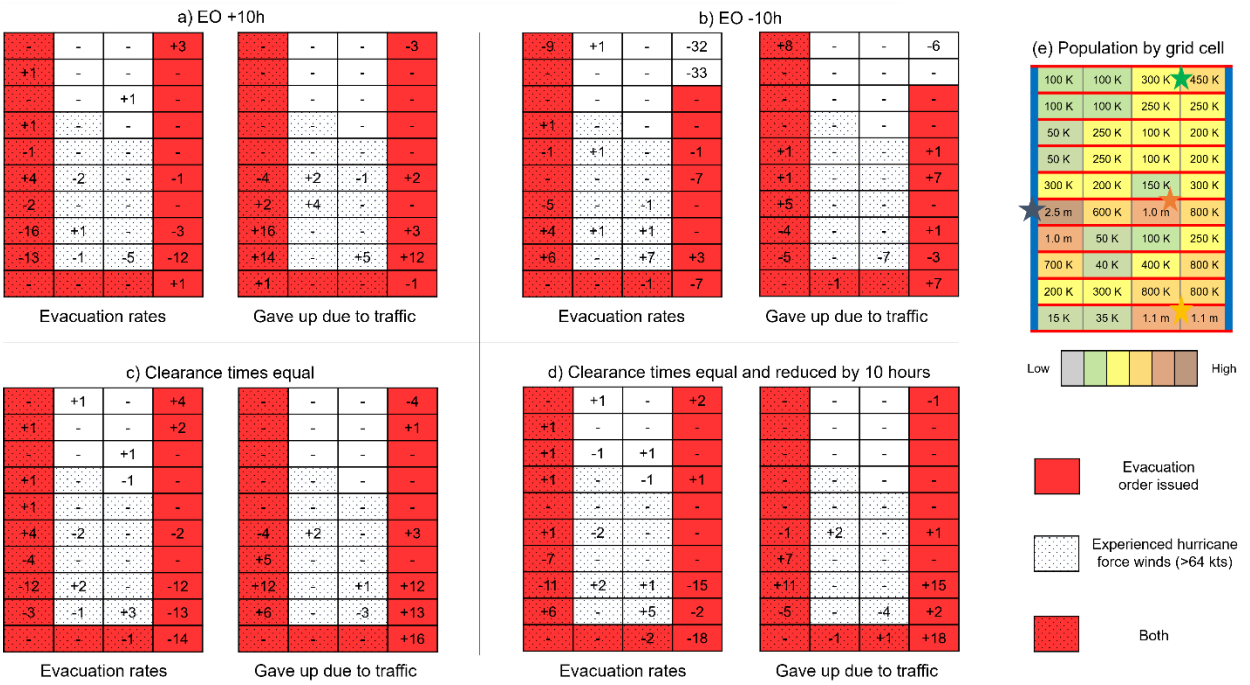

Figure 9: Effects of EO timing on evacuations across grid cells. Results are presented for the experiments modifying the timing of evacuation orders, specifically by a) shifting evacuation orders 10 h earlier than default, b) shifting evacuation orders 10 h later than default, c) equalizing the clearance times, making the storm's arrival time the only influence on evacuation order timing, causing evacuation orders to be issued linearly from south to north as the storm approaches, and d) reducing clearance times by 10 hours than in experiment c. Values are expressed as the departure from the default settings in section 4.1 and in Figure 7f. Also expressed is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (e). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). Note, run-to-run variability due to stochastic elements in the model ranges from 0–2% in grid cells for both evacuation rates and percent giving up due to traffic. Therefore values of -2 to 2 lie within that variability and should be ignored.

Figure 10 shows the evolution of evacuation rates (and rates giving up due to traffic) with time for the different experiments. Shifting evacuation orders 10 hours earlier (green lines) than default (black lines) simply causes evacuation rates to increase earlier in the simulation, and does not meaningfully change the evacuation "shape" otherwise. Similar effects are observed with the uniform clearance time experiments (orange/red lines). This information suggests the model behaves as expected, and in general, the experiments demonstrate how the model can quantify and explore, in a simplified context, the effects of varying evacuation order decisions by EMs. This includes simulating the tradeoffs

between waiting on evacuation orders and its effect on evacuation success, which cannot be quantified using empirical methods. In addition, the results suggest the modeling system is capable of exploring the effects of evacuation strategies such as phased evacuations, which may be helpful to emergency management (Chiu et al. 2008; Chen and Zhan 2004; Zhang et al. 2014).

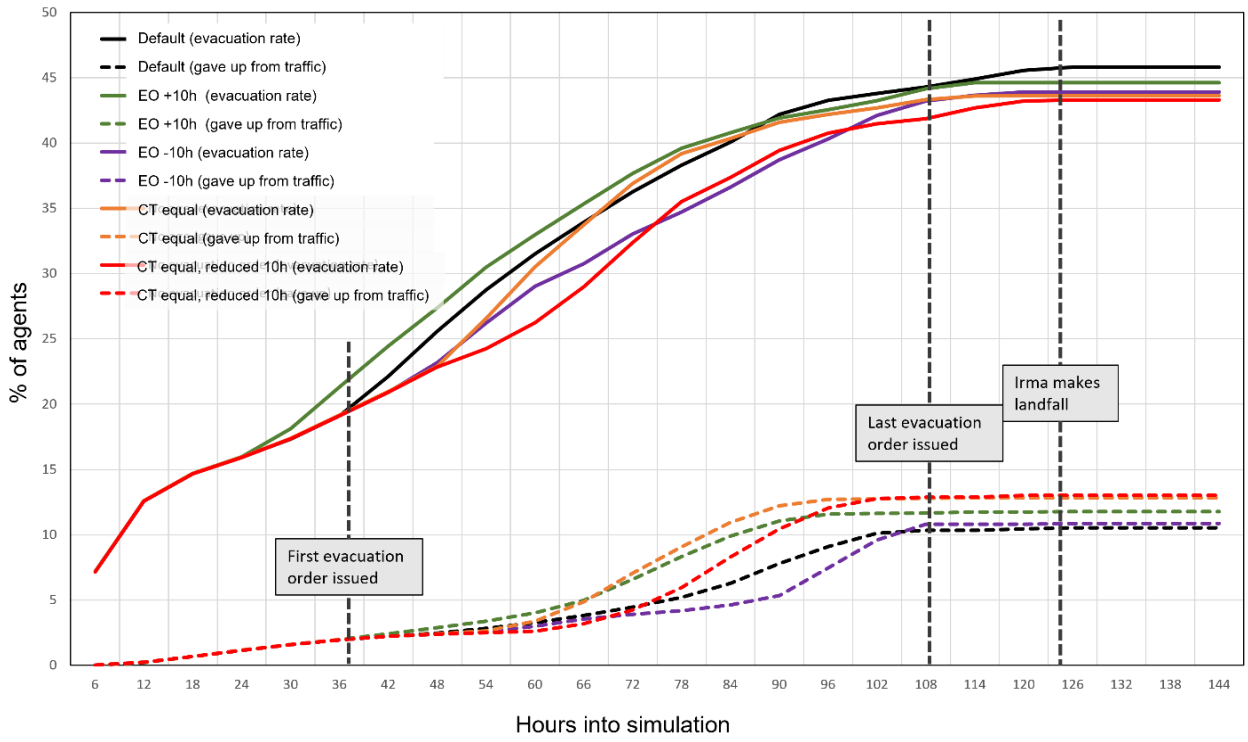

Figure 10: Temporal effects of changing evacuation order timing on evacuation rates (solid lines) and numbers giving up due to traffic (dashed lines), averaged across all grid cells, throughout the 144 hour simulation. The default simulation (Table 3a; section 4.1) is expressed (black lines), as are experiments modifying the timing of evacuation orders, specifically by a) shifting evacuation orders 10 h earlier than default (green lines), b) shifting evacuation orders 10 h later than default (purple lines), c) equalizing the clearance times, making the storm's arrival time the only influence on evacuation order timing, causing evacuation orders to be issued linearly from south to north, and (orange lines) d) shifting evacuation orders 10 hours earlier than in experiment c (red lines).

4.3 Implementing interstate contraflow

Next, we investigate the effects of adding contraflow to lessen evacuation traffic and improve evacuation rates in FLEE. For the experiments, we add one contraflow lane on

I-95, one contraflow lane on I-75, and one contraflow lane on both interstates (Figure 3c).

The results in Table 6 suggest adding contraflow lanes does improve evacuation rates and reduces traffic overall. For example, evacuation rates improve from 45.1% in the default simulation (top row) to 48.0, 47.6, and 49.8% when adding contraflow onto I-95, I-75, and both interstates, respectively. This equates to an increase of 475,600, 410,000, and 770,800 evacuees. Meanwhile, rates giving up from traffic decrease from 10.5% to 6.6–8.3%, which is a decrease of 639,700—360,800 people. The improvements in evacuation rates – and reduction in traffic – are not limited to particular times in the simulation; rather the improvements are uniform throughout (Figure 11).

Experiment	% Successfully evacuated							
	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 knot zone	Inland < 64 knot zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5
+1 I-95	48.0	52.3	25.3	62.6	36.9	57.0	28.0	8.3
+1 I-75	47.6	56.8	22.4	58.0	37.0	57.3	25.9	8.8
+1 I-95, I-75	49.8	56.8	25.7	62.5	36.9	59.3	28.4	6.6

Table 6: Evacuations by impact zone when implementing contraflow. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

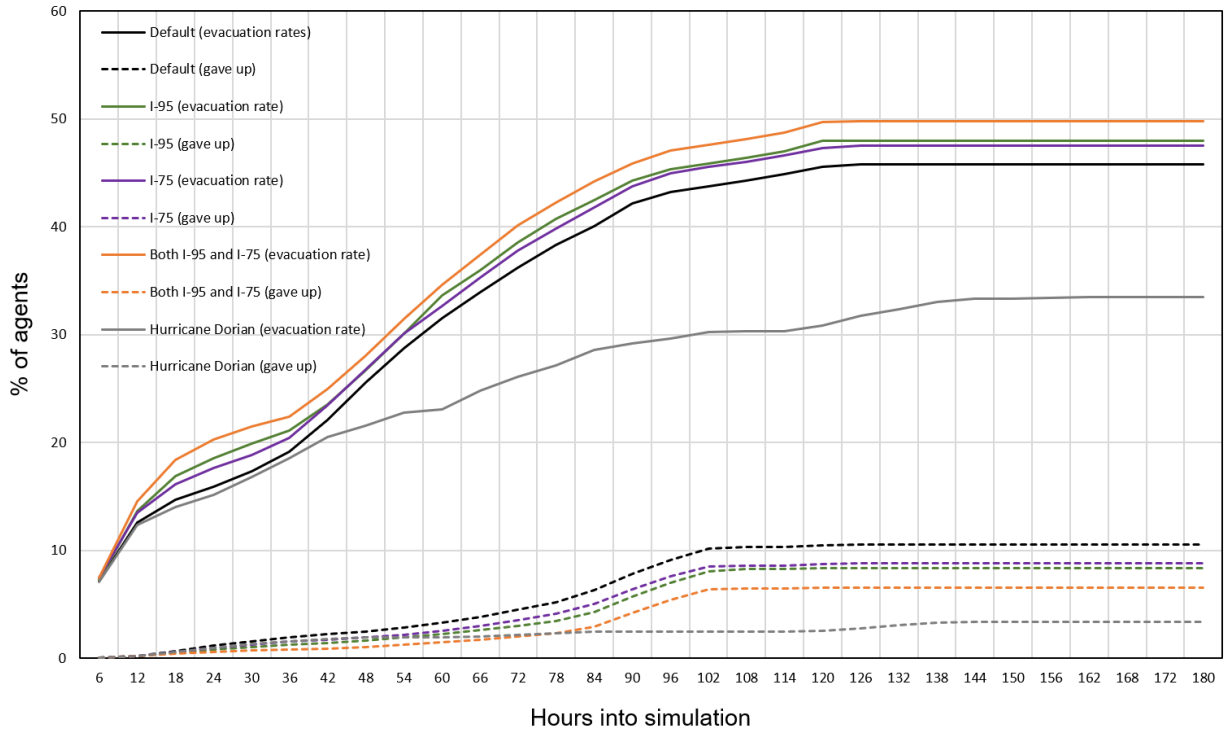

Figure 11: Temporal effects of implementing contraflow on evacuation rates (solid lines) and numbers giving up due to traffic (dashed lines), averaged across grid cells, throughout the 144 hour simulation. The default simulation is expressed (black lines), as are experiments adding one lane of contraflow onto I-95 (green lines), I-75 (purple lines), both I-95 and I-75 (orange lines). The default run for Hurricane Dorian (Table 3d; Section 4.4) is also expressed (grey lines). Note, Dorian's simulation extends to 184 hours while Irma's ends after 144 hours.

When comparing the impact of the different experiments on various grid cells (Figure 12a-d), the targeted effect of contraflow becomes clear. For example, adding contraflow onto I-95, which is located along the eastern coastline, improves evacuation rates (and reduces traffic) along the eastern half of the model grid. Adding contraflow onto I-75, which is found along the western coastline, improves evacuation rates (and reduces traffic) along the western half of the model grid. These improvements in evacuation rates are large locally, ranging from 3–14% along the southwest coast and 5–12% along the southeast coast.

The results suggests that, if given accurate forecasts, implementing contraflow in the modeling system reduces traffic and thus increases successful evacuation in targeted regions, which is what contraflow is designed to do. This provides evidence that the model can be used to investigate the potential impacts of modifying different parts of the system, such as implementing contraflow or other evacuation management strategies, and determine its influence on the hurricane evacuation in its full context (e.g., supporting studies by Zhang et al. 2014; Dixit and Radwan 2009; Chen 2012; Sbayti and Mahmassani 2006; Mitchell and Radwan 2006; Chen and Zhan 2004, 2008; Ballard and Borchardt 2006; Wolshon and Lambert 2004; Fang and Edara 2014; Chiu et al. 2008; Wolshon 2001; Yi et al. 2017).

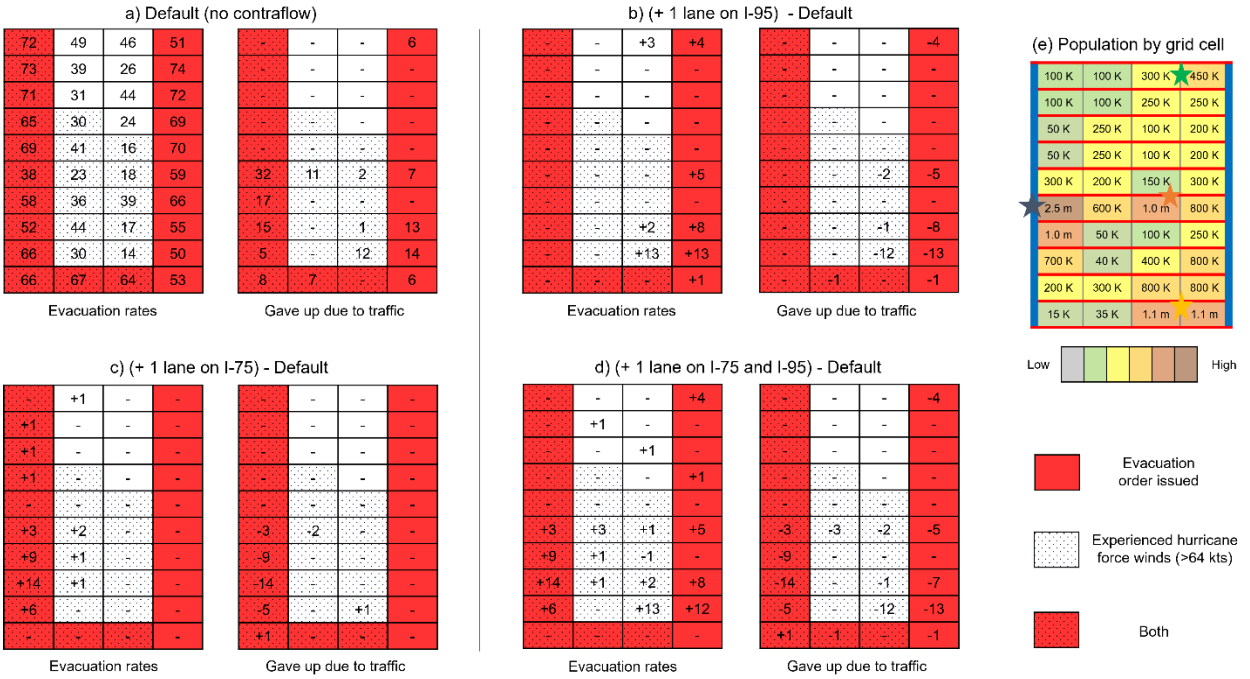

Figure 12: Influence of contraflow for all grid cells. Evacuation rates (left) and the percent of households unable to evacuate due to traffic (right) are shown. Results are presented for the default experiment without contraflow (a), when adding one lane of contraflow to I-95 (b), when adding one lane of contraflow on I-75 (c), and when adding one lane of contraflow onto both interstates (d). These results (b-d) are compared to the default simulation (a) where values are expressed as the percent difference from the default settings (a). Also expressed is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (e). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). Note, run-to-run variability due to stochastic elements in the model ranges from 0–2% in grid cells for both evacuation rates and percent giving up due to traffic. Therefore values of -2 to 2 lie within that variability and should be ignored.

4.4 Hurricane Dorian

Finally, we explore the modeling system's behavior when a different scenario, with a different storm and a different set of evolving forecasts, is simulated. This experiment (Table 3d), with Hurricane Dorian (2019), uses the same set of parameters as in the default Irma simulations. This experiment should be of interest to meteorologists and emergency managers, by exploring how differences in storm characteristics and forecast information can propagate through the agent-based system and translate into different patterns in evacuations.

Figure 13 shows the evolution of Dorian along with the NHC and light system forecasts. The early forecasts (0–72h into the simulation) predict the most likely scenario as a landfalling major hurricane along Florida's east coast. However, the forecasts shift northward (96–120h), significantly reducing areas under threat. After remaining nearly-stationary over the Bahamas (120–144h), the storm re-accelerates northward (>168h) narrowly missing Florida's east coast. As with Irma, the light system captures the spatial and temporal shifts in threats with Dorian. Because of the forecasts, EMs issue evacuation orders along the central east-coast by 72 hours (Figure 14; red cells). The evacuation orders spread along the coastline over the next several days, generally matching what was observed (Roache 2019; Cangialosi 2019).

Compared to Irma, this is a fundamentally different storm with different areas at risk and less people under evacuation orders. As a result, evacuation rates were less with Dorian (33.5%) than with Irma (45.1%), which is 2 million less evacuees (Table 7). Similarly, fewer households give up on evacuating due to traffic with Dorian (6.1%) than Irma (10.5%). This reduction in evacuation rates in FLEE generally matches existing observational data for Dorian (Mongold et al. 2020).

During the first 24–72 hours, evacuation rates are increasing everywhere, as most areas are under threat (Figure 14a-b). As with Irma, we suspect the model is producing evacuation rates higher than realistic during this period, especially in the northern part of Florida and inland. However, this observational data (Mongold et al. 2020) is quite limited and cannot confirm this. Beyond 48 hours, however, evacuation rates only increase along the eastern-most portions of the grid where evacuation orders are issued (Figure 14c-f). By the end, the highest evacuation rates occur in areas where you would expect (i.e.,

along the east coast where risk is highest, and where evacuation orders are issued), which is consistent with real-world evacuation behaviors (e.g., Baker 1991). With the exception of the Tampa-Bay–St. Petersburg area where evacuations occurred early in the simulation, evacuation traffic was primarily confined along the southeast coast (Figure 14h).

The evolution of Dorian's evacuation rates with time, averaged across the model grid, is shown in Figure 11 (grey lines). Similar to Irma's default run (black lines), evacuation rates during Dorian quickly increase due to the dire initial forecasts. Once the forecasts shift northward, Dorian's evacuation rates slows significantly but with some increases due to the issuance of evacuation orders between 60–120h. The evacuation stops by 140 hours because, at this point, the storm is expected to remain offshore. The results again suggest that Dorian's evacuation is, in many respects, different than Irma's.

Robust empirical data on Dorian's evacuation rates is not publicly available. However, the available data (Mongold et al. 2020) suggests the model is, to first order, generating reasonable evacuation behaviors e.g., it captures the inland versus coastal differences in evacuations, the correct issuance of evacuation orders, and the prolonged, linear increases in evacuation rates observed for several days (Mongold et al. 2020). When combined with the results from Irma (section 4.1), the results provide further evidence that the model reasonably simulates the integrated hurricane evacuation system, and can be used to study various storm scenarios, real or imagined. Furthermore, the differences in the spatial and temporal patterns of evacuation between the two hurricanes confirm the

importance of forecast information to the evacuation dynamics (Huang et al. 2016; Baker 1991).

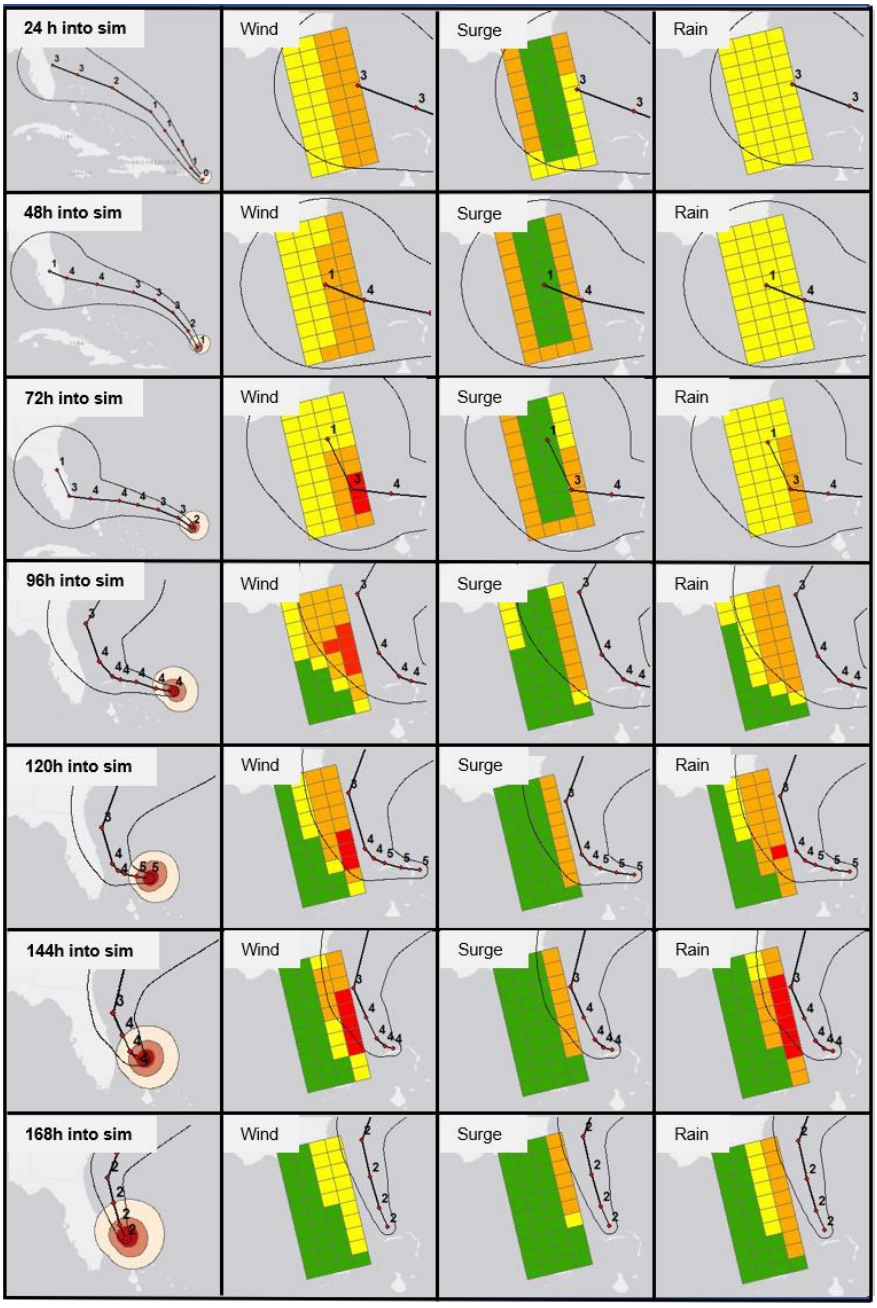

Figure 13: Light system forecast for Hurricane Dorian (2019). Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. Right three columns: The light-system threats for wind, surge, and rain are shown for equivalent times with the forecast track (center black line) and cone of uncertainty (outer black lines) included for reference. Note: threats are highest when near the center of the forecast cone and when hazards are most imminent, among other factors.

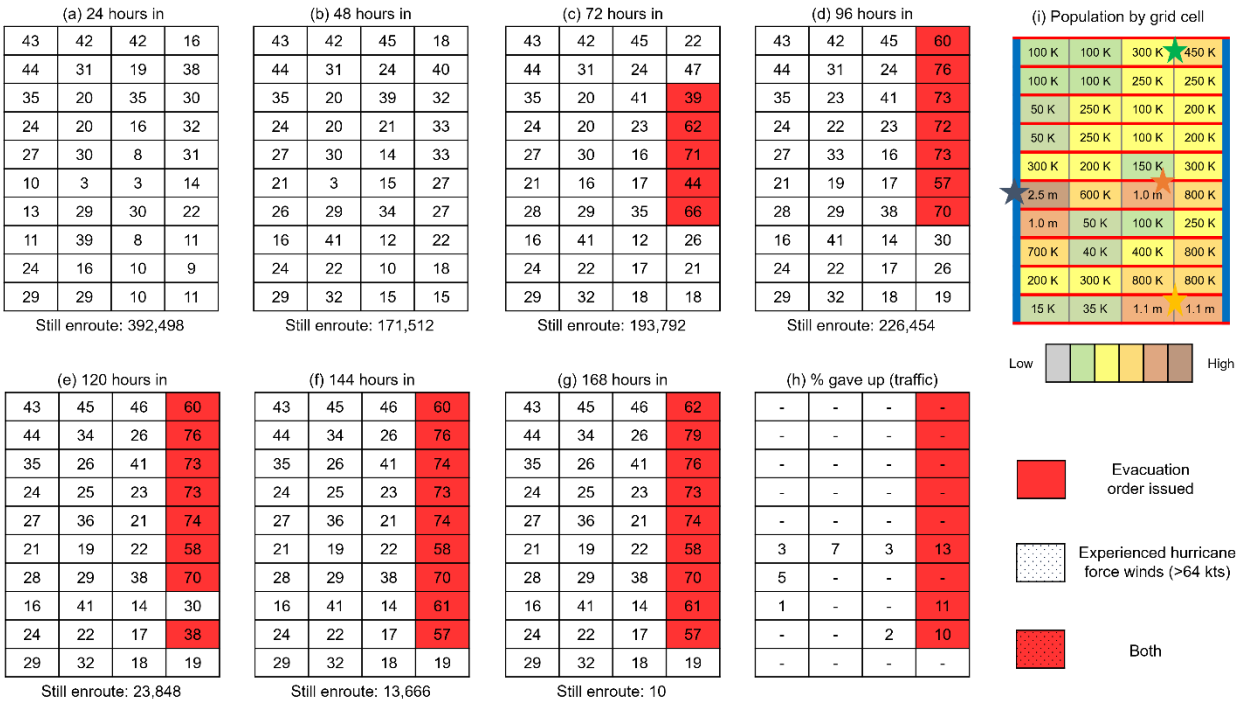

Figure 14: Evacuation rates for grid cells during Dorian (2019). Rates are expressed every 24 hours (a-g). The percentage of each grid cell which intended to evacuate but could not due to traffic is also expressed (h), as is the spatial and temporal patterns of evacuation orders (red cells). In addition, the number of evacuees still enroute at the various times is shown (bottom of panels a-f). Note, the hurricane force winds (>64kts) did not impact the model grid. Also expressed is the population by grid cell (i) which provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

Experiment	% Successfully evacuated							
	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 knot zone	Inland < 64 knot zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5
Dorian Default	33.5	-	-	67.6	24.8	64.4	23.3	6.1

Table 7: Evacuation behaviors by impact zone for Dorian. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

5. Summary and Discussion

This article conceptualizes and implements a modeling framework for studying the dynamics of the hurricane-forecast-warning system. The modeling framework, called FLEE, integrates models of the natural hazard, the human system, the built environment, and connections between systems. It includes millions of agents – with behaviors and characteristics informed by empirical research – who interact with each other, with their physical environments, and with evolving, uncertain forecast information to produce evacuation decisions and generate evacuation traffic. After describing FLEE, we validate the model framework by comparing its evacuation behaviors to observations, mainly for Hurricane Irma (2017), and present a set of proof-of-concept experiments illustrating its behaviors when key parameters are modified. In doing so, we show FLEE is capable of examining the dynamics of the hurricane-forecast-evacuation system from a new perspective.

We propose several areas for future work. First, FLEE can explore how changes in forecast track, intensity, storm size, forward speed, uncertainty, and different forecast scenarios influence evacuations (see, e.g., Fossell et al. 2017). This provides meteorologists with a societally-relevant alternative to traditional measures of forecast accuracy (need described by Morss 2005, Murphey 1993, Roebber and Bosart 1996), by measuring the impact of forecasts elements and uncertainties on how people receive and process the information, make evacuation decisions, and physically evacuate. Second, the model can be used to address behavioral science questions, such as how future projections of population density, socioeconomic status, inequality, and car access may affect hurricane evacuations. Third, FLEE can further determine the relative effectiveness of evacuation management strategies such as contraflow, adding public transportation, evacuation order timing, and phased evacuations (building on, e.g., Urbina and Wolshon 2003, Madireddy et al. 2011) and how forecasts influence evacuation order decisions (Davidson et al. 2018). This benefits researchers, practitioners, and policy-makers in hazard risk management.

FLEE is intentionally abstracted to explore the broader evacuation dynamics. However, additional layers of complexity can be added, depending on research goals e.g., to account for family composition, social circle's evacuation status, social-media influence, and house/building strength in evacuation decisions. FLEE can be extended to study other regions or hazards, such as hurricanes followed by flooding, loss of power networks, damage to roads, and other cascading failures. Additional in-depth comparisons with observational data can improve FLEE’s realism, and subsequently, its capability to answer questions of interest. But given the sparse availability of empirical data on

hurricane evacuations, new data sets are likely needed. Nevertheless, in its current form, FLEE can significantly advance our understanding of the integrated hurricane-forecast-warning system. This new knowledge is informed by and feeds back into empirical research, and can ultimately support researchers, practitioners, and policy-makers in a variety of disciplines, thereby offering the promise of direct applications to save lives and mitigate hurricane losses.

ARTICLE 2

WHAT IMPROVES EVACUATIONS? EXPLORING THE HURRICANE-FORECAST-EVACUATION SYSTEM DYNAMICS USING AN AGENT-BASED FRAMEWORK

1. Introduction

In the mainland US, the hurricane-forecast-evacuation system is dynamic, complex, and difficult to predict, owing to interacting physical-social factors and uncertainties that change as the storm approaches. Take Hurricane Irma (2017) and Rita (2005), cases where accurate but uncertain forecasts triggered mass evacuations — and severe traffic jams — in Florida and Texas, respectively. As the forecasts shifted and traffic worsened, some evacuees became more exposed to hazardous conditions than had they remained in-place (Cangialosi et al. 2018; Wong et al. 2018; Zhang et al. 2007; Knabb et al. 2006). Despite forecasts being as accurate as one can expect given current forecast skill, these cases illustrate the complexities of people using inevitably imperfect information to make pre-storm evacuation decisions, and they demonstrate how evacuations involve many intersecting physical-social parts and uncertainties that evolve over time (e.g., Morss et al. 2017; Barton 2014; Miller and Page 2007; Watts et al. 2019). Because of these complex dynamics, evacuations are a formidable challenge.

Empirical studies help explain aspects of evacuations, such as how forecasts, warnings, and other factors influence evacuation decisions (e.g., Huang et al. 2016; Lindell and Perry 2012; Baker 1991). However, it's impossible to empirically study all aspects of evacuation across multiple cases. Computational models, on the other hand, provide a complementary tool where empirical knowledge can be codified and used to run virtual

experiments for many hurricane scenarios, real or synthetic (e.g., Morss et al. 2017, Watts et al. 2019; Harris et al. 2021; Blanton et al. 2018; Davidson et al. 2018). In this context, modeling studies often optimize evacuations through strategies such as contraflow (opening additional lanes for traffic), public transportation, and evacuation order timing (e.g., summary of strategies in Murray-Tuite et al. 2019). However, the full hurricane-forecast-evacuation system — including these different evacuation management strategies — has not yet been represented in one framework, meaning a holistic exploration of the system's dynamics has not been achieved.

To address this need, Harris et al. (2021), which is also Article 1 of this dissertation, developed a modeling framework capable of exploring the system's dynamics holistically. Called FLEE (Forecasting Laboratory for Exploring the Evacuation-system), the agent-based framework models key aspects of real-world hurricane evacuations: the natural hazard (hurricane), the human system (information flow, evacuation decisions), the built environment (road infrastructure), and connections between systems (forecasts and warning information, traffic, impact zones). By coupling the models into one framework for the first time, FLEE is capable of simultaneously exploring interactions across these sub-systems. In the current version, FLEE's agent-based model grid is a 10 x 4 abstracted representation of the north-south axis of Florida (e.g., see conceptualization, implementation, assumptions, and proof-of-concept experiments in Harris et al. 2021). Note that FLEE is designed to represent key aspects of evacuations established in the literature, while remaining sufficiently idealized to build fundamental and practical knowledge (e.g., see Watts et al. 2019; Sun et al. 2016, discussion in Section 2). As a result, it is capable of providing a high-level, first-order look at the system's dynamics.

This paper's objective is to begin exploring the system's dynamics using FLEE. Specifically, through sensitivity analyses on the model, we assess how evacuations change with evacuation management strategies and policies (public transportation, contraflow, evacuation order timing), evolving population characteristics (population growth, urbanization), and forecast scenarios impacting the Florida peninsula (Irma, Dorian, rapid-onset version of Irma). Throughout the analysis, we ask:

1) How do the factors impact evacuation success?
2) How does the factor's impact vary across forecast-population scenarios?
3) How are the impacts distributed geographically?

In answering these questions, we demonstrate how these agent-based, computational frameworks can build our understanding of evacuations across many different real and imagined scenarios. Such knowledge would support researchers, practitioners, and policy-makers in a variety of disciplines including hazard risk management (building on, e.g., Madireddy et al. 2011), evacuation modeling (e.g., building on Watts et al. 2017, Davidson et al. 2018), and meteorology (e.g., need for measures of forecast impacts described by Morss 2005, Murphy 1993, Roebber and Bosart 1996).

2. Methodology

2.1 Experimental Design

We start by performing a set of experiments with FLEE that involve changing the storm and corresponding forecast scenarios. To do so, we run FLEE using forecasts from Hurricane Irma and Dorian, which triggered evacuations across Florida in 2017 and 2019,

respectively. These simulations were compared to observational evacuation data for validation (Section 3). We then run FLEE with a hypothetical version of Irma in which Irma's forecasts are condensed into a shorter timeframe to assess how rapid onset (and similarly, rapid intensification) impacts evacuation success. Together, these three experiments are used to provide a baseline of evacuation behaviors to compare additional experiments against.

Next, starting from these "default" simulations for each storm, we systematically modify FLEE's parameters one-by-one, while holding other variables constant, to explore FLEE's behaviors and sensitivities to evacuation management strategies, evacuation policies, and evolving population characteristics. This includes simulating contraflow by adding one, two, and three lanes across important highways; public transportation/carpooling by adjusting the number of evacuees per vehicle from 4 people per car to 2, 3, 5, 6, and 8 people per car; modifications to evacuation order timing by shifting the issuance of evacuation orders earlier/later by 10 and 20 hours; and the impact of population growth and urbanization by projecting the population forward to 2030 and 2040 or making population density uniform across the model grid, respectively. By comparing the experiments to the default simulations and to each other, we investigate the relative importance of different factors and key interactions, and we demonstrate how FLEE can explore the system's dynamics across a variety of scenarios.

2.2 Data Analyses

To compare evacuation behaviors quantitatively across simulations, we track evacuation statistics for each of FLEE's 40 grid cells, which represent the north-south axis of Florida

(see full description of the model's virtual world in Harris et al. 2021, hereafter HRM21). The primary model output analyzed here are the percent of households that successfully evacuated (i.e., evacuation rates), and the percent who intended to evacuate but gave up due to traffic. The latter statistic provides insight to where excessive traffic prevented successful evacuations. In addition to displaying data by grid cells and averaged across the domain, these two statistics are sometimes broken down into multiple impact zones, designed as first-order approximations of areas likely to experience different levels of impacts based on the actual meteorological conditions produced by the storm. Here, we use four impact zones, defined by whether the grid cells are: a) coastal or inland, and b) experience winds that are greater than 64 knots (hurricane-force) or less than 64 knots during the storm of interest. Using the impact zones, we can determine who evacuated from locations that did not end up experiencing hazardous conditions. In addition, we examine the percentage of residents under evacuation orders in the model who evacuated (sometimes called compliance rates), and the percentage of residents who evacuated from areas not under evacuation orders (sometimes called shadow evacuations; McGhee and Grimes 2006; Murray-Tuite et al. 2019).

In interpreting the results for different experiments, we compare multiple metrics that might indicate successful outcomes in different ways. For example, high evacuation rates may not be "good" if the storm ends up not having much impact in those areas, and unnecessary evacuations may not matter if those in areas experiencing the greatest impacts were able to evacuate successfully.

Because FLEE includes stochastic elements, it can exhibit some run-to-run variability. Therefore, when interpreting results, changes less than 0–2% in evacuation statistics across experiments are considered insignificant (e.g., see experiments in HRM21).

2.3 Model Updates and Validation

Based on comparison with recent empirical data on evacuation rates and traffic for the storms studied here, several updates have been made to FLEE since its conceptualization and implementation in HRM21. First, to reduce evacuation rates early in the simulations, making the temporal patterns of evacuation closer to data for Irma and Dorian, we made two changes: 1) integrating households into the decision-making process gradually rather than assuming instantaneous communication of risk information and 2) making evacuation barriers time dependent, i.e., decrease as the storm approaches. Second, the initial formulation of FLEE exhibited unrealistic variability between grid cells in some situations; to address this, we reduced the influence of mobile home and socioeconomic status in the evacuation decision-making algorithm. In addition, FLEE's initial formulation overestimated evacuation rates in coastal grid cells compared to inland cells. Thus, we changed the formulation of evacuation orders so they are only issued to a percentage of households in coastal grid cells, reflecting the approximate percentage of population in different regions of Florida that live in evacuation zones, rather than the entire grid cell. Details regarding the implementation of the updates is provided in the Appendix B (Table B1, Figure B1), as are full descriptions and sources of the empirical data used for validation (Table B2).

3. Results and Discussion

First, we examine results from the default simulations for Hurricane's Irma and Dorian. Comparing these simulations against observations provides a first-order assessment that the model behaves reasonably for these two forecast scenarios based on the processes implemented. Then, we examine results from the default simulation using forecasts from a rapid-onset version of Irma (hereafter called Irma-RO). Together, the experiments provide a baseline for interpreting results from subsequent experiments.

3.1 Hurricane Irma

Irma's 3-10–day forecasts, as shown in Figure 2 of HRM21 (Article 1 of this dissertation), place the entire model under significant threat, with the most likely outcome being a landfalling major hurricane near Miami. However, forecasts shifted westward as the storm approached, with the storm eventually making a first mainland U.S. landfall as a Category 4 in the Florida Keys and a second landfall as a Category 3 in southwest Florida. Irma's hurricane-force winds (Figure 15a–b; dotted cells) impacted the western two-thirds of the model – particularly the southwest coastlines – while leaving the east-coast unscathed. Evacuation orders were issued along both coasts in FLEE (Figure 15a–b; red cells), similar what was observed during Irma (Wong et al. 2018; Darzi et al. 2020; model description in Figure 3 of HRM21). The comparison with empirical data increase our confidence the evacuation order algorithm – and the synthesized National Hurricane Center forecasts on which its based – behaves sufficiently realistically for the purpose of the subsequent experiments.

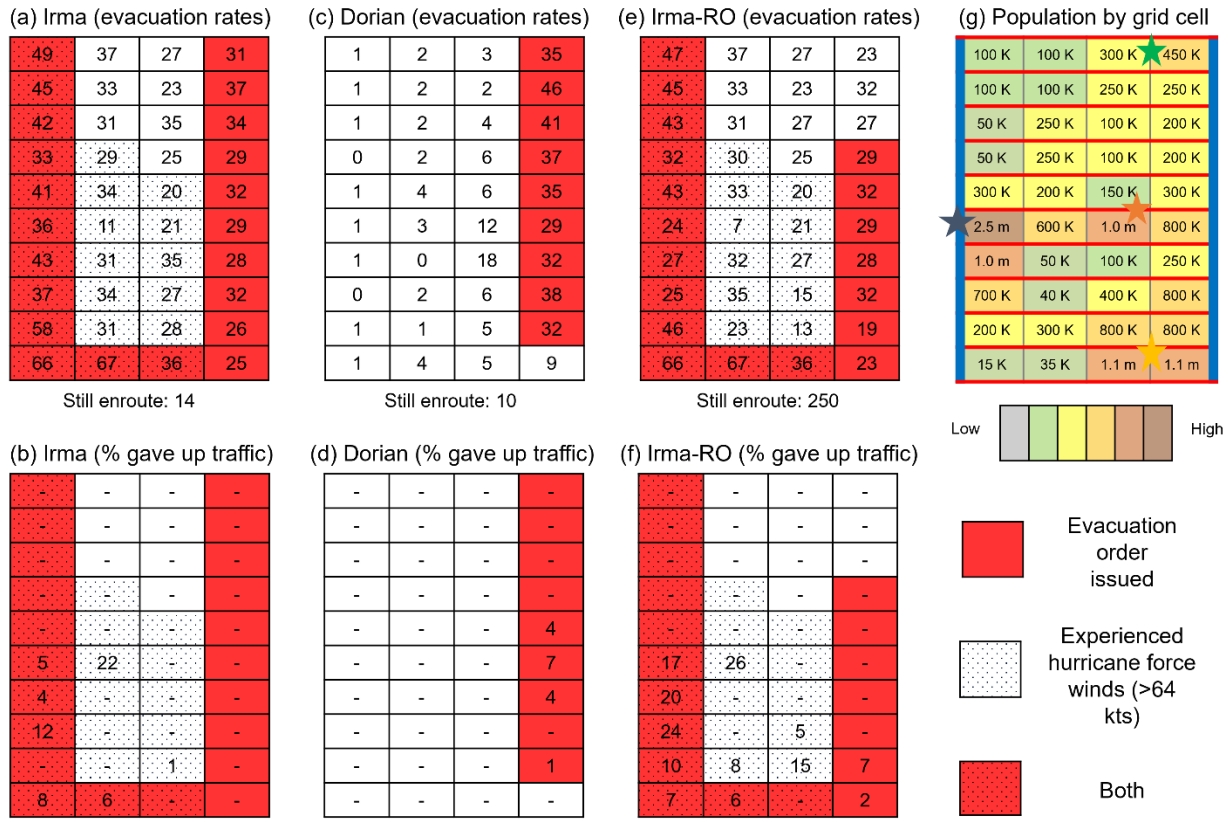

Figure 15: Evacuation rates and traffic by grid cell for the default hurricane scenarios. Also shown is the population by grid cell (g) to provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). In addition, the number of evacuees still enroute at the end of the simulations are shown.

Irma's simulated evacuation rates (Figure 15a) are similar to observational data, which suggest evacuation rates vary from 20–40% along Florida's east coast, to 40–70% across the south and west coasts, and around 10–30% inland (Wong et al. 2018; Long et al. 2020; Martin et al. 2020; Feng and Lin 2021). This was the largest evacuation in US history, meaning many households failed to evacuate due to excessive traffic. This occurs most frequently around Tampa Bay-St. Petersburg in FLEE (Figure 15b), again matching observations (Feng and Lin 2021; States et al. 2021).

Table 8 depicts evacuation rates across impact zones. 32.0% of households on the model grid evacuate for Irma, which equals 5.4 million people. Estimates from the Florida Department of Emergency Management (2017) suggest actual evacuation numbers totaled 6.9 million. When considering households evacuating to local shelters in FLEE (not shown), the simulated evacuation rates are similar to observations. For a given level of wind impact, evacuation rates are higher along the coasts than inland, which typical for real-world evacuations (Feng and Lin 2021; Martin et al. 2020).

Experiment	% Successfully evacuated							
	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 knot zone	Inland < 64 knot zone	% Under evacuation orders who evacuated	% Not under evacuation orders who evacuated	Gave up to traffic
Irma	32.0	39.3	24.5	29.6	29.4	34.9	25.6	2.5
Dorian	12.0	-	-	14.7	0.6	35.4	4.3	0.7
Irma-RO	26.3	30.2	18.8	26.8	28.9	28.8	25.3	7.7

Table 8: Evacuation rates and traffic by impact zones for the default storm scenarios. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but “gave up” due to excessive amounts of traffic.

The evolution of evacuation rates and traffic is shown across time in Figure 16. Despite not receiving evacuation orders, some households (black solid line) evacuate in the first 0–72 hours. Evacuation rates increase linearly between 72–108 hours as evacuation orders expand. Just before the storm moves ashore around 126 hours, evacuation rates decrease, while the number of households giving up due to excessive traffic (black dashed line) increase. These temporal patterns of evacuations, as with the spatial

patterns, generally match existing empirical data (e.g., Wong et al. 2018). Therefore, we believe FLEE's simulated evacuations provide a realistic baseline for interpreting results from Irma's other experiments.

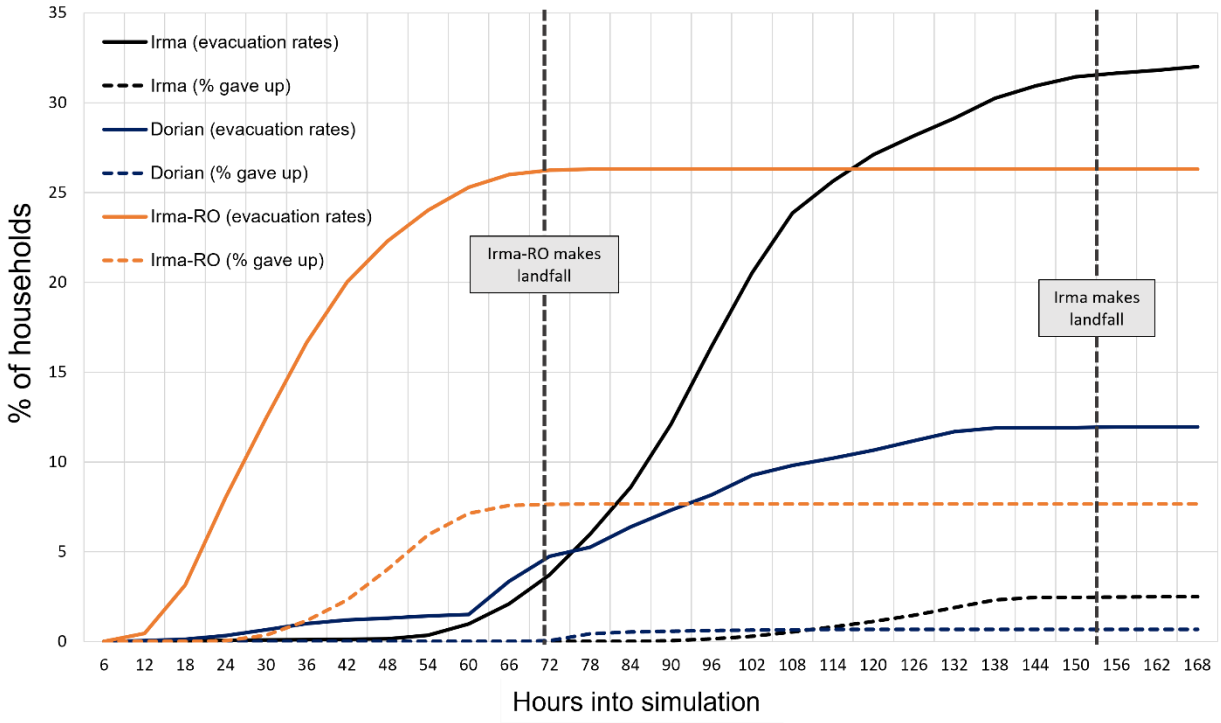

Figure 16: Evacuation rates and traffic over time for the default hurricane scenarios. Evacuation rates are shown (solid lines), averaged across all grid cells, as are the percentage giving up due to traffic (dashed line). The times of landfall for the Irma and Irma-RO scenarios are indicated (vertical dashed lines).

3.2 Hurricane Dorian

Dorian's early forecasts (Figure 13 of HRM21) place the entire Florida peninsula under threat, with the most likely scenario as a landfalling major hurricane along Florida's east coast. Because of the forecasts, evacuation orders were issued along the majority of FLEE's east coast (Figure 15c–d; red cells), matching what was observed (TIME 2019). However, after remaining nearly-stationary over the Bahamas for many hours, the storm accelerates northward, missing Florida's east coast entirely.

Dorian is a fundamentally different storm with less people at risk. As a result, evacuation rates were less with Dorian (12.0%) than Irma (32.0%), which is 3.3 million less evacuees (Table 8). Due to fewer evacuees overall – and the evacuation being spread over a longer time (Figure 16; blue lines) – less people give up due to traffic (0.7%) than Irma (2.5%). Evacuation rates and traffic are confined to the east coast with Dorian (Figure 15c–d), matching our understanding of what occurred (Mongold et al. 2020).

Though robust empirical data on Dorian's evacuation rates is not available, the available data suggests the model is generating reasonable evacuation behaviors i.e., the Dorian simulation appears accurate as possible given current empirical data, and is reasonable for experimentation.

3.3 Hurricane Irma-RO

The hypothetical Irma-RO forecasts are shown in Figure 17. For each grid cell, the peak magnitudes of forecast risk are identical to Irma's forecasts. However, the forecasts – and subsequently, the simulation – are compressed from 168 hours to 72 hours. By comparing the evacuation response between the real and hypothetical storms, we aim to explore the potential effects of a storm that exhibits rapid onset (and also, to a degree, rapid intensification) on the hurricane evacuation dynamics.

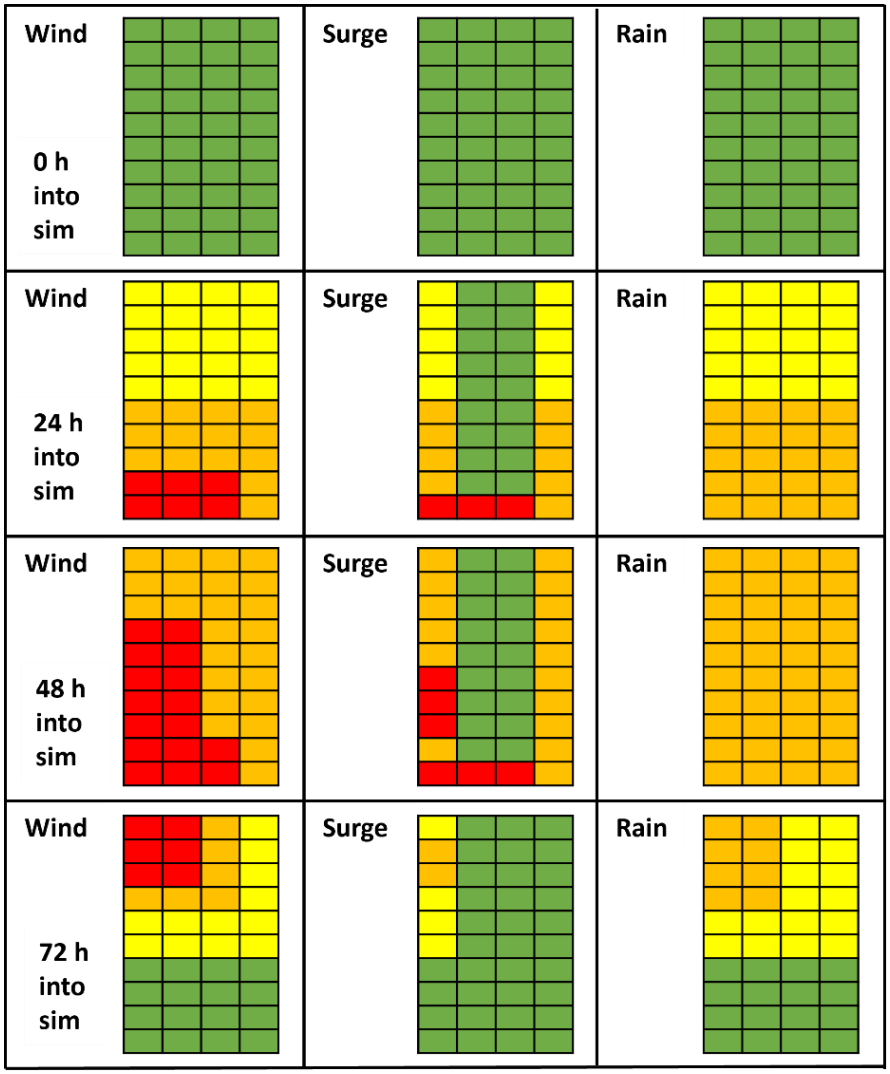

Figure 17: Forecasts for Hurricane Irma-RO approaching FLEE's 10 x 4 version of the north-south axis of Florida. Forecasts are shown for every 24 hours but update every 6 hours (not shown). For a detailed explanation of the forecasts' conceptualization and implementation, see HRM21.

FLEE's response to the Irma-RO forecast is similar to Irma's, but with key differences. First, evacuation orders (Figure 15e–f, red cells) are not issued around Jacksonville (Figure 15g, green star). This occurs because the westward shift in forecasts happens more quickly than with Irma, meaning the area was removed from risk before evacuation order decisions were made. As a result, Irma-RO has 300,000 fewer people intending to evacuate in these regions, despite having identical peak magnitudes of risk.

In addition, a lower percentage of the Florida population evacuates Irma-RO (26.3%) than Irma (32.0%). This is partially because fewer evacuation orders are issued, but mostly because more people gave up evacuating due to excessive traffic in Irma-RO (7.7%) than Irma (2.5%), an increase of 852,000 people (Table 8). The increased traffic is mostly confined to the Tampa Bay-St. Petersburg metropolitan areas and surrounding southwest coastlines. However, unlike with Irma, residents around Miami-Ft. Lauderdale also experience significant traffic (Figure 15f). Therefore, we believe the Irma-RO case suggests that rapid-onset – and possibly, rapid intensification – can worsen evacuation rates and traffic in higher-risk areas. This is consistent with our conceptual understanding of these situations, as there is less time to evacuate safely.

3.4 Public Transportation and Carpooling

In the following subsections, we modify key model parameters and compare their evacuation behaviors to the Irma, Dorian, and Irma-RO default, baseline experiments discussed earlier. Specifically, we assess how evacuation outcomes change with evacuation management strategies and policies (public transportation, contraflow, evacuation order timing) and evolving population characteristics (population growth, urbanization). The idea is to demonstrate how models like FLEE can explore the role of these factors in improving evacuations, including exploring 1) how the factors impact evacuation success, 2) how their impacts vary across forecast scenarios, and 3) how their impacts are distributed geographically.

Relative to evacuation management strategies like contraflow, public transportation and carpooling is studied less frequently and is rarely provided (e.g., summary in Murray-Tuite

et al 2019; Bullard and Wright 2011; exceptions in Swamy et al. 2017; Zhang and Chang 2014). To our knowledge, additional public transportation was not offered for Irma or Dorian, nor was carpooling widely used across households. The goal of this section is to demonstrate how models like FLEE can explore the role of these strategies in improving evacuations.

Experiments were conducted by changing the number of people evacuating per car from 4 (default) to 2, 3, 5, 6, and 8. For a given number of evacuees, this experiment changes the number of cars on the road. For example, going from 4 to 2 people per car doubles the number of cars evacuating. On the other hand, going from 4 to 8 people per car halves the number of cars evacuating. Though the car-length is kept the same in FLEE (in reality, if simulating buses, it should be longer), this is intended as a first-order exploration of the effects of strategies such as public transportation and carpooling using FLEE.

Evacuation rates and traffic are shown in Table 9. Across all scenarios, results indicate that evacuation rates increase when public transportation and carpooling are used as fewer people get stuck in traffic and give up evacuating. These improvements are significant. For example, with Irma, going from 4 to 6 people per car improves evacuation rates by 2.2%, an increase of 500,000 people. The magnitude of improvements is scenario-dependent, with the largest improvements happening with Irma-RO followed by Irma and Dorian (this is because Irma-RO has the most traffic issues, followed by Irma and then Dorian).

Experiments		Evacuation Rates	% Change relative to default	Evacuated	% Under evacuation orders who evacuated	% Not under evacuation orders who evacuated	% Gave up due to traffic	Gave up due to traffic
Irma	2/car	22.5	-9.5	3.67m	25.5	16.0	12.0	1.97m
	3/car	28.5	-3.5	4.67m	32.0	20.9	6.0	980k
	4/car (default)	32.0	0	5.24m	34.9	25.6	2.5	410k
	5/car	33.7	+1.7	5.52m	26.4	27.6	0.8	130k
	6/car	34.2	+2.2	5.61m	37.1	27.9	0.3	49k
	8/car	34.4	+2.4	5.64m	37.1	28.3	0.1	16k
Dorian	2/car	9.5	-2.5	1.56m	26.7	3.9	3.1	508k
	3/car	11.2	-0.8	1.84m	32.9	4.2	1.4	229k
	4/car (default)	12.0	0	1.97m	35.4	4.3	0.7	115k
	5/car	12.2	+0.2	1.99m	36.4	4.3	0.4	66k
	6/car	12.3	+0.3	2.02m	36.8	4.3	0.3	49k
	8/car	12.5	+0.5	2.05m	37.8	3.7	0.1	16k
Irma-RO	2/car	16.9	-9.4	2.77m	17.4	16.0	17.1	2.81m
	3/car	21.5	-4.8	3.52m	23.1	18.6	12.4	2.03m
	4/car (default)	26.3	0	4.31m	28.8	25.3	7.7	1.26m
	5/car	29.7	+3.4	4.87m	32.3	25.0	4.3	704k
	6/car	31.0	+4.7	5.08m	34.2	25.6	2.9	475k
	8/car	33.2	+6.9	5.44m	37.0	26.4	0.9	147k

Table 9: Evacuation rates and traffic for the public transportation and carpooling experiments. Also shown are the total numbers evacuated, numbers giving up due to traffic, compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), and shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did).

The experiments also reveal non-linearities in the system (Table 9). For example, in all forecast scenarios, doubling the number of evacuating cars (i.e., going from 4 to 2 people per car) results in significantly larger changes to evacuation rates than halving the number of evacuating cars (i.e., going from 4 to 8 people per car). We believe this suggests excessive traffic can worsen evacuations exponentially. Future work may consider furthering our understanding of this relationship, including identifying whether tipping points exist, or whether targeted carpooling and bussing helps, supporting practitioners in risk management.

Figure 18 shows which areas are impacted by the public transportation and carpooling experiments. In the 8 people per car experiment, evacuation rates and traffic improve areas heavily trafficked in the default simulations (default in Figure 15). In the 2 people per car experiment, evacuations worsen considerably, particularly across at-risk areas where evacuees are dependent on traffic downstream i.e., southern Florida.

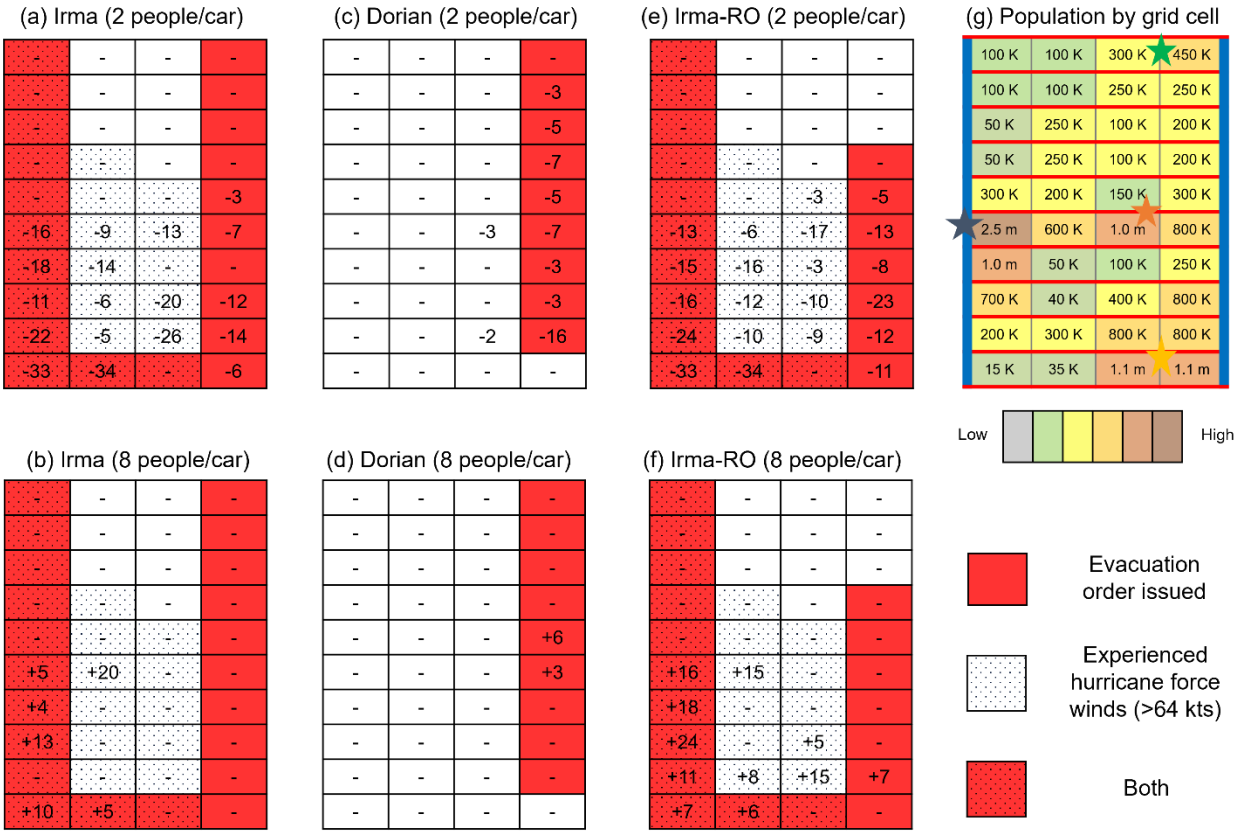

Figure 18: Evacuation rates and traffic by grid cell for select carpooling and public transit experiments. Values are expressed as changes relative to default, as shown in Figure 15. Also expressed is the population by grid cell (g). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

By and large, results suggest that public transportation and carpooling improve evacuations significantly. This is especially true for the most heavily trafficked scenarios and regions in these scenarios (i.e., Irma-RO, south Florida). More broadly, results confirm FLEE behaves appropriately for this type of modification. It also demonstrates how this type of modeling framework can be used to explore public transportation and carpooling use across many scenarios, real and synthetic, and to identify regions to target using the strategy.

3.5 Contraflow

Though used before Hurricane's Floyd (1999) and Katrina (2005), contraflow was not implemented, to our knowledge, on a wide-scale for Irma or Dorian (Wong et al., 2018; Wolshon and Lambert 2004), as it requires considerable personnel resources and must be planned days in advance (e.g., summary in Murray-Tuite et al. 2019). Here we aim to explore how contraflow might have influenced evacuations in these cases, and more broadly, to examine how strategies focused on improving road capacity compare with other evacuation management strategies and factors in the forecast-evacuation system.

As described in HRM21, FLEE's road network consists of two northbound, five-lane interstates situated on the model's "coasts" (i.e., Florida's I-75 and I-95). Meanwhile, two east-west running, three-lane interstates move evacuees across the grid horizontally (i.e., I-75 and I-4), while eight two-lane highways move inland residents onto the "coastal" interstates where they flee to safety. Here we examine the influence of adding contraflow on these highways and interstates.

Evacuation rates and traffic are shown when adding one lane on I-95 and I-75, and one, two, and three lanes across the entire road network (Table 10). Contraflow improves evacuation rates across all scenarios, confirming the model behaves appropriately. Adding lanes everywhere is more effective than adding lanes on one highway, consistent with other studies (e.g., see Murray-Tuite et al 2019). Opening three lanes of contraflow before Irma increases evacuation rates by 1.9%. For comparison, the public transportation experiments of 5 and 6 people per car increase evacuation rates by 1.7% and 2.2%, respectively.

Experiments		Evacuation Rates	% Change relative to default	Evacuated	% Under evacuation orders who evacuated	% Not under evacuation orders who evacuated	% Gave up due to traffic	Gave up due to traffic
Irma	Default	32.0	0	5.24m	34.9	25.6	2.5	409k
	+1 I-95	32.1	+0.1	5.26m	35.0	25.8	2.4	393k
	+1 I-75	32.0	0	5.24m	34.9	25.6	2.5	409k
	+1 all	32.7	+0.7	5.36m	35.7	26.0	1.9	311k
	+2 all	33.5	+1.5	5.49m	36.3	27.2	1.0	163k
	+3 all	33.9	+1.9	5.56m	36.8	27.4	0.7	115k
Dorian	Default	12.0	0	1.97m	35.4	4.3	0.7	115k
	+1 I-95	12.2	+0.2	2.0m	36.3	3.8	0.5	82k
	+1 I-75	12.0	0	1.97m	35.5	4.3	0.7	115k
	+1 all	12.1	+0.1	1.98m	36.2	4.3	0.5	82k
	+2 all	12.3	+0.3	2.02m	36.7	4.3	0.3	49k
	+3 all	12.3	+0.3	2.02m	36.9	4.3	0.3	49k
Irma-RO	Default	26.3	0	4.31m	28.8	25.3	7.7	1.26m
	+1 I-95	27.1	+0.8	4.44m	29.2	23.2	6.9	1.13m
	+1 I-75	26.3	0	4.31m	28.8	22.1	7.6	1.25m
	+1 all	28.0	+1.7	4.59m	30.6	23.5	5.9	967k
	+2 all	29.3	+3.0	4.80m	31.8	25.0	4.7	770k
	+3 all	29.8	+3.5	4.88m	32.6	25.0	4.2	688k

Table 10: Evacuation rates and traffic for the contraflow experiments. Also shown are the total numbers evacuated, numbers giving up due to traffic, compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), and shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did).

The magnitude of improvements from contraflow are forecast-scenario-dependent (Table 10). For example, relative to the default simulations, adding three lanes increases evacuation rates by 1.9% with Irma, 0.3% with Dorian, and 3.5% with Irma-RO. In terms of areas impacted (Figure 19), contraflow improves heavily trafficked areas in the default scenarios (i.e., urban areas and areas dependent on traffic "upstream" like south Florida). This is consistent with the public transportation and carpooling experiments, suggesting evacuation management strategies are most effective in heavily-trafficked forecast scenarios and regions.

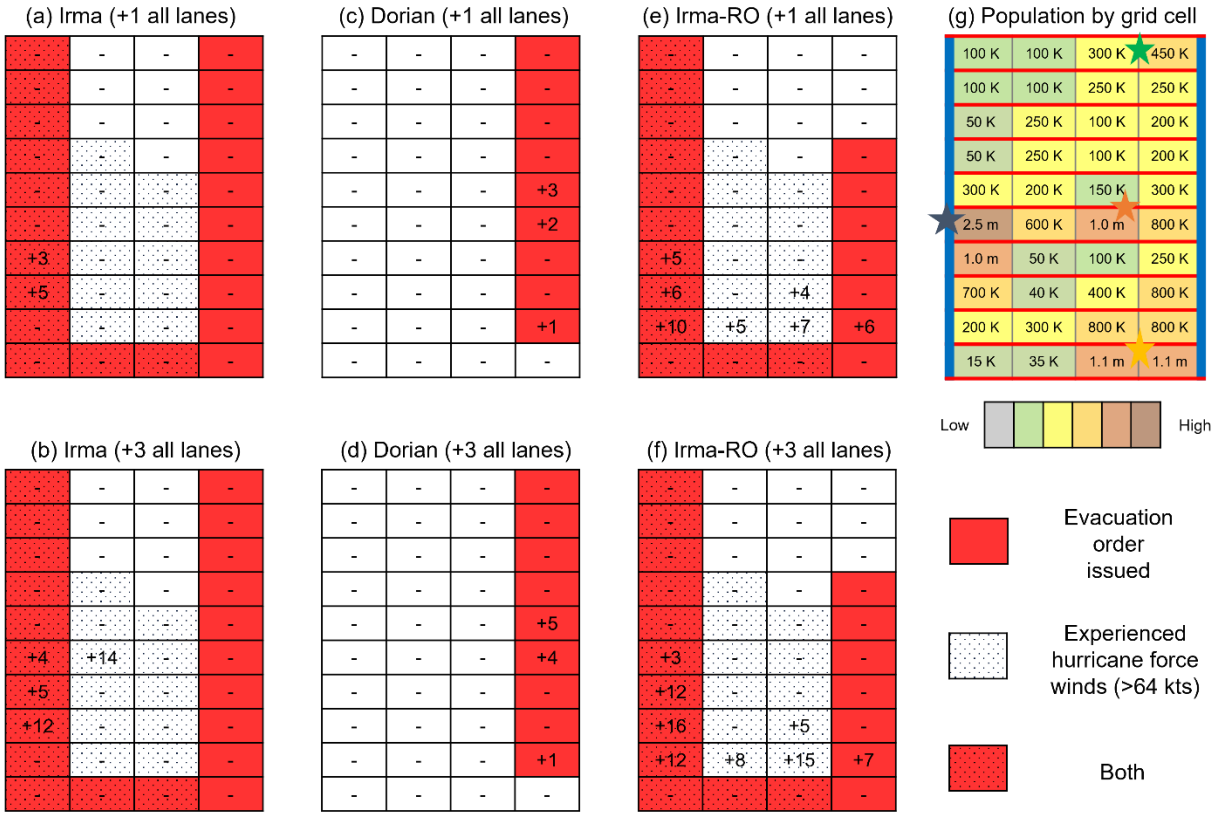

Figure 19: Evacuation rates and traffic by grid cell for select contraflow experiments. Values are expressed as changes relative to default, as shown in Figure 1. Also expressed is the population by grid cell (g). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

Overall, results suggest that adding road capacity through strategies such as implementing contraflow can improve evacuations significantly. Results also show how FLEE can compare the relative importance of factors to evacuation success and identify areas benefiting most from contraflow across scenarios, thus potentially supporting decision-makers in their cost-benefit analyses.

3.6 Evacuation Order Timing

Regarding evacuation orders, there are modeling studies investigating how to optimize evacuation order timing to support decision-makers (e.g., see summary in Yi et al. 2016; examples in Dixit and Radwan 2009, Davidson et al. 2019; clearance time studies in Florida Statewide Regional Evacuation Study Program 2019). Building upon this work, we explore the importance of evacuation order timing in FLEE, specifically by issuing orders 10 and 20 hours earlier and later than the default simulations (note: this was done by adjusting the clearance times for each grid cell, which subsequently shifts evacuation order decision-making thresholds by these amounts; description in HRM21).

When averaged across all grid cells, shifting the timing of evacuation orders did not meaningfully impact evacuation rates and traffic. The largest changes occurred with Irma-RO; however, evacuation rates only improved by 0.8%, 0.8%, 0.2%, and 0.1% for the 20 hours earlier, 10 hours earlier, 10 hours later, and 20 hours later experiments, respectively (not shown). The effects are much smaller than implementing public transportation or contraflow.

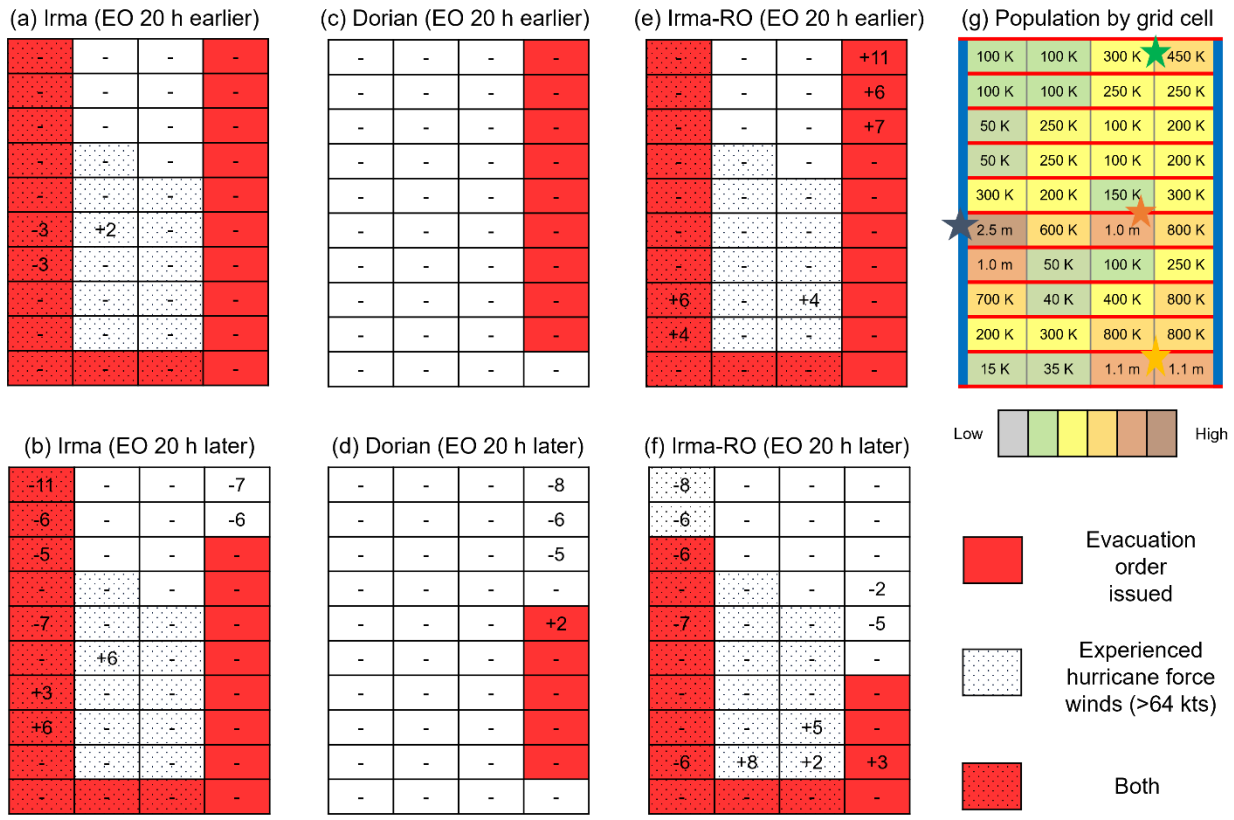

Figure 20: Evacuation rates and traffic by grid cell for select evacuation order timing experiments. Values are expressed as changes relative to default, as shown in Figure 1. Also expressed is the population by grid cell (g). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

However, between grid cells, evacuation order timing has complex – but sometimes significant – impacts (Figure 20). For example, during Irma, later evacuation orders improve evacuation rates around Tampa Bay-St. Petersburg by 3-6%, and reduce evacuation rates along the northwest coastline and Jacksonville metroplex by 5-11%. The latter is actually a positive, however, as the later evacuation order thresholds caused evacuation orders to not be issued in this area, as the forecast risk shifted westward, removing Jacksonville from harm's way. In this way, FLEE captures the tradeoff between issuing evacuation orders earlier (earlier warnings) versus waiting (reduction in false alarms). Similar impacts occur with Dorian, where later evacuation orders reduced

evacuation rates, and subsequently, false alarms. With Irma-RO, earlier orders helps generally (but also increases false alarms), while later evacuation orders again exhibits localized effects (e.g., it improves evacuation rates inland but worsens evacuations along coastal grid cells).

Relative to public transportation, carpooling, and contraflow, evacuation order timing has little impact on evacuation rates/traffic overall. That's not to say it's unimportant, however, as evacuation order timing has significant but localized effects on evacuations that varies considerably across scenarios. Perhaps this is to be expected, as evacuation orders are earlier in FLEE's "order of operations" between sub-systems. Future work may consider studying these complex effects by shifting evacuation order timing at specific areas in FLEE, studying more forecast scenarios, and/or by implementing phased evacuations (e.g., building on Chiu et al. 2008; Chen and Zhan 2008; Zhang and Chang 2014).

3.7 Population Growth and Urbanization

To our knowledge, few studies explore how US hurricane evacuations may change with population growth and urbanization. In this section, we begin to explore these effects by increasing FLEE's Florida peninsula population from 16.9 million in 2020 to 2030 and 2040 projections of 19.3 million and 22.3 million, respectively (see Figure 21g–h for population increase by grid cell; projections in Florida Department of Transportation 2020). Meanwhile, a third experiment representing the potential effects of urbanization is conducted by making the population density uniform across the model grid. Together, these experiments begin to look at how evacuations may change with population

characteristics, including how it impacts evacuation success relative to evacuation management strategies.

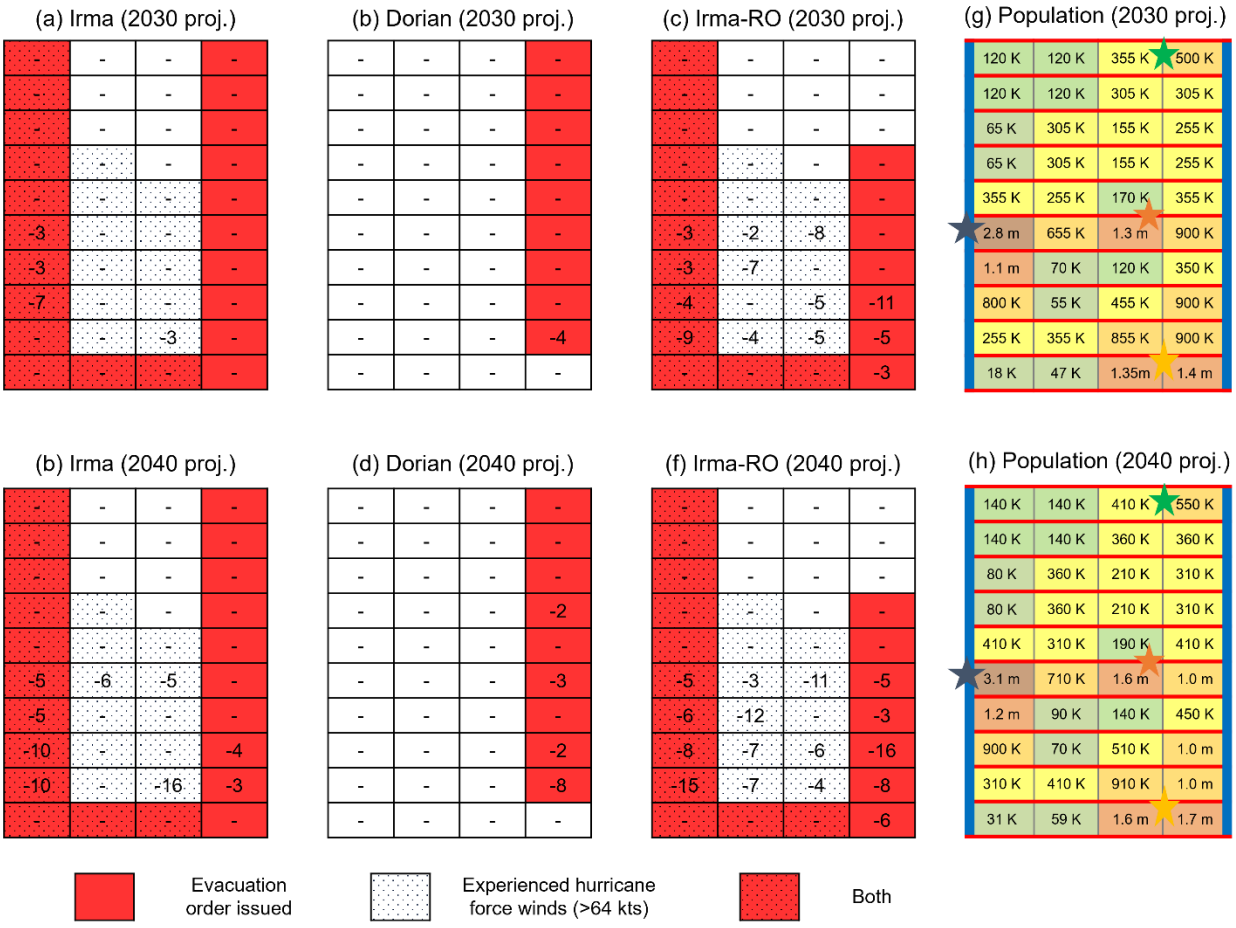

Figure 21: Evacuation rates and traffic by grid cell for the 2030 and 2040 projected populations. Values are expressed as changes relative to default, as shown in Figure 1. Also expressed is the 2030 population by grid cell (g) and the 2040 population by grid cell (h). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

Evacuation statistics are shown for this set of experiments in Table 11. Results suggest that evacuations may worsen in the future due to projected population changes. For example, relative to the default simulation, Irma's evacuation rates decrease by 1.3% and 3.3% in 2030 and 2040, respectively. Relative to the evacuation management strategies

tested earlier, the changes are significant, e.g., adding three lanes of contraflow increases evacuation by 1.9%, and going from 4 to 3 people/car decreases evacuation rates by 3.5%. Lastly, making population density uniform improves evacuations overall. To a first order, this provides evidence that urbanization can worsen evacuations rates.

Experiments		Evacuation Rates	% Change relative to default	Evacuated	% Under evacuation orders who evacuated	% Not under evacuation orders but evacuated	% Gave up due to traffic	Gave up due to traffic
Irma	Default	32.0	0	5.24m	34.9	25.6	2.5	409k
	Uniform	33.6	+1.6	5.51m	36.7	29.8	2.6	426k
	2030	30.7	-1.3	5.93m	33.3	24.9	3.6	695k
	2040	28.7	-3.3	6.38m	31.8	22.1	5.5	1.22m
Dorian	Default	12.0	0	1.97m	35.4	4.3	0.7	115k
	Uniform	11.2	-0.8	1.84m	36.4	3.8	0.6	98k
	2030	11.6	-0.4	2.24m	34.1	4.4	1.0	193k
	2040	11.2	-0.8	2.49m	32.7	4.5	1.5	333k
Irma-RO	Default	26.3	0	4.31m	28.8	25.3	7.7	1.26m
	Uniform	29.2	+2.9	4.79m	32.0	26.7	6.0	983m
	2030	23.7	-2.6	4.58m	28.1	20.0	10.0	1.93m
	2040	21.9	-4.4	4.87m	23.6	21.7	11.7	2.61m

Table 11: Evacuation rates and traffic for the population growth and urbanization experiments. Also shown are the total numbers evacuated, numbers giving up due to traffic, compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), and shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did).

As with public transportation, carpooling, and contraflow, the impacts of population growth and urbanization are scenario dependent. For example, the 2040 experiments result in a 3.3% decrease in evacuation rates with Irma, a 0.8% decrease with Dorian, and a 4.4% decrease with Irma-RO, i.e., the heavily-trafficked scenarios are most sensitive to changes in the evacuation dynamics.

Figure 21 shows the change in evacuation rates and traffic by grid cell for these population change experiments. In the Irma and Irma-RO scenarios, evacuation worsen in 2030/2040 across the southern half of the model, with notable impacts surrounding Tampa Bay-St. Petersburg and Miami-Ft. Lauderdale. This also occurs in Dorian, though to a lesser extent, mainly in the areas at risk (Figure 21c–d; red cells) that are most "upstream" i.e., southern portions of the areas at risk.

Results suggest that, in the absence of other changes, population growth and urbanization may worsen future evacuations. Its impacts may be most significant in heavily-trafficked, rapid-onset forecast scenarios, and across south Florida, which is further "upstream" with respect to traffic flow. More broadly, the value of the experiments is they show how agent-based frameworks like FLEE can explore evacuations across many different forecast-population scenarios, both real and hypothetical, and in doing so, emphasize the need for additional implementation of evacuation management strategies moving forward.

4. Summary and Conclusions

This paper explores the hurricane-forecast-evacuation system dynamics using an agent-based framework positioned to study interactions across the forecast-human-built

environment sub-systems. By changing key model parameters one-by-one, we investigate how evacuations change with public transportation, contraflow, evacuation order timing, population growth, urbanization, and different forecast scenarios affecting the Florida peninsula (e.g., Irma, Dorian, rapid-onset version of Irma). In viewing the results, we ask the following questions:

1) **How do the changes impact evacuation success?** Results suggest that forecasts, public transportation and carpooling, contraflow, and population growth significantly impact evacuation rates and traffic. Evacuation order timing is less important overall, but has significant, localized effects.

2) **How does this vary across forecast-population scenarios?** We find evidence that evacuations are less successful with population growth and rapid onset scenarios. We also find that evacuation management strategies are most effective in these heavily-trafficked scenarios.

3) **How are impacts distributed geographically?** Heavily-trafficked areas (e.g., urban areas at-risk and/or places “upstream” of traffic) benefit from evacuation management strategies the most. Population growth and urbanization particularly negatively impact southern Florida.

Additionally, we find evidence that non-linearities exist in the traffic portion of FLEE, where excessive traffic can worsen evacuations significantly. These results demonstrate how agent-based frameworks like FLEE are powerful virtual laboratories capable of investigating the system dynamics across many forecast-population-infrastructure

scenarios, real or synthetic, including non-linear effects. The findings presented are not intended to be definitive, but rather to improve understanding of the system dynamics and to provide a foundation for this type of work.

FLEE and/or models with additional detail provide several, immediate opportunities for future work. One avenue is to further investigate the system's non-linearities, including identifying saturation points to be avoided. Another is to unpack the role of evacuation order timing in evacuation success. A third area of immediate future work is to assess how changes in forecast track, intensity, storm size, forward speed, and uncertainty influence evacuation success in FLEE (see, e.g., Fossell et al. 2017). This would demonstrate to meteorologists how agent-based models offer a societally-relevant alternative to traditional measures of forecast accuracy (a need described by Morss 2005, Murphy 1993, Roebber and Bosart 1996), by measuring the impact of forecasts on different aspects of evacuation.

Looking further ahead, as empirical data on evacuation behaviors and traffic improves, the information can be codified into virtual laboratories like FLEE, thus increasing their realism, and subsequently, their ability to answer questions of interest. In this way, empirical and modeling studies feedback into each other and provide many future opportunities to advance our understanding of the hurricane forecast-evacuation system.

The research supports practitioners, policy-makers, and scholars in hazard risk management and related disciplines, thereby offering the promise of direct applications to save lives and mitigate hurricane losses. For example, practitioners can use these types of models to explore different scenarios and build understanding about which

evacuation strategies may be most effective, when, and where. Policymakers can use the information to identify resource needs in future forecast-population scenarios. And researchers particularly benefit from studying the system holistically, where cross-disciplinary understanding can be cultivated (Bostrom et al. 2016).

ARTICLE 3:

A NEW VERIFICATION APPROACH? USING COUPLED NATURAL-HUMAN MODELS TO EVALUATE FORECAST IMPACT ON EVACUATIONS

1. Introduction

Murphy (1993), Roebber and Bosart (1996), and Morss (2005) show how "accurate" forecasts are not necessarily useful ones. Hurricanes Irma (2017) and Rita (2005) exemplify this distinction. Irma's 3-10-day forecasts called for landfall as a major hurricane in southeast Florida near Miami–Ft. Lauderdale; this triggered the largest evacuation in US history (FDEM 2017), with Tampa Bay–St. Petersburg in west Florida as a common evacuation destination. However, as the storm approached, forecasts shifted westward, and the storm eventually made a landfall near Tampa Bay–St. Petersburg, while avoiding major impacts in Miami–Ft. Lauderdale altogether (Cangialosi et al. 2018; Wong et al. 2018). Meanwhile, uncertainties in Rita's track and intensity forecasts, combined with the aftermath of Hurricane Katrina, led to severe traffic in Houston–Galveston. The worst of the storm missed the city, but had it struck Houston–Galveston directly, the consequences could have been severe, as many evacuees were stranded on highways (Zhang et al. 2007; Knabb et al. 2006).

Irma and Rita's forecasts were accurate by meteorological standards, with Irma's westward shift falling within the National Hurricane Center's (NHC) cone of uncertainty (Cangialosi et al. 2018), and with Rita's track errors being less than average errors at the time (Knabb et al. 2006). However, the above discussion illustrates several ways in which the forecasts were less successful in providing useful guidance for evacuation decisions.

Thus, in addition to measuring skill in terms of errors in a hurricane's track and other meteorological characteristics, the National Academy of Sciences recommends the weather enterprise "measure the impact of forecasts" (p. 105; NASEM 2018) i.e., to increase forecast usefulness, traditional accuracy metrics should be supplemented with a new approach.

Coupled natural-human models are increasingly being used to model aspects of hurricane evacuations (e.g., Davidson et al. 2018; Blanton et al. 2018), including warning communication (e.g., Morss et al. 2017; Watts et al. 2018), evacuation-related decision-making (e.g. Yin et al. 2014; Widener et al. 2013; Zhang et al. 2009; Davidson et al. 2018), and evacuation traffic (e.g., Yang et al. 2019; Yi et al. 2017). These models provide a tool where empirical knowledge can be codified and used to run virtual experiments for many evacuation scenarios, real or synthetic. This suggests that coupled natural-human models could be used to study the impacts of forecast elements and uncertainties on evacuations.

With this in mind, Harris et al. (2021) (hereafter HRM21) designed a coupled-model framework, called FLEE (Forecasting Laboratory for Exploring the Evacuation-system), that models the key, interwoven aspects of hurricane evacuations: the natural hazard (hurricane), the human system (information flow, evacuation decisions), the built environment (road infrastructure), and connections between systems (forecasts and warning information, traffic, impact zones). Hurricane and forecast information are represented using archived National Hurricane Center (NHC) forecast products. Two agent-based models replicate the information flows, evacuee decision-making, built infrastructure, and evacuation traffic. By integrating the systems into a unified framework, FLEE becomes a "virtual laboratory" positioned to advance fundamental knowledge of

the system's dynamics and explore the role of hurricane forecasts in the forecast-evacuation system (e.g., see FLEE's conceptualization, implementation, validation, and proof of concept experiments in HRM21; use in studying evacuation management strategies and policies, evolving population characteristics, and forecast scenarios in Article 2 of the dissertation).

Using FLEE, this paper's objective is to explore how tropical cyclone forecast elements impact evacuations, and in doing so, to build towards the development of new verification approaches. Within the coupled-model framework, we perturb forecasts of track amounts typical of errors today (2021) – and in the past (2001) – and evaluate their impact on evacuations across both real and hypothetical forecast scenarios (e.g., Hurricane Irma, and Hurricane Dorian making landfall across east Florida). For these storms, we compare forecast track errors with intensity and forward speed errors characteristic of rapidly intensifying/onset scenarios, which are widely studied in meteorology.

Throughout the analysis, we ask:

1) What is the relative influence of changes in forecast elements (e.g., track vs. intensity)?

2) Do improvements in forecast accuracy over time (e.g., 2001–2021) translate to improved evacuations?

In answering these questions, we demonstrate how coupled natural-human models, and specifically agent-based models, offer a societally-relevant alternative to traditional metrics of forecast accuracy, by exploring the *impact* of forecasts elements and

uncertainties on how people receive and process the information, make evacuation decisions, and physically evacuate.

2. Design and Approach

To answer questions 1–2, we first determine what "typical" errors are for different elements of tropical cyclone forecasts, both today and in the past. For track and intensity, average forecast errors – and their trends over time – are documented and available at 0–120 hour lead times on the NHC website (https://www.nhc.noaa.gov/verification/index.shtml). Forward speed forecast errors are not readily available to the public, though they are noted to be "slightly larger than cross-track errors" (Fossell et al. 2017). Storm size forecast errors are also unavailable, as it is impossible to verify wind radii forecasts because they "are likely to have errors so large as to render a verification of official radii forecasts unreliable" (Cangialosi 2021). Based on data availability, we focus on NHC track and intensity forecasts errors.

To explore the role of forecast track on evacuations, our approach is to introduce perturbed tracks to the left and right of NHC's official track forecasts by distances equaling 2001 and 2021 average errors at 0–120 hour lead times (Table C1). Then, the left and right perturbed forecasts are ingested into FLEE and used to run evacuation simulations. By comparing the evacuation response from perturbed forecasts to those from the official forecasts, we assess the relative importance of track errors on evacuations.

There are several details to note. First, we compare average errors in 2021 with 2001, as this represents 20 years of improvements in track accuracy (also, 96–120 hour lead time errors were not readily available before 2001). Second, since forecast information

updates every 6 hours in FLEE (and in the real-world), the left/right perturbed tracks update every 6 hours. Third, the official forecasts were downloaded at https://www.nhc.noaa.gov/gis/ for Irma and Dorian; ArcGIS was used for data visualization and track perturbations.

To explore the role of forecast intensity on evacuations, we intended to introduce perturbed intensities higher/lower than original forecasts by amounts representing average errors in 2001 and 2021. However, average errors are less than 20 kts, even at long lead times. As a result, intensity errors are too small to effectively resolve in FLEE, where forecasts are synthesized into a green-yellow-orange-red “light system” forecast for all grid cells (see HRM21). Because of the limitation, we instead create rapid intensity/onset (RI/RO) forecast scenarios where intensity and forward speed errors are significant. In these scenarios, we shorten the forecast timeline of the original NHC forecasts (e.g., from 168 hours to 84 and 72 hours), while keeping the peak magnitudes of risk the same, and determine its effect on evacuations. By comparing the evacuation response from RI/RO forecast scenarios to those from the official forecasts, we begin to assess the relative importance of RI/RO on evacuations (note: in these RI/RO scenarios, track, intensity, and forward speed elements are intertwined; see Figures C5–6, C12–13).

To compare evacuation behaviors quantitatively, we track evacuation statistics for FLEE’s grid cells (note: FLEE’s grid is a 10 x 4 abstracted representation of the north-south axis of Florida; full description of FLEE’s virtual world and built environment in HRM21). The primary model outputs analyzed here are the percent of households that successfully evacuated (i.e., evacuation rates), and the percent who intended to evacuate but gave up due to traffic. The latter statistic provides insight to where the excessive traffic may be

preventing successful evacuations. In addition to displaying these output data by grid cells, we aggregate data into multiple impact zones, designed as first-order approximations of areas likely to experience different levels of impacts based on the actual meteorological conditions produced by the storms. In interpreting results, we compare metrics that might indicate successful outcomes in different ways e.g., high evacuation rates may not be “good” if the storm ends up not having much impact in those areas, and unnecessary evacuations may not matter if those at highest risk can get out safely.

The list of experiments is provided in Table 12. Along with a default simulation, simulations are run for the four perturbed track forecasts and two RI/RO forecasts across two storm scenarios: one real (Irma) and one hypothetical (Hurricane Dorian making landfall across east Florida). In the hypothetical case, Dorian and its forecasts were shifted west of the original track by 70 km so the storm impacts eastern Florida (Section 3.2). Together, these experiments allow us, to first-order, to 1) compare the relative influence of track forecast errors to RI/RO cases and 2) assess whether improvements in forecast accuracy over time translate to improved evacuations across different cases, real and synthetic.

Irma	Dorian (landfalling)
1. Default	8. Default
2. 2001 Left	9. 2001 Left
3. 2001 Right	10. 2001 Right
4. 2021 Left	11. 2021 Left
5. 2021 Right	12. 2021 Right
6. RI/RO	13. RI/RO
7. RI/RO – 12 hours	14. RI/RO – 12 hours

Table 12: Experiments in the study. The Default Irma and Dorian (landfalling) case uses official NHC forecasts for Irma and Dorian. However, Dorian's later forecasts are intentionally shifted westward to create a hypothetical scenario where the storm impacts eastern Florida. RI/RO cases contain default forecasts shortened from 168 hours to 84 hours, while the RI/RO-12 hours case contains forecasts shortened further to 72 hours. Track errors are introduced by shifting the forecast left/right from default by amounts equivalent to average errors in 2001 and 2021.

3. Results

3.1. Hurricane Irma

Irma's official (Table 12– default) forecasts are shown in Figure 22 at 24 hour intervals. Early forecasts place the entire model grid under threat, with the most likely outcome as a landfalling major hurricane near Miami. However, forecasts shifted westward as the storm approached, with the storm eventually making one mainland U.S. landfall as a Category 4 in the Florida Keys and a second landfall as a Category 3 in southwest Florida. Irma's hurricane-force winds (Figure 23a–b; dotted cells) impacted the western two-thirds of the model – particularly the southwest coastlines – while leaving the east-coast unscathed. Evacuation orders were issued along both coasts in FLEE (Figure 23a–b; red cells), matching what was observed, thus increasing our confidence that the evacuation order algorithm – and the synthesized NHC forecasts on which it is based – behave

realistically in the model context (Wong et al. 2018; Darzi et al. 2020; additional verification provided in HRM21).

Also shown on Figure 22 are perturbed left and right tracks based on average errors in 2001 (blue arrows) and 2021 (green arrows). Compared to default, the left perturbed tracks place more of south/western Florida under threat, while the right perturbed forecasts place more of north/eastern Florida under threat. Because average forecast errors decrease closer to landfall, the perturbed track forecasts eventually converge to the default forecast by 144 hours into the simulation. Note: full light system forecasts for perturbed tracks are provided in Figures C1–4; light system forecast for RI/RO cases are provided in Figures C5–6.

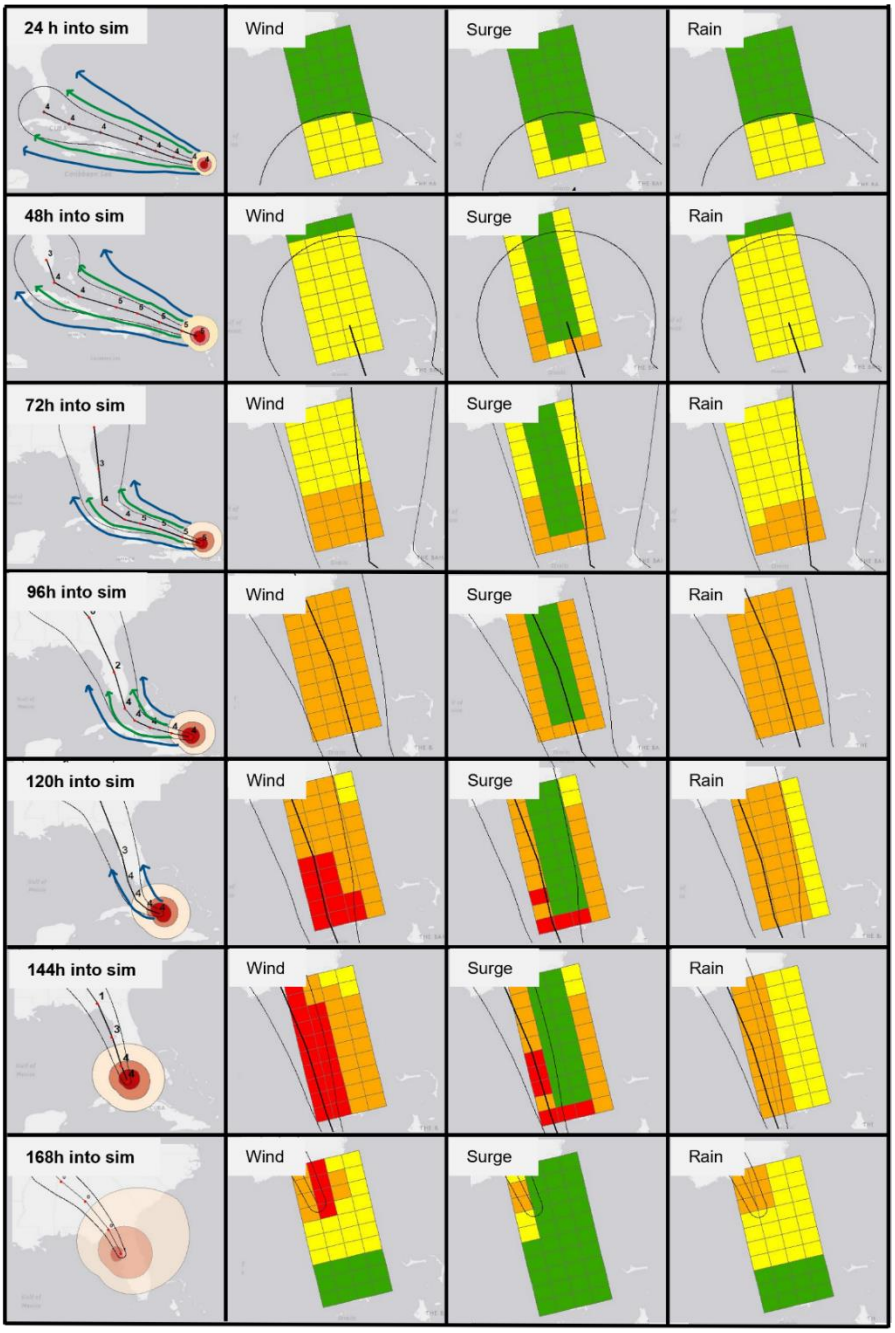

Figure 22: Default forecasts for Hurricane Irma (2017) as the storm approaches and travels through the Florida-like, model grid. Forecasts are shown at 24 hour intervals, but update every 6 hours in the model simulations (not shown). Left column: Evolving, official NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts are shown, including the 2001 left/right errors (blue arrows) and the 2021 left/right errors (green arrows). Right three columns: The light-system threats for wind, surge, and rain are shown for equivalent times in the simulation, with the forecast track (center black line) and cone of uncertainty (outer black lines) included for reference. Note: threats are highest when near the center of the forecast cone and when hazards are most imminent, among other factors.

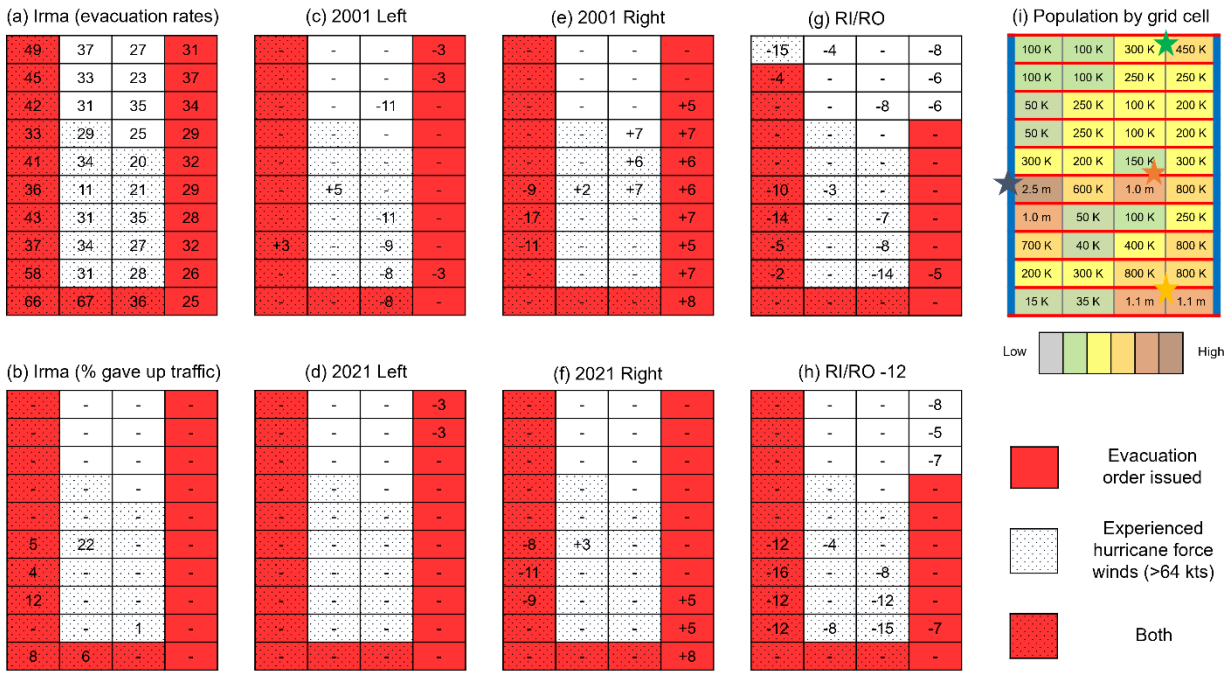

Figure 23: Irma's evacuation rates across grid cells. Evacuation rates are provided for the (a) default forecast and the (b) percent of households giving up due to excessive traffic. Evacuation rates are also presented for (c) 2001 left track, (d) 2021 left track, (e) 2001 right track, (f) 2021 right track, (g) RI/RO, and (h) RI/RO – 12 hour perturbed cases. Values (c–h) are expressed as the departure from the (a) default forecast evacuation rates. Also expressed is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (i). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

Using the default forecasts, Irma's simulated evacuation rates (Figure 23a) are similar to observational data, which suggest evacuation rates varied from 20–40% along Florida's east coast, to 40–70% across the south and west coasts, and around 10–30% inland (Wong et al. 2018; Long et al. 2020; Martin et al. 2020; Feng and Lin 2021). This was the largest evacuation in US history, meaning many households failed to evacuate due to excessive traffic. This occurs most frequently around Tampa Bay-St. Petersburg (Figure 23b), again similar to observations (Feng and Lin 2021; States et al. 2021).

Table 13 shows evacuation rates across impact zones. 32.0% of households on the model grid evacuate for Irma, which equals 5.24 million people. Estimates from the Florida Department of Emergency Management (2017) suggest actual evacuation numbers totaled 6.9 million. When considering households evacuating to local shelters in FLEE (not shown), the simulated evacuation rates match observations. For a given level of wind impact, evacuation rates are higher along the coasts than inland, which is to be expected (Feng and Lin 2021; Martin et al. 2020). Therefore, we believe FLEE's simulated evacuations provide a realistic baseline for interpreting results from Irma's other experiments.

Experiments	(a) Evacuation rates	(b) % Change relative to default	(c) People evacuated	(d) Coastal >64 knot zone	(e) Inland >64 knot zone	(f) Coastal < 64 knot zone	(g) Inland < 64 knot zone	(h) % Gave up to traffic	(i) Gave up to traffic
Irma Default	32.0	0	5.24 m	39.3	24.5	29.6	29.4	2.5	410 K
2001 Left	30.2	-1.8	5.10 m	37.7	22.2	27.8	27.4	3.2	541 K
2021 Left	31.9	-0.1	5.39 m	39.4	24.7	29.0	29.4	2.4	406 K
2021 Right	30.9	-1.1	5.22 m	33.2	25.1	32.8	29.5	4.7	794 K
2001 Right	31.4	-0.6	5.31 m	31.6	26.7	35.1	30.0	5.5	929 K
RI/RO	27.3	-4.7	4.61 m	32.0	20.3	27.2	28.4	6.3	1.06 m
RI/RO -12	26.3	-5.7	4.44 m	27.7	18.2	26.5	25.9	7.6	1.28 m

Table 13: Irma's evacuation behaviors averaged across all grid cells for the different experiments. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts) and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

How do the experiments with Irma's perturbed tracks translate to evacuations? The 2001 left errors increase evacuation rates across southern and western Florida and reduce evacuation rates further east (Figure 23c). This results in fewer evacuees overall (Table 13a–c), including fewer evacuees across both >64 knot and <64 knot zones (Table 13d–g). The 2021 left errors had considerably smaller changes. Relative to default, slightly fewer people evacuated unnecessarily across northeast Florida (Figure 23d). The 2001 right errors (Figure 23e) resulted in many unnecessary evacuations along the east coast, and a reduction in evacuation rates across southern and western Florida. The 2001 right errors had the worst outcomes of the track experiments, with a 7.7% reduction in evacuation rates across the most impacted coastal >64 knot zone (Table 13d). The 2021 right errors show a similar pattern, but the outcomes are not quite as bad, with a 6.1% reduction in evacuation rates across the most impacted zones (Table 13d; Figure 23f). When considering the results together, 2021 errors had better outcomes than 2001 – suggesting the value of smaller forecast errors – while left errors had better outcomes than right. The latter is not a general result – left errors are not necessarily better than right generally, but are a function of areas affected by any particular error i.e., the impact of errors is very much case/event dependent.

How does this compare with Irma's RI/RO cases? In both cases, RO/RI leads to a larger reduction in evacuation rates (Table 13a–b) and a greater increase in evacuation traffic than with the track errors (Table 13h–i). This includes a significant reduction in evacuation rates across more-impacted and less-impacted areas (Table 13d–g). The latter is partially because evacuation orders (Figure 23g–h, red cells) are not issued around Jacksonville in these cases (Figure 23i, green star), as the forecast's westward shift happens quicker

than in default, removing the area from risk before evacuation orders were issued. When comparing the two RI/RO cases, the case that was 12 hours shorter resulted in worse evacuation outcomes than RI/RO, particularly when considering the decreases across the >64 knot zone, suggesting the extra time for evacuation makes a difference (Table 13b–i; Figure 23g–h). Compared to track errors, the two RI/RO cases more negatively affect evacuation outcomes than the perturbed track cases, suggesting RI/RO is a critical feature for evacuation success.

3.2. Hurricane Dorian (landfalling)

Dorian's (landfalling) forecasts are shown in Figure 24. Early forecasts place the entire Florida peninsula under threat, with the most likely scenario as a landfalling major hurricane along Florida's east coast. Because of the forecasts, evacuation orders were issued along the majority of the east coast (Figure 25a–b; red cells). The exception is in Miami-Ft. Lauderdale, which avoided evacuation orders, matching what was observed (TIME 2019). In the actual Dorian case, the storm remains nearly-stationary over the Bahamas for many hours, before accelerating northward and missing Florida's east coast. However, in this hypothetical case, we shift the forecasts westward by one grid cell (70 km) 120–168 hours into the simulation. This creates a scenario where Dorian's hurricane-force winds impact Florida's east coast as the storm accelerates northward (e.g., see Figure 25a–b; dotted cells).

Also shown on Figure 24 are perturbed left and right tracks based on average errors in 2001 (blue arrows) and 2021 (green arrows). Compared to forecasts in the default case, left tracks place southern and western Florida under threat, while right tracks place more

of northern and eastern Florida under threat. Because average forecast errors decrease as the storm approaches landfall, the perturbed track forecasts eventually converge to default by 144 hours into the simulation. Note: light system forecasts for perturbed tracks is provided in Figures C8–11; light system forecast for RI/RO cases are provided in Figures C12–13.

Evacuation rates were lower with Dorian (landfalling) (16.8%) than Irma (32.0%), which is 2.4 million less evacuees (Table 14a–c). Due to fewer evacuees overall – and the evacuation being spread over a longer time (Figures C7, C14) – fewer people give up due to traffic (0.0%) in Dorian than Irma (2.5%). These outcomes seem reasonable, given that both the forecasts and the storm itself influence a smaller portion of the model domain in Dorian (landfalling) compared to Irma.

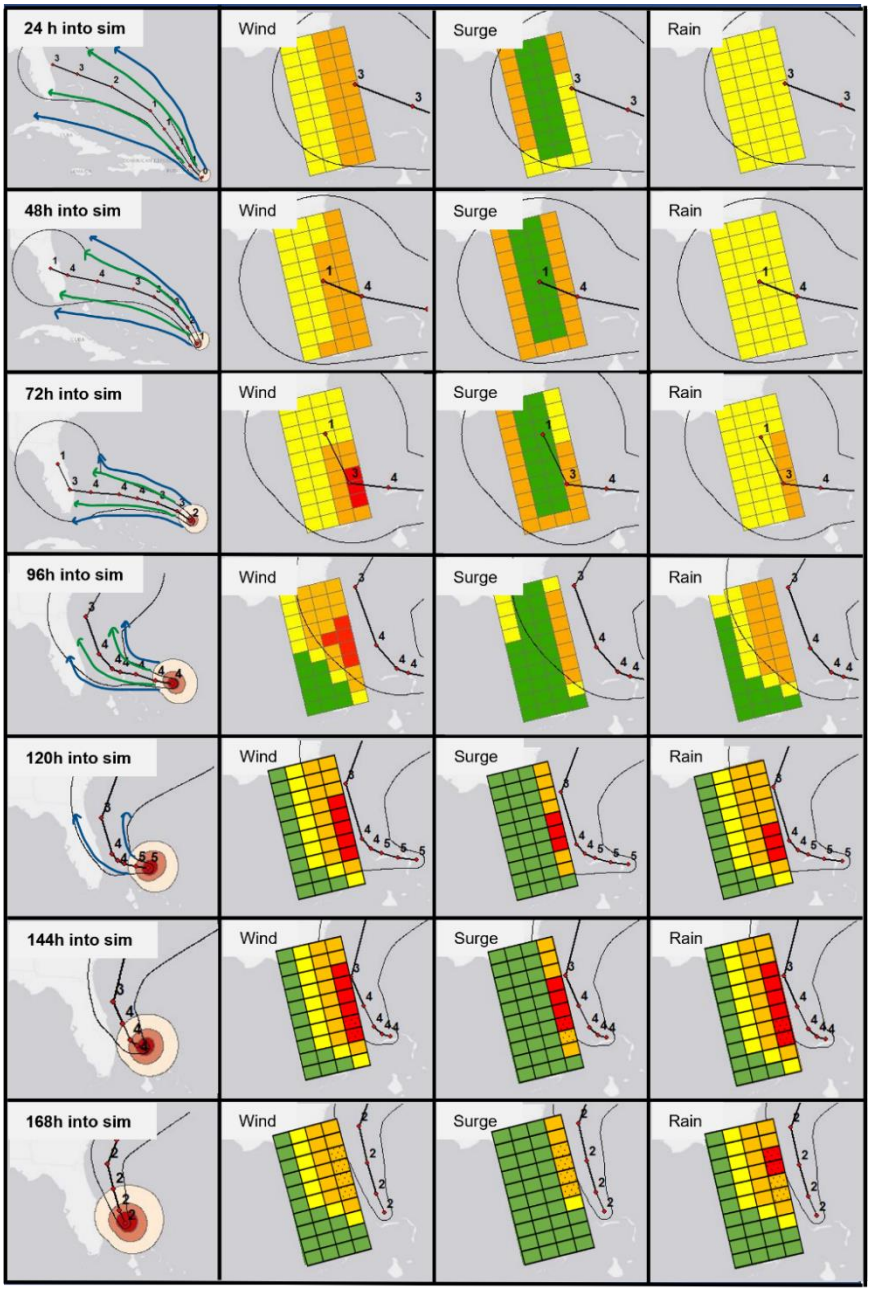

Figure 24: Default forecasts for Hurricane Dorian (2019) (landfalling) as the storm approaches and travels through the Florida-like, model grid. Forecasts are shown at 24 hour intervals, but update every 6 hours in the model simulations (not shown). Left column: Evolving, official NHC forecast track for the original Dorian (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts are shown, including the 2001 left/right errors (blue arrows) and the 2021 left/right errors (green arrows). Right three columns: The light-system threats for wind, surge, and rain are shown for equivalent times in the simulation, with the forecast track (center black line) and cone of uncertainty (outer black lines) included for reference. Note: threats are highest when near the center of the forecast cone and when hazards are most imminent, among other factors. Note: the 120-168 h hour forecasts have been shifted westward by 70 km to create a hypothetical scenario where Dorian makes landfall along the east coast.

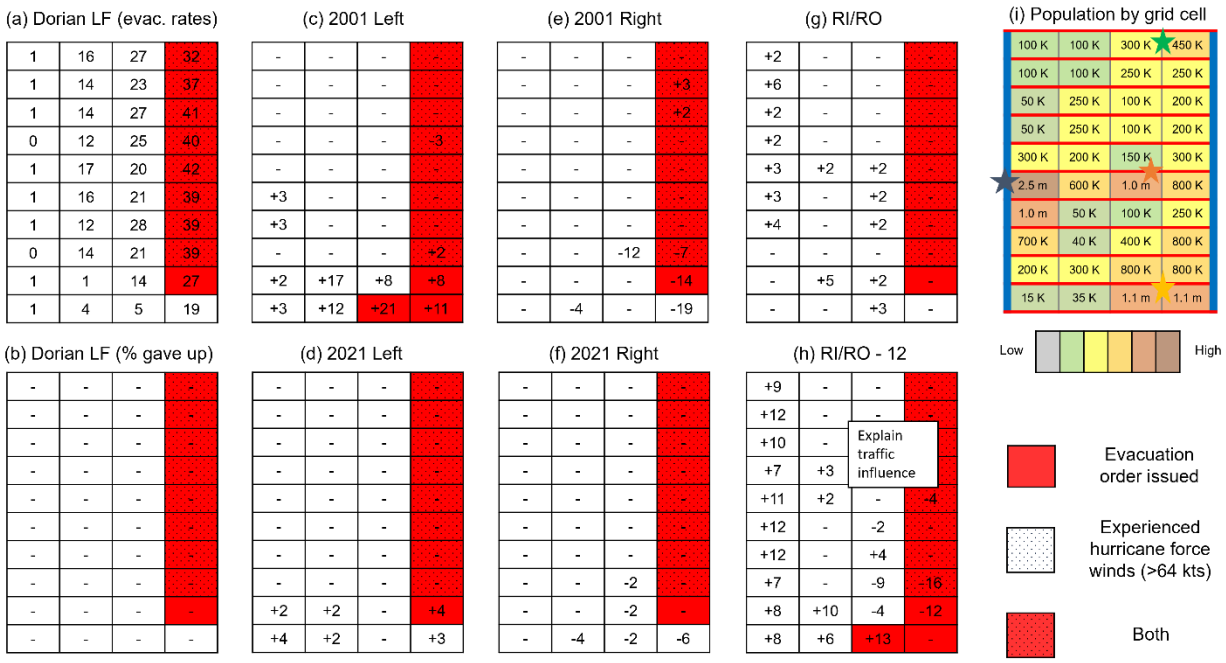

Figure 25: Dorian LF's evacuation rates across grid cells. Evacuation rates are provided for the (a) default setting and (b) the percent of households giving up due to excessive traffic. Evacuation rates are also presented for (c) 2001 left track, (d) 2021 left track, (e) 2001 right track, (f) 2021 right track, (g) RI/RO, and (h) RI/RO – 12 hour cases. Values (c–h) are expressed as the departure from the (a) default forecast evacuation rates. Also expressed is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (i). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star).

Experiments	(a) Evacuation rates	(b) % Change relative to default	(c) People evacuated	(d) Coastal >64 knot zone	(e) Inland >64 knot zone	(f) Coastal < 64 knot zone	(g) Inland < 64 knot zone	(h) % Gave up to traffic	(i) Gave up to traffic
Dorian LF default	16.8	0	2.84 m	39.1	-	7.0	18.2	0.0	32
2001 Left	20.6	+ 3.8	3.48 m	39.4	-	13.5	19.8	0.1	10 K
2021 Left	17.4	+ 0.6	2.94 m	39.1	-	8.3	17.8	0.0	460
2021 Right	15.5	- 1.3	2.61 m	39.0	-	5.2	16.9	0.2	35 K
2001 Right	12.2	- 4.6	2.06 m	37.3	-	1.3	13.4	1.1	187 K
RI/RO	18.3	+ 1.5	3.09 m	39.0	-	9.2	19.3	0.1	17 K
RI/RO -12	19.3	+ 2.5	3.26 m	33.9	-	14.7	17.3	3.0	486 K

Table 14: Dorian LF's evacuation behaviors averaged across all grid cells for the different experiments. Evacuation rates, percent change in evacuation rates relative to the default simulation, and the total number of people evacuated are shown. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts) and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

How do perturbed tracks translate to evacuations in this scenario? 2001 left errors trigger more evacuations than default (Table 14b), with significant increases around Miami–Ft. Lauderdale where additional evacuation orders were issued (Figure 25c; red cells). With the 2021 left errors, evacuation orders were not issued in these regions (Figure 25d), leading to evacuation rates that more closely resembled the default forecast scenario. The simulation with 2001 right errors resulted in lower evacuation rates (Table 14a–b), with fewer evacuees across southeast Florida, and slightly more evacuees up north (Figure 25e). Of the perturbed tracks, 2021 right errors arguably had the best evacuation outcomes overall, with a reduction in unnecessary evacuations across southern Florida (Figure 25f). When considering the results together, it appears the directionally of errors

matters to evacuations, particularly in how population centers are affected. Furthermore, experiments suggest that 2001 errors were large enough to shift the issuance of evacuation orders around Miami–Ft. Lauderdale – which had significant impacts on evacuation rates and traffic – whereas 2021 errors did not have such a substantial impact.

The RI/RO cases create significant increases in evacuation rates (Table 14a–b) and evacuation traffic (Table 14h–i) compared to the default simulation. The former is surprising, as there is less time to evacuate everyone safely, and is opposite to Irma's results. Upon closer examination, the increase in evacuation rates occurs primarily across western Florida (Figure 25g-h) where the shorter timeline forced evacuation decisions early when forecasts were uncertain. As with the Irma case, the 12 hour difference between the RI/RO scenarios is significant. For example, in addition to creating unnecessary evacuations across western Florida, the RI/RO – 12 scenario decreases evacuation rates along the east coast (Figure 25h), reducing evacuation rates across the most impacted zones (Table 14d). In summary, compared to track errors, the RI/RO cases more negatively impacted evacuation outcomes than perturbed tracks, suggesting RI/RO is a critical feature, and that NHC's efforts to improve RI/RO forecasting are important (e.g., DeMaria et al. 2021).

4. Summary and looking ahead

This paper demonstrates how coupled natural-human models can be employed to explore how changes in hurricane forecasts affect evacuations. Specifically, within one such framework (FLEE), we create RI/RO cases and track scenarios representative of accuracy errors today (2021) – and in the past (2001) – and evaluate their impact on

evacuations across real and hypothetical forecast scenarios (e.g., Hurricane Irma, Hurricane Dorian making landfall across east Florida).

Our analysis of the results provides first-order assessment of the following questions:

1) **What is the relative influence of changes in forecast elements (e.g., track vs. RI/RO)?** Our results confirm that RI/RO scenarios – which are already believed to be a problem in meteorology (e,g., see DeMaria et al. 2021) – are a significant feature with respect to evacuation success, as outcomes from the RI/RO simulations are generally worse than those from perturbed track experiments. RI/RO scenarios emphasize how track, intensity, and forward speed elements are intertwined.

2) **Do improvements in forecast accuracy over time (2001–2021) translate to improved evacuation outcomes?** We provide evidence suggesting average track errors in 2001 more negatively impact evacuations than average track errors in 2021. For example, 2001 errors in the hypothetical Dorian scenario changed the issuance of evacuation orders in some areas, triggering many unnecessary evacuations, whereas the 2021 errors did not. Furthermore, an additional 12 hours of forecast lead time between the RI/RO cases and the RI/RO – 12 cases can significantly improve evacuations in these scenarios as well.

Additionally, we provide evidence suggesting that hurricane evacuation outcomes are sensitive to the forecast-population-infrastructure scenarios, e.g., evacuation outcomes appear especially sensitive when metropolitan areas – which require extra time for evacuations and evacuation orders – are on the edges of the cone of uncertainty. Results

from this study are not intended to be definitive; rather they demonstrate, to a first-order, how coupled natural-human models, and specifically agent-based models, offer a societally-relevant alternative to traditional metrics of forecast accuracy, by exploring the *impact* of forecasts elements/uncertainties on evacuations.

Coupled natural-human and agent-based models provide many opportunities for future work. First, models with a more sophisticated representation of forecast intensity could be used to further tease out the different effects of track, intensity, and forward speed errors on evacuations e.g., to determine which elements are most important and should be the focus of research and forecasting efforts. Second, the models could explore additional questions such as: are there diminishing returns in improving aspects of forecast accuracy on evacuation? How much does human input over models and ensembles translate to evacuation success? And do forecasts help some groups evacuate more than others? Similarly, models can be extended to additional weather phenomenon such as tornadoes, potentially transforming public warning and protection scenarios in these areas as well.

The models show promise for helping meteorology in the long term. As computing power increases, and as empirical data on hurricane evacuation behaviors and traffic improves, additional information can be codified into computational models, thus increasing their realism, and subsequently, their ability to answer questions of interest. This emphasizes the importance of funding social and behavioral science research within the weather enterprise, and shows how, together, empirical and modeling studies provide many opportunities to advance our understanding of the forecast's role in the hurricane forecast-evacuation system.

ARTICLE 4

A LITERATURE REVIEW ON INEQUITIES IN HURRICANE EVACUATIONS

1. Introduction

Throughout this dissertation, we introduce an ABM built to investigate the complex dynamics of the hurricane-forecast-evacuation system i.e., to determine which factors are most important, key connections, and how factors interact across a range of real or synthetic scenarios. The ABM framework, called FLEE, includes models of the natural hazard (hurricane), the human system (information flow, evacuation decisions), the built environment (road infrastructure), and connections between systems (forecasts and warning information, traffic). Carless households are excluded from the modeling framework, as these groups contribute less to evacuation traffic. However, FLEE, and subsequent models like it, can be designed to include these groups, depending on the research goals.

In this brief literature review, we investigate who carless households are in the US, whether inequities exist in hurricane evacuations as a result, and how future hurricane evacuation models can use the information to account for the inequities. The idea is to help 1) ensure hurricane evacuation models like FLEE do not contribute to any inequities by excluding minoritized groups, and 2) demonstrate how models can explore inequities in the system across many forecast-population-infrastructure scenarios.

Specifically, we ask the following questions regarding inequities in hurricane evacuations:

1. What research has been done regarding *who* can (and cannot) safely evacuate before a hurricane i.e., who are the carless?

2. If inequities exist, how are current evacuation plans and policies addressing (or contributing to) these inequities?

3. How can computational modeling frameworks help be a solution?

2. Inequities in evacuation (the carless)

The question “who evacuates and why?” has motivated 40+ empirical studies investigating how people respond to hurricane risks for actual hurricane events and in hypothetical hurricane scenarios (see, e.g., reviews in Huang et al. 2016; Bowser and Cutter 2015; Lazo et al. 2015; Dash and Gladwin 2007). Although predictors of evacuation decisions vary across studies, a metaanalysis by Huang et al. (2016) found some common factors. These include risk perceptions, official evacuation orders, storm characteristics, and personal/situational characteristics such as mobile home residence. Socioeconomic and demographic variables such as race, income, and disabilities have effects on evacuations that are not entirely clear from these empirical studies (Bowser and Cutter 2015; Huang et al. 2016). However, *car ownership is a common predictor of evacuation,* with the carless being considerably less likely to evacuate (Huang et al. 2016). That said, we ask, who are the carless?

Answering this question is important to hurricane evacuations, as a non-insignificant number of people are carless. Across the US, 8% of households are carless, a value that has been fairly constant over the last decade. According to the 2017 census, the percentage of carless households is highest in US coastal cities where hurricanes

frequent (e.g. New York City: 56%, Washington D.C.: 37%, Baltimore: 36%, Boston: 35%, New Orleans: 27%, Providence: 22%, Miami: 20%, Tampa Bay: 11%, Orlando: 9%, Houston: 8%).

When the carless population is stratified by demographics, huge racial disparities become apparent. For example, 7% of white households are carless, compared to 24% of African-American households, 17% of Latino households, and 13% of Asian-American households (Sanchez et al. 2003). This makes African Americans and Latinos nearly 3.5 times and 2.5 times more likely to be carless than whites, respectively. People of color make up a greater proportion of the nation's largest coastal cities (e.g., Miami: 79%; New Orleans: 73%; New York City: 65%; Houston: 58%). Of all 11 major cities that have seen 5+ hurricanes in the last 100 years (Houston, Miami-Ft. Lauderdale, Orlando, Jacksonville, Tampa Bay-St. Petersburg, New York City, Providence, Boston, New Orleans), those without a car are disproportionately people of color (Lui et al. 2006). As described by Bullard and Wright (2011) a "transportation apartheid, a two-tiered system of people with cars and people without cars, is alive and well in most metropolitan regions" and is "firmly and nationally entrenched in American society." These inequities are apparent in who has access to vehicles for hurricane evacuations.

3. Existing evacuation policies (or lack thereof)

Disasters such as hurricanes — and the inequities which result — have social, political, and economic roots (Smith 2006) i.e., it has been emphasized by many in the hazards community that there is no such thing as a natural disaster. Thus, it is important to examine hurricane evacuation policies to identify the roots of inequities in car-lessness,

its effects on hurricane evacuations, and whether sufficient “transportation alternatives” are provided.

Following Katrina, several nationwide evaluations of government evacuation plans were conducted. These reports suggest that most plans – at the federal, state, and local levels – assume citizens can evacuate on their own and provide little assistance to those who cannot. For example:

- A U.S. Government Accountability Office report determined that “state and local governments are generally not well prepared to evacuate transportation disadvantaged populations” (US GAO 2006).
- The Department of Homeland Security’s assessment of evacuation plans for each state and the 75 largest urban areas found “low-mobility and special-needs groups, while included in most state emergency operation plans, has been largely unaddressed by state DOTs” (DHS 2006).

A recent CDC report suggests the lack of government assistance for the transportation disadvantaged remains a systematic, nationwide problem. According to Kruger et al. (2020),

> “Analysis of evacuation policies in eight southern U.S. coastal states in 2018 found that all have laws to execute evacuation orders. However, only four have laws that require informing racially and ethnically diverse populations and persons with disabilities and functional needs of emergency evacuation plans. Only one state

(Florida) authorized creation of a registry for persons with access and functional needs for the purposes of evacuation and sheltering."

In other words, the personal automobile remains the primary means of emergency evacuation in most hurricane evacuation plans with little to no assistance provided otherwise. These plans effectively privilege the middle-to upper-class, able-bodied, non-elderly, predominantly white households more likely to own cars, leaving behind less mobile residents of American society.

4. Implications for computational models

The literature demonstrates inequities in car ownership, and subsequently, inequities in who can easily evacuate. This is important context when interpreting results from modeling studies, such as those in Articles 1–3 of this dissertation. Moving forward, models aiming to improve the forecast-evacuation-system should include representations of the carless, or at the very least, acknowledge their existence. Doing so ensures the modeling frameworks do not contribute to inequities by excluding these groups when studying the evacuation-system. Ideally, the models should be used to demonstrate the advantages of using buses or providing cars for the carless (e.g., experiments in Article 2). In other words, modeling studies should quantify the *value* of government assistance for various hurricane scenarios and thus encourage the development of equitable policies across the country.

We are entering new territory where we can quantify forecast value through computational models (e.g., experiments in Article 3). As a result, we need to establish a conception of forecast and evacuation success which places the wellbeing of minoritized and

disadvantaged peoples at the forefront. Ideas to help achieve this goal include 1) explicitly considering the impact of forecasts on traditionally underserved communities and making this the primary means of measuring forecast/evacuation success in research and operational settings, and 2) emphasizing and supporting inclusive research techniques such as community-based participatory research and co-production which empower communities. With these types of considerations in mind, coupled natural-human models of hurricane evacuation can be used to improve equity across the hurricane forecast-evacuation system.

CONCLUSIONS

The aim of this research was to holistically examine the dynamics of the hurricane-forecast-warning system from a novel, agent-based perspective that compliments and builds-upon empirical work. This was accomplished through a series of manuscripts discussing the following:

1. The agent-based model (FLEE)'s conceptualization, implementation, and validation.

2. How FLEE's evacuations change with evacuation management strategies and policies, evolving population characteristics, and real and synthetic forecast scenarios.

3. How changes in forecast elements (e.g., track and intensity) impact FLEE's evacuations, and whether accuracy improvements translate to evacuation success.

4. The carless households in the US, the inequities which result, and ideas for inclusive evacuation modeling moving forward.

Together, the studies provide a first-order assessment of the system's dynamics, and demonstrate the ability of coupled natural-human models to study – and improve – hurricane evacuations. In other words, this dissertation serves as a starting point for examining the forecast-evacuation system's dynamics *holistically* i.e., to establish the relative importance of factors, key interactions between systems, and the broader, potentially non-evident emergent patterns.

Moving forward, coupled natural-human modeling – and specifically agent-based modeling – offers a promising research tool for the weather enterprise and hazards communities. Firstly, *models will improve as computing power increases*. As things currently stand, FLEE takes several days to run individual cases. One can envision, however, models being used to run evacuation simulations in real or quasi-real-time. Increased computing power also allows for increased level of detail in the models, and subsequently, ability to answer questions of interest. Secondly, *models can be extended to study other regions or hazards*, e.g., tornadoes, hurricanes followed by flooding, loss of power networks, damage to roads, and other cascading failures. For meteorology, this continues the process of improving forecast evaluation across the weather enterprise, which helps make forecasts more useful to society. Thirdly, *additional in-depth comparisons with observational data can improve FLEE's realism*, and subsequently, its capability to study evacuations i.e., computational models improve as empirical data improves, and vice versa, effectively creating a feedback relationships between the two, and an area of research that is promising long-term. For this reason, I intend to continue agent-based hazards modeling, while simultaneously broadening expertise to include empirical social science knowledge and methods (e.g., surveys, interviews etc.) over the next few years. More broadly, I aim to be someone who can effectively straddle multiple scholarly worlds (e.g., empirical vs modeling worlds, meteorology vs hazards worlds), and in turn, continue to offer innovative perspectives to support the weather enterprise.

REFERENCES

Alfieri, L., P. Salamon, F. Pappenberger, F. Wetterhall, and J. Thielen, 2012: Operational early warning systems for water-related hazards in Europe. *Env. Sci. Policy.*, **21**, 35–49. https://doi.org/10.1016/j.envsci.2012.01.008.

Baker, E., 1991: Hurricane Evacuation Behavior. *Int. J. Mass Emerg. Disasters,* **9**, 287–310.

Ballard, A. J. and D. W. Borchardt, 2006: Recommended practices for hurricane evacuation traffic operations. College Station, Texas: Texas Transportation Institute. [Available online at https://static.tti.tamu.edu/tti.tamu.edu/documents/0-4962-P2.pdf].

Barton, C. M., 2014: Complexity, social complexity, and modeling. *J. Archaeol. Method Theory*, **21**, 306–324. https://doi.org/10.1016/j.ancene.2015.12.004.

Bazghandi, 2012: Techniques, Advantages, and Problems of Agent-Based Modeling for Traffic Simulation. *IJCSI*, **9**, 115-119. https://core.ac.uk/download/pdf/25749192.pdf

Blanton, B., and coauthors, 2018: An integrated scenario ensemble-based framework for hurricane evacuation modeling: Part 2–Hazard modeling. *Risk Anal*., **40**, 117–133. DOI: 10.1111/risa.13004.

Bostrom, A., R. E. Morss, J. K. Lazo, J. L. Demuth, H. Lazrus, and R. Hudson, 2016: A Mental Models Study of Hurricane Forecast and Warning Production, Communication, and Decision-Making. *Wea. Climate Soc.*, **8**, 111–129. https://doi.org/10.1175/WCASD-15-0033.1.

Bostrom, A., R. E. Morss, J. K. Lazo, J. Demuth, and H. Lazrus, 2018: Eyeing the storm: How residents of coastal Florida see hurricane forecasts and warnings. *Int. J. Disaster Risk Reduct*., **30**, 105–119. https://doi.org/10.1016/j.ijdrr.2018.02.027.

Bowser, G. C., and S. L. Cutter, 2015: Stay or Go? Examining Decision Making and Behavior in Hurricane Evacuations. *Env: Sci. Pol. Sust. Dev*., **57**, 28–41. https://doi.org/10.1080/00139157.2015.1089145.

Brommer, D. M., and J. C. Senkbeil, 2010: Pre-landfall evacuee perception of the meteorological hazards associated with Hurricane Gustav. *Nat. Hazards*, **55**, 353–369. https://doi.org/10.1007/s11069-010-9532-7.

Bullard R. & Wright B. (2011). Race, Place, and Environmental Justice After Hurricane Katrina: Struggles to Reclaim, Rebuild, and Revitalize New Orleans and the Gulf Coast.

Cangialosi, J. P., A. S. Latto, and R. Berg, 2018: Tropical Cyclone Report. Hurricane Irma 30 August–12 September 2017. National Hurricane Center. [Available online at https://www.nhc.noaa.gov/data/tcr/AL112017_Irma.pdf].

Cangialosi, J. P., 2019: National Hurricane Center forecast verification report: 2019 hurricane season. NOAA/NHC Tech. Rep. 7-10 pp. [Available online at https://www.nhc.noaa.gov/verification/pdfs/Verification_2019.pdf].

Cava, D. L., 2018: Hurricane Irma Emergency Evacuation Report and Recommendations. Florida Department of Transportation. [Available at https://www.miamidade.gov/district08/library/irma-after-report.pdf].

Chen, X., 2008: Microsimulation of Hurricane Evacuation Strategies of Galveston Island. *The Professional Geographer,* **60**, 160-173. https://doi.org/10.1080/00330120701873645.

Chen, X., 2012: Agent-based micro-simulation of staged evacuations. *International Journal of Advanced Intelligence Paradigms,* **4***.* DOI: 10.1504/IJAIP.2012.046964

Chen, X., J. W. Meaker, and F. B. Zhan, 2006: Agent-Based Modeling and Analysis of Hurricane Evacuation Procedures for the Florida Keys. *Nat. Hazards,* **38,** 321. https://doi.org/10.1007/s11069-005-0263-0.

Chen, X., and F. B. Zhan, 2004: Agent-based modeling and simulation of urban evacuation: relative effectiveness of simultaneous and phased evacuation strategies. *83rd Annual Meeting of the Transportation Research Board*, Washington, DC.

Chen, X., and F. B. Zhan, 2008: Agent-based modelling and simulation of urban evacuation: relative effectiveness of simultaneous and staged evacuation strategies, *J Oper. Res. Soc*., **59**, 25–33. https://doi.org/10.1057/palgrave.jors.2602321.

Chiu, Y-C., H. Zheng, J. A. Villalobos, W. G. Peacock, and R. Henk, 2008: Evaluating regional contra-flow and phased evacuation strategies for Texas using a largescale dynamic traffic simulation and assignment approach. *J. Homel. Secur. Emerg. Manag.,* **5**. 1547–7355. DOI: 10.2202/1547-7355.1409.

Cuite, C. L., R. L. Shwom, W. K. Hallman, R. E. Morss, and J. L. Demuth, 2017: Improving coastal storm evacuation messages. *Wea. Climate Soc*., **9**, 155–170. https://doi.org/10.11 75/WCAS-D-16-0076.1.

Cutter, S. L., 2019: The Science, Skill, and Luck Behind Evacuation Order Calls. *The Conversation*. https://theconversation.com/the-science-skill-and-luck-behind-evacuation-order-calls-103209.

Darzi, A., and coauthors. 2021. “Constructing Evacuation Evolution Patterns and Decisions Using Mobile Device Location Data: A Case Study of Hurricane Irma.” *ArXiv*.

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

Davidson, R., and coauthors, 2018: An integrated scenario ensemble-based framework for hurricane evacuation modeling: Part 1—decision support system. *Risk Anal*., **40**, 97–116. https://doi.org/10.1111/risa.12990.

DeMaria M, Franklin JL, Onderlinde MJ, Kaplan J. Operational Forecasting of Tropical Cyclone Rapid Intensification at the National Hurricane Center. *Atmosphere*. 2021; 12(6):683. https://doi.org/10.3390/atmos12060683

Demuth, J. L., B. H. Morrow, and J. K. Lazo, 2009: Weather forecast uncertainty information: An exploratory study with broadcast meteorologists. *Bull. Amer. Meteor. Soc*., **90**, 1614–1618. ttps://doi.org/10.1175/2009BAMS2787.1

Demuth, J. L., R. E. Morss, B. Hearn-Morrow, and J. K. Lazo, 2012: Creation and communication of hurricane risk information. *Bull. Amer. Meteor. Soc.*, **93**, 1133–1145. . https://doi.org/10.1175/BAMS-D-11-00150.1.

Demuth, J. L., R. E. Morss, J. K. Lazo, and C. Trumbo, 2016: The effects of past hurricane experiences on evacuation intentions through risk perception and efficacy beliefs: A mediation analysis. *Wea. Climate Soc*., **8**, 327–344. https://doi.org/10.1175/WCASD-15-0074.1.

Demuth, J. L., R. E. Morss, L. Palen, K. M. Anderson, J. Anderson, M. Kogan, K. Stowe, M. Bica, H. Lazrus, O. Wilhelmi, and J. Henderson, 2018: “Sometimes da #beachlife ain't always da wave”: Understanding people’s evolving hurricane risk communication, risk assessments, and responses using Twitter narratives. *Weather, Climate, and Society*, **10**, 537–560. https://doi.org/10.1175/WCAS-D-17-0126.1.

Dixit, V.V., and E. Radwan, 2009: Optimal scheduling of evacuation orders for cities. *88th Annual Meeting of the Transportation Research Board*, Washington DC. DOI: 10.3141/1964-26.

Dixon, D.S., P. Mozumder, W. F. Vasquez, and H. Gladwin, 2017: Heterogeneity within and across households in hurricane evacuation response. *Netw. Spat. Econ*., **17**, 645–680. https://doi.org/10.1007/s11067-017-9339-0.

Dow, K., and S. L. Cutter, 2000: Emerging Hurricane evacuation issues: Hurricane Floyd and South Carolina. *Nat. Hazards Rev*., **3**, 12–18. https://doi.org/10.1016/S1464-2867(01)00014-6.

Dye, K. C., J. P. Egers, and Z. Shapira, 2014: Trade-offs in a tempest: Stakeholder influence on hurricane evacuation decisions. *Organization Science,* **25**, 1009–1025. https://doi.org/10.1287/orsc.2013.0890.

Fang, L., and P. Edara, 2014: Mobility benefits of intermediate crossovers on contraflow facilities during hurricane evacuation. *Trans. Res. Rec.,* **2459**, 37–46. https://doi.org/10.3141/2459-05.

Feng, K., and N Lin, 2021: Reconstructing and analyzing the traffic flow during evacuation in Hurricane Irma (2017). *Trans. Res. Part D,* **94***,* https://doi.org/10.1016/j.trd.2021.102788.

Flanagan, B. E., E. W. Gregory, E. J. Hallisey, J. L. Heitgerd, and B. Lewis, 2011: A social vulnerability index for disaster management. *J. Homel. Secur. Emerg. Manage.,* **8***,* 1–22. DOI: 10.2202/1547-7355.1792.

FDEM (Florida Division of Emergency Management), 2018: Regional Emergency Management Liaison Team. [Available at: https://www.floridadisaster.org/dem/directors-office/regions/]

Florida Department of Transportation. 2020. "Technical Memorandum Projections of Florida Population by County, 2020-2070." https://www.fdot.gov/planning/demographic/default.shtm

Florida Statewide Regional Evacuation Study Program, 2019: Regional Evacuation Transportation Analysis. Technical Data Report South Florida Region. [Available at: http://sfregionalcouncil.org/wp-content/uploads/2019/08/Vol1-11_ChVI.pdf].

Fossell, K. R., D. Ahijevych, R. E. Morss, C. Snyder, and C. Davis, 2017: The practical predictability of storm tide from tropical cyclones in the Gulf of Mexico. *Monthly Weather Review*, **145**, 5103–5121. . https://doi.org/10.1175/MWR-D-17-0051.1.

Gehlot, H., X. Zhan, X. Qian, and C. Thompson, 2019: A-RESCUE 2.0: A High-Fidelity, Parallel, Agent-Based Evacuation Simulator. *Journal of Computing in Civil Engineering,* **33.**

Ghorbanzadeh M., and coauthors. 2021. "Spatiotemporal Analysis of Highway Traffic Patterns in Hurricane Irma Evacuation." *Trans. Res. Rec.* **9**, 2625, https://doi.org/10.1177/03611981211001870

Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S.A., Milles, A., Nabe-Nielsen, J, Polhill, J. Gareth, R, Viktoriia, R, Marie-Sophie, S, Richard A., T, Jan C. and Ayllón, D., 2020: The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism. *Journal of Artificial Societies and Social Simulation,* **23**, 7 doi: 10.18564/jasss.4259

Hammond R. A., 2015: Considerations and Best Practices in Agent-Based Modeling to Inform Policy. *Institute of Medicine*. [Available at: https://www.ncbi.nlm.nih.gov/books/NBK305917/]

Harris, A. R., P. J., Roebber, and R. E., Morss, 2021: An agent-based modeling framework for examining the dynamics of the hurricane-forecast-evacuation system*. Int. Jour. Dis. Risk Reduct.* 67. https://doi.org/10.1016/j.ijdrr.2021.102669

Hasan, S., R. Mesa-Arango, and S. Ukkusuri, 2013: A random-parameter hazard-based model to understand household evacuation timing behavior. *Trans. Res. C*, **27**, 108-116. https://doi.org/10.1016/j.trc.2011.06.005.

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

Huang, S.-K., M. K. Lindell, and C. S. Prater, 2017: Multistage Model of Hurricane Evacuation Decision: Empirical Study of Hurricanes Katrina and Rita. *Nat. Haz. Rev*., **18**, 1527–6996. . https://doi.org/10.1061/(ASCE)NH.1527-6996.0000237.

Huang, S.-K., M. K. Lindell, and C. S. Prater, H-C. Wu, and L. K. Siebeneck, 2012: Household evacuation decision making in response to Hurricane Ike. *Nat. Hazards Rev*., **13**, 283–296. . https://doi.org/10.1061/(ASCE)NH.1527-6996.0000074.

Keim, B. D., R. A. Muller, and G. W. Stone, 2007: Spatiotemporal Patterns and Return Periods of Tropical Storm and Hurricane Strikes from Texas to Maine. *J. Climate*, **20**, 3498–3509. DOI: 10.1175/JCLI4187.1.

Knabb, R. D., D. P. Brown, and J. R. Rhome, 2006: Tropical Cyclone Report. Hurricane Rita 18-26 September 2005. National Hurricane Center. [Available at https://www.nhc.noaa.gov/data/tcr/AL182005_Rita.pdf].

Kruger J, Smith MJ, Chen B, et al. Hurricane Evacuation Laws in Eight Southern U.S. Coastal States — December 2018. MMWR Morb Mortal Wkly Rep 2020;69:1233–1237. DOI: http://dx.doi.org/10.15585/mmwr.mm6936a1

Lazrus, H., O. Wilhelmi O, R. E. Morss, J. Henderson, and A. Dietrich, 2020: Information as Intervention: How Hurricane Risk Communication Interacted with Vulnerability and Capacities in Superstorm Sandy. *Int. J. Mass Emerg. Disasters*, **38**, 89–120.

Lazo, J. K., A. Bostrom, R. E. Morss, J. L. Demuth, and H. Lazrus, 2015: Factors affecting hurricane evacuation intentions. *Risk Anal*., **35**, 1837–1857. https://doi.org /10.1111/risa.12407.

Liang, W., N. N. Lam, X. Qin, and W. Ju, 2015: A Two-level Agent-Based Model for Hurricane Evacuation in New Orleans, *J. Homel. Secur. Emerg. Manag.*, ***12***, 407-435.

Lindell, M. K., J-C. Lu, C. S. Prater, 2005: Household decision making and evacuation in response to Hurricane Lili. *Nat. Hazards Rev*., **6**, 171–179. DOI: 10.1061/(ASCE)1527-6988(2005)6:4(171).

Lindell, M. K., and R. W. Perry, 2012: The Protective Action Decision Model: Theoretical modifications and additional evidence. *Risk Anal*., **32**, 616–632. https://doi.org/10.1111/j.1539-6924.2011.01647.x.

Long, E. F., Chen M. K., and Rohla R, 2020: Political storms: emergent partisan skepticism of hurricane risks. *Sci Adv*., **6**, 37. DOI: 10.1126/sciadv.abb7906.

Lui, M., E. Dixon, and B. Leondar-Wright. (2006). Stalling the Dream: Cars, Race and Hurricane Evacuation. Boston: United for a Fair Economy. January 10.

Madireddy, M., D. J. Medeiros, and S. Kumara, 2011: An agent based model for evacuation traffic management. *Proceedings of the 2011 Winter Simulation Conference (WSC)*, Phoenix, AZ, pp. 222-233. DOI: 10.1109/WSC.2011.6147753.

Martin R., and M. Schlüter, 2015: Combining system dynamics and agent-based modeling to analyze social-ecological interactions—an example from modeling restoration of a shallow lake. *Frontiers in Environmental Science,* **3**, 66. https://doi.org/10.3389/fenvs.2015.00066.

McGhee, C.C., and M. C. Grimes, 2006: An Operational Analysis of the Hampton Roads Hurricane Evacuation Traffic Control Plan. Richmond VA: Virginia Department of Transportation. [Available online at https://rosap.ntl.bts.gov/view/dot/19729].

Meyer, R. J., J. Baker, K. Broad, J. Czajkowski, and B. Orlove, 2014: The Dynamics of Hurricane Risk Perception: Real-Time Evidence from the 2012 Atlantic Hurricane Season. *Bull. Amer. Meteor. Soc.*, **95**, 1389–1404. https://doi.org/10.117 5/BAMS-D-12-00218.1.

Mileti, D. S., and J. H. Sorensen, 1990: Communication of emergency public warnings: A social science perspective and state-of-the-art assessment. Oak Ridge National Laboratory Manuscript, 166 pp. [Available online at https://www.osti.gov/servlets/purl/6137387].

Miller, J. H., and S. E. Page, 2007: *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press, 284 pp.

Mitchell, S., and E. Radwan, 2006: Heuristic priority ranking of emergency evacuation staging to reduce clearance time. *Trans. Res. Rec*.,**1964**, 219–228. DOI: 10.3141/1964-24.

Mongold E., Davidson R. A., Trivedi J., DeYoung S., Wachtendorf T., and Anyidoho P, 2020: Hurricane evacuation beliefs and behaviour of inland vs. coastal populations, *Env. Haz.* https://doi.org/10.1080/17477891.2020.1829531.

Morss, R. E., 2005: Problem Definition in Atmospheric Science Public Policy: The Example of Observing-System Design for Weather Prediction. *Bull. Amer. Meteor. Soc.*, **86**, 181–192. https://doi.org/10.1175/BAMS-86-2-181.

Morss, R. E., J. L. Demuth, H. Lazrus, L. Palen, C. M. Barton, C. A. Davis, C. Snyder, O. V. Wilhelmi, K. M. Anderson, D. A. Ahijevych, J. Anderson, M. Bica, K. R. Fossell, J. Henderson, M. Kogan, K. Stowe, and J. Watts, 2017: Hazardous weather prediction and communication in the modern information environment. *Bulletin of the American Meteorological Society*, **98**, 2653–2674. https://doi.org/10.1175/BAMS-D-16-0058.1.

Morss, R. E., J. L. Demuth, J. K. Lazo, K. Dickinson, H. Lazrus, and B. H. Morrow, 2016: Understanding public hurricane evacuation decisions and responses to hurricane risk messages. *Wea. Forecasting*, **31**, 395–417. https://doi.org/10.1175/WAFD-15-0066.1.

Morss, R. E., and M. H. Hayden, 2010: Storm surge and "certain death": Interviews with Texas coastal residents following Hurricane Ike. *Wea. Climate Soc.*, **2**, 174–189. https://d oi.org/10.1175/2010WCAS1041.1.

Morss, R. E., H. Lazrus, A. Bostrom, J. Demuth, 2020: The influence of cultural worldviews on people's responses to hurricane risks and threat information, *J. Risk Res.* https://doi.org/10.1080/13669877.2020.1750456.

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281–293. https://doi.org/10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2

Murray-Tuite, P., M. K. Lindell, B. Wolshon, E. J. Baker, 2019*: Large-Scale Evacuation: The Analysis, Modeling, and Management of Emergency Relocation from Hazardous Area: 1st Edition*. Taylor & Francis.

NASEM (National Academies of Sciences, Engineering, and Medicine) 2018. Integrating Social and Behavioral Sciences Within the Weather Enterprise. Washington, DC: The National Academies Press.

NOAA, 2021: What is storm surge? Storm surge is the rise in seawater level caused solely by a storm. National Ocean Service website, https://oceanservice.noaa.gov/facts/stormsurge-stormtide.html, accessed on 6/3/21.

Peacock, W. G., S. D. Brody, and W. Highfield, 2005: Hurricane risk perceptions among Florida's single family homeowners. Landscape Urban Plann., **73**, 120–135. https://doi.org/10.1016/j.landurbplan.2004.11.004

Petrolia, D. R., S. Bhattacharjee, and T. R. Hanson, 2011: Heterogeneous evacuation responses to storm forecast attributes. *Nat. Hazards Rev*., **12**, 117–124. DOI: 10.1061/(ASCE)NH.1527-6996.0000038.

Rand, W., and Rust, R. T., 2011: Agent-Based Modeling in Marketing: Guidelines for Rigor. *International Journal of Research in Marketing*. Available at SSRN: https://ssrn.com/abstract=1818543.

Rezapour, M., and T. E. Baldock, 2014: Classification of Hurricane Hazards: The Importance of Rainfall. *Wea. Forecasting*, **29**, 1319–1331. https://doi.org/10.1175/WAF-D-14-00014.1

Roache, M. 2019. "At Least 1 Million People in These Counties Are Under Evacuation Orders As Hurricane Dorian Approaches." TIME. https://time.com/5666941/hurrican-dorian-evacuation-orders-counties/

Robinson, R. M., A. J. Khattak, J. A. Sokolowski, P. Foytik, and X. Wang, 2009: Role of Traffic Incidents in Hampton Roads Hurricane Evacuations. *88th Annual Meeting of the Transportation Research Board*, Washington, DC.

Roebber, P. J., and L. F. Bosart, 1996: The Complex Relationship between Forecast Skill and Forecast Value: A Real-World Analysis. *Wea. Forecasting*, **11**, 544–559. https://doi.org/10.1175/1520-0434(1996)011<0544:TCRBFS>2.0.CO;2.

Sadri, A. M., S. V. Ukkusuri, and H. Gladwin: 2017: The role of social networks and information sources on hurricane evacuation decision making. Nat. Haz. Rev., **18**, 04017005.

Sadri, A. M., S. V. Ukkusuri, P. Murray-Tuite, and H. Gladwin, 2014: How to Evacuate: Model for Understanding the Routing Strategies during Hurricane Evacuation. J. Trans. Eng., **140**, 61-69. DOI: 10.1061/(ASCE)TE.1943-5436.0000613.

Sanchez, T. A., R. Stolz, and J. Ma. (2003). Moving to Equity: Addressing the Inequitable Effect of Transportation Policies on Minorities. Cambridge, MA: *The Civil Rights Project*, Harvard University

Sbayti, H., and H. S. Mahmassani, 2006: Optimal scheduling of evacuation operations. *85th Annual Meeting of the Transportation Research Board*, Washington, D.C. https://doi.org/10.1177/0361198106196400126.

Senkbeil, J., J. Collins, and J. Reed, 2019: Evacuee Perception of Geophysical Hazards for Hurricane Irma. *Wea. Climate Soc.*, **11**, 217–227. https://doi.org/10.1175/WCAS-D-18-0019.1.

Smith, N. (2006): There's no such thing as a natural disaster. Insights from the Social Sciences.https://items.ssrc.org/understanding-katrina/theres-no-such-thing-as-a-natural-disaster/

Staes B., N, Menon, R. L., Bertini. 2021." Analyzing transportation network performance during emergency evacuations: Evidence from Hurricane Irma." *Trans. Res. D*., 102841, https://doi.org/10.1016/j.trd.2021.102841

Sun, Z., I. Lorscheid, J. D., Millington, S. Lauf, N. R. Magliocca, J. Groeneveld, S. Balbi, H. Nolzen, B. Müller, J. Schulze, and C. M. Buchmann, 2016: Simple or complicated agent-based models? A complicated issue. *Environ. Model. Softw.,* **86**, 56–67. https:// doi.org/10.1016/j.envsoft.2016.09.006.

Swamy R., J. E. Kang, R. Batta, and Y Chung. 2017. "Hurricane evacuation planning using public transportation." *Socio-economic planning sciences.* **59**, 43-55. https://doi.org/10.1016/j.seps.2016.10.009

Trainor, J. E., P. Murray-Tuite, P. Edara, S. Fallah-Fini, and K. Triantis, 2012: Interdisciplinary approach to evacuation modeling. Nat. Haz. Rev., **14**, 151-16. DOI: 10.1061/(ASCE)NH.1527-6996.0000105.

Ukkusuri, S. V., S. Hasan, B. Luong, K. Doan, X. Zhan, P. Murray-Tuite, and W. Yin., 2017: A-RESCUE: An Agent based Regional Evacuation Simulator Coupled with User Enriched Behavior. *Networks and Spatial Economics*. 17, 197–223.

U.S. Department of Transportation and U.S. Department of Homeland Security (2006). Catastrophic Hurricane Evacuation Plan Evaluation: A Report to Congress. Washington, DC: U.S. DOT and U.S. DHS. June 1.

U.S. Government Accountability Office (2006). Transportation-Disadvantaged Populations: Actions Needed to Clarify Responsibilities and Increase Preparedness for Evacuations. Washington, DC: U.S. GAO. December.

Watts, J., R. E. Morss, C. M. Barton, and J. L. Demuth, 2019: Conceptualizing and implementing an agent-based model of information flow and decision making during hurricane threats. *Environ. Model. Softw*, **122**, 104524. https://doi.org/10.1016/j.envsoft.2019.104524.

Widener, M. J., M. W. Horner, and S. S. Metcalf, 2013: Simulating the effects of social networks on a population's hurricane evacuation participation. *J. Geogr. Syst*., **15**, 193–209. https://doi.org/10.1007/s10109-012-0170-3.

Wolshon, B. 2001. 'One-way-out': contraflow freeway operation for hurricane evacuation. *Nat. Hazards Rev.,* **2**, 105–112. https://doi.org/10.1061/(ASCE)1527-6988(2001)2:3(105).

Wolshon, B., and L. Lambert, 2004: Convertible Roadways and Lanes. NCHRP Synthesis of Highway Practice 340. Washington DC. https://doi.org/10.17226/23331.

Wong, S., S. Shaheen, and J. Walker, 2018: Understanding evacuee behavior: a case study of hurricane Irma. Transportation Sustainability Research Center, Berkeley. [Available online at: https://escholarship.org/uc/item/9370z127].

Yang, K., R. A. Davidson, B. Blanton, B. Colle, K. Dresback, R. Kolar, and T. Wachtendorf, 2019: Hurricane evacuations in the face of uncertainty: Use of

integrated models to support robust, adaptive, and repeated decision-making. *International Journal of Disaster Risk Reduction*, 101093.

Yi, W., L. K. Nozick, R. A. Davidson, B. Blanton, and B. A. Colle, 2017: Optimization Of The Issuance Of Evacuation Orders Under Evolving Hurricane Conditions. *Trans. Res. Part B,* ***95***, 285-304. https://doi.org/10.1016/j.trb.2016.10.008.

Yin, W., P. Murray-Tuite, S. V. Ukkusuri, and H. Gladwin, 2014: An agent-based modeling system for travel demand simulation for hurricane evacuation. *Transp. Res. Part C.,* **42**, 44–59. https://doi.org/10.1016/j.trc.2014.02.015.

Zhan, F. B., and X. Chen, 2008: Agent-Based Modeling and Evacuation Planning. *Geospatial Technologies and Homeland Security,* **99**, 189–208.

Zhang, F., R. E. Morss, J. A. Sippel, T. K. Beckman, N. C. Clements, N. L. Hampshire, J. N. Harvey, J. M. Hernandez, Z. C. Morgan, R. M. Mosier, S. Wang, and S. D. Winkley, 2007: An in-person survey investigating public perceptions of and response to Hurricane Rita forecasts along the Texas Coast. *Weather and Forecasting*, **22**, 1177–1190. https://doi.org/10.1175/2007WAF2006118.1/

Zhang, B., W. K. Chan and S. V. Ukkusuri, 2009: Agent-based modeling for household level hurricane evacuation. *Proceedings of the 2009 Winter Simulation Conference*, Austin, TX, pp. 2778-2784. DOI: 10.1109/WSC.2009.5429248.

Zhang, X., and G. Chang. 2014. “A Transit-Based Evacuation Model for Metropolitan Areas.” *Jour. Pub. Trans*. **17**, 129-147. 10.5038/2375-0901.17.3.9

Zhang, Z., K. Spansel, and B. Wolshon, 2014: Effect of phased evacuations in megaregion highway networks. Transp. Res. Rec., **2459**, 101–109. https://doi.org/10.3141/2459-12.

Zhu, Y. J., Y. Hu, and J. M. Collins, 2020: Estimating road network accessibility during a hurricane evacuation: A case study of hurricane Irma in Florida. *Transp. Res. Part D*, **83**, 102334. https://doi.org/10.1016/j.trd.2020.10233

APPENDIX A: Article 1

Wind risk	*Forecast category	Forecast wind field	Location in cone of uncertainty	*Expected arrival time
1	<TS	None	Fully outside	Outside cone
2	TS-1	>34 kts	Mostly outside	>96 h
3	2-3	>50 kts	Mostly inside	48-96 h
4	4-5	>64 kts	Fully inside	<48 h
Weight	15%	50%	20%	15%
*Conditional upon being in the cone of uncertainty and/or within forecast wind radii (=1 otherwise)				

Table A1: Wind risk is calculated at each grid cell by assigning a risk score (1-4) based on the storm's forecast category at that location, its location in the forecast wind field (34, 50, 64+ knot intervals) which depicts the size of the storm, location in the cone of uncertainty, and expected arrival time of tropical storm force winds. The scores are weighted, summed, and rounded to the nearest integer to provide an overall wind threat score (1-4) expressed as green-yellow-orange-red, respectively. Note: scores for the forecast category and expected arrival time are set to 1 if the grid cell is not situated within the cone of uncertainty and/or any forecast wind radii. When taken together, the products capture the wind's critical forecast elements (e.g., storm's track, intensity, size, forward speed, amount of uncertainty, evolution with time, imminency).

Surge risk	*Inundation potential	*Forecast category	*Forecast wind field	*Approach angle	*Location in cone of uncertainty	*Expected arrival time
1	None	<TS	None	Outside cone	Fully outside	Outside cone
2	Weak	TS-1	>34 kts	Left of track	Mostly outside	>96 h
3	Moderate	2-3	>50 kts	Right of track, track parallel to shore	Mostly inside	48-96 h
4	High	4-5	>64 kts	Right of track, track perpendicular to shore	Fully inside	<48 h
Weight	12%	12%	25%	16%	20%	15%
*Conditional upon being along shoreline (i.e., =1 inland) and upon being inside cone of uncertainty and/or within forecast wind radii (=1 otherwise)						

Table A2: Surge risk is determined at each grid cell by assigning a risk score (1-4) based on the cell's inundation potential (estimated using NHC's potential storm surge inundation products), expected category at that location, location within the forecast wind field (34, 50, 64+ knot intervals) which depicts the size of the storm, the storm's approach angle, the location in the cone of uncertainty, and the expected arrival time of tropical storm force winds. The scores are weighted, summed, and rounded to the nearest integer to provide an overall surge threat score (1-4) expressed as green-yellow-orange-red, respectively. Note: scores for the expected category and expected arrival time are set to 1 if the grid cell is not situated within the cone of uncertainty and/or the forecast wind radii. Likewise, the values are only calculated for areas along the shoreline, as storm surge does not occur inland.

Rain risk	*Storm speed	Forecast wind field	Location in cone of uncertainty	*Expected arrival time
1	Fast	None	Fully outside	Outside cone
2	Medium	>34 kts	Mostly outside	*>96 h
3	Slow	>50 kts	Mostly inside	*48-96 h
4	Nearly stationary	>64 kts	Fully inside	*<48 h
Weight	30%	35%	20%	*15%
*Conditional upon being inside cone of uncertainty and/or within forecast wind radii (=1 otherwise)				

Table A3: Rain risk is calculated for each grid cell by assigning a risk score (1-4) based on the storm speed (>15 knots, 10-15 knots, 5-10 knots, and <5 knots), location within the forecast wind field (34, 50, 64+ knot intervals) which estimates the size of the rain field, location in the cone of uncertainty, and the expected arrival time of tropical storm force winds. The scores are weighted, summed, and rounded to the nearest integer to provide an overall rain threat score (1-4) expressed as green-yellow-orange-red, respectively. Note: scores for the expected category and forecast period are set to 1 if the grid cell is not situated within the cone of uncertainty and/or the forecast wind radii. When taken together, the products capture the rain's critical forecast elements (e.g., storm's track, size, forward speed, amount of uncertainty, evolution with time, imminency).

Variable	Variability in time	Definition	Values
Wind	Dynamic	Score from light system normalized to a 0-100 scale	Green = 0, Yellow = 33, Orange = 66, Red = 100
Surge	Dynamic	Score from light system normalized to a 0-100 scale	Green = 0, Yellow = 33, Orange = 66, Red = 100
Rain	Dynamic	Score from light system normalized to a 0-100 scale	Green = 0, Yellow = 33, Orange = 66, Red = 100
Forecast	Dynamic	The highest score from the wind, rain, and surge threats. This is used in a household's risk assessment	Values range from 0-100
Evacuation Orders	Dynamic	Is an evacuation order issued for the household's grid cell (yes/no)? This is used in a household's risk assessment	If yes = 100. If no = 0
Mobile Home Ownership	Static	Is the household in a mobile home (yes/no)? This is used in a household's risk assessment	If yes = 100. If no = 0
Age	Static	1-5 score from the household's grid cell (Figure A1) normalized to 0-100 scale. This is used in a household's risk assessment	If 1=20, If 2=40, If 3=60, If 4=80, If 5=100
Household risk assessment	Dynamic	The sum of the forecast, evacuation orders, mobile home ownership, and age factors	Values range from 0-400
Evacuation barrier	Static	If household has a car and household's risk assessment > socioeconomic barrier, household will evacuate	Car ownership and socioeconomic barrier in Table A5

Table A4: Key variables in the household evacuation decision-making algorithm. The algorithm's inputs (i.e., forecast, evacuation orders, mobile home ownership, age) are normalized onto a 0-100 scale and summed to produce household risk assessment, which is then weighed against evacuation barriers to produce a decision.

Socioeconomic status				Car ownership				Age				Mobile home ownership			
5	4	4	3	4	3	3	4	5	3	3	3	5	5	5	1
5	4	4	1	3	4	3	1	5	3	3	5	5	4	3	3
4	4	2	1	2	2	4	1	3	3	3	3	4	3	4	2
4	4	3	2	2	1	2	3	1	1	5	3	3	3	2	3
3	1	3	2	3	3	3	2	3	3	1	3	3	3	1	3
2	3	3	3	4	4	2	2	5	5	1	3	2	3	2	3
1	5	5	3	2	2	2	3	1	1	5	3	3	4	4	3
3	5	5	1	2	5	4	3	1	3	3	3	2	5	2	1
1	2	2	2	2	2	4	4	3	5	3	3	1	1	1	1
1	1	3	3	3	4	5	5	5	5	3	3	2	2	1	1

Figure A1: Cell-by-cell distribution of agent characteristics identified by Huang et al. (2016) as being important determinants of hurricane evacuations. These characteristics are spatially distributed by subjectively projecting the county-level social vulnerability data (see Flanagan et al. 2011) onto the abstracted, Florida-like agent-based model grid. Note, for reference, grid cells are 69 km by 69 km each. Higher values for socioeconomic status and car ownership increase the evacuation barriers and thus reduce the likelihood of evacuation. Higher values for age and mobile home ownership increase evacuation intentions.

Household characteristics	Variability in time	Definition	Values
Socioeconomic status	Static	Establishes the evacuation barrier threshold. Low (high) values indicate grid cell has less (more) financial obstacles to evacuate	If = 1, barrier = random between 5-105 If = 2, barrier = random between 10-110 If = 3, barrier = random between 15-115 If = 4, barrier = random between 20-120 If = 5, barrier = random between 25-125
Car ownership	Static	Establishes whether a household owns a vehicle. Carless households do not evacuate	If = 1, 96% of households own car If = 2, 94% of households own car If = 3, 93% of households own car If = 4, 91% of households own car If = 5, 89% of households own car
Mobile home ownership	Static	Establishes whether a household lives in a mobile home. If home is mobile, will increase risk perception	If = 1, 5% of houses are mobile If = 2, 10% of houses are mobile If = 3, 20% of houses are mobile If = 4, 33% of houses are mobile If = 5, 46% of houses are mobile

Table A5: Prescribing agent characteristics to individual households. At the beginning of the simulation, FLEE checks the agent's location and subsequent values in Figure A1, then stochastically assigns household characteristics at the values established above. These variables are static, meaning they are assigned at the beginning of the simulation and do not change, but serve as inputs into the agent decision-making algorithm as detailed in Table A4.

Variable weight	Variability in time	Definition	Values
Evacuation order	Static	Trust in evacuation orders from EMs	Random between 0-1
Forecast	Static	Trust in forecast information i.e., the light system	Random between 0-0.8
Mobile home	Static	Agent belief in whether their housing type influences perceived risk	Random between 0-10
Age	Static	Agent belief in whether household age influences perceived risk	Random between 0-0.1
Wind	Static	Household's perceived vulnerability to wind	Random between 0.1-1
Surge	Static	Household's perceived vulnerability to surge	Random between 0-1
Rain	Static	Household's perceived vulnerability to rain	Random between 0-0.9

Table A6: Weighting of key variables in a household's risk assessment. Weights are designed to reflect the relative importance of each factors (e.g., evacuation orders, forecast information, mobile home ownership, and age, in that order) as established in Huang et al. (2016). For the individual hazards, studies suggest most households perceive wind and surge as the primary threat over rain (e.g., Senkbeil et al. 2019). But in general, the relative weighting is not well known.

Variable	Variability in time	Definition	Default values
Departure times	Static	Time between when an agent decides to evacuate and when they actually leave	Random between 0-12 hours
Destinations (out-of-state)	Static	The number of evacuees who evacuate out-of-state	50% of the bottom 4 rows of grid cells; 100% of top 6 rows
Destinations (in-state)	Static	The number of accommodations available in each grid cell for in-state-evacuees	½ of grid cell's population (i.e., metros have more accommodations)
Patience threshold	Dynamic	Household patience i.e., the amount of time a household is willing to spend waiting to get onto a heavily trafficked road	Random between 0 and the estimated time of arrival of tropical storm force wind
Left/right	Static	Agents in the bottom row of grid cells can choose between moving westward/eastward on the lower interstate	40% westward, 60% eastward
Erratic drivers	Static	Percent of time steps (1.2 seconds) in which a driver may act "erratically" by randomly slowing down	0.05%
Random accident frequency	Static	The frequency of accidents along the two outer interstates i.e., I-95 and I-75. These stop traffic for 10 minutes.	1-3 random accidents per hour

Table A7: Key variables for the traffic agent-based model. These parameters are the default settings for the experiments detailed in Section 4.1. Static variables are assigned once a vehicle decides to evacuate and does not change, whereas dynamic variables do change throughout the simulation.

Supplementary Results 1 - Varying household's weighting of different types of information

Next, we investigate the effects of changing household agent's weightings of the four factors that influence their hurricane risk assessment: the forecast, evacuation orders, mobile home ownership, and age. For each experiment, we set the information weights to zero, effectively "turning off" each parameter, one-by-one, while holding the others constant (Table 3e). When comparing the results to the default settings in Section 4.1, the experiments demonstrate the specific influence of the different information on the evacuation behaviors, both spatially and temporally.

In Irma's default simulation (Section 4.1), 45.1% of households evacuate. However, turning off the information for evacuation orders, the forecast, mobile home, and age, one-by-one, results in evacuation rates of 28.3%, 33.2%, 40.6%, and 44.8%, respectively. Similarly, in the default simulation, where 10.5% of households give up due to traffic, turning off the inputs reduces the rate to 2.6%, 8.1%, 9.8%, and 9.3%, respectively (Table A8). In other words, the results indicate that, in the model's current formulation, evacuation rates are generally more sensitive to evacuation orders than they are to forecast information, mobile home ownership, and age. However, this is zone dependent e.g., evacuation orders has a greater influence in coastal zones, and mobile homes have a greater influence upstate/inland. The former is due to model formulation (evacuation orders are limited to coastal zones) and the latter due to the geographic distribution of mobile homes (e.g., as shown in Figure A1). That said, we cannot draw conclusions (or interpret the model dynamics) based on these findings. Rather, we can say the relative importance of these factors is generally consistent with the metaanalysis of Huang et al.

(2016), which we used to prescribed the information weightings, thus adding confidence that the model behaves reasonably.

Breaking down the experiments by impact zones shows that, as expected, evacuation orders primarily impact evacuations along the coast. For example, turning off the evacuation order parameter decreases evacuation rates in the coastal >64 knot zone from 52.3% to 31.9%, while inland evacuation rates remain the same (Table A8).

	% Successfully evacuated							
Experiment	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 kts zone	Inland < 64 kts zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5
EO = 0	28.3	31.9	24.0	25.5	36.9	29.0	26.9	2.6
Forecast = 0	33.2	38.2	11.7	45.3	26.5	41.4	17.5	8.1
MH = 0	40.6	48.5	17.3	54.9	14.2	51.4	16.6	9.8
Age = 0	44.8	51.3	22.3	56.4	35.0	53.6	25.3	9.3

Table A8: Evacuation behaviors by impact zone when varying household weighting of information. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

Figure A2 shows evacuation rates and traffic broken down by grid cell. Note, in this figure, rates are expressed as the departure from the default settings in Figure 7. The results further show how evacuation orders are a strong determinant of evacuation rates, as turning off the parameter reduces evacuation rates from 7% to 40% in places along the coast (Figure A2b). Note: turning off evacuation orders increases evacuation rates in the inland Miami suburbs, as traffic is reduced in the surrounding coastal areas. This highlights how evacuation rates in a given grid cell are also influenced by those in other grid cells. Unlike evacuation orders, the other three parameters (Figure A2 a, c-d) exhibit a more uniform influence on evacuation rates across FLEE's grid. Areas most influenced by mobile home and age information occur in grid cells where rates of mobile home ownership are highest, and where age is expected to play a larger role (Figure A1, see cells with higher ranking). Though such information does not provide any new behavioral insights, it does verify that FLEE behaves as expected given the model's current configuration, and is capable of capturing complex processes (e.g., evacuation behaviors in one part of the model influencing those in other areas). These results increase our confidence that FLEE adequately represents real-world evacuations and is suitable for further experimentation.

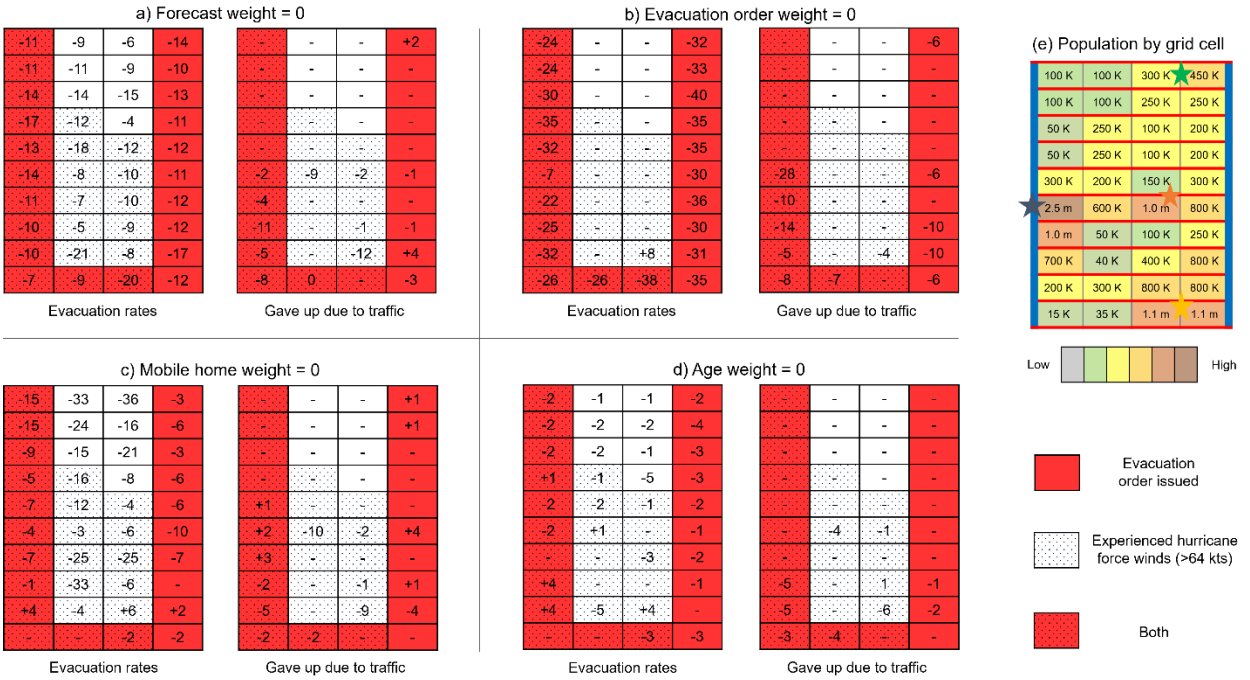

Figure A2: The spatial effects of "turning off" information inputs on evacuations rates and percent "giving up" from traffic for grid cells. Results are presented for experiments "turning off" the forecast information (a), evacuation orders (b), mobile home ownership (c), and age (d), one-by-one while holding the other parameters constant. Values are shown as the departure from the default settings in section 4.1 and in Figure 7f-g. Also presented is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (e) which provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). Note, run-to-run variability due to stochastic elements in the model ranges from 0–2% in grid cells for both evacuation rates and percent giving up due to traffic. Therefore values of -2 to 2 lie within that variability and are insignificant.

Figure A3 shows the importance of the different information on certain periods during the evacuation. For example, turning off evacuation orders (red lines) causes a reduction in evacuation rates compared to the default simulation (black lines), especially during the 36–102 hour period when evacuation orders were issued. Forecast information (purple lines) most influences evacuation rates between 30–60 hours, as forecasts indicated significant risk throughout Florida during this period. Unlike evacuation orders and forecast information, modifying the age (orange lines) and mobile home (green lines) factors do not impact any specific periods of time, but simply reduces the evacuation rates overall. This is to be expected, as these parameters are defined at the start of the simulation and are not updated.

In summary, the simulations in this section illustrate how modifying the factors that influence households' evacuation decisions in the human system agent-based model propagate through FLEE's full modeling system to influence the spatial and temporal patterns of evacuation. In general, the results suggest FLEE behaves as expected given the model's current configuration, and matches patterns seen in empirical studies which suggest forecast/warning information is a key driver for evacuations (e.g., Wong et al. 2018; Huang et al. 2016). Additionally, the results illustrate how modeling laboratories such as this can build our understanding of the evacuation decision-making processes and how they intersect with other factors (e.g., the evolving forecast information, traffic) to produce evacuations.

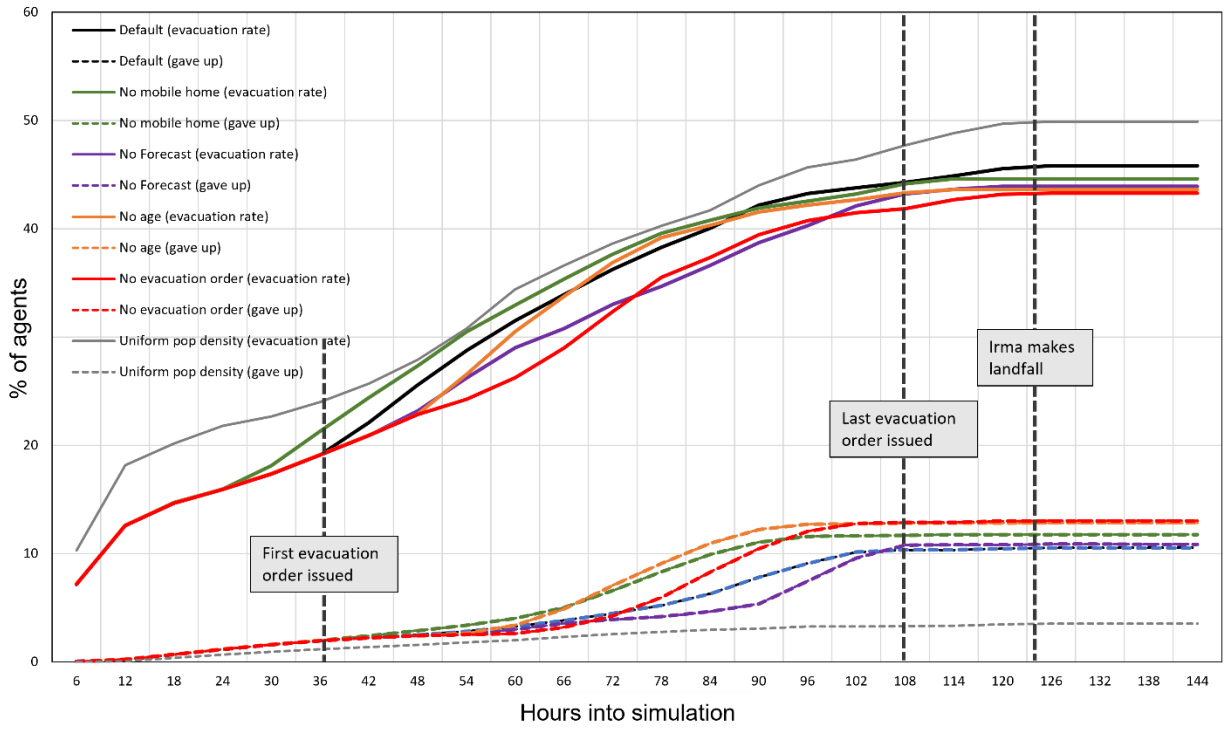

Figure A3: The temporal effects of "turning off" information inputs on the timing of evacuation rates (solid lines) and numbers giving up due to traffic (dashed lines), averaged across all grid cells. The default simulation (Table 3a; section 4.1) is expressed (black lines), as are experiments turning off the four main types of information used to assess risk: no forecast information (purple lines), no evacuation orders (red lines), no mobile home ownership (green lines), and no age (orange lines). Comparing the experiments to the default experiment (black lines) provides a general sense of the relative importance of the parameter on the overall evacuation behaviors. Also shown is the simulation where the population density is uniform (grey lines), which is further described in Supplementary Results 2.

Supplementary Results 2 - Varying geographical distribution of households

In this section, we investigate FLEE's behavior when the non-uniform geographical distribution of households in the default settings is changed to a uniform population distribution (Table 3f) i.e., where the 16.4 million residents (4.1 million households) are spread evenly across grid cells. As a result, the experiment is a first attempt to explore the effects of population density on evacuations, as this cannot be done empirically, and it demonstrates how FLEE can be used to run different scenarios with population shifts, e.g., times of year when there are a lot of tourists in certain areas, looking 10+ years out for how evacuations may change as the population grows.

In total, evacuation rates increase from 45.1% in the default simulation to 49.9% when the population distribution is uniform, which is an increase of 786,720 people (Table A9). Meanwhile rates of households unable to evacuate due to excessive traffic decrease from 10.5% in the default simulation to 3.5%, a decrease of 1,147,300 people. Thus, the experiments suggest the real-world, non-uniform population density substantially increases evacuation traffic and reduce evacuation rates.

A more in-depth look reveals an interesting pattern in the spatial distribution of evacuation behaviors. In most places, evacuation rates are higher than in default while traffic is minimal (Figure A4, bottom panel); the exception is the southern "coastal" cells where rates unable to evacuate due to traffic increase 12–17%, particularly around Miami, which reduces evacuation rates by 9–17%. One possible explanation is that the southern cells have 1) more evacuees than in the default run and 2) more evacuees downstream i.e., the area is "last in line" to evacuate based on the available road network. It is also possible

that we are seeing the impacts of clearance times being out of balance with what the clearance times would be in a world with this revised population density.

The results quantify contributions of the built environment to the evacuations. Furthermore, they illustrate the significant and potentially complex effects of population density on the evacuation success, which should be explored further. The experiment also shows how, in a modeling laboratory such as this, different components can be modified systematically to isolate influences which are impossible to do empirically, and highlights the potential value of this type of modeling laboratory to increasing our fundamental understanding of the system dynamics, and our understanding how evacuations may change as the population grows.

	% Successfully evacuated							
Experiment	Total (all cells)	Coastal >64 knot zone	Inland >64 knot zone	Coastal < 64 kts zone	Inland < 64 kts zone	Compliance rates	Shadow evacuation	Gave up to traffic
Irma Default	45.1	52.3	22.2	58.1	36.7	55.0	25.6	10.5
Uniform pop.	49.9	65.3	29.7	62.3	36.9	64.1	32.5	3.5

Table A9: Evacuation behaviors by impact zones when making the population uniform across the grid. Successful evacuation rates are broken down into impact zones (coastal vs. inland, and areas experiencing vs. not experiencing hurricane force winds of 64+ kts), compliance rates (i.e., those instructed to evacuate via evacuation order who did evacuate), shadow evacuation rates (i.e., percentage of people not instructed to evacuate who did), and the percentage of evacuees who attempted to evacuate but "gave up" due to excessive amounts of traffic.

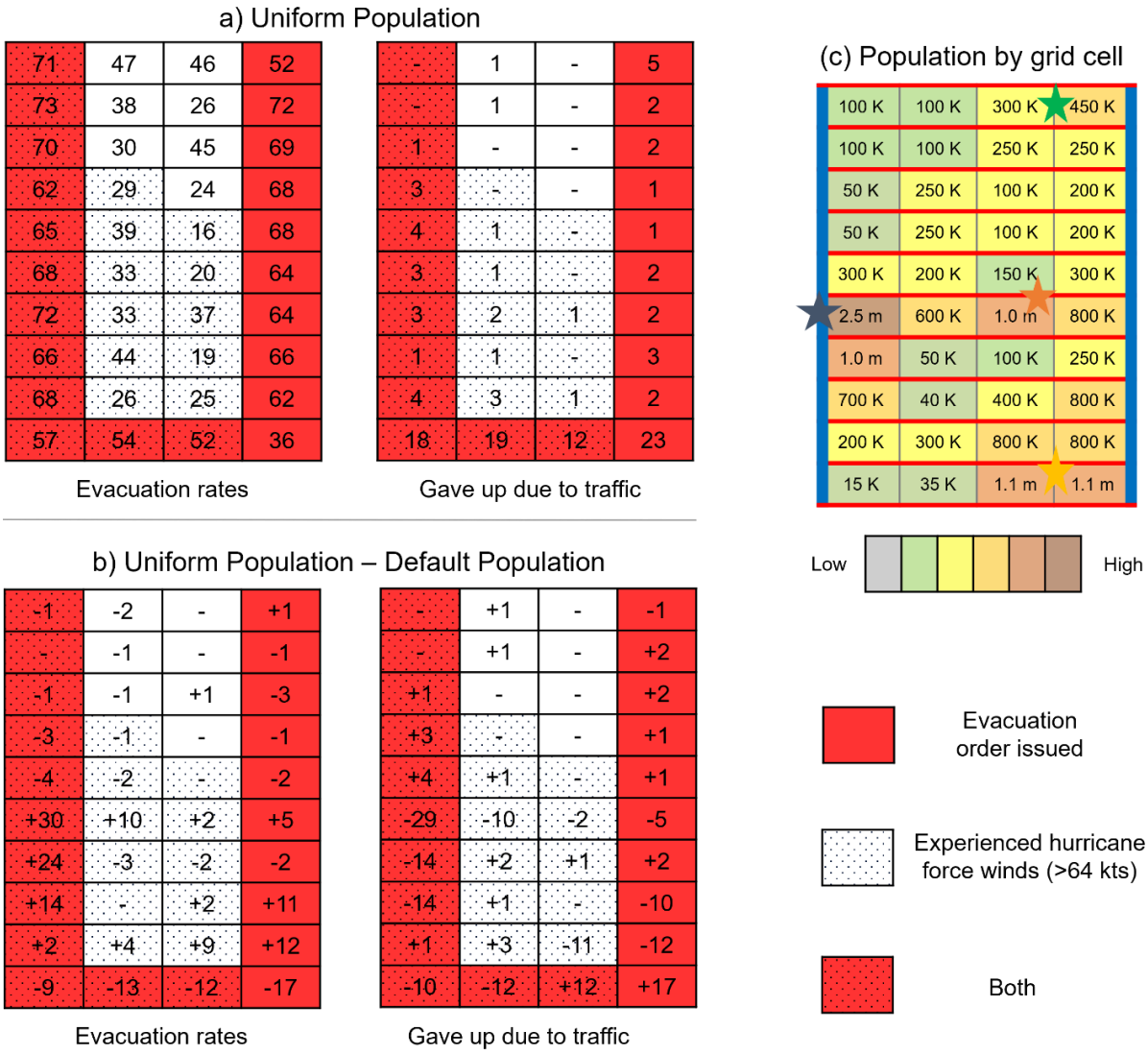

Figure A4: Influence of population density on evacuation for grid cells. Evacuation rates (left) and the percent of households unable to evacuate due to traffic (right) are shown. Results are presented for the experiment where population density is even across all grid cells (top panel). These results are compared to the default simulation with non-uniform population (bottom panel) where values are expressed as the difference from the default settings in section 4.1 and in Figure 7. Also expressed is the swath of hurricane force winds (dotted cells), evacuation orders (red cells), and the population by grid cell (c) which provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). Note, run-to-run variability due to stochastic elements in the model ranges from 0–2% in grid cells for both evacuation rates and percent giving up due to traffic. Therefore values of -2 to 2 lie within that variability and should be ignored.

APPENDIX B: Article 2

Original parameters	Updated parameters	Intended effect
Instantaneous communication – forecasts and warning information were received immediately by households to begin the evacuation decision-making process	Households are randomly assigned a value between 0-24 hours. This is the amount of time needed to receive information before making evacuation decisions	To provide a more realistic account of information diffusion (e.g., see Morss et al. 2017)
Evacuation barriers are static and do not change	Evacuation barriers are raised by 36 initially and decrease linearly to their original values in HRM21 by 54 hours before landfall	To prevent too many people from evacuating early in the simulations e.g., upstate and inland in the Irma and Dorian simulations
Mobile home ownership weighting in household evacuation decision-making algorithm: Random between 0-10	Mobile home ownership weighting in household evacuation decision-making algorithm: Random between 0-0.5	To reduce unsensible variability between grid cells, particularly inland
Socioeconomic status weighting in household evacuation decision-making algorithm: Random between 5-125	Socioeconomic status weighting in household evacuation decision-making algorithm: Random between 5-109	To reduce unsensible variability between grid cells, particularly inland

Table B1: Updates to the model since HRM21. This includes describing the original model parameters, their updates, and the reasoning behind them.

(a) Percentage of grid cell receiving evacuation orders

30	-	-	20
15	-	-	15
20	-	-	5
5	-	-	5
20	-	-	10
15	-	-	5
30	-	-	5
45	-	-	15
65	-	-	5
100	100	15	5

(b) Population by grid cell

Figure B1: The percentage of coastal grid cells receiving evacuation orders, if issued for the cell. Also expressed is the population by grid cell (b). These provide a frame of reference e.g., major cities depicted include Miami-Ft. Lauderdale (yellow star), Tampa Bay-St. Petersburg (blue star), Jacksonville (green star), and Orlando (orange star). Values are based on Figure 2 of Wong et al. (2018), Florida Department of Emergency Management (personal communication), Figure 1 of Darzi et al. (2020).

Information category	Storm	Data specifics	Sources
Evacuation orders	Irma	Counties, evacuation zones	Figure 2 of Wong et al. (2018), Florida Department of Emergency Management (personal communication), Figure 1 of Darzi et al. (2020)
	Dorian	Counties, evacuation zones	TIME (2019)
Evacuation rates	Irma	By region	Figure 4 of Wong et al. (2018)
	Irma	By voting precinct	Figure 1c of Long et al. (2020)
	Irma	By county	Table 10 of Martin et al. (2020)
	Irma	By state	Florida Department of Emergency Management (2017), Figure 6 of Wong et al. (2018), Figure 2c of Long et al. (2020)
	Irma	By city	Feng and Lin (2021)
	Dorian	Storm total	Mongold et al (2020)
Evacuation traffic	Irma	Areas/times of congestion	Page 15 of Wong et al. (2018), Ghorbanzadeh et al. (2021), States et al. (2021)
	Irma	Total numbers stuck	Feng and Lin (2021)

Table B2: Empirical information used to validate FLEE. This includes a description of the information, which storm the data is form, and the exact source.

APPENDIX C: Article 3

Hours until landfall	Irma Advisory No.	Dorian-LF Advisory No.	2001 average errors (km)	2021 average errors (km)
120	25	17	601	319
114	26	18	564	299
108	27	19	527	279
102	28	20	490	260
96	29	21	453	240
90	30	22	433	223
84	31	23	412	207
78	32	24	392	190
72	33	25	370	174
66	34	26	342	160
62	35	27	314	146
54	36	28	287	132
48	37	29	259	118
42	38	30	229	104
36	39	31	200	90
30	40	32	170	78
24	41	33	138	67
18	42	34	116	55
12	43	35	94	43
6	44	36	72	27
0	45	37	50	13

Table C1: Average track errors in 2001 and 2021, expressed in km for every 6 hours leading up to landfall. These values were taken from the NHC website (https://www.nhc.noaa.gov/verification/index.shtml), were extrapolated from 24 to 6 hour intervals, and matched to the equivalent advisory numbers from the official NHC forecasts for Irma and Dorian. Based off this information, track errors were perturbed left/right of the original NHC forecast track by amounts in this table.

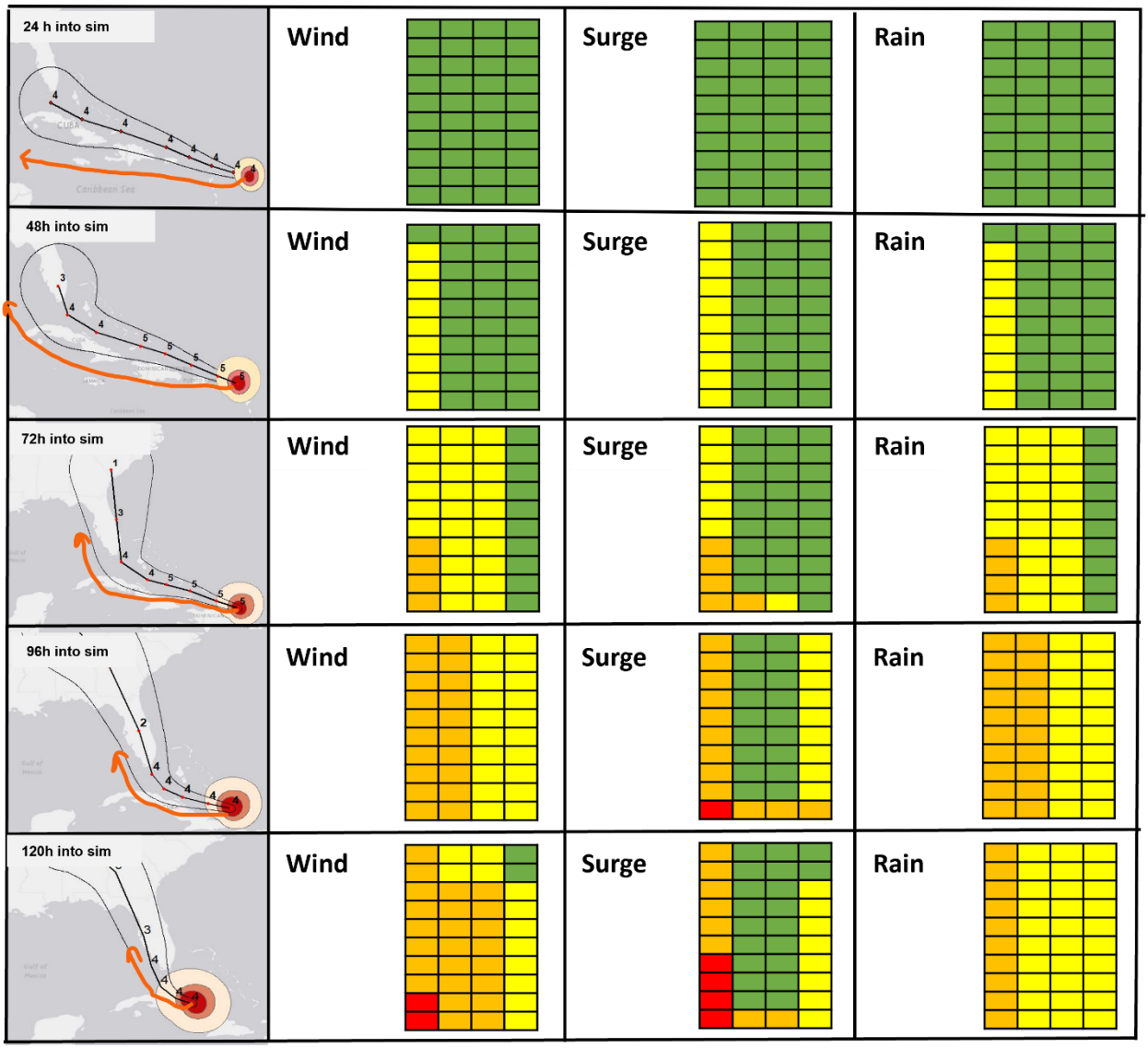

Figure C1: Irma's 2001 left track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

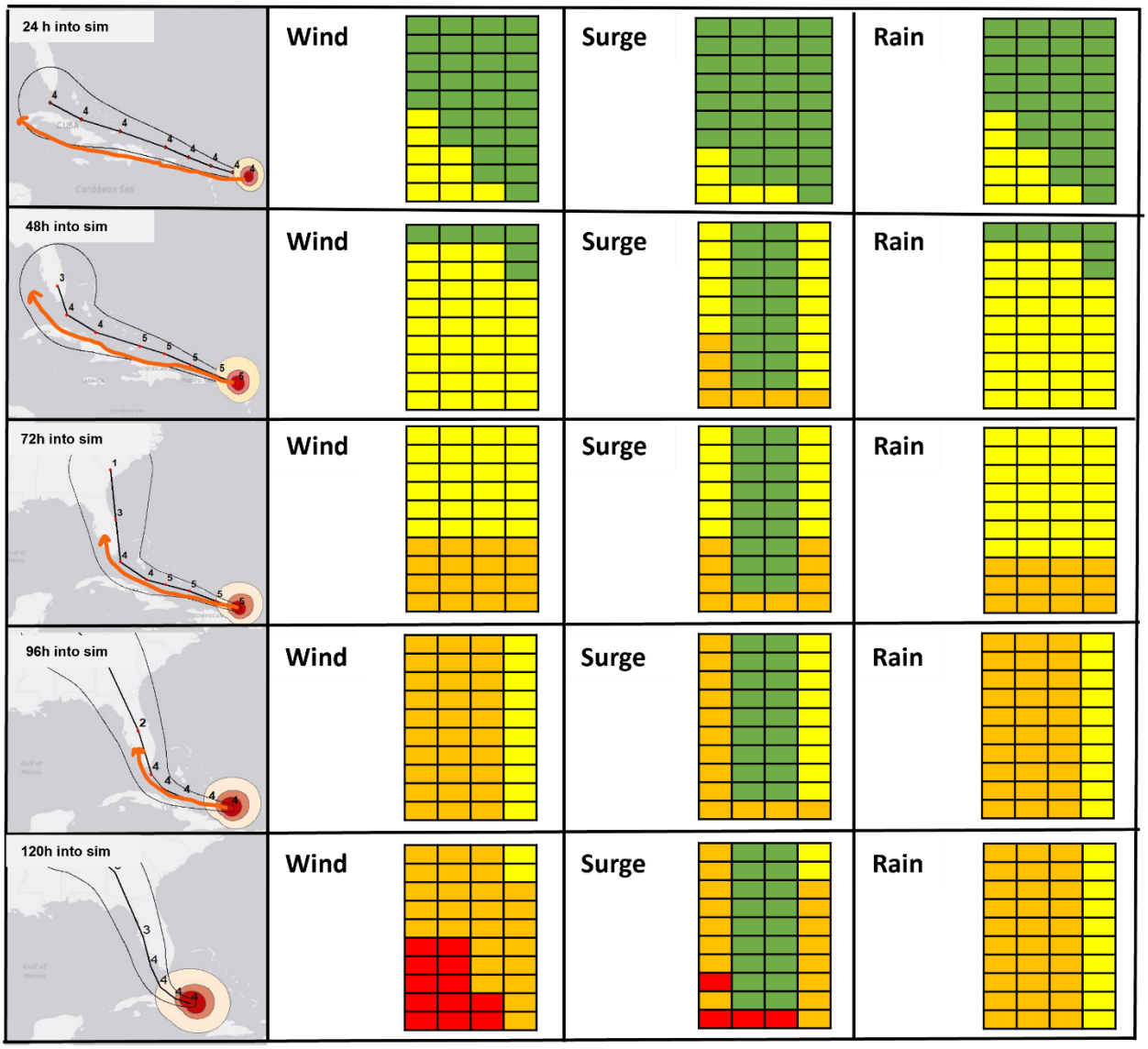

Figure C2: Irma's 2021 left track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

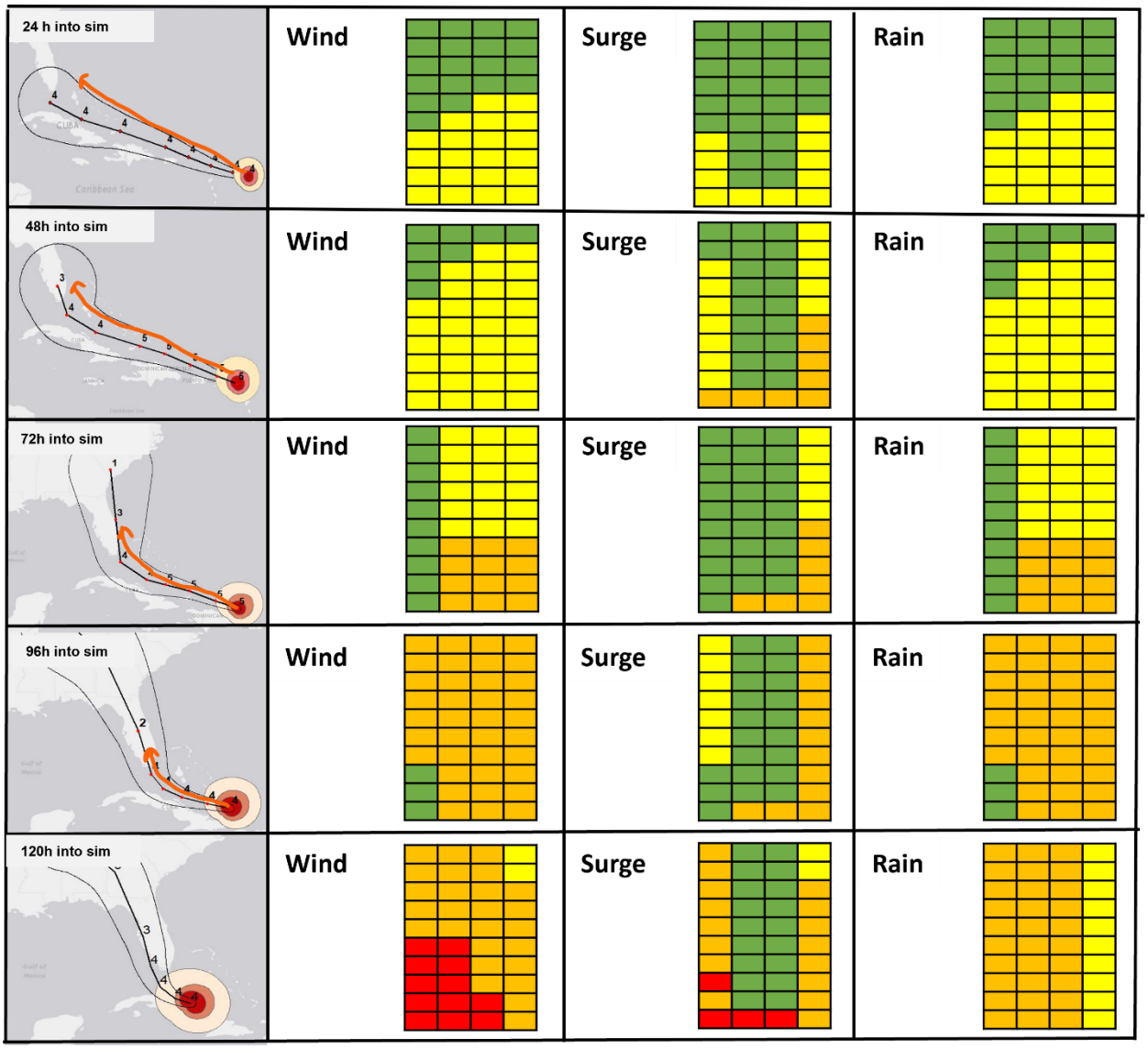

Figure C3: Irma's 2021 right track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

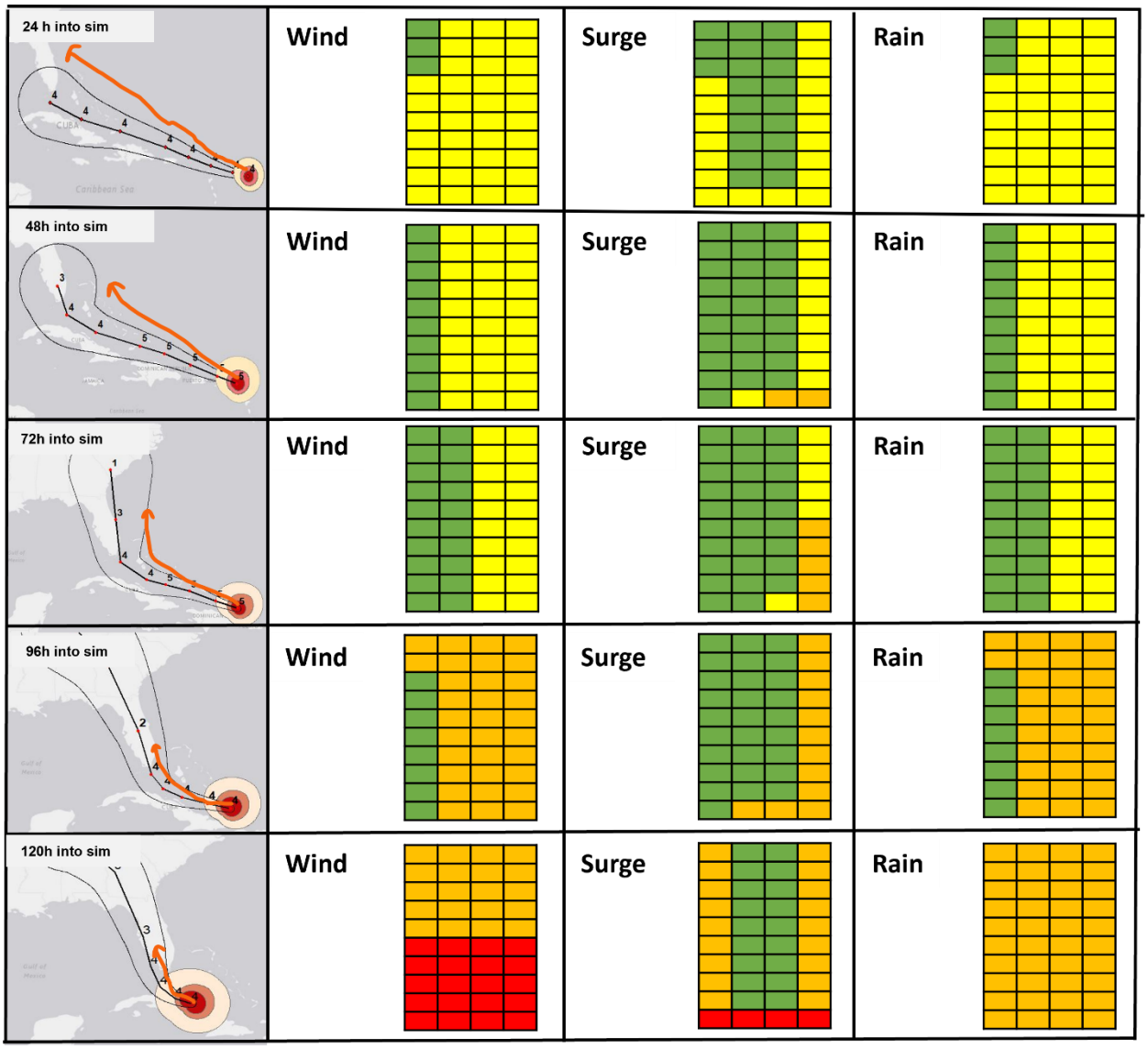

Figure C4: Irma's 2001 right track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

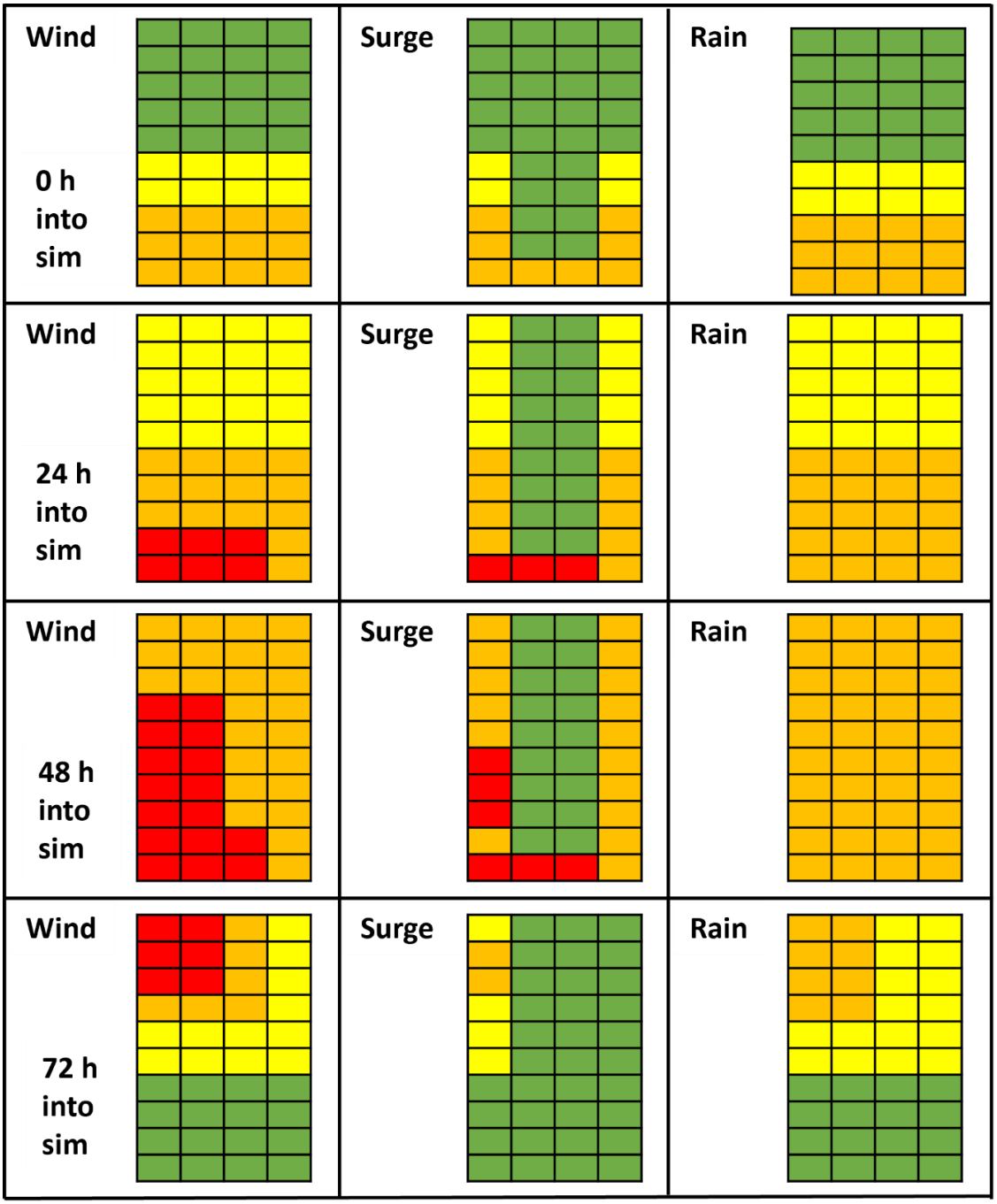

Figure C5: Irma's RI/RO forecasts. Here we shorten the forecast timeline of the original NHC forecasts shown in Figure 22 from 168 hours to 84 hours. Forecasts are shown for every 24 hours but update every 6 hours (not shown). The peak magnitudes of risk (wind, surge, rain) for each grid cell are the same as Irma; however, the timelines are condensed, effectively creating large and intertwined forward speed, intensity, and track errors meant to emulate the effects of rapid onset/intensity cases.

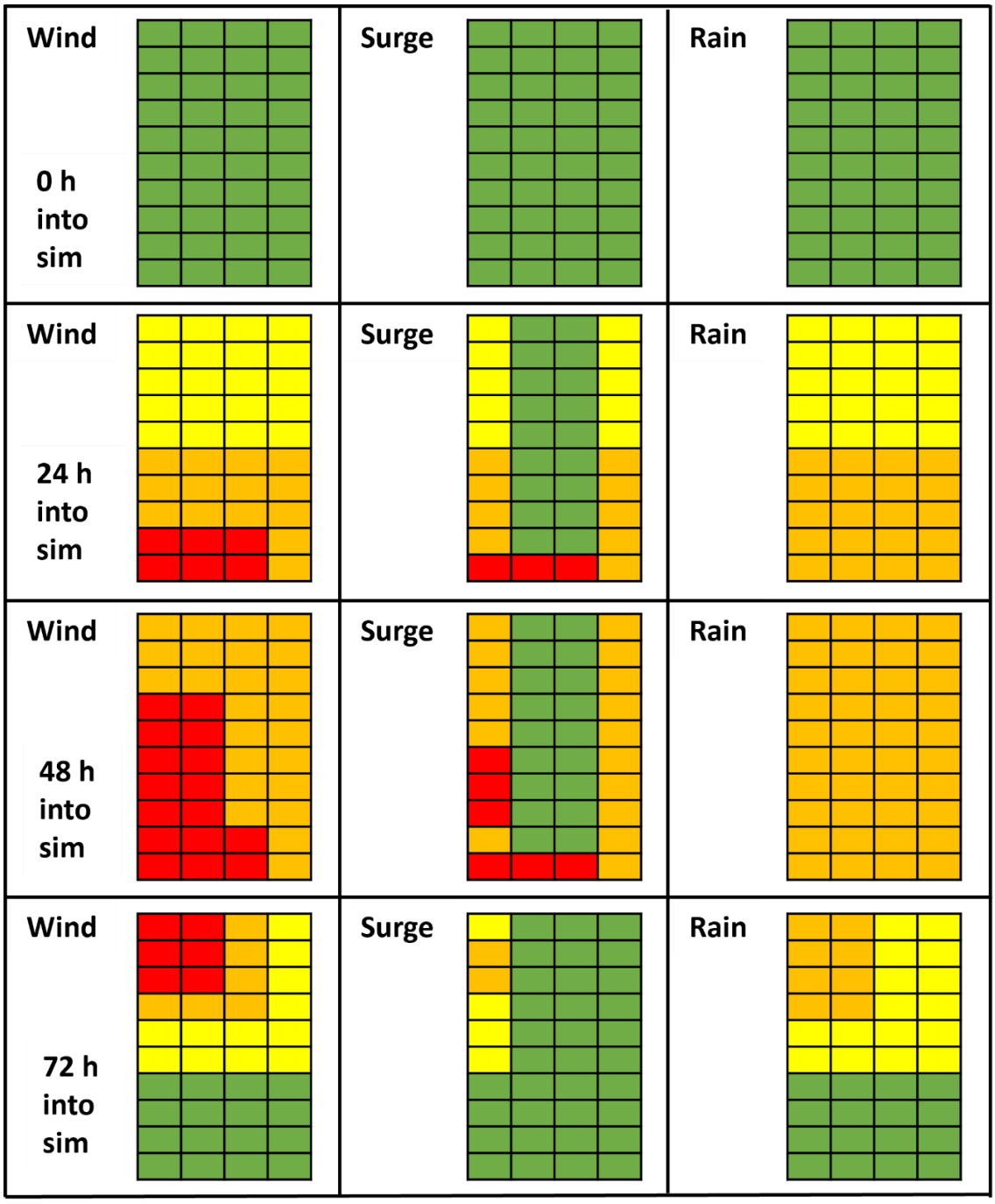

Figure C6: Irma's RI/RO – 12 forecasts. Here we shorten the forecast timeline of the original NHC forecasts shown in Figure 22 from 168 hours to 72 hours. Forecasts are shown for every 24 hours but update every 6 hours (not shown). The peak magnitudes of risk (wind, surge, rain) for each grid cell are the same as Irma; however, the timelines are condensed, effectively creating large and intertwined forward speed, intensity, and track errors meant to emulate the effects of rapid onset/intensity cases.

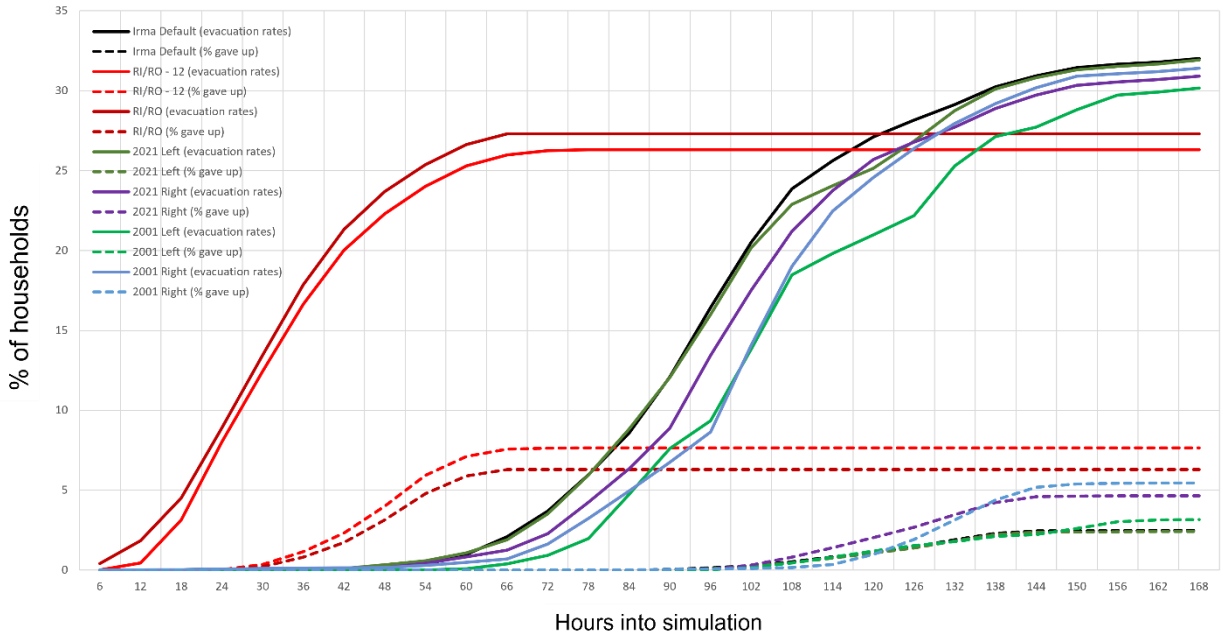

Figure C7: Evacuation rates over time for Irma's experiments. The temporal effects of the experiments on evacuation rates (solid lines) and numbers giving up due to traffic (dashed lines), averaged across grid cells, throughout the simulations. The default simulation is expressed (black lines), as are RI/RO (dark red lines), RI/RO – 12 (light red lines), 2021 left (dark green lines), 2001 left (light green lines), 2021 right (purple lines), 2001 right (blue lines).

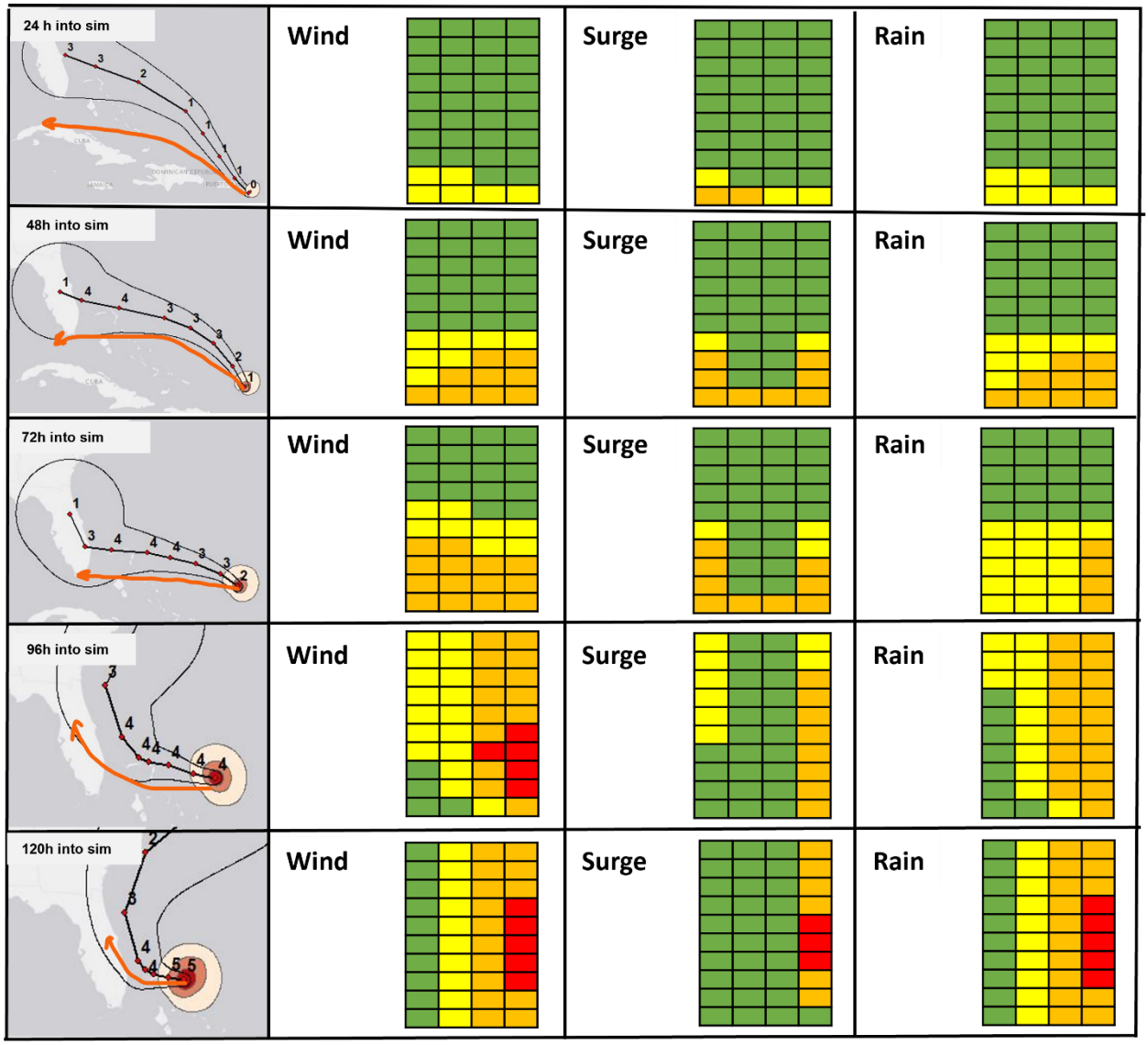

Figure C8: Dorian LF's 2001 left track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

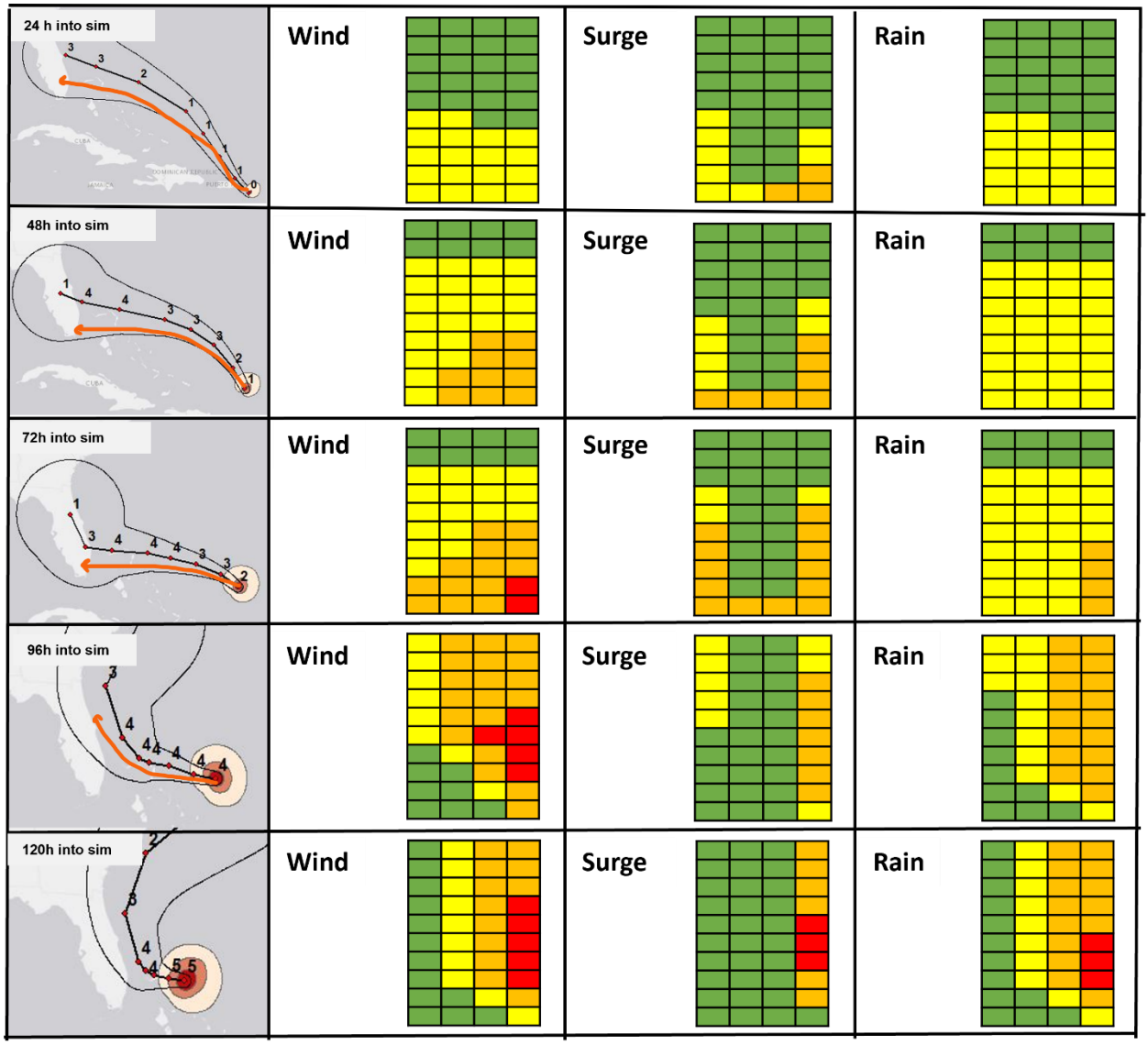

Figure C9: Dorian's 2021 left track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

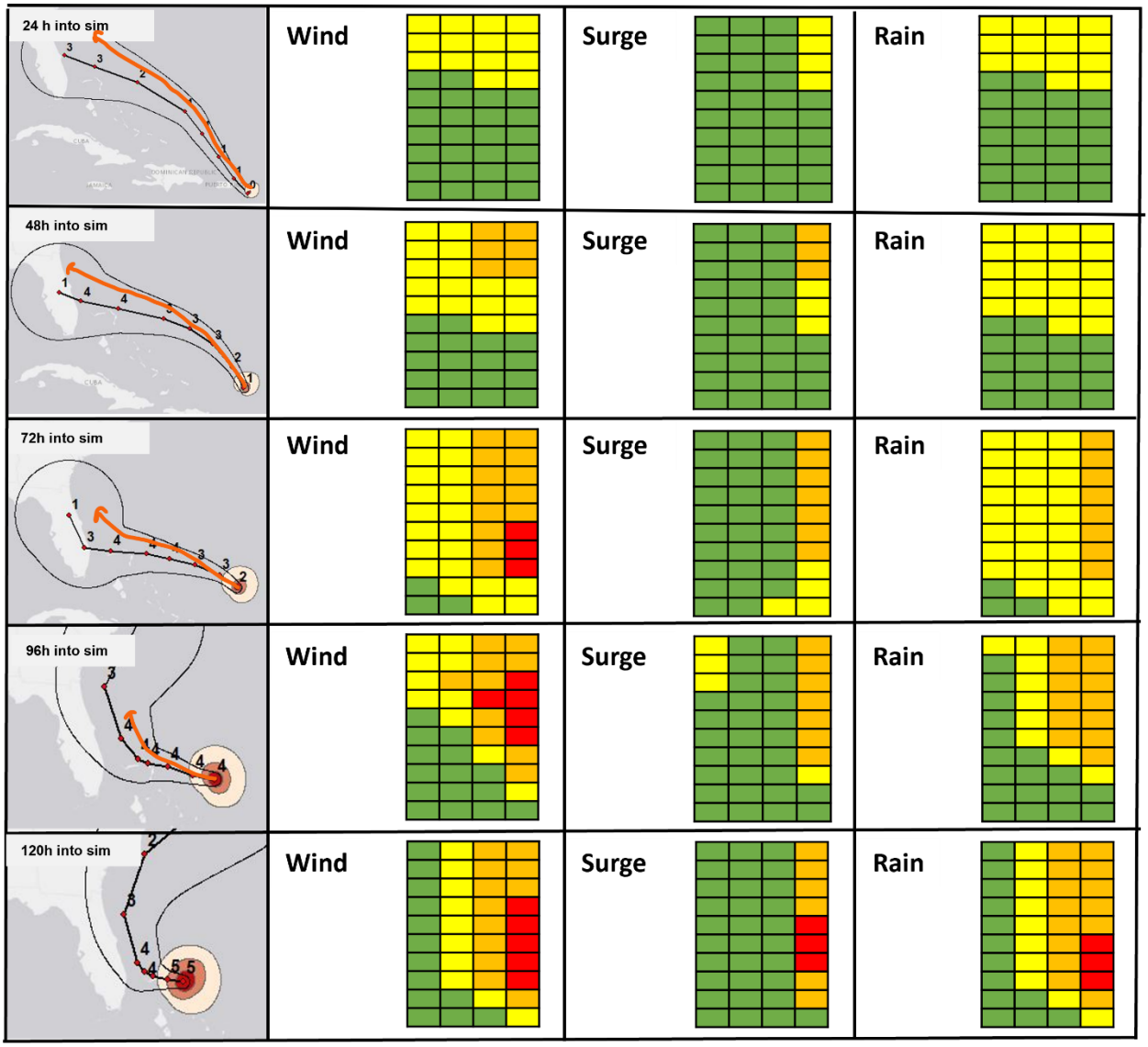

Figure C10: Dorian's 2021 right track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

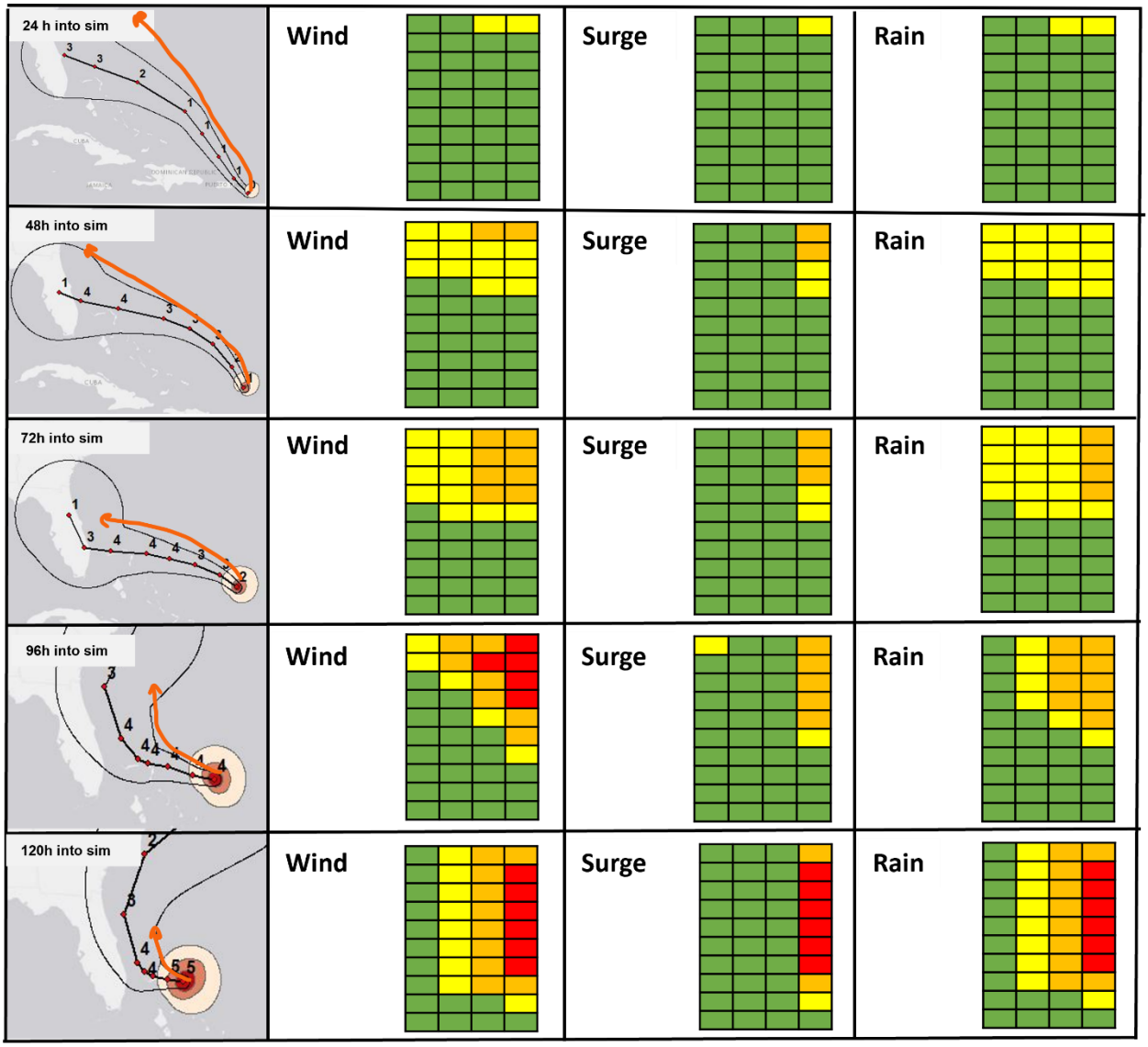

Figure C11: Dorian LF's 2001 right track forecasts. Forecasts are shown for every 24 hours but update every 6 hours (not shown). Left column: Official evolving NHC forecast track (black center line), category (numbers), cone of uncertainty (edges are outer black lines), and current wind radii at 34 (white), 50 (pink), and 64+ (red) knot intervals. The approximate track of the perturbed forecasts is shown (orange arrow). Right three columns: The light system threats corresponding to the perturbed forecasts for wind, surge, and rain are shown for equivalent times.

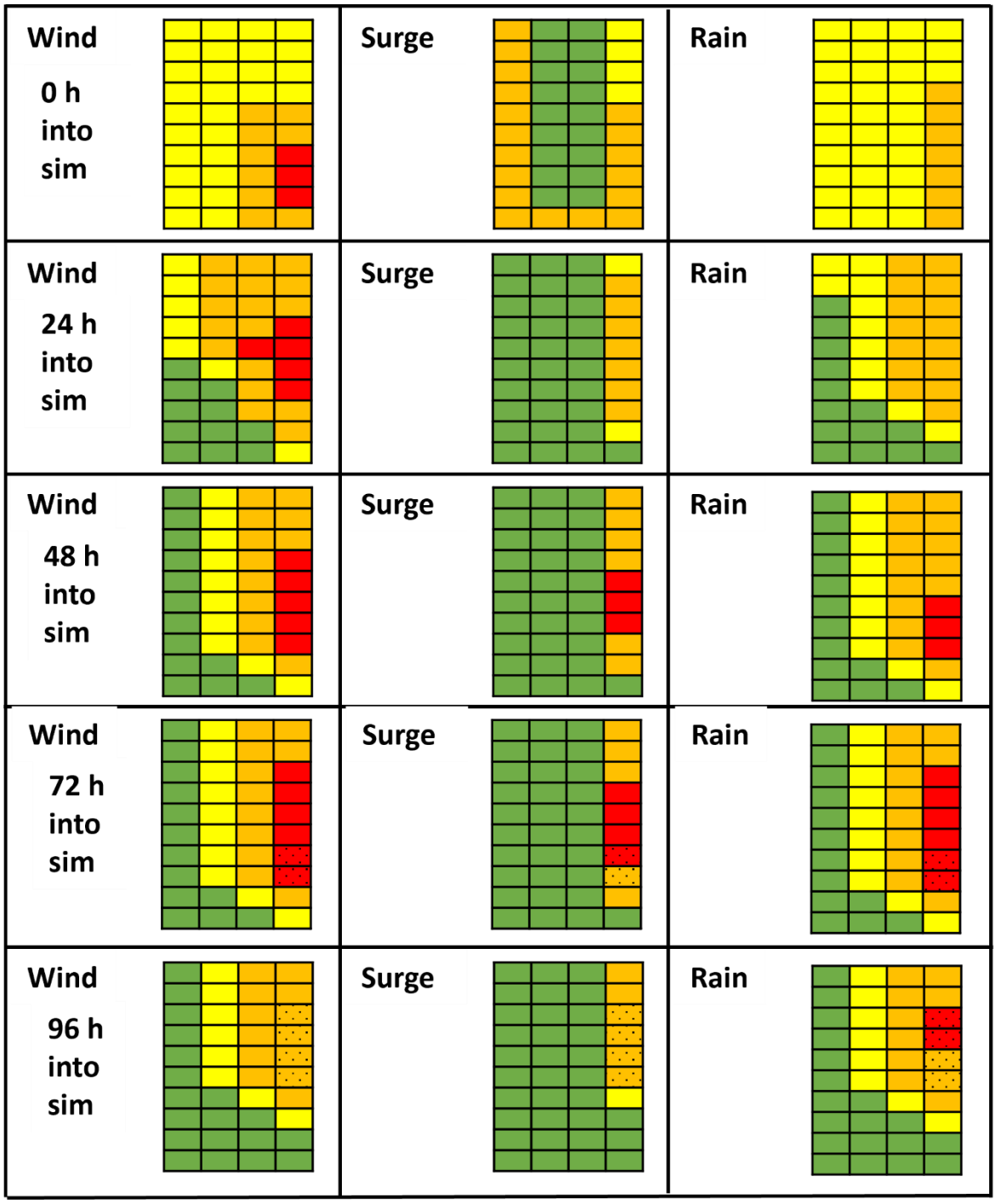

Figure C12: Dorian LF's RI/RO forecasts. Here we shorten the forecast timeline of the original NHC forecasts shown in Figure 22 from 168 hours to 84 hours. Forecasts are shown for every 24 hours but update every 6 hours (not shown). The peak magnitudes of risk (wind, surge, rain) for each grid cell are the same as Dorian LF; however, the timelines are condensed, effectively creating large and intertwined forward speed, intensity, and track errors meant to emulate the effects of rapid onset/intensity cases.

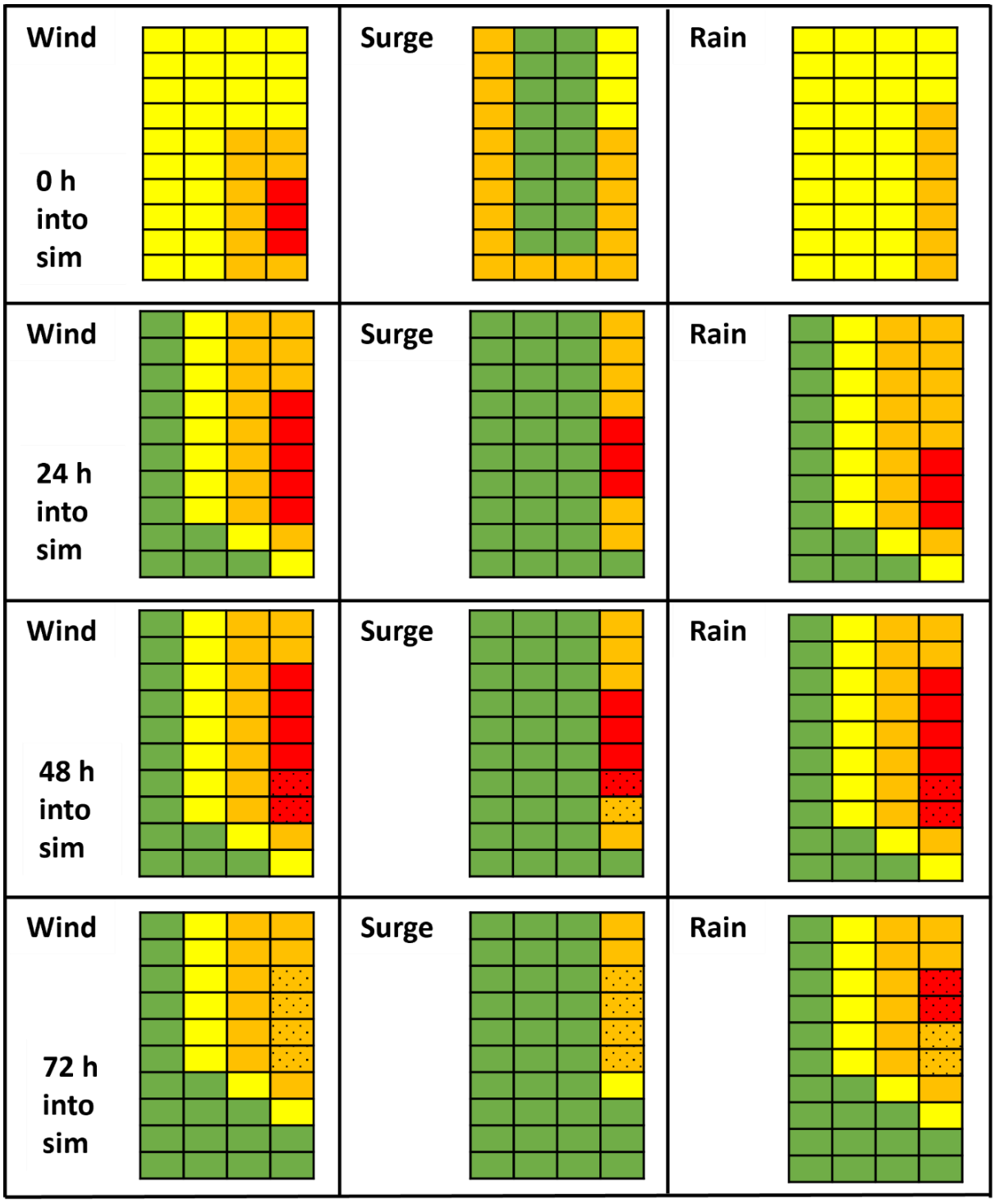

Figure C13: Dorian LF's RI/RO – 12 forecasts. Here we shorten the forecast timeline of the original NHC forecasts shown in Figure 22 from 168 hours to 72 hours. Forecasts are shown for every 24 hours but update every 6 hours (not shown). The peak magnitudes of risk (wind, surge, rain) for each grid cell are the same as Dorian-LF; however, the timelines are condensed, effectively creating large and intertwined forward speed, intensity, and track errors meant to emulate the effects of rapid onset/intensity cases.

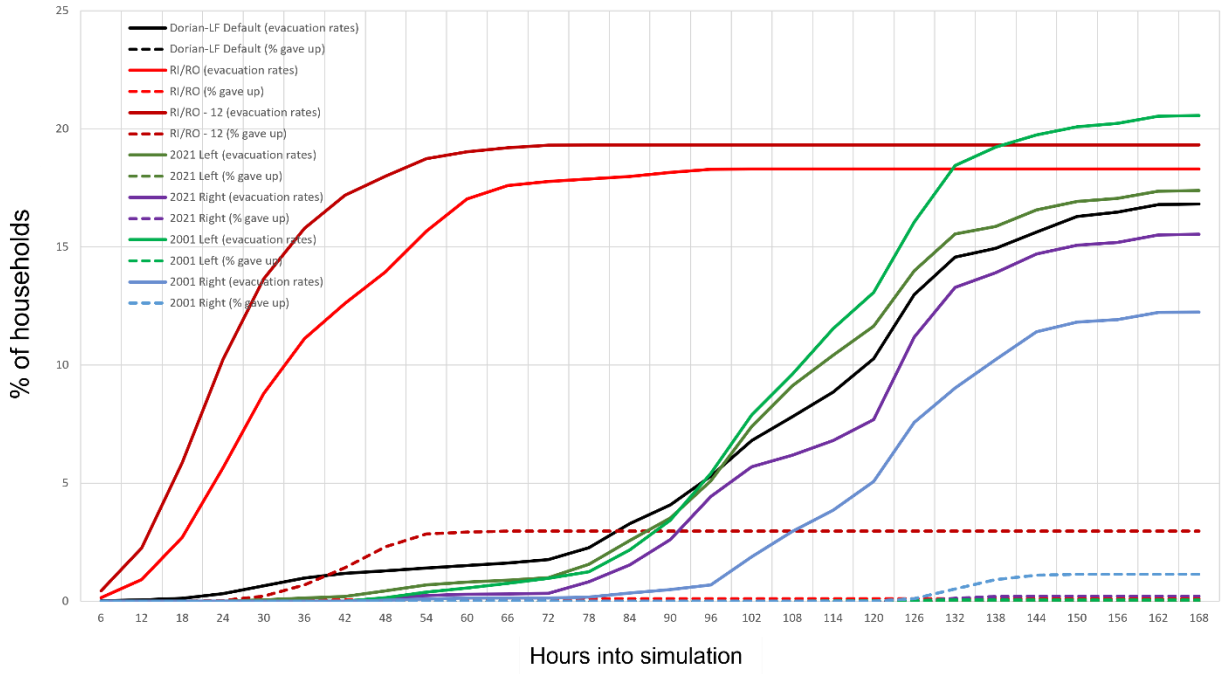

Figure C14: Evacuation rates over time for Dorian LF's experiments. The temporal effects of the experiments on evacuation rates (solid lines) and numbers giving up due to traffic (dashed lines), averaged across grid cells, throughout the simulations. The default simulation is expressed (black lines), as are RI/RO (light red lines), RI/RO – 12 (dark red lines), 2021 left (dark green lines), 2001 left (light green lines), 2021 right (purple lines), 2001 right (blue lines).