Vol. 00, No. 0, Xxxxx 0000, pp. 000–000 ISSN 1523-4614 | EISSN 1526-5498 | 00 | 0000 | 0001

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

The Hurricane Decision Simulator: A Tool for Marine Forces in New Orleans to Practice Operations Management in Advance of a Hurricane

Eva D. Regnier

 $Naval \ Postgraduate \ School, \ eregnier@nps.edu, \ faculty.nps.edu/eregnier@nps.edu \ faculty.nps.edu/eregnier@nps.edu \ faculty.nps.edu/eregnier@nps.edu \ faculty.nps.edu \ faculty.nps.ed$

Cameron A. MacKenzie

 $Iowa\ State\ University,\ camacken@iastate.edu,\ www.imse.iastate.edu/directory/faculty/cameron-a-mackenzie/$

The U.S. Marine Forces Reserve (MFR) in New Orleans is frequently threatened by hurricanes. To protect the safety of personnel and their families while maintaining mission capability, the Commander must make timely decisions to set up an alternate headquarters and allow for an orderly evacuation. The MFR relies on forecasts from the National Hurricane Center, but these forecasts are uncertain, are updated frequently, and can be difficult to interpret in the context of the MFR's decision timeline. In addition, there are few opportunities to learn from experience. We developed the Hurricane Decision Simulator (HDS) to allow MFR personnel to practice making preparation decisions in the context of many realistic simulated storms and forecasts, and to develop a better understanding of the decision sequence, the forecast products and their relationship. The HDS has improved MFR's training and readiness for hurricane preparation operations and decision making by enabling more and better focused training and is being extended to other facilities.

Key words: OM Practice, Simulation, Risk Management *History*:

1. Introduction

When hurricanes are active or forecast in the Gulf of Mexico, the Commander of the Marine Forces Reserves (MFR) must make timely decisions to prepare for a possible evacuation of the headquarters in New Orleans. While preparation and evacuations are costly, the consequences of failing to prepare—or of preparing too late—are even more serious. In this high-stakes decision context, there are many layers of uncertainty about the eventual outcome. The Commander and his top military staff are replaced every two to three years. Seven Commanders have served MFR since 2000, and they frequently arrive during the peak of hurricane season, sometimes without experience living in a hurricane-prone region. The MFR maintains detailed hurricane operations plans and trains annually for operations during a hurricane threat from decision making through potential evacuation and operation out of an alternate headquarters. While the details of the implementation of the plan are spelled out, nowhere is it specified what forecasts should prompt the Commander to make the decision to proceed to more serious—and costly—preparation actions.

In response to a request from MFR, we developed the Hurricane Decision Simulator (HDS), an online training tool that lets users step through the MFR's hurricane decision process in the context of simulated storms, with forecasts that resemble recent National Hurricane Center (NHC) forecasts. Unlike other hurricane-preparation training, the HDS allows the user to make decisions throughout a storm and receive outcome feedback that depends on her choices and the storm's impacts.

The HDS includes 339 simulated storms designed to capture both typical and unusual storm behavior consistent with the historical record, including both hits and misses. The simulated storms were generated such that the threat to New Orleans is high enough that they would be monitored by the MFR Commander. The HDS allows an individual user to step through a storm in about 15 minutes. Since real storms posing a similar threat level occur on average approximately once per year, users can gain decades' worth of simulated experience in a day.

The MFR has used the HDS in individual training and for group tabletop exercises, and the emergency management staff and senior decision makers find it very useful. It provides more training opportunities, saves time in preparing and executing trainings, provides a more realistic experience of uncertainty, and improves their understanding of forecast products. The MFR has requested that we create versions for the Marines' Reserve Training Centers in other vulnerable locations on the Gulf and Atlantic coasts.

The prior literature on hurricane preparedness and response decision making can be generally divided into three areas: (i) helping authorities determine if and when to evacuate communities, (ii) descriptive models of people's decisions to evacuate and return, and (iii) guidance and optimization models pre-positioning and delivering emergency relief supplies. Operations research models can provide useful guidance to emergency managers (Larson et al. 2006) and help public officials determine how to trade off risks to life against the cost of false alarms when considering whether to evacuate a community (Lindell and Prater 2007, Regnier 2008, Kailiponi 2010, Apivatanagul et al. 2012). Even if an evacuation is ordered, a number of geographic, sociodemographic, and communication and risk factors influence behavior regarding if, when, and how to leave (Baker 1991, Dow and Cutter 1998, Gladwin et al. 2001, Kang et al. 2007, Smith and McCarty 2009, Hasan et al. 2010, Lindell et al. 2011, Lindell and Perry 2012) and when to return after a hurricane (Siebeneck and Cova 2008, 2012, Siebeneck et al. 2013). Optimization can help inform officials where to best pre-position emergency relief supplies (Salmerón and Apte 2010, Taskin and Lodree 2011, Uichanco 2016), but damaged infrastructure due to the hurricane can impede the delivery of those supplies (Holguín-Veras et al. 2007, Horner and Widener 2011). The HDS and this article contribute to this literature by providing decision training for a single organization, the MFR. Rather than optimizing the decision for officials, this article focuses on how officials can benefit from practicing making preparedness decisions in a simulated environment.

This article describes the HDS and how it is being used by the MFR to improve its operations management during a hurricane threat. Section 2 provides background on the MFR's decision context, the need for training aids, and a review of relevant literature. Section 3 describes the HDS and its development. The HDS has three primary components: decisions selected by the user, simulated storms, and outcomes based on the user's decisions and the storm's behavior. Section 4 describes the benefits that MFR has experienced by using the HDS. Concluding remarks and the potential future uses of such a simulation-based training environment for managing hurricane preparation operations are discussed in Section 5.

2. Background

When a hurricane—or a tropical storm that is forecast to strengthen into a hurricane is active in the Gulf of Mexico, the MFR Commander monitors the storm to determine whether and when to prepare to evacuate and transfer operations to an alternate headquarters. While the consequences of failing to act or acting too late can be deadly, false alarms are also costly, both financially and in terms of mission disruption. The high stakes are coupled with multi-dimensional and multi-layer uncertainty about the impacts to the MFR, which pose a challenge to judgemental decision making make it difficult to gain expertise.

2.1. Hurricanes and MFR

A hurricane is a tropical cyclone with sustained (1-minute) winds exceeding 74 mph, but in some cases reaching 190 mph.¹ Hurricane-force winds can cause structural damage and are dangerous to people in inadequate shelter,² but flooding due to precipitation and storm surge is more deadly than winds (Rappaport 2014). Hurricanes can also disable local services such as power and water, and streets can become impassable due to flooding, fallen trees, and other debris.

The topography of southeast Louisiana makes the region particularly vulnerable to flooding. In 2005, Hurricane Katrina killed approximately 1,200 people, damaged or destroyed about 300,000 houses (Townsend 2006), and caused a total of \$198 billion in damages (Blake et al. 2011) in Louisiana and nearby parts of Mississippi. While Katrina is the most widely known Louisiana hurricane, three other hurricanes that hit Louisiana are also among the top ten most deadly in U.S. history (Blake et al. 2011).

The MFR is headquartered in New Orleans, adjacent to the Mississippi River and less than 10 feet above sea level. The facility is also the headquarters for the Marine Forces Northern Command. The MFR must be ready to deploy to reinforce and support the active duty Marine forces, and the Northern Command conducts homeland security operations and supports U.S. civilian authorities. A single general serves as the Commander of the two organizations and they are headquartered in the same facility. For simplicity, we refer to them collectively as the MFR. The headquarters facility was built to withstand hurricane conditions, has backup power systems and supplies, and is located within the protection of the levee system. However, in the event of severe impacts such as flooding and the loss of services, the MFR would transfer its operations to an alternate headquarters to ensure full mission capability. Moreover, the MFR is responsible for its personnel and their families' safety, and many of them live outside the levee system.

The MFR maintains a Hurricane Evacuation and Continuity of Operations Plan that formally documents the decisions and the actions necessary to evacuate the current headquarters and relocate to an alternate headquarters at an inland facility. Three timelines (matrices) are included in this document. The Decision Support Matrix (DSM) and the Hurricane Condition Matrix recommend key decisions for the Commander. These decisions

¹ http://www.aoml.noaa.gov/hrd/tcfaq/E1.html, accessed August 25, 2015.

² http://www.nhc.noaa.gov/aboutsshws.php, accessed August 25, 2015.

are tied to approximately 150 implementation actions that MFR must take to maintain continuity of operations and protect personnel and their families, documented in the Execution Matrix. The implementation actions take time, which drives the decision timeline.

Like many military organizations, the MFR commonly uses backwards planning with a deterministic timeline (e.g., D-Day, with D-1 indicating one day before D-Day, etc.) to plan its operations. Similarly, the DSM is based on the number of hours before the arrival of hurricane-force winds. According to the plan, key decisions should be made 48 to 96 hours before tropical-storm force winds reach the New Orleans area. The plan treats the time until storm impacts arrive as deterministic. Although the Commander's decisions about whether to implement more serious—and costly—levels of preparation are based on the best available forecasts, neither the DSM nor the Execution Matrix specifies the precise forecast conditions that should prompt each level of preparation.

Although the hurricane operations plans provide important information to Commander and his emergency management staff and enable rapid implementation of hurricane preparation actions when ordered, they do not explicitly account for the uncertainty in storm forecasts. While NHC forecast errors have improved dramatically in recent years,³ high-probability hurricane forecasts are very rare for the key decision lead times of 48 to 96 hours before the onset of tropical-storm force winds specified in the DSM.

For example, in 2012, Hurricane Isaac made landfall in Louisiana and at the same time tropical-storm force winds reached the city. However, 66 hours earlier, the storm was forecast to make landfall in the Florida panhandle, about 230 miles from its eventual landfall location. During the next 18 hours, the forecast track moved over Louisiana, which led the MFR to deploy the advance emergency relocation staff (ERS) just 36 hours before landfall, and the team was caught in severe traffic.

Figure 1 shows the probabilities of Hurricane Isaac's winds exceeding various thresholds at New Orleans coupled with the DSM timeline for four of the key decisions. Understandably, the MFR's after-action report indicated that in the future, they would like to be able to initiate action with more lead time. However, as in the case of Isaac, the threat may not appear sufficient to justify costly preparation at the longer lead times. Hurricane

³ Average NHC Atlantic Track Forecast Errors for tropical storms and hurricanes, http://www.nhc.noaa.gov/ verification/pdfs/1970-present_OFCL_ATL_annual_trk_errors_noTDs.pdf, accessed September 8, 2016. Note distances are in nautical rather than statute miles.



Figure 1 National Hurricane Center forecasts of the probability of winds exceeding three thresholds at New Orleans during the 120 hours following the forecast for Hurricane Isaac. The horizontal axis shows the issue time of the forecast, in hours before 7 p.m. August 28, 2012, when the storm made landfall in Louisiana. The planned timing of four of the key decisions are shown, together with timing of the late movement of the track. The ERS is the Emergency Response Staff and C2 is the abbreviation for Command and Control.

preparation timelines can also be compressed because storms can form very quickly. In 2007, Humberto formed in the Gulf of Mexico just south of Louisiana and made landfall as a hurricane approximately 16 hours after the first public advisory from NHC.

2.2. Decision context

In the 20th century, forecasts with lead times sufficient for evacuation have dramatically lowered the hurricane death rate. Forecasting organizations within the National Oceanic and Atmospheric Administration (NOAA), in particular the NHC, generate most of the underlying forecast information for the North Atlantic, the Gulf of Mexico, and the eastern North Pacific. Local National Weather Service offices and commercial forecasters get much of their information from the NHC.

When a tropical storm is active in its area of responsibility, the NHC forecasts its track (center positions), intensity (maximum 1-minute sustained winds), and size (wind radii). An example of the forecast track for 2012's Hurricane Isaac is shown in Figure 2c, together

with a track forecast cone that encompasses the likely future positions of the storm center.





Hurricane Force Wind Speed Probabilities For the 120 hours (5 days) from 8 PM EDT Sat Aug 25 to 8 PM EDT Thu Aug 30

EST TH

٨

≳

2

OW

Ø

Tropical Storm Force Wind Speed Probabilities For the 120 hours (5 days) from 8 PM EDT Sat Aug 25 to 8 PM EDT Thu Aug 30

M

₹

3

OW

0

N

N



Figure 3 Schematic representation of layers of uncertainty in hurricane forecast interpretation.

Several layers of uncertainty make the hurricane-preparation decision context challenging. As represented in the center (white) rectangle in Figure 3, the variables that NHC forecasts do not match those of direct consequence to MFR. The impacts that directly threaten MFR personnel and operations are flooding in the immediate area and disruption of local services, such as power, drinking water, and security. Of these variables, the NHC only forecasts flooding, and currently it issues no storm-specific surge or flooding forecasts with more than 48 hours' lead time, at which point, according to the DSM, the MFR should have completed its evacuation and decided to transfer command and control (C2) to the alternate headquarters.⁴ Because the plan is driven by the storm's lead time, the time until impacts is also very important. However, the NHC does not officially forecast landfall. Their text products sometimes (but not always) provide a window for landfall or for the arrival of tropical-storm force winds, but rarely more than 24 hours before impact.

The NHC does forecast several variables highly related to the direct impacts—the storm's center positions, its size, and its intensity are highly related to winds experienced at New Orleans. However, the relationship is multidimensional and includes variability, limiting the decision maker's ability to accurately judge the local impact (Stewart and Lusk 1994). There is evidence that many users have trouble translating the forecast products into an assessment of their direct impacts.

For example, in conversations with local officials, we have found that many prefer to use the track forecast cone as the basis for evacuation decisions, ordering an evacuation if

⁴ http://www.nhc.noaa.gov/surge/products.php accessed September 8, 2016.

their region is inside the track forecast cone. Broad et al. (2007) cites many instances of forecast users relying on the cone (see Figure 2c) to encompass the entire area that may experience impacts. This is consistent with our experience with local officials throughout the southeast. However, the cone does not show the extent of the storm's impact. The cone's size depends only on average track errors, and would be similar for a large, intense hurricane like Katrina, whose maximum intensity was 175 mph, and for a small tropical storm.

Another example of imperfect interpretation of forecasts is the belief that wind-speed is a highly valid predictor of storm surge. As described earlier, storm surge is a more common killer than high winds, yet many people expect the relationship between storm intensity and storm surge to be much closer than it is. While the NHC emphasizes that storm intensity is only loosely associated with storm surge,⁵ many users including local emergency managers refer to storm surge as a direct function of wind speed (e.g. "expect a Category 2 storm surge") and there are many instances of similar advice being offered.⁶

A second layer of uncertainty is represented in the striped rectangle in Figure 3—the storm's actual outcome differs from what is forecast. The difference is called forecast error. Based on the 2010-2015 seasons, average track errors range from about 85 miles for the 48-hour lead time to 181 miles for the 96-hour lead time. Average intensity errors range from 15 mph at the 48-hour lead time to 17 mph for the 96-hour lead time.⁷

Like most hurricane preparation plans, the MFR's DSM effectively treats lead time as deterministic. Although the NHC does not forecast landfall and therefore does not systematically provide information about landfall timing uncertainty, Powell and Aberson (2001) found that implied time-to-landfall forecasts were in error by an average of over 12 hours in the 55-72 hour lead time.

In 2006, the NHC began issuing a probability forecast for winds exceeding 39 mph (tropical-storm force), 58 mph (destructive force), and 74 mph (hurricane-force) (DeMaria et al. 2009). This product integrates the uncertainty associated with track, intensity, and size error to estimate the probability of winds exceeding each threshold at each location

⁵ http://www.nhc.noaa.gov/surge/surge_intro.pdf, accessed January 25, 2017

⁶ For example https://www.hurricanezone.net/articles/saffirsimpsonscale.html, accessed January 17, 2017.

⁷ http://www.nhc.noaa.gov/verification/pdfs/1990-present_OFCL_ATL_annual_int_errors.pdf, accessed September 8, 2016.

within the next 120 hours. Examples from 2012's Hurricane Isaac are shown for the 39 and 74-mph thresholds in Figures 2a and 2b.

A third layer of uncertainty, represented in the gray shaded rectangle in Figure 3, is the dynamic nature of the forecasts. The NHC's forecasts change at least every six hours because they are updated to reflect the most recent available information. Generally, their accuracy improves with each update. However, this means that the forecast issued six or more hours earlier does not match the current forecast. The psychology literature (Gneezy 1996, Johnson and Busemeyer 2001, Kunreuther et al. 2002) indicates that people are uncomfortable with multistage probabilities (called ambiguity), and that forecast changes may introduce additional biases, such as anchoring and other order effects in hurricane risk judgements and preparation decisions (Hogarth and Einhorn 1992).

Finally, as represented in the outer, dotted rectangle in Figure 3, the information available and its quality changes from one hurricane season to the next. The NHC's forecast accuracy has been improving throughout its history. For example, the average track error for the 48-hour forecast has declined from about 119 miles in the 2000-2009 period to 85 miles in the 2010-2015 period. Therefore the distribution of forecast errors changes. Even the types of products available change regularly. For example, in the 2017 season, the NHC will add time-of-arrival contours to its wind-speed probability product, and extent-of-wind swaths to its cone graphic.⁸ The real-world experience with forecast products that decision makers gain can become outdated by the following season.

A large body of psychology research shows that people have trouble assessing uncertainty due to cognitive biases (Tversky et al. 1982), other psychological factors such as affect (Slovic 1987), and motivational biases such as optimism (Armor and Taylor 2002). Moreover, the more predictive information is available and the more distant the predictors' relationship to the target judgement, the worse people perform (Stewart and Lusk 1994). People often do not process probabilistic information well (Tversky et al. 1982), and their interpretation of probabilities depends on how the information is presented (Ibrekk and Morgan 1987, Harding et al. 1982, Halpern et al. 1989).

Experience can improve intuitive judgement (Hogarth 2001, Klein 2015), but the factors that make the hurricane preparation decision challenging—many environmental variables,

⁸ J. Franklin, Branch Chief, Hurricane Specialist Unit, NOAA/NWS/National Hurricane Center, personal communication, January 11, 2017.

frequent forecast updates, substantial irreducible environmental uncertainty, high-stakes decisions, and time pressure—also make it more difficult for people to learn through experience (Shanteau 1992, Brehmer 1980, Ericsson 2006, Kahneman and Klein 2009). Experience with prior storms can also affect hurricane preparation decision making (Trumbo et al. 2011), possibly for the worse (Dillon et al. 2011, Meyer et al. 2013). Hindsight bias in evaluating predictability (Fischhoff and Beyth 1975), outcome bias in evaluating decision quality (Baron and Hershey 1988) and recency bias (Hogarth and Einhorn 1992) could influence individuals to overemphasize these few examples—effectively, as Marines would put it, preparing for the last war. Yet every storm is different, and MFR cannot count on perfect forecasts.

There are far too few tropical storms to allow emergency managers to develop highperforming intuitive judgement based on personal experience alone. On average, fewer than 10 tropical cyclones per year arise in the North Atlantic, and only about 60% of those will make landfall (Landsea 2007). Moreover, most of these will not become hurricanes, and only about one or two will be Category 3 or higher (Landsea 1993), which is considered the threshold for evacuating the city of New Orleans. On average one Category 3 or higher hurricane makes landfall at New Orleans every 23 years (Blake et al. 2011). Even a manager nearing the end of a 30-year career would not experience enough storms to develop a highly reliable intuitive judgement about hurricane preparation decisions.

2.3. Computer-Based Simulations for Training

Given the uncertainties involved with making decisions prior to a hurricane and the infrequency of a hurricane event, the MFR requested assistance in developing training tools. Computer simulations are increasingly being used to train decision makers in operations management. Simulation is used to teach important principles such as the bull whip effect (Pasin and Giroux 2011), Kanban systems (Baranauskas et al. 2000), capacity utilization (Johnson and Drougas 2002), lean manufacturing (Wan et al. 2008), and logistics (Battini et al. 2009). Simulation seems to be a valuable learning tool because users can connect learning to a real-world experience (Faria and Dickinson 1994, Haapasalo and Hyvnen 2001, Lainema and Hilmola 2005), can receive feedback on their decisions (Bakken et al. 1992, Faria and Dickinson 1994), and can learn from their mistakes in a risk-free environment (Faria and Dickinson 1994, Baker et al. 2005, Adobor and Daneshfar 2006, Salas et al. 2009). Emergency response teams are increasingly training via simulations (Crichton et al. 2000), and integrating gaming and simulation provides a means for emergency responders to train together as a team (Jain and McLean 2005).

Project Ensayo (Becerra-Fernández et al. 2007, 2008, Nikolai et al. 2009) has developed a virtual emergency operations center to train Florida officials to respond to disasters such as hurricanes. It is focused on implementation—knowing what to do and how to do it rather than on the decisions, and its experience is team-based, not individual. In addition, its scope extends beyond hurricanes.

MetEd⁹ provides educational resources on tropical-storm forecasting that are designed for forecasters but are sometimes used by emergency managers. Some of these resources contain a scenario in which a user steps through a chronological sequence of forecast updates. Since there is only one scenario and one way to progress through the scenario, the user receives immediate feedback on the correct answer. There is only one path through the experience.

The HDS, on the other hand, is designed for an individual user or a group to train on making critical hurricane preparation decisions using many simulated storms. Like a choose-your-own adventure book or a video game, it allows the user to take many paths through any storm, with decision-specific feedback. The HDS offers several unique contributions to the operations management and decision-training literature, in particular the possibility of multiple experiences of a dynamic, multistage decision under uncertainty in a realistic context.

3. Hurricane Decision Simulator

The MFR reached out to the Naval Postgraduate School in 2014 looking for tools to help with hurricane preparation decision making. The previous August, a new Commander had taken over responsibility for the lives and mission of the MFR and Marine Forces Northern Command on the eighth anniversary of Hurricane Katrina. Like many Commanders, he had not previously lived in the hurricane-prone Gulf region. At that time, their senior leaders and emergency managers prepared for hurricanes with annual tabletop exercises, using one simulated storm in each exercise. We proposed developing a simulator to give MFR personnel practice in understanding the NHC weather forecasts, interpreting the probabilistic information, and making decisions in advance of a hurricane.

⁹ https://www.meted.ucar.edu/, accessed September 7, 2016.

The HDS is a computer-based tool, available online,¹⁰ that steps the user through the evolution of a simulated storm, showing forecasts and storm impacts. The forecast is updated at time steps representing six hours of real-world time, which corresponds to the NHC update schedule.

The user is asked to make decisions regarding preparation and evacuation based on the updated forecasts. The HDS provides a realistic learning environment for a Commander, senior officer, or emergency manager, who can practice on several simulated storms in only one hour.

3.1. User Decisions

A simulation starts with the formation of a storm in the Gulf or approximately 120 hours before the expected landfall for storms forming far from New Orleans. The HDS first displays a window describing the history of the storm to that point. Once the user dismisses the history window, the HDS shows the first forecast and asks the user the first of six decision questions as seen in Figure 4. The six key decisions are shown in Table 1 with DSM-specified lead times. Although the DSM specifies that the first decision should be made 96 hours before the arrival of hurricane-force winds, the user is given the opportunity to make the first decision at every update starting as much as 120 hours before predicted landfall.

If the user replies "No," to any question, the HDS provides the next forecast update and repeats the decision question. As illustrated in Figure 5, if she chooses "Yes", she proceeds through the sequence of decisions until she chooses "No" or takes all preparatory actions, including transferring C2. The HDS records user decisions and incorporates their timing in determining the outcomes of the simulation.

While the decisions must be taken in sequence, the preparation actions can be made at any time during the simulation. This can result in a timeline that is compressed relative to the DSM. As in the case of Hurricane Isaac, changes in the forecast indicating an increased risk to New Orleans due to changing forecast track, speed, or intensity can create this compressed timeline. While the HDS provides a forecast time-to-landfall, this time does not decline by precisely six hours in every time step, reflecting the uncertainty in track forecasts, which causes uncertainty in the critical time-to-landfall variable.

¹⁰ http://eddy.nps.edu/hurricaneSim/simulation, accessed September 10, 2016. An earlier, laptop-based version is described in MacKenzie and Regnier (2015).

Table 1

Decision	Recommended hours before landfa	.11
Deploy the advance ERS to the alternate he	eadquarters 9)6
Deploy liaison officers to local municipalitie	s' emergency operations centers 9)6
Deploy the rest of the ERS to the alternate	headquarters 7	'2
Activate personnel who will remain behind	after the evacuation 7	'2
Order evacuation or shelter in place at 60 h	ours before landfall 7	'2
Transfer C2 to alternate headquarters	4	8

Key decisions according to the MFR DSM



Figure 4 Decision prompt in the Hurricane Decision Simulator. ADVON stands for advance ERS.

The Record of Events at the left of the user interface (see Figures 4 and 7) details actions that must be taken by MFR and Marine Forces Northern Command. Some of these actions are triggered by the user's decisions, and some are triggered by storm conditions. This information provides a rich and realistic context and helps and familiarize users with the hurricane operations procedures. For example, if the user decides to deploy the advance ERS, 12 simulated hours later, the HDS announces with a popup message that the ERS arrived at the alternate headquarters and has begun to establish communications.

The HDS also provides information on actions taken by local and state authorities that may affect MFR outcomes. If the city begins an evacuation, it plans to begin evacuating its less mobile citizens about 60 hours before landfall, assuming the decision can be made in time. The state of Louisiana's evacuation plan is also time dependent. Areas south



Figure 5 Decision tree representation of the dynamic decision problem represented in the HDS.

of the Intracoastal Waterway outside the levee protection plan to begin evacuation 50 hours before landfall. Areas south of the Mississippi River which are levee protected but remain vulnerable plan to begin evacuation 40 hours before landfall. The rest of New Orleans begins to evacuate 30 hours before landfall, and contraflow is implemented. The HDS provides immediate announcements when these evacuations are ordered, and provides outcomes dependent on state and local decisions. For example, if the user waits to order an evacuation of MFR until the time that New Orleans orders the evacuation, the post-simulation feedback reports that MFR personnel are caught in traffic, and their arrival to the alternate headquarters may be delayed by 24 hours or more. Local and state actions and consequences were developed based on interviews with local officials and planning documents.

3.2. Simulated Storms

In order to give users the opportunity to experience decades' worth of storms in a few hours, the HDS requires hundreds of distinct simulated storms. They must be realistic enough to increase users' confidence in the learning value of the tool. As described in Section 2.2, uncertainty in the future behavior of the storm is an important feature of the decision context. Therefore, the simulated storms need to span the diversity of possible storm and

17

forecast behavior so that users encounter not just typical storms but also realistic storms with unexpected behavior.

The objectives of realism and diversity can be conflicting, which poses a modeling challenge. The goal is to reflect typical behavior with a realistic degree of unexpected behavior. Figure 6 shows real and simulated storms, side-by-side with actual track and forecast tracks shown. The storm in Figure 6a shows 2011's Tropical Storm Lee, which, like Humberto, formed close to shore. Unlike Humberto, it did not quite reach hurricane strength, but it did pass close to New Orleans. Many residents of the Gulf coast are surprised that this can happen, and the long-lead hurricane preparation plans do not account for this eventuality.



(e) Hurricane Noel, 2007
 (f) HDS simulated storm
 Figure 6 Real (left) and simulated (right) storms. Actual track is shown in red, forecast tracks are shown in blue, and notional 39, 58, and 74-mph wind swaths are shown in shades of gray.

3.2.1. Storm model Statistical storm models generally summarize typical or central behavior. For example, linear models such as CLIPER (for climatology and persistence) models are among those used to predict track (Aberson 1998), intensity (Knaff et al. 2003), and radius (Knaff et al. 2007). They estimate the expected value of the target variable. The estimation of such a model also produces error statistics, which could be used to generate simulated storm tracks by generating random deviations from the central forecast. However, deviations from centrality are correlated over time, and such correlations would need to be captured to produce realistic tracks. Moreover, the correlations are themselves a function of predictor variables such as latitude, intensity, and whether the storm center is over land.

Rather than specify a complex error model, we follow Regnier and Harr (2006) and Emanuel et al. (2006), in using a first-order discrete-time Markovian model to generate the simulated storm's track (its center positions). The model is fit using NHC historical track data from 14,882 observations of 542 storms from 1980 to 2014 as found in the NHC's best-track database (Jarvinen et al. 1984). Observations are in six-hour time increments, and each consists of the storm's

center position (latitude and longitude in 0.1° increments), intensity (velocity of maximum sustained winds in 5-kt increments), and size (radius of maximum winds in nautical miles). Size is not always available.

The states in the Markov model are defined using a k-means clustering algorithm with the following normalized variables for each observation: (i) intra-cluster great-circle distances in the current period, (ii) intra-cluster great-circle distances six hours later, (iii) the speed of forward motion, and (iv) a binary variable indicating if the storm center is over land. Each cluster defines a state in the Markov model. The mean latitude, longitude, intensity, and radius of maximum winds are calculated for each cluster and used to determine the wind speed at New Orleans for a storm in this modeled state.

The transition probabilities among states are proportional to the observed frequency of transitioning from one state to another state in the data set, or dissipating in the next time period. The transition matrix is used to calculate the probability that hurricane-force (79-mph) winds will eventually affect New Orleans, conditional on the state.

Although the vast majority of tropical cyclones tracked by NHC never threaten New Orleans, the purpose of the HDS is to provide experience with storms that would require active attention of the Commander of MFR. Tracks of interest to MFR are generated by selecting states in which the probability of the storm with hurricane-force winds affects New Orleans is at least 5%. Because they are designed to be much more likely than an arbitrary storm to affect New Orleans, they tend to be larger and more intense. However, their paths and forecast errors are representative of real storms in the Gulf area, as seen in Figure 6. Consistent with reality, most of the storms in the HDS do not produce hurricane-force winds at New Orleans.

The storm's intensity and size are generated using linear models. The storm's intensity begins with the cluster mean intensity, and six-hour changes are generated from a normal distribution with a mean and standard deviation derived from a linear regression with the current latitude, current intensity, the prior period's change in intensity, and the overland indicator variable as predictors. The storm's size is similarly modeled via linear regression of six-hour changes predicted on current intensity, current size, and the overland indicator, with size change applied with a 50% probability in each period.

3.2.2. Forecasts At each update, the HDS displays forecast products that replicate many of the NHC's products:¹¹

• a map showing the forecast track, overlaid with a cone showing the region where the storm's center position is likely to fall;

• maps showing the probability of winds exceeding each of three thresholds (39 mph, 58 mph, and 74 mph) in the next 120 hours following the update (Examples are shown in Figures 7a - 7c);

• a text-based update, shown in summary form overlaid on the maps, which may be expanded to show a narrative description of the forecast (see Figure 7d), similar to NHC's text-based advisories.

To produce these products, for every six hours of the storm's progress it needs new forecasts. Each forecast includes lead times of 12, 24, 36, 48, 72, 96, and 120 hours,¹² unless the storm is forecast to dissipate earlier. Figure 7 shows updated forecasts—7c is identical to 7a except it reflects a later update. Each forecast must be consistent with the

¹¹ In addition to the examples in Figure 2, the remaining forecasts for 2012's Hurricane Isaac can be viewed at http: //www.nhc.noaa.gov/archive/2012/ISAAC_graphics.shtml for the maps and http://www.nhc.noaa.gov/archive/ 2012/ISAAC.shtml for the text products.

 $^{^{12}}$ Longer lead forecasts are not simulated as MFR decisions are made with less than 120 hours' lead time, and the error data are very sparse at longer lead times.

simulated storm's actual events and presented in a format similar to recent NHC forecast products. The simulated forecast errors should be similar to their real counterparts so that the intuition users gain about risky decisions in the HDS is relevant to a real storm in the upcoming season.

The simulated forecasts are generated by starting with the simulated actual values of the four variables for the storm, and adding errors generated from distributions similar to NHC forecast error distributions during the most recent five years for which data were available (2010-2014).¹³ This is consistent with NHC's practice of reporting error statistics for the most recent five years to balance the value of reflecting the latest forecasting processes against the value of a larger data set in order to reduce the impact of inter-storm and inter-seasonal variability. As NHC forecast accuracy improves over time, simulated forecasts in the HDS should be replaced with forecasts whose errors correspond to the most recent NHC errors in order to maintain consistency between simulated forecasts and real forecasts. Table 2 summarizes the track errors for the last ten years from NHC,¹⁴ together with the HDS track errors for comparison. While track errors are generated using a Monte Carlo approach, with distributional parameters matching those of NHC's along-track and cross-track errors from the 2010-2014 period, the HDS track errors are reasonably close in both mean and standard deviation to errors observed during the 2015 and 2016 seasons.

To make the forecasts realistic, the correlations among forecast errors must be realistic. For a given forecast update, track errors tend to be positively correlated across lead times. For example, if the 36-hour forecast speed is very fast (slow), the 48-hour forecast speed is also likely to be very fast (slow). Errors are also positively correlated across updates. For example, if Monday morning's forecast of Wednesday morning's position is too far west, it is likely that Tuesday morning's forecast of Wednesday morning's position is also too far west. Since a given storm could have 20 or more updates before it makes landfall or dissipates, and each update has forecasts at seven lead times, there are well over 100 intercorrelated errors per variable in each storm. The error data set is too small to estimate a good covariance matrix for all the forecasts for all updates are simulated from a multivariate normal distribution with a covariance matrix correlated at 0.9 for adjacent updates,

¹³ downloaded March 6 and 7, 2015 from http://www.nhc.noaa.gov/verification/verify7.shtml.

¹⁴ calculated using NHC error data downloaded September 22, 2017 from http://www.nhc.noaa.gov/verification/ verify7.shtml.

 Table 2
 Comparison of recent NHC track errors with HDS track errors. Average (standard deviation) of errors are in nautical miles.

Forecast lead time (hours)

year	12	24	36	48	72	96	120		
2007	33(22)	51 (32)	71(43)	92(47)	146(74)	167 (110)	258(153)		
2008	28(18)	48 (32)	69(46)	88 (61)	127 (79)	160 (97)	192(111)		
2009	30(22)	45(29)	62(36)	73(33)	119(54)	198 (79)	292~(116)		
2010	34(22)	53(32)	70(42)	88(55)	129(76)	166 (104)	187(127)		
2011	28(17)	43(28)	57(36)	71 (43)	110(78)	167(106)	245 (146)		
2012	25(16)	40(25)	54(35)	69(48)	101 (79)	143 (87)	194~(103)		
2013	28(19)	49 (31)	72(39)	103 (58)	141 (123)	166 (125)	165~(69)		
2014	26(20)	38(23)	51(32)	65(41)	100(70)	162 (91)	279(213)		
2015	25(16)	40(25)	56(37)	77~(59)	125 (107)	188 (178)	241 (191)		
2016	24(14)	36(23)	48(32)	62 (43)	89(53)	133~(67)	168 (91)		
HDS	27 (14)	36 (19)	35(20)	54(28)	75~(56)	141 (78)	207 (118)		

and each variable's variance matches the data set. For each update, the forecast errors (from which forecasts are calculated) are generated from a multivariate normal distribution matching the previously generated 72-hour errors, using the covariance matrix of the historical forecast errors. Intensity errors are generated similarly. Some features, such as estimated time-to-landfall, require forecasts for smaller time intervals, which are interpolated between the above-listed lead times. Each update includes a forecast of the expected time-to-landfall for the storm, which is estimated by assuming the current forecast track is correct and calculating the time steps until the track crosses the U.S. coastline along the Gulf of Mexico.

The wind-speed probability forecasts are one of the most important of NHC's forecast products. For the HDS, wind-speed probability forecasts were generated by Charles R. Sampson of the Naval Research Laboratory in Monterey using the same code used by the NHC. Three maps—for the 39-mph, 58-mph, and 74-mph thresholds—are shown in the HDS, and the user can switch among maps using the buttons at the top of the map (see Figure 7). The wind-speed probability forecasts for New Orleans are also given in the text product, and shown in the detail view in Figure 7d.

















3.3. User Feedback

After the storm makes landfall or dissipates, the HDS provides narrative feedback in a Results/Feedback window. At the top, a Results section describes direct storm impacts, such as the flood level and wind speeds, as well as indirect impacts such as damage to physical infrastructure. This section also summarizes the decisions taken and their timing.

Next, a Feedback section, as shown in Figure 8, describes outcomes resulting from the user's decisions, including the effects of the relative timing of the decisions. The outcome model is based on our discussions with MFR subject matter experts, including representatives from the personnel, operations, logistics, and communications departments in MFR. Each decision can incur monetary costs, disrupt normal operations, or create difficulties for MFR in completing other operations. The benefits of preparation are mitigating the consequences of a hurricane if it does reach New Orleans. Subject matter experts at MFR were asked to describe how decisions would impact costs and benefits for different size storms. Outcomes are a function of the decisions taken by the user and the severity of the storm.

One of the outcomes displayed in the Feedback section is the monetary cost of the user's decisions. The subject matter experts revealed that three decisions incur monetary costs: (i) deploying the advance ERS, (ii) deploying the rest of the ERS, and (iii) evacuating the MFR headquarters. The decisions to deploy the advance ERS and the rest of the ERS cost relatively little and are represented as fixed costs in the HDS: \$125,000 total. The estimated cost of evacuating the MFR headquarters is \$300,000 for each day that MFR is evacuated. MFR subject matter experts estimated that if a hurricane does not reach New Orleans, MFR personnel could return about 3 days after the storm ends. If a hurricane reaches New Orleans, MFR personnel could return in 3-7 days, but it could be as many as 14 days before personnel return. In the Feedback section, the HDS provides verbal feedback "A return to New Orleans will likely be scheduled in 3-7 days, but it could be as many as 14 days before people can return." to give the user a sense of the variability and estimates the cost of a 10-day post-hurricane evacuation period. For example, if a user orders an evacuation 72 hours (i.e., 3 days) before a storm dissipates and no hurricane strikes New Orleans, the HDS calculates that the total monetary cost is 125,000 + 300,000(3+3) =\$1,925,000. If the user makes the same decision but a hurricane strikes New Orleans, the cost is \$125,000 + \$300,000(3 + 10) = \$4,025,000.



Figure 8 Results/Feedback window in the Hurricane Decision Simulator, scrolled down to show Feedback.

Some outcomes depend only on the relative timing of decisions. For example, the advance ERS is supposed to be deployed to the alternate headquarters 36 hours before an evacuation is ordered. If the advance ERS is deployed late, the staff might not have time to set up the alternate headquarters, and there is a risk that some mission-essential functions may not be performed, at least initially. This information is provided to the user if the user compresses the timeline in the HDS, as seen in the Feedback section shown in Figure 8.

The Feedback section also describes outcomes that are a function of the user's decisions, their timing, and the storm impacts. For example, if the decision to evacuate is taken too late, the user receives feedback stating that personnel were in danger on the roads during tropical-storm force winds or flooding. Thus, the user can begin to understand the impact of uncertainty in time-to-landfall.

The most important outcomes that depend on both decisions and the storm are the consequences of severe weather that could have been avoided by preparation. Eliciting information about some of these outcomes was perhaps the most difficult part of building the HDS because MFR leadership was hesitant to describe what might happen if the DSM and Execution Matrix were not followed with the intended lead times.

The MFR emergency manager did describe the principal infrastructure damage and difficulties that a hurricane would cause. Infrastructure damage would include loss of power, communication failures, flooded roadways, and a shortage of water, fuel, and food. His description revealed that even if the MFR headquarters is operable, MFR personnel would find it difficult to come to work if no evacuation is ordered. The ability of the MFR to continue operations if no evacuation occurs would depend on MFR's ability to resupply its generators and the status of local infrastructure. The Feedback section describes this scenario if a user does not evacuate and hurricane strikes New Orleans. In this manner, the HDS provides a qualitative description of the non-monetary consequences.

4. Benefits of Simulation Tool

The MFR has begun to use the HDS in its individual and team-based hurricane preparation exercises. Their experiences show the value of the HDS in improving MFR's hurricane preparation training and decision processes by:

- increasing the number of simulated storms they experience, at lower cost;
- focusing training on uncertainty and ambiguity inherent in the decision process; and
- improving MFR personnel's understanding of the forecast products.

These improvements in the training process, provided in a realistic environment with rich, specific, decision-dependent feedback, can be expected to reduce known biases in decision-making under uncertainty. This section also discusses the challenges to quantifying decision quality in this context and the opportunities that the HDS provides for studying dynamic decisions under uncertainty.

4.1. Experience more simulated storms

The MFR holds two group hurricane preparation exercises per year. In the spring, a functional exercise focuses on implementation of hurricane-preparation operations. The 40+ participants are mid-level to senior-level officers responsible for functional areas including personnel, operations, and logistics. In this exercise, MFR units perform some of the actions that they would need to undertake during a hurricane. An August exercise brings together more senior participants—principals and deputies from various directorates—for a tabletop exercise focusing on decision making. The primary purposes of the August exercise are to validate the decision timelines and to demonstrate the tension between wanting to evacuate Marines as soon as possible in the event of a hurricane and the high financial and other costs of evacuation.

Before the development of the HDS, each exercise centered around one simulated storm which the MFR's Emergency Management Program Specialist developed using composites of historical storms or transpositions of historical storms to different locations. He created each storm and its forecasts and injected updates periodically during the exercise to represent the progress of the storm. In 2016, the MFR used the HDS for its August exercise in place of a hand-created scenario, saving approximately 30 hours of preparation time.¹⁵

In addition, the HDS enabled the participants to experience two storms in the exercise, rather than just one as in prior years. MFR leadership plans to continue using the HDS for both of its annual group exercises (J. Garcia, personal communication, September 14, 2016).

Because the HDS provides a scenario in seconds, it creates the opportunity for individual training. Individual users can experience a career's worth of threatening storms in a day, with the current forecast products and in the context of the MFR preparation timeline. MFR leadership has asked its Crisis Action Team and ERS, approximately 70 people, to continue to practice with the HDS on an individual basis to become more familiar with the forecasting products, the uncertainty, and the evolution of the uncertainty as lead time declines (J. Garcia, personal communication, September 14, 2016). This is additional training that was not available before the HDS.

4.2. Focus on decision making with uncertainty

In addition to providing the opportunity to experience many more storms at a lower cost, the feedback from MFR indicates that the HDS focused users on why and when a decision might be taken while considering a realistic level of uncertainty. Marines are accustomed to making decisions that follow a prescribed timeline based on "cold, hard facts,"¹⁶

but the HDS demonstrated the inherent uncertainty in hurricane forecasting and the difficulty in balancing the trade-offs:

"...in 2015 ... the exercise focused more on reviewing operations and the timeline in our decision support matrix. This year, with the simulator, we focused the exercise more on the difficult decisions and how the uncertainty and ambiguity reflects real-life situations." (R.W. Mantzel, personal communication, October 4, 2016)

The hand-generated storms used in prior years frequently followed a pattern of increasing threat over the course of the storm, which the MFR exercise participants came to anticipate. Therefore, they were not exposed to the degree or type of uncertainty inherent

¹⁵ J. Garcia, MFR Emergency Management Program Specialist, personal communication, September 14, 2016.

 $^{^{16}}$ Col. R.W. Mantzel, MFR Operations and Plans Officer, who directs the Crisis Action Team, personal communication, September 1, 2016.

in real storms. In the HDS, storms sometimes change direction, lose intensity as they reach the coast, and form close to shore in the Gulf—all behavior that is realistic but infrequent enough that most staff have never experienced it. The HDS

"provided insight into the ambiguity and uncertainty inherent in hurricanes. We would like a hurricane to follow a predictable path, but hurricanes in reality have unpredictable paths. ... The simulator reflects this real-life behavior ... and improves our understanding of what can happen with storms in the Gulf of Mexico." (R.W. Mantzel, personal communication, October 4, 2016)

Experiencing many realistic storms within the HDS can be expected to reduce common biases that occur in decision making with uncertainty. For example, outcome bias tends to cause judgements of decision quality to depend too much on outcome, rather than on the appropriateness of judgements under uncertainty (Baron and Hershey 1988). For lowprobability events, such as a hurricane striking New Orleans, people tend to underestimate the likelihood because they are unlikely to have experienced the event (Hertwig et al. 2004). More experience—provided by the HDS—increases the likelihood that the low-probability event will be experienced, making the judgement of its likelihood more realistic.

4.3. Understand forecast products

Training with the HDS and the elevation of uncertainty as an important focus affected the way participants in the MFR exercises used and interpreted the forecast products. They relied more on the wind-speed probability forecasts rather than on the track forecast cone:

"We tend to view the forecast with the shaded cone as certain, meaning that we think the storm will never move outside the shaded cone. Relying completely on the cone gives a false sense of security because if New Orleans is outside of the cone, we might think New Orleans is safe. But the hurricane can deviate outside of that cone and destructive weather can occur outside the cone. ... For example, even though New Orleans might be outside of the cone, the simulator might depict a 20% chance that hurricane-force winds will reach New Orleans. This provided an important point of discussion and clarification for us during the table-top exercise." (R.W. Mantzel, personal communication, October 4, 2016)

As discussed in Section 2.2, misinterpretations of the NHC cone are common. Perception and interpretation of information is one of the layers of error that commonly reduces the quality of judgements, and reducing the mismatch between the informational products and their interpretation improves judgements (Stewart 2001). The HDS improved users' interpretation of the forecast track, cone and wind-speed probability products, and therefore can be expected to improve their judgements about hurricane risk.

The participants in the August tabletop exercise noted learning other important relationships between the forecast products and their hurricane preparation operations. Tropicalstorm force winds commonly arrive 12 or more hours before the center of the storm, which was not universally understood among the key staff at MFR before the HDS. Tropicalstorm force winds pose a hazard for personnel who are on the roads; therefore it is important to complete all travel well before the arrival of 39-mph winds. In addition, the timing of the storm relative to normal waking hours is very important. If the storm's winds are forecast to reach New Orleans in the early evening, the Commander and the Crisis Action Team can meet first thing in the morning, give the order to evacuate, and ensure the Marines are evacuated safely in the morning and early afternoon. If the storm's winds are forecast to reach New Orleans in the early afternoon, evacuating the headquarters is logistically more difficult because the order to evacuate would be given in the late afternoon of the previous day at which time many people would have already returned home. The MFR needs to be prepared for both possibilities.

4.4. Evaluating the effectiveness of HDS

The MFR senior staff's feedback and the authors' own observations and informed conclusions indicate that the HDS has been highly effective in improving hurricane training, increasing decision-relevant learning by the participants. Providing further evidence of its value, the MFR plans to make the HDS available to its Reserve Training Centers, many of which are located in hurricane-prone regions, including Jacksonville, Tampa, and Hialeah, Florida; and Galveston and Corpus Christi, Texas (J. Garcia, personal communication, September 14, 2016).

Ideally, we could assess the effectiveness of the HDS by comparing the MFR's decisions during tropical cyclones before using the HDS and real decisions made after using the HDS. However, real-world decision quality before and after using the HDS cannot be meaningfully compared for a number of reasons. The sample size of real storms is tiny and highly variable with respect to storm impacts and forecasts, as detailed in Section 2.2. Other sources of variability that cannot be controlled for include ongoing changes in NHC forecast products and the characteristics of the individuals participating in the decision such as their prior tropical-storm experiences.

A further challenge in measuring the effectiveness of the HDS is defining good decision performance or even good outcomes. The relative importance of financial impacts, disruption of operations, and personnel safety can change year-to-year, month-to-month, and even day-to-day as the threats and budget change. For example, the impact of spending \$2 million of the operations and support budget changes depending on the timing in the fiscal year and upcoming demands. The tolerance for risking reduced mission capability is sensitive to changing threats that the MFR Commander may be tracking. Using expected monetary equivalent of all consequences, as in Taylor (2007) is unrealistic. Decisions that are appropriate in one instance may not be appropriate in every similar storm.

The above-described complexity in hurricane preparation at MFR creates obstacles to assessing users' decision performance. Reducing the HDS's complexity would reduce its effectiveness. The HDS was designed to respond to the MFR's requirement for a rich and realistic environment, and as noted in Ward et al. (2006)'s review, high performance associated with expertise is highly task-specific, and the performance benefit associated with expertise is smaller in less realistic environments. In other words, the more realistic the task environment, the less we can learn from experiments, while the more stylized the simulated environment, the less applicable the learning is to the real environment.

A stylized version of the HDS could be used as a tool for studying dynamic decision making under uncertainty. This requires constructing a scenario within which it is possible to identify better and worse choices and/or outcomes. Results using such a stylized simulator may not be generalizable to real decisions, however. One reason is that the stakes are very high in real decisions, while in the simulator they are low. Levitt and List (2007) cite numerous studies showing that the magnitude of the stakes in laboratory experiments has a significant impact on behavior in studies of pro-social behavior. Moreover, outcome bias in assigning blame (McGraw et al. 2011) would tend to increase the stakes relative to the direct utility impact of financial costs, mission disruption, and health and welfare impacts in the real-world scenario. Therefore decisions that users make in the low-stakes HDS environment may not reflect their behavior in real, high-stakes decisions. The documented improved understanding of forecast products and their associated uncertainty as a result of using the HDS will allow MFR to make better-informed risk judgments.

5. Conclusions

Louisiana has a long hurricane history, and the two critical Marine Commands headquartered in New Orleans need to be prepared to continue to operate and protect their personnel in the event of a hurricane. In August 2016, Tropical Cyclone Hermine—later Hurricane Hermine—appeared in the Gulf of Mexico, demonstrating the ongoing relevance of the hurricane threat to the MFR.

Managing MFR operations in preparing for a hurricane requires a good understanding of forecasts and forecast uncertainty, and how they interact with the organization's decision and implementation processes. The HDS enables MFR leaders to develop that ability and make better operations management decisions when threatened by a hurricane, as evidenced by the MFR's experiences:

"The Hurricane Decision Simulator provides the Marines with an important tool that we will continue to use to practice our preparation for hurricanes. ... The simulator helps us to manage our operations more effectively both in terms of training and in readiness in the event that a hurricane hits New Orleans." (R.W. Mantzel, personal communication, October 4, 2016)

Because the simulated storms' forecasts replicate recent NHC forecast error distributions, the HDS provides experience that is literally impossible to gain through real-world experience, even for career hurricane experts. The fact that hurricanes occur relatively rarely prevents emergency managers and decision makers from gaining experience with the range of possible storm behavior and with the latest forecast products. Before the development of the HDS, individuals would experience one or two simulated storms per year during group exercises, about one hurricane threat per year, a hurricane approximately every ten years and a major hurricane (Category 3 or higher) once every 23 years (Blake et al. 2011). As the NHC updates its products, the HDS can be updated so that simulated versions of the new products are available, providing the possibility of training in advance on new NHC products.

The HDS provides a unique tool to help senior leaders practice making decisions in a dynamic and uncertain environment. The HDS can also be modified to become a valuable tool for civilian emergency responders and city and state officials who also need to make high-stakes preparation decisions in advance of a hurricane. Other situations which require quick decisions in dynamic and uncertain environments include responding to emergencies and natural disasters, military conflicts, and business disruptions. Developing simulations that adequately capture both uncertainty and how it changes over time for these rare but high-impact events can help senior decision makers be better prepared to deal with these situations and successfully manage their organizations through a crisis.

While the MFR Commander does not face concerns regarding the public's response to evacuation orders, the dynamic decision with ever-changing considerations is very similar to that faced by public officials and private-sector planners. Rather than seeking to model all aspects of teh decision and develop optimized policy recommendations that are unlikely to be directly prescriptive, the HDS serves to prepare the Commander and his staff to operate in this challenging environment. Building on prior descriptive work on hurricane preparation behavior and optimization models for preparation and evacuation timing, our experience indivates that further development of tools to train judgmental decision making could improve decisions and outcomes in this important and challenging context.

Acknowledgments

The authors gratefully acknowledge the support of the Marine Forces Reserve, in particular Bob McGuiness and the Information/Knowledge Management team and Jose Garcia, the Center for Design, Development and Distribution and Andy Hernandez at the Naval Postgraduate School's, and Charles R. Sampson of the Naval Research Laboratory, Monterey

References

- Aberson, S. D. (1998). Five-day tropical cyclone track forecasts in the North Atlantic basin. Weather and Forecasting 13(12), 1005–1015.
- Adobor, H. and A. Daneshfar (2006). Management simulations: Determining their effectiveness. Journal of Management Development 25(2), 151–168.
- Apivatanagul, P., R. A. Davidson, and L. K. Nozick (2012). Bi-level optimization for risk-based regional hurricane evacuation planning. *Natural Hazards* 60(2), 567–588.
- Armor, D. A. and S. E. Taylor (2002). When predictions fail: The dilemma of unrealistic optimism, Chapter 19, pp. 334–347. Cambridge University Press.
- Baker, A., E. O. Navarro, and A. Van Der Hoek (2005). An experimental card game for teaching software engineering processes. *Journal of Systems and Software* 75(1), 3–16.
- Baker, E. J. (1991). Hurricane evacuation behavior. International Journal of Mass Emergencies and Disasters 9(2), 287–310.
- Bakken, B., J. Gould, and D. Kim (1992). Experimentation in learning organizations: A management flight simulator approach. European Journal of Operational Research 59(1), 167–182.

- Baranauskas, M., N. G. Gomes Neto, and M. A. Borges (2000). Gaming at work: A learning environment for synchronized manufacturing. *Computer Applications in Engineering Education* 8(3–4), 162–169.
- Baron, J. and J. C. Hershey (1988). Outcome bias in decision evaluation. Journal of Personality and Social Psychology 54(4), 569–579.
- Battini, D., M. Faccio, A. Persona, and F. Sgarbossa (2009). Logistic game: Learning by doing and knowledgesharing. Production Planning and Control 20(8), 724–736.
- Becerra-Fernández, I., G. Madey, M. Prietula, and D. Rodríguez (2007). Project Ensayo: A virtual Emergency Operations Center for disaster management research, training and discovery. In Second International Conference on Internet Monitoring and Protection, pp. 31–31.
- Becerra-Fernández, I., G. Madey, M. Prietula, D. Rodríguez, R. Valerdi, and T. Wright (2008). Design and development of a virtual Emergency Operations Center for disaster management research, training, and discovery. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences*, pp. 27–27.
- Blake, E. S., C. W. Landsea, and E. J. Gibney (2011, August). The deadliest, costliest, and most intense United States tropical cyclones from 1851 to 2010 (and other frequently requested hurricane facts). NOAA Technical Memorandum NWS NHC-6, NOAA/National Weather Service, National Hurricane Center. NOAA Technical Memorandum NWS NHC-6.
- Brehmer, B. (1980). In one word: Not from experience. Acta Psychologica 45(1), 223–241.
- Broad, K., A. Leiserowitz, J. Weinkle, and M. Steketee (2007). Misinterpretations of the "Cone of Uncertainty" in Florida during the 2004 hurricane season. Bulletin of the American Meteorological Society 88(5), 651–667.
- Crichton, M. T., R. Flin, and W. A. Rattray (2000). Training decision makers-tactical decision games. Journal of Contingencies and Crisis Management 8(4), 208–217.
- DeMaria, M., J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria (2009). A new method for estimating tropical cyclone wind speed probabilities. Weather and Forecasting 24(6), 1573–1591.
- Dillon, R. L., C. H. Tinsley, and M. Cronin (2011). Why near-miss events can decrease an individual's protective response to hurricanes. *Risk Analysis* 31(3), 440–449.
- Dow, K. and S. Cutter (1998). Crying wolf: Repeat responses to hurricane evacuation orders. Coastal Management 26, 237–252.
- Emanuel, K., S. Ravela, E. Vivant, and C. Risi (2006). A statistical deterministic approach to hurricane risk assessment. Bulletin of the American Meteorological Society 87(3), 299–314.
- Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance, Chapter 38, pp. 685–704. In K. A. Ericsson and Feltovich K. A. Ericsson and Feltovich (2006).

- Faria, A. J. and J. R. Dickinson (1994). Simulation gaming for sales management training. Journal of Management Development 13(1), 47–59.
- Fischhoff, B. and R. Beyth (1975). "I knew it would happen": Remembered probabilities of once-future things. Organizational Behavior and Human Performance 13(1), 1–16.
- Gladwin, C. H., H. Gladwin, and W. G. Peacock (2001). Modeling hurricane evacuation decisions with ethnographic methods. *International Journal of Mass Emergencies and Disasters* 19(2), 117–143.
- Gneezy, U. (1996). Probability judgments in multi-stage problems: Experimental evidence of systematic biases. Acta Psychologica 93(1), 59–68.
- Haapasalo, H. and J. Hyvnen (2001). Simulating business and operations managementa learning environment for the electronics industry. *International Journal of Production Economics* 73(3), 261–272.
- Halpern, D. F., S. Blackman, and B. Salzman (1989). Using statistical risk information to assess oral contraceptive safety. Applied Cognitive Psychology 3(3), 251–260.
- Harding, C. M., J. R. Eiser, and C. M. Kristiansen (1982). The representation of mortality statistics and the perceived importance of causes of death. *Journal of Applied Social Psychology* 12(3), 169–181.
- Hasan, S., S. Ukkusuri, H. Gladwin, and P. Murray-Tuite (2010). Behavioral model to understand householdlevel hurricane evacuation decision making. *Journal of Transportation Engineering* 137(5), 341–348.
- Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science* 15(8), 534–539.
- Hogarth, R. M. (2001). Educating Intuition. Chicago: University of Chicago Press.
- Hogarth, R. M. and H. J. Einhorn (1992). Order effects in belief updating: The belief-adjustment model. Cognitive Psychology 24(1), 1–55.
- Holguín-Veras, J., N. Prez, S. Ukkusuri, T. Wachtendorf, and B. Brown (2007). Emergency logistics issues affecting the response to Katrina: A synthesis and preliminary suggestions for improvement. *Trans*portation Research Record (2022), 76–82.
- Horner, M. W. and M. J. Widener (2011). The effects of transportation network failure on people's accessibility to hurricane disaster relief goods: a modeling approach and application to a Florida case study. *Natural Hazards* 59(3), 1619–1634.
- Ibrekk, H. and M. G. Morgan (1987). Graphical communication of uncertain quantities to nontechnical people. Risk Analysis 7(4), 519–529.
- Jain, S. and C. R. McLean (2005). Integrated simulation and gaming architecture for incident management training. In *Proceedings of the Winter Simulation Conference*, pp. 904913. IEEE.
- Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis (1984). A tropical cyclone data tape for the North Atlantic: Contents, limitations, and uses. NOAA Technical Memorandum NWS NHC 22.

- Johnson, A. C. and A. M. Drougas (2002). Using Goldratt's game to introduce simulation in the introductory operations management course. *INFORMS Transactions on Education* 3(1), 20–33.
- Johnson, J. G. and J. R. Busemeyer (2001). Multiple-stage decision-making: The effect of planning horizon length on dynamic consistency. *Theory and Decision* 51(2-4), 217–246.
- K. A. Ericsson, N. Charness, R. R. H. and P. J. Feltovich (Eds.) (2006). The Cambridge Handbook of Expertise and Expert Performance. New York: Cambridge University Press.
- Kahneman, D. and G. Klein (2009). Conditions for intuitive expertise: A failure to disagree. American Psychologist 38(6), 515–526.
- Kailiponi, P. (2010). Analyzing evacuation decisions using multi-attribute utility theory (MAUT). Procedia Engineering 3, 163–174.
- Kang, J. E., M. K. Lindell, and C. S. Prater (2007). Hurricane evacuation expectations and actual behavior in Hurricane Lili. *Journal of Applied Social Psychology* 37(4), 887–903.
- Klein, G. (2015). A naturalistic decision making perspective on studying intuitive decision making. Journal of Applied Research in Memory and Cognition 4(3), 164–168.
- Knaff, J. A., M. DeMaria, C. R. Sampson, and J. M. Gross (2003). Statistical, 5-day tropical cyclone intensity forecasts derived from climatology and persistence. Weather and Forecasting 18(1), 80–92.
- Knaff, J. A., C. R. Sampson, M. DeMaria, T. P. Marchok, J. M. Gross, and C. J. McAdie (2007). Statistical tropical cyclone wind radii prediction using climatology and persistence. Weather and Forecasting 22(4), 781–791.
- Kunreuther, H., R. Meyer, R. Zeckhauser, P. Slovic, B. Schwartz, C. Schade, M. F. Luce, S. Lippman, D. Krantz, B. Kahn, and R. Hogarth (2002). High stakes decision making: Normative, descriptive and prescriptive considerations. *Marketing Letters* 13(3), 259–268.
- Lainema, T. and O.-P. Hilmola (2005). Learn more, better and faster: Computer-based simulation gaming of production and operations. International Journal of Business Performance Management 7(1), 34–59.
- Landsea, C. W. (1993). A climatology of intense (or major) Atlantic hurricanes. Monthly Weather Review 121(6), 1703–1713.
- Landsea, C. W. (2007). Counting Atlantic tropical cyclones back to 1900. Eos 88(18), 197-202.
- Larson, R. C., M. D. Metzger, and M. F. Cahn (2006). Responding to emergencies: Lessons learned and the need for analysis. *Interfaces* 36(6), 486–501.
- Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21(2), 153–174.
- Lindell, M. K., J. E. Kang, and C. S. Prater (2011). The logistics of household hurricane evacuation. *Natural Hazards* 58(3), 1093–1109.

- Lindell, M. K. and R. W. Perry (2012). The protective action decision model: Theoretical modifications and additional evidence. *Risk Analysis 32*(4), 616–632.
- Lindell, M. K. and C. S. Prater (2007). A hurricane evacuation management decision support system (EMDSS). *Natural Hazards* 40(3), 627–634.
- MacKenzie, C. A. and E. Regnier (2015). A hurricane decision simulator for the U.S. Marine Corps Reserve Forces in New Orleans. In *Proceedings of the* 44th Annual Meeting of the Western Decision Sciences Institute. Western Decision Sciences Institute.
- Meyer, R., K. Broad, B. Orlove, and N. Petrovic (2013). Dynamic simulation as an approach to understanding hurricane risk response: Insights from the *Stormview* lab. *Risk Analysis* 33(8), 1532–1552.
- Nikolai, C., I. Becerra-Fernandez, and M. Prietula (2009). Project Ensayo: Designing a virtual emergency operations center. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics.
- Pasin, F. and H. Giroux (2011). The impact of a simulation game on operations management education. Computers and Education 57(1), 1240–1254.
- Powell, M. D. and S. D. Aberson (2001). The accuracy of United States tropical cyclone landfall forecasts in the Atlantic basin: 1976-2000. Bulletin of the American Meteorological Society 82(12), 2749–2767.
- Rappaport, E. N. (2014). Fatalities in the United States from Atlantic tropical cyclones: New data and interpretation. Bulletin of the American Meteorological Society 95(3), 341–346.
- Regnier, E. (2008). Public evacuation decisions and hurricane track uncertainty. *Management Science* 54(1), 16–28.
- Regnier, E. and P. Harr (2006). A dynamic decision model applied to hurricane landfall. *Weather and* Forecasting 21(5), 764–780.
- Salas, E., J. L. Wildman, and R. F. Piccolo (2009). Using simulation-based training to enhance management education. Academy of Management Learning and Education 8(4), 559–573.
- Salmerón, J. and A. Apte (2010). Stochastic optimization for natural disaster asset prepositioning. Production and Operations Management 19(5), 561–574.
- Shanteau, J. (1992). Competence in experts: The role of task characteristics. Organizational behavior and human decision processes 53(2), 252–266.
- Siebeneck, L. K. and T. J. Cova (2008). An assessment of the return-entry process for Hurricane Rita 2005. International Journal of Mass Emergencies and Disasters 26(2), 91–111.
- Siebeneck, L. K. and T. J. Cova (2012). Spatial and temporal variation in evacuee risk perception throughout the evacuation and return-entry process. *Risk Analysis 32*(9), 1468–1480.
- Siebeneck, L. K., M. K. Lindell, C. S. Prater, H.-C. Wu, and S.-K. Huang (2013). Evacuees' reentry concerns and experiences in the aftermath of Hurricane Ike. *Natural Hazards* 65(3), 2267–2286.

Slovic, P. (1987). Perception of risk. Science 236(4799), 280-285.

- Smith, S. K. and C. McCarty (2009). Fleeing the storm(s): An examination of evacuation behavior during Florida's 2004 hurricane season. *Demography* 46(1), 127–145.
- Stewart, T. R. (2001). Improving reliability of judgmental forecasts. In J. S. Armstrong (Ed.), Principles of Forecasting: A Handbook for Researchers and Practitioners, Volume 30 of International Series in Operations Research and Management Science, Chapter 4, pp. 81–106. Springer Science & Business Media.
- Stewart, T. R. and C. M. Lusk (1994). Seven components of judgmental forecasting skill: Implications for research and the improvement of forecasts. *Journal of Forecasting* 13(7), 579–599.
- Taskin, S. and E. J. Lodree (2011). A Bayesian decision model with hurricane forecast updates for emergency supplies inventory management. *Journal of the Operational Research Society* 62(6), 1098–1108.
- Taylor, B. J. (2007, 6). Eastern North Carolina Marine Corps Forces and Installations (ENCMCFI) high intensity hurricane (HIH) evacuation decision support – a dynamic decision model under uncertainty and risk. Master's thesis, Naval Postgraduate School, 1 University Circle, Monterey CA. Finalist for MORS-Tisdale Award.
- Townsend, F. F. (2006). The federal response to Hurricane Katrina: Lessons learned.
- Trumbo, C., M. Lueck, H. Marlatt, and L. Peek (2011). The effect of proximity to Hurricanes Katrina and Rita on subsequent hurricane outlook and optimistic bias. *Risk analysis* 31(12), 1907–1918.
- Tversky, A., P. Slovic, and D. Kahneman (1982). Judgment under Uncertainty: Heuristics and Biases. Cambridge: Cambridge University Press.
- Uichanco, J. (2016). A robust model for pre-positioning emergency relief items before a typhoon with an uncertain trajectory.
- Wan, H.-D., F. F. Chen, and C. Saygin (2008). Simulation and training for lean implementation using web-based technology. *International Journal of Services Operations and Informatics* 3(1), 1–14.
- Ward, P., A. M. Williams, and P. A. Hancock (2006). Simulation for performance and training, Chapter 14, pp. 243–262. In K. A. Ericsson and Feltovich K. A. Ericsson and Feltovich (2006).