

The Effect of Tropical Cyclone Characteristics on U.S. Landfall Probability

Julie Lederer

July 28, 2008

Abstract

When a tropical cyclone threatens the coastline, decision makers can take preparatory actions designed to mitigate the damage caused by landfall. Those in this situation must decide whether and when to begin their preparations. Regnier and Harr have developed a dynamic decision model in which the decision maker has the option of delaying preparation and waiting for an updated, more accurate forecast. Regnier and Harr combine their decision model with a Markov model of tropical cyclone motion derived from fifty-three years of Atlantic hurricane data. We examine the state space Ω used in the Markov model with the goal of including additional state variables for storm characteristics such as direction of travel and wind speed. Logistic regression analysis is used to examine characteristics that had a statistically significant effect on landfall probability for Atlantic tropical cyclones between 1950 and 2007. The revised state space developed here will improve the Markov model, enabling Regnier and Harr's dynamic decision model to provide decision makers with more valuable information.

Contents

1	Introduction	2
2	Review of Regnier and Harr	2
2.1	The Dynamic Decision Model	2
2.2	Markov Model of Tropical Cyclone Motion	3
2.2.1	Stochastic Modeling	3
2.3	Transition Probabilities	5
2.4	Instantaneous Strike Probabilities	5
2.5	Modeling Tropical Cyclone Preparations	6
2.5.1	The Alternatives	6
2.5.2	Preparation Cost Profile	6
2.6	The Forecast	7
2.7	Dynamic Decision Making with the Markov Model	7
2.8	Expected Total Cost	8
2.9	Monte Carlo Simulation	8
2.10	Real-time Decision Making	8
2.11	Possible Extensions	9
3	Extending the Work of Regnier and Harr	9
3.1	Examining the Markov State Space	9
3.2	Methodology	9
3.2.1	Logistic Regression	9
3.2.2	HURDAT Database	10
4	Results of Logistic Regression Analysis	10
4.1	Direction of Travel	10
4.2	Wind Speed	12
4.3	Speed of Forward Motion	14
4.4	Climatological Year Type	15
4.5	Month of Origination	17
4.6	Refining the Markov State Space	17
4.7	Improved Decision Making with a Revised State Space	19
5	Conclusion	20

1 Introduction

Tropical cyclones are low pressure systems made of clusters of rotating thunderstorms. These storms form in tropical and subtropical regions, usually between 5° and 20° of the equator [1]. Tropical cyclones in the Atlantic and Eastern Pacific are labeled as tropical depressions, tropical storms, or hurricanes based on maximum sustained wind speed.¹ A tropical storm becomes a hurricane when its wind speeds reach 74 mph (64 knots).

Tropical cyclones that make landfall bring high winds and heavy rain and can spawn tornados. Even locations 100 miles or more from the place of landfall can experience massive flooding. Advanced preparation, such as moving ships from a harbor or shuttering windows, can mitigate some of these dangers. But preparation is costly and often requires a certain amount of lead time in order to be effective. One oft-cited study lists the cost of civilian evacuations as \$1M per mile of coastline [2]. One way to reduce the cost of preparation is to develop more accurate forecasts that give threatened locations greater time to prepare. A second tactic is to optimize the decision-making process of those who must decide whether and when to begin preparatory actions. This second approach is the one Regnier and Harr adopt in “A dynamic decision model applied to hurricane landfall” (2006) [3].

Regnier and Harr (hereafter referred to as RH) use a Markov model of cyclone motion in tandem with a dynamic decision model. They believe their decision model can lead to a reduction in cost for decision makers with assets at a location that is threatened by an impending cyclone. By considering the value of waiting for an updated forecast, instead of beginning preparatory actions immediately, the decision maker can avoid making costly preparations that later turn out to be unnecessary.

In this paper, we refine the state space used in the Markov model of cyclone motion. By analyzing historical data from the National Oceanic and Atmospheric Administration and examining tropical cyclone observations from 1950 through 2007, we identify storm features that affected the probability a tropical cyclone would make landfall in the U.S. Logistic regression analysis is used to determine which characteristics in which regions of the Atlantic Ocean had a statistically significant effect on whether or not a cyclone struck the coastline. Combining the results of these analyses enables us to formulate a revised Markov state space that includes such variables as direction, wind speed, and speed of forward motion. Considering a smaller region of the Atlantic while adding additional state variables enhances the Markov state space without significantly increasing the complexity of performing simulations with the Markov model. An improved model of cyclone motion can lead to better decision making when combined with the RH dynamic decision model.

2 Review of Regnier and Harr

2.1 The Dynamic Decision Model

Traditional decision models describe the hurricane preparation scenario as a series of static decisions. At each decision point, a decision maker with assets at a threatened target location chooses whether or not to prepare based on the instantaneous probability that the hurricane will strike the target. Regnier and Harr (RH) devise a dynamic decision model which purportedly leads to a reduction in the expected cost of a hurricane strike.

¹The maximum sustained wind speed is defined by the National Weather Service as the highest one-minute surface winds occurring within the circulation of the system at a height of 10 m.

In the RH dynamic decision model, the decision maker decides at each decision point to begin preparations immediately or wait for an updated, more accurate forecast. The choice at each decision point is now “prepare or wait,” as opposed to “prepare or do not prepare” with the static decision model. RH combine their dynamic decision model with a stochastic model of cyclone motion derived from historical Atlantic cyclone tracks. This cyclone model provides an indication of how the uncertainty of the forecast and the instantaneous strike probability at a particular target will evolve as the lead time declines.

RH believe that decision makers using their dynamic decision model, in tandem with the cyclone motion model, can avoid undertaking irreversible preparations that later turn out to be unnecessary. Testing their model with cyclones that hit Norfolk, VA, and Galveston, TX, RH calculate that the expected cost of the cyclone strikes is less when the dynamic, rather than static, decision model is used.

2.2 Markov Model of Tropical Cyclone Motion

2.2.1 Stochastic Modeling

First-order Markov chain models describe stochastic processes in which the state of a system at time $t + 1$ depends only upon its state at time t , and not on its state at any time before t [4].

More formally, let S_t be a random variable that describes the state of some process at time t . For $t = 1, 2, \dots$ and for each possible sequence of states s_1, s_2, \dots, s_{t+1} , then

$$Pr(S_{t+1} = s_{t+1} \mid S_1 = s_1, S_2 = s_2, \dots, S_t = s_t) = Pr(S_{t+1} = s_{t+1} \mid S_t = s_t) \quad (1)$$

RH use a first-order Markov chain model to describe cyclone motion. In this model, information on how the cyclone reached its present state (i.e., its historical evolution through time) has no bearing on the probabilities of transitioning to future states. Two cyclones in state g have the same probability of moving to state h in the next time period, regardless of how each reached state g .

In the Markov model, the *state* of a cyclone is given by the location of its center. This location is defined as the 1° latitude $\times 1^\circ$ longitude cell within the region $0^\circ - 70^\circ\text{N}$ and $0^\circ - 100^\circ\text{W}$ that contains the center of the storm. Within the region $10^\circ - 25^\circ\text{N}$ and $55^\circ - 80^\circ\text{W}$ (hereafter, Region RH), the state of a cyclone also includes its dominant direction of motion, since direction changes in this area have a critical influence on the potential landfall location. Region RH is outlined in Figure 1.

The direction of motion is defined as “north,” “west,” or “other” and is calculated through observing the change in position that occurs between discrete 6-hour time steps. Figure 2 shows the cutoffs between the three directions.

Of the 7750 possible states (70×100 position cells + 2 additional directions within Region RH $\times 375$ position cells in Region RH), only 3333 states were observed in the 538 Atlantic tropical cyclones that occurred between 1950 and 2002 — the set of storms RH used to make their model. RH added the state $j = 0$, denoting the termination of a cyclone, to the 3333 observed states to get a Markov state space Ω containing 3334 states.

The Markov chain model of cyclone motion is a *discrete-time* model because the state of the hurricane is observed only at discrete points in time, and not continuously in time. The discrete time interval used is 6 hours because the information provided in HURDAT, the National Hurricane Center’s historical database, is given in 6-hour time steps.

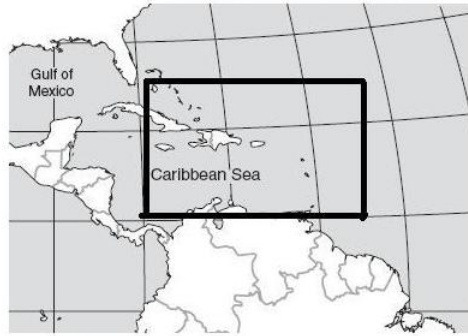


Figure 1: Map with Region RH outlined in black. Adapted from <http://www.eduplace.com/ss/maps/pdf/americas.pdf>.

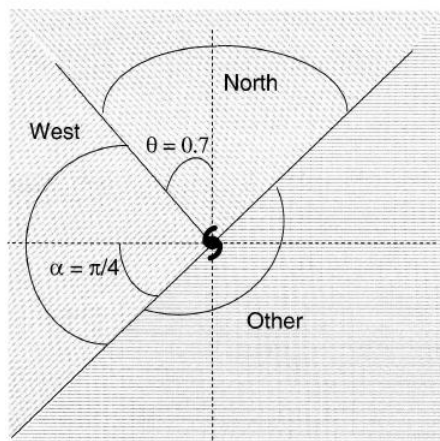


Figure 2: Depiction of the directions classified as "north," "west," and "other" in the Markov model of tropical cyclone motion. From Regnier and Harr, 2006.

In addition to being discrete-time, the Markov model employed by RH is also *finite*. That is, there are a finite number of possible states.

2.3 Transition Probabilities

Transitions between states are described by the changes in the location of the storm's center and, for storms within Region RH ($10^\circ - 25^\circ\text{N}$ and $55^\circ - 80^\circ\text{W}$), changes in the direction of motion. The transition probability q_{jk} is the probability that the cyclone is in state k at time $t + 1$, given that its state at time t is j . That is,

$$q_{jk} = Pr(s_{t+1} = k \mid s_t = j) \quad (2)$$

Each transition probability q_{jk} was calculated using the historical HURDAT database as the fraction of storms in state j that moved into state k from one 6-hour time interval to the next. The total number of transition probabilities is 3334^2 , but only 9445 are nonzero, as cyclones rarely change location significantly in one 6-hour time step.

The transition probabilities can be arranged in a 3334×3334 square matrix \mathbf{Q} , called the *transition probability matrix* of the Markov chain. The jk^{th} entry of \mathbf{Q} is q_{jk} for $j = 0, 1, \dots, 3334$ and $k = 0, 1, \dots, 3334$. That is,

$$\mathbf{Q} = \begin{bmatrix} q_{0,0} & q_{0,1} & \cdots & q_{0,3334} \\ q_{1,0} & q_{1,1} & & \vdots \\ \vdots & & \ddots & \vdots \\ q_{3334,0} & \cdots & & q_{3334,3334} \end{bmatrix} \quad (3)$$

The matrix \mathbf{Q} is a *stochastic matrix* because all of its entries are nonnegative and each of its rows sums to 1. That is, $q_{jk} \geq 0$ for all states j and k and $\sum_{k=0}^{3334} q_{jk} = 1$ for $j = 0, 1, \dots, 3334$, since a hurricane in state j must move to some state k in the state space Ω during the next 6-hour time step.

The transition probabilities q_{jk} are derived solely from the historical cyclone tracks recorded in the HURDAT database for storms occurring between 1950 and 2002. Therefore, this stochastic model does not have the forecast accuracy of the National Hurricane Center prediction models.

2.4 Instantaneous Strike Probabilities

A cyclone is considered to strike a particular target if its center moves through the 1° latitude $\times 1^\circ$ longitude cell containing the target or any of the six $1^\circ \times 1^\circ$ cells to the immediate north, south, east, west, southeast, or southwest of the target (see Figure 3).

The number of cells within the strike zone ranges from 7 (if the target cell and the six other cells of interest are outside Region RH) to $7 \times 3 = 21$ (if the target cell and the six other cells of interest are all within Region RH where direction of travel is also a state variable). The set of states in the strike zone is denoted κ .

For each state $j \in \Omega$ (the state space), the instantaneous strike probability for a particular target, denoted p_j , is the probability that a cyclone passing through state j will eventually hit the strike zone. The value of p_j is calculated as the solution to the set of simultaneous equations:

$$\begin{cases} p_j = 1 & \forall j \in \kappa \\ p_j = \sum_{k \in \Omega} q_{jk} p_k & \forall j \in \Omega \setminus \kappa \end{cases} \quad (4)$$

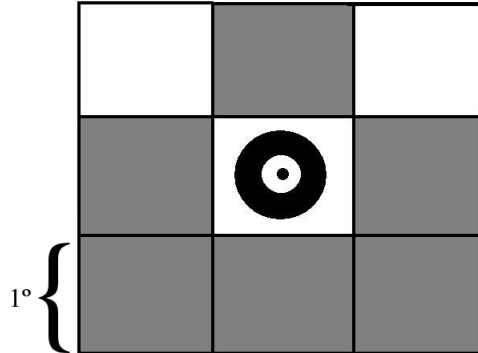


Figure 3: The middle square with the bull’s eye and the six surrounding shaded cells make up the strike zone.

The first of equations 4 says that for all states j in the strike zone κ , the probability that a cyclone in state j will hit the strike zone is 1. The second equation indicates that the probability that a cyclone passing through state j (where j is not in the target’s strike zone) will strike the target depends on the transition probabilities from j to all other states k and the instantaneous strike probabilities p_k .

2.5 Modeling Tropical Cyclone Preparations

2.5.1 The Alternatives

In the RH model, the decision maker with assets at the threatened target location has one type of preparatory action, denoted as a , available to him. For example, a fleet commander at the naval base in Norfolk, VA, has the option of ordering a sortie of ships from port. At each decision point, the decision maker may choose to take action (in which case $a = 1$) or to delay action and wait for an updated forecast (in which case $a = 0$).

2.5.2 Preparation Cost Profile

RH define τ as the minimum possible remaining time before a cyclone strikes a target. The lead time τ_j is calculated as 6 hours times the minimum number of forward transitions (with nonzero probabilities) necessary for a storm in state j to strike the target [5]. The cost of preparation, C , is a function of the remaining lead time, τ . The critical lead time, τ_{crit} , is the lead time required to complete a certain preparation before the arrival of the storm at the target location; the critical lead time is different for different preparatory actions.

RH normalize costs and losses and define $L = 1$ as the maximum mitigable loss, that is, the fraction of mitigable damage caused when a cyclone strikes an unprepared target. The cost function $C(\tau)$ increases from $C = C_{crit}$ (the cost of undertaking the preparatory action at or before the critical lead time τ_{crit}) and approaches $L = 1$ as τ decreases to zero. The decision maker can still make preparations even if the critical lead time has already passed (if $\tau \leq \tau_{crit}$), but these actions will probably be more costly and less effective. Here, the cost of preparation is assumed to be constant and at its minimum at all times at and before τ_{crit} . The equation $C(\tau = 0) = L$ shows that the maximum mitigable loss is incurred if the

decision maker made no advanced preparations and the cyclone strikes the target. $C = 0$ if no action is taken and the cyclone does not strike the target. The cost function is specific to the decision maker and to the preparatory action under consideration.

RH chose 0.1 as the cost:loss ratio at $C = C_{crit}$. That is,

$$C_{crit}/L = C_{crit}/1 = C_{crit} = 0.1 \text{ at and before } \tau = \tau_{crit}. \quad (5)$$

The decision maker's goal is to minimize the expected cost of the cyclone, which depends on the cyclone's track and the preparation actions undertaken.

2.6 The Forecast

Track forecasts determine the instantaneous strike probabilities, p_j for $j \in \Omega$, which are used in the Markov model of cyclone motion and in both the static and the dynamic decision models. Though the track forecasts themselves are not included as parameters in the dynamic decision model, this model takes into account the value of waiting for updated forecasts.

2.7 Dynamic Decision Making with the Markov Model

RH define a policy π as a description of the action a decision maker will take in any possible state j of the Markov cyclone model. In state j , the decision maker will choose action $a_j = \pi_j \in \{0, 1\}$, where $a_j = 0$ if no action is taken and $a_j = 1$ if action is taken. In both the static decision model and the RH dynamic decision model, the decision maker has only one type of preparatory action available to him, such as boarding up windows or ordering an evacuation.

In the static decision model, with a cyclone in state j , the decision maker follows the static policy π_s and will prepare if and only if $C(\tau_j) \leq p_j L$ for $\tau_j \leq \tau_{crit}$, where $C(\tau_j)$ is the cost of the preparatory action and $p_j L$ is the expected mitigable loss if no action is taken and the cyclone strikes the target. Starting at time τ_{crit} , the decision maker applies this static policy rule at each decision point. The policy π_s is considered static because this decision rule does not take into account how the instantaneous strike probabilities will evolve as updated forecasts become available.

The dynamic policy is denoted π_D and illustrates that, in certain situations, a decision maker can benefit by delaying action until more accurate forecasts are released.

Each state j in the Markov cyclone model is assigned a value, V_j , which quantifies the expected total cost to the decision maker of a cyclone in state j . The decision maker's goal, of course, is to minimize V_j . V_j is defined in the following way:

$$V_j = 1 \text{ for all } j \in \kappa, \text{ where } \kappa \text{ is the set of states in the strike zone.} \quad (6)$$

$$V_j = C(\tau_j) \text{ for all } j \in \Omega \setminus \kappa \text{ and } a_j = \pi_D(j) = 1. \quad (7)$$

$$V_j = \sum_{k \in \Omega} q_{jk} V_k \text{ for all } j \in \Omega \setminus \kappa \text{ and } a_j = \pi_D(j) = 0. \quad (8)$$

Statement 6 says that if a cyclone reaches the strike zone without any preparations being undertaken, then the decision maker incurs the maximum mitigable loss.² Statement 7 says that if the preparation is undertaken, the value associated with state j is the cost

²Recall that $L = 1$ is the normalized maximum mitigable loss.

of preparation. Statement 8 says that if the decision maker chooses to delay preparation, the value associated with state j is the expected total cost associated with the state of the cyclone at the next decision point.

A decision maker following the dynamic decision model will prepare if the cost of preparation associated with a cyclone in state j , $C(\tau_j)$, is less than or equal to $\sum_{k \in \Omega} q_{jk} V_k$ for $\tau_j \leq \tau_{crit}$. Otherwise, the decision maker will delay taking action and reevaluate at the next decision point.

2.8 Expected Total Cost

RH compare the expected total cost of a cyclone to a decision maker under both the static decision model and the dynamic decision model. For each of two targets – one at Norfolk, VA and one in Galveston, TX – they vary the critical lead time τ_{crit} from 120 hours to 6 hours and test both a linear and exponential cost function.³ The savings gained from using the dynamic instead of static model range from 0% to 6% for Norfolk and 0% to 8% for Galveston, depending on τ_{crit} and the cost function used. The savings are highest when the critical lead time is between 24 and 60 hours. In this window, updated forecasts often bring valuable information, so decision makers can benefit by delaying action for 6 to 12 hours and waiting for more accurate forecasts. Improving the decision making process may be less costly than reducing the critical lead time for a preparatory action.

2.9 Monte Carlo Simulation

The RH dynamic decision model can help prevent costly false alarm preparations when a decision maker delays action and an updated forecast shows preparation to be unnecessary. This situation might occur if a cyclone that appears to be heading for the coastline later recurves and begins moving out to sea. However, if the updated forecast shows a cyclone strike to be more likely, the decision maker must now undertake expedited, more costly preparations and faces the risk that the necessary actions might not be completed before the target is hit.

In order to study whether the net benefit of the dynamic model over the static model is positive or negative, RH use Monte Carlo simulation to generate 10,000 storm tracks using their Markov cyclone model. They find that using the dynamic over the static decision model for simulated cyclones decreases the total expected cost. False alarms are reduced by about 25%, but the number of delayed preparatory actions and strikes on unprepared targets increases slightly.

2.10 Real-time Decision Making

RH name several problems with using their dynamic decision model:

- The decision model must be adapted to the individual decision maker’s cost function and the preparatory action under consideration.
- The stochastic Markov model of cyclone movement, necessary for dynamic optimization in the RH decision model, is based purely on the historical cyclone tracks as recorded in the HURDAT database. Therefore, this model does not have the forecast accuracy of the National Hurricane Center prediction models.

³Recall that if a certain preparatory action’s critical lead time is 120 hours, this means that 120 hours are required to complete the preparation before the arrival of the storm at the target location.

2.11 Possible Extensions

RH mention several possibilities for further research:

- Repeat the analysis for typhoons in the North Pacific.
- Expand the model to include preparatory actions that can occur in stages.
- Expand the state space of the Markov model of cyclone motion by including additional atmospheric parameters, such as wind speed.

3 Extending the Work of Regnier and Harr

3.1 Examining the Markov State Space

As mentioned in the previous section, RH list the expansion of the Markov state space as a possible extension to their research. Our objective was to refine the state space and thereby improve the quality of the Markov model.

However, adding state variables to the state space quickly leads to increasing complexity. For instance, enlarging the size of the state space from n to $n + 1$ increases the size of the transition matrix \mathbf{Q} by $(n + 1)^2 - n^2 = 2n + 1$. This makes it prudent to be systematic in the determination of which factors – such as wind speed and pressure – to include in the state space. If the value of a particular characteristic in a particular region of the Atlantic affected the probability a tropical cyclone in that region with that value would eventually make landfall in the U.S., we wanted to add this characteristic in this region to the state space. If not, we would disregard this characteristic in this region in order to keep the model from growing unwieldy.

For example, if it was found that cyclones traveling west in the region $20^\circ - 25^\circ\text{N}$ and $60^\circ - 65^\circ\text{W}$ were significantly more likely to make landfall than cyclones traveling north or otherwise in this region, then direction of travel in $20^\circ - 25^\circ\text{N}$ and $60^\circ - 65^\circ\text{W}$ would be added as a state variable to the state space Ω . Logistic regression was used to perform this analysis.

3.2 Methodology

3.2.1 Logistic Regression

Logistic regression is used to predict the probability an event will occur given the values of one or more predictor variables. The response, or dependent, variable equals 0 if the event does not occur and 1 if the event occurs. The predictor, or independent, variables may be quantitative (numerical) or qualitative (categorical).

The logistic curve P gives the probability the response variable equals 1 (the event occurs) given the values of the n predictor variables:

$$P = \frac{e^{b_0 + b_1 x_1 + \dots + b_n x_n}}{1 + e^{b_0 + b_1 x_1 + \dots + b_n x_n}} \quad (9)$$

The numbers b_0, b_1, \dots, b_n are the regression coefficients and indicate the relationship between the predictor variables and the probability the response variable equals 1. For example, if b_i is positive for some i between 1 and n , then an increase in the value of x_i increases

the probability the event in question will occur. Note that the value of P is between 0 and 1, since P represents a probability.

In this study, landfall was used as the response variable, with landfall equaling 1 if the tropical cyclone made landfall in the U.S. and 0 otherwise. The effect of various predictor variables in different regions on the probability of landfall was examined. The objective was to determine which of the following factors affected the probability of landfall at the 0.05 level of significance:

- Direction of travel
- Wind speed
- Speed of forward motion
- Climatological year type (El Niño or La Niña)
- Month of origination

The statistical software package STATA was used to carry out the logistic regression analysis.

3.2.2 HURDAT Database

The HURDAT dataset, provided by the National Oceanic and Atmospheric Administration, was used in this analysis. Recall that this database provides observations in 6-hour time intervals; each observation gives information on the storm's position and intensity. Only tropical cyclones that occurred between 1950 and 2007 were considered. Reconnaissance aircraft, radar, and satellite technology were not widely used before 1950, making observations of earlier cyclones less reliable.

Furthermore, properly addressing the question *Did characteristic X in region Y have a statistically significant effect on whether or not a tropical cyclone made landfall in the U.S.?* required the post-landfall observations to be disregarded. To do this, 27 points along the coastline from Veracruz, Mexico, to Cape Cod, Massachusetts, were connected using 26 line segments (see Figure 4). For each observation t of a particular cyclone, a vector connecting the cyclone's position at time $t - 1$ to its position at time t was formed. Landfall was said to occur if the position vector intersected any of the 26 coastal line segments. All of a cyclone's post-landfall observations were eliminated from the analysis.

4 Results of Logistic Regression Analysis

4.1 Direction of Travel

For each observation between 1950 and 2007, the landfall response variable was assigned to equal 1 if the cyclone eventually made landfall in the U.S. and 0 otherwise. Each observation was also assigned a direction of travel. The direction of travel at time t for a particular cyclone was calculated using the change in the cyclone's position from time $t - 1$ to time t . The direction of travel at time $t = 1$ for each cyclone (that is, at the time of the first observation) was assigned to be the same as the direction of travel at time $t = 2$. The same cutoffs between north, west, and other used by RH were used in this study. The histogram in Figure 5 shows the relative frequencies of each direction.



Figure 4: Map showing the 26 coastal line segments used to eliminate the post-landfall observations. Adapted from <http://www.worldatlas.com/webimage/countrys/namerica/naoutl.htm>.

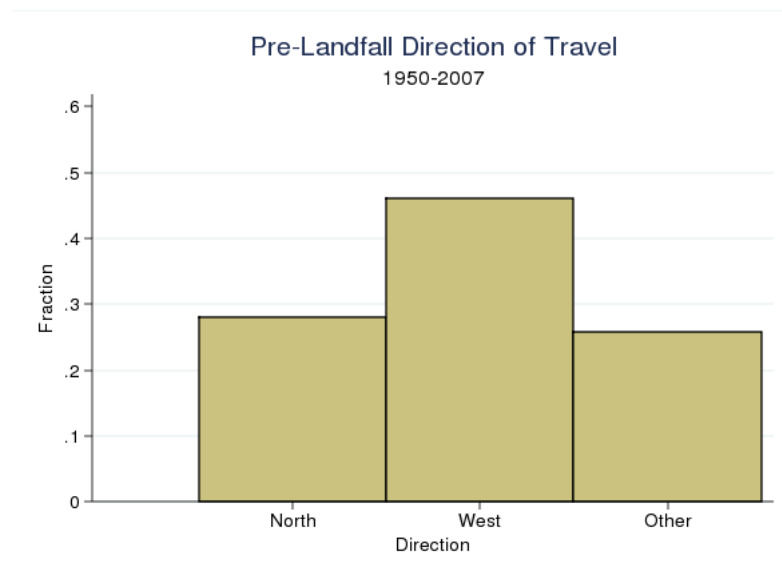


Figure 5: Histogram showing the relative frequencies of “north,” “west,” and “other” observations in pre-landfalling cyclones from 1950 to 2007.

Region RH RH included direction of travel as a state variable for tropical cyclones within the region $10^{\circ} - 25^{\circ}\text{N}$ and $55^{\circ} - 80^{\circ}\text{W}$ (Region RH). Logistic regression was used to determine whether a cyclone’s direction of travel within Region RH had a statistically significant effect on whether or not the cyclone made landfall in the U.S. The independent variables “north,” “west,” and “other” were treated as categorical variables that, for each observation, equaled 1 if true and 0 otherwise. For example, if a cyclone that eventually made landfall in the U.S. was moving northward during a particular observation, then north = 1, west = 0, other = 0, and landfall = 1 for this observation.

The results of the regression analysis are shown in Figure 6. Cyclones traveling west in Region RH were significantly more likely to make landfall in the U.S. than cyclones moving northwards or otherwise, all else equal.

$\#N$	$Pr(L N)$	$\#W$	$Pr(L W)$	$\#O$	$Pr(L O)$
583	0.178	2316	0.288	186	0.081

Figure 6: Chart displaying the number of observations and the landfall probabilities for the three possible directions of travel – north (N), west (W), and other (O). “ $\#N$ ” represents the number of “north” observations in Region RH between 1950 and 2007. “ $Pr(L | N)$ ” is the probability a cyclone traveling north in Region RH made landfall in the U.S.

5° by 5° Regions in $0^{\circ} - 70^{\circ}\text{N}$ and $0^{\circ} - 100^{\circ}\text{W}$ Next, the region from $0^{\circ} - 70^{\circ}\text{N}$ and $0^{\circ} - 100^{\circ}\text{W}$ was divided into 280 5° by 5° regions. Logistic regression was performed in each of these regions. Figure 7 illustrates the results of this analysis.

The shaded boxes represent regions in which direction of travel had a significant effect on the probability of landfall in the U.S. The color of the shading denotes the direction(s) of travel that was associated with a higher probability of landfall. For example, cyclones traveling north in the area bounded by 20°N , 25°N , 90°W , and 95°W were significantly more likely to make landfall in the U.S. than cyclones moving westward or otherwise in this region, all else equal. The area bounded by 20°N , 25°N , 80°W , and 85°W is shaded purple to show that cyclones in this region moving north or west were more likely to make landfall than cyclones traveling in the direction called “other,” all else equal. The results provide support for including the direction of travel within these regions in the Markov state space.

Region RH consists of 15 5° by 5° cells, while our analysis identified 19 such cells, including 5 in Region RH, in which direction of travel was significant.

4.2 Wind Speed

As mentioned previously, a cyclone’s maximum sustained wind speed is defined as the highest one-minute surface winds occurring within the circulation of the system at a height of 10 m. For each observation between 1950 and 2007, the cyclone’s wind speed was categorized as “depression-force” (wind speed < 33 kts), “storm-force” ($33 \text{ kts} \leq \text{wind speed} < 64$ kts), or “hurricane-force” (wind speed ≥ 64 kts). Figure 8 shows the relative frequencies of wind speed observations categorized as depression-force, storm-force, and hurricane-force.

Logistic regressions were performed in the same 280 5° by 5° regions used in the direction analysis. The results are depicted in Figure 9.

Wind speed had a significant effect on the probability of landfall in the shaded regions. For example, the red shading in the cell bounded by 25°N , 30°N , 65°W , and 70°W illustrates

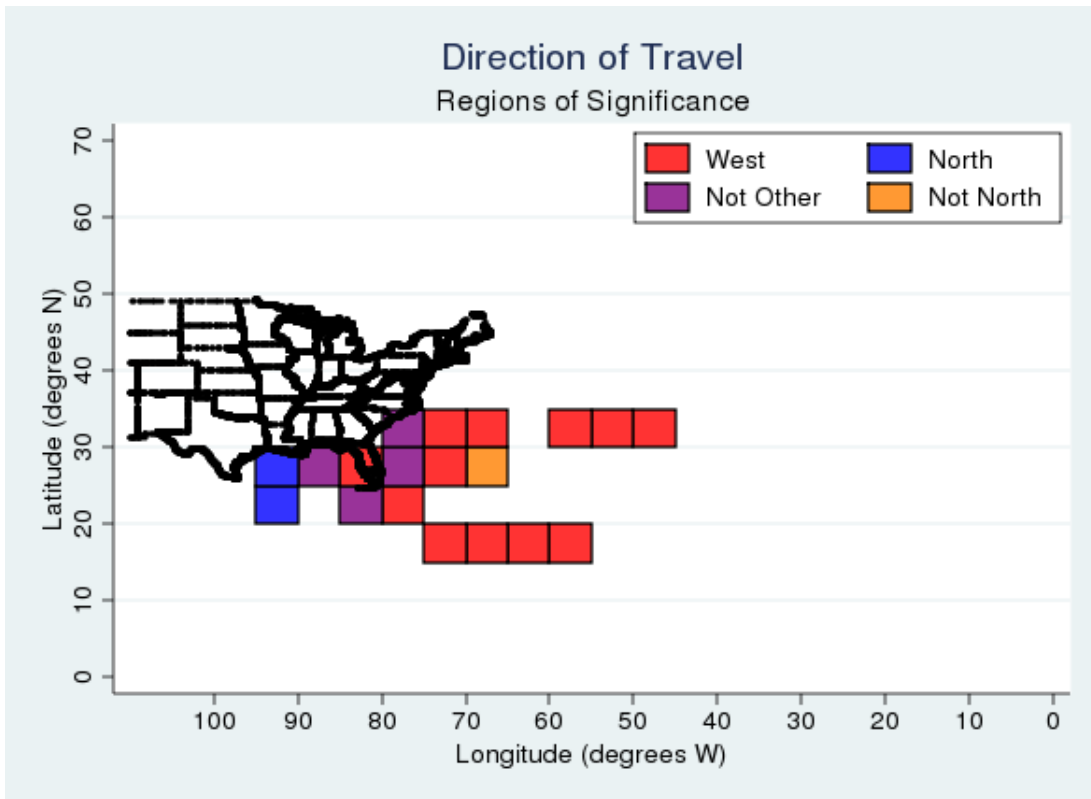


Figure 7: Map displaying the direction of travel associated with a higher probability of U.S. landfall.

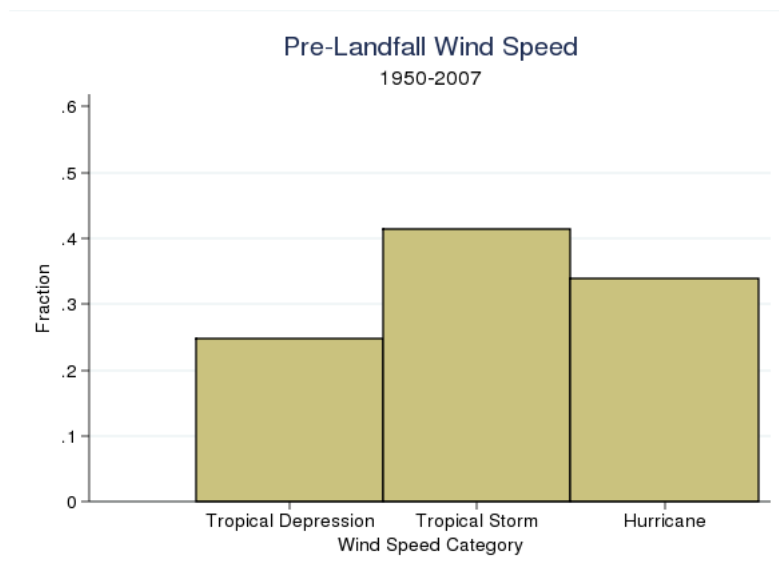


Figure 8: Histogram showing the relative frequencies of tropical depression-, tropical storm-, and hurricane-force winds in pre-landfalling cyclones from 1950 to 2007.

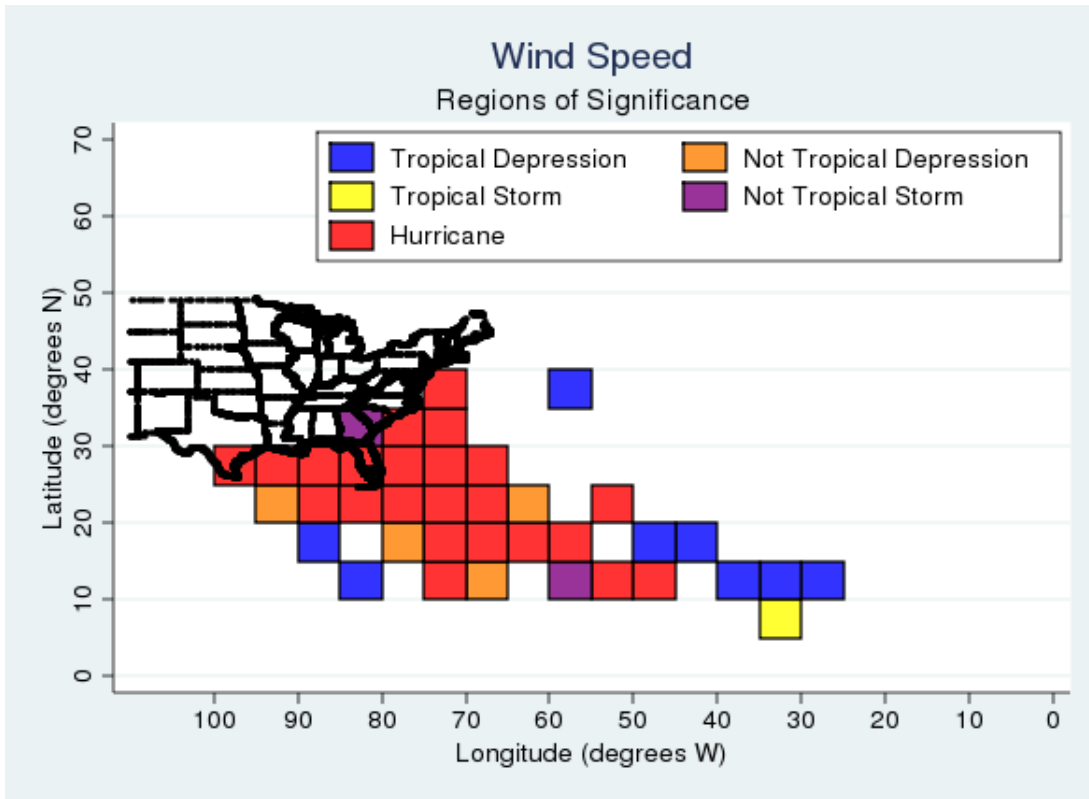


Figure 9: Map displaying the wind speed associated with a higher probability of U.S. landfall.

that a hurricane-force storm in this region was more likely to make landfall in the U.S. than a cyclone with lesser wind speed, all else equal.

4.3 Speed of Forward Motion

While wind speed measures the wind within the circulation of the storm system, the speed of forward motion considers the motion of the system itself. The speed of forward motion in kilometers per hour at time t for a particular cyclone was calculated by dividing the change in km in the cyclone's position from time $t - 1$ to time t by the length of time between observations (6 hrs). The speed at time $t = 1$ for each cyclone (that is, at the time of the first observation) was assigned to be the same as the speed at time $t = 2$. Care was taken when converting the change in longitude from degrees to km, since the length of 1° longitude varies depending on the latitude. The formula $1^\circ \text{ longitude} = \cos(\text{latitude}) \times 111.325 \text{ km}$ was used in these calculations [6].⁴

The histogram in Figure 10 shows the frequencies of cyclone speeds.

The speed of forward motion for each observation was categorized as “slow” (speed ≤ 15 km per hr), “medium” ($15 \text{ kph} < \text{speed} \leq 30 \text{ kph}$), or “fast” (speed $> 30 \text{ kph}$). These cutoffs were selected because 15 kph and 30 kph approximate the 25th and 75th percentiles, respectively, of cyclone speeds between 1950 and 2007.

⁴The length of 1° latitude also varies from the equator to the poles, but only by 1.13 km. Therefore, the standard value of 111.325 km was used here.

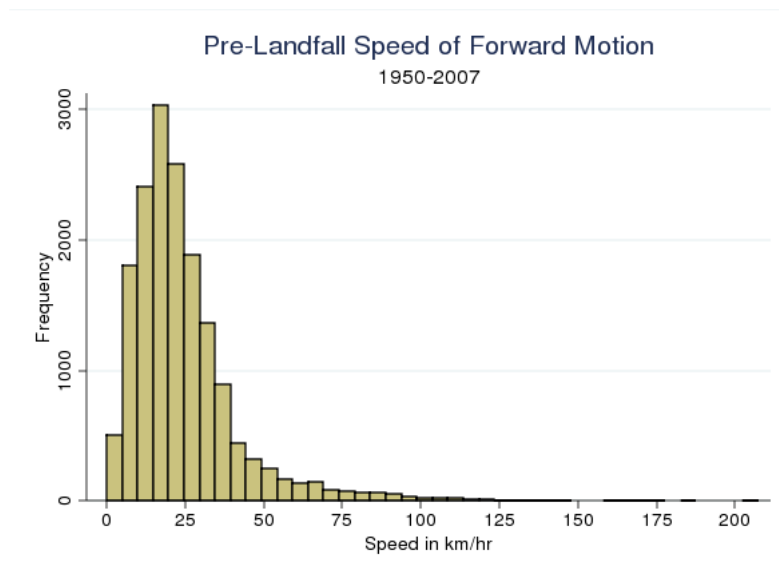


Figure 10: Histogram showing the distribution of pre-landfalling cyclone speeds between 1950 and 2007.

Logistic regressions were run in each of the 280 5° by 5° regions to examine the relationship between forward speed and landfall probability. Figure 11 illustrates the results. In the shaded regions, a cyclone’s speed of forward motion had a statistically significant effect on whether or not the cyclone eventually made landfall in the U.S. For example, the orange shading in the region bounded by 15°N , 20°N , 50°W , and 55°W shows that a tropical cyclone moving at a speed greater than 15 kph in this region was more likely to make landfall than one moving at a slower speed, all else equal.

4.4 Climatological Year Type

El Niño and La Niña episodes refer to an abnormal warming or cooling, respectively, of ocean surface temperatures in the eastern equatorial Pacific [7]. El Niño episodes are usually associated with an increased frequency of tropical cyclones in the Pacific Ocean and a decrease in Atlantic hurricane activity. La Niña events produce opposite effects, and tropical cyclone activity in the Atlantic is thought to increase during La Niña years [8]. The chart in Figure 12 shows the years between 1950 and 2007 that were categorized as either El Niño or La Niña years. Years that do not appear in the chart exhibited neutral weather conditions.

Logistic regression was performed to determine whether the year classification – El Niño, La Niña, or neither – had a statistically significant effect on whether or not a tropical cyclone made landfall in the U.S. This analysis suggests that, all else being equal, the year type did not have a significant effect on landfall. That is not to deny, of course, that the year classification had an effect on storm activity in the Atlantic. There were an average of 9 cyclones per season during El Niño years and 11 during La Niña years.

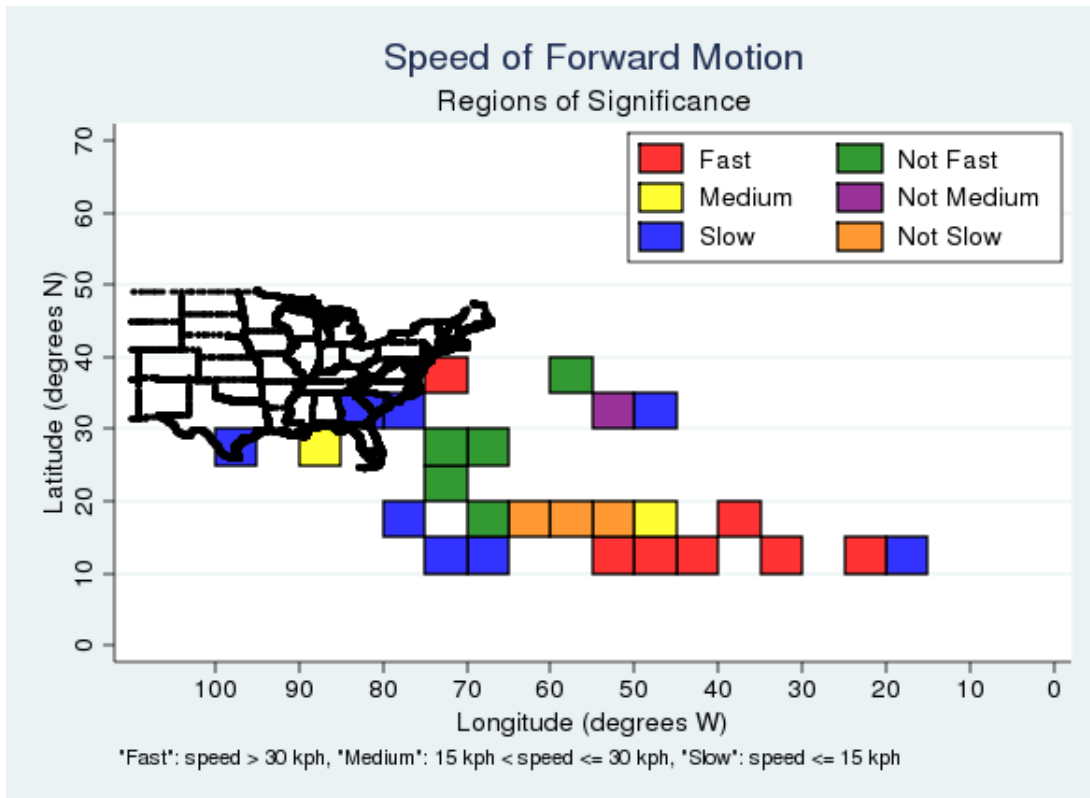


Figure 11: Map displaying the speeds associated with a higher probability of U.S. landfall.

El Niño		La Niña	
1951	1982	1950	1974
1957	1986	1954	1975
1963	1987	1955	1983
1965	1991	1956	1984
1968	1994	1961	1988
1969	1997	1964	1995
1972	2002	1967	1998
1976	2004	1971	1999
1977	2006	1973	2000

Figure 12: El Niño and La Niña years between 1950 and 2007.

4.5 Month of Origination

The Atlantic hurricane season officially begins on June 1 and ends on November 30. The histogram in Figure 13 shows the relative frequency of different months of origination for tropical cyclones that occurred between 1950 and 2007.

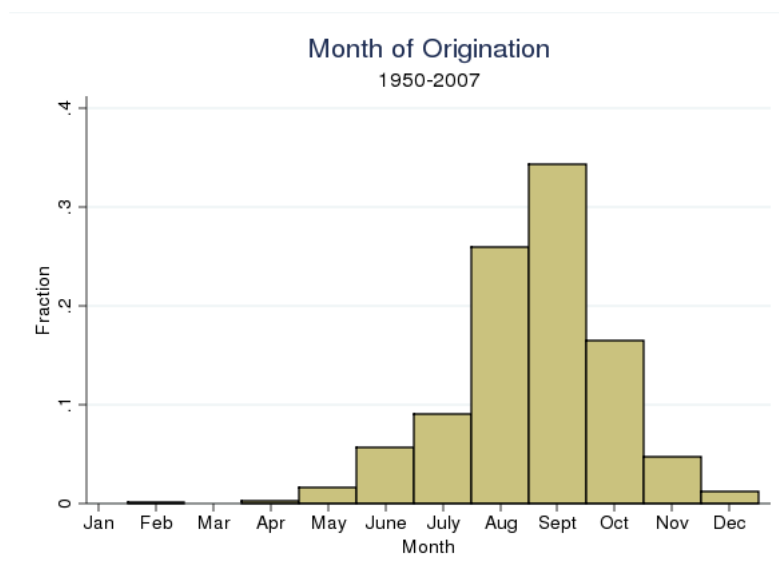


Figure 13: Histogram showing relative frequencies by month of origination.

The majority of tropical cyclones (77%) originated in August, September, or October. Seventy-three out of 92 landfalling cyclones (79%) formed in one of these three months.

In this study, the months of the year were divided into three categories: “early” (January through July), “middle” (August through October), and “late” (November and December). Each tropical cyclone was assigned to a category, and logistic regression analysis was used to study the relationship between the time of origination and the probability of landfall in the U.S. The regression results suggest that the time of year did not have a statistically significant effect on the probability of landfall. That is, while the frequency of tropical cyclones during August through October was higher than during other months, a cyclone that originated in September was not significantly more likely to make landfall than one that originated in January, assuming that the two cyclones’ other characteristics were identical.

4.6 Refining the Markov State Space

As mentioned previously, enlarging the size of the Markov state space Ω to include additional state variables increases the complexity of performing the Monte Carlo simulation, especially when starting with the 7000 1° by 1° position cells within the region bounded by 0°N , 70°N , 0°W , and 100°W .

However, the majority of tropical cyclones that made landfall in the U.S. between 1950 and 2007 first passed through the region bounded by 15°N , 40°N , 55°W , and 90°W (hereafter referred to as Region X). Furthermore, the northwestern part of Region X lies over the U.S., so that a cyclone in this area will have already made landfall. Therefore, two boxes – one defined by 30°N , 35°N , 80°W , and 90°W ; and the other bounded by 35°N , 40°N , 75°W ,

and 90°W — can be removed from Region X to give a new area, Region X*. Region X* contains only 30°S by 5° regions, or 750° by 1° position cells. The map in Figure 14 shows Region X*

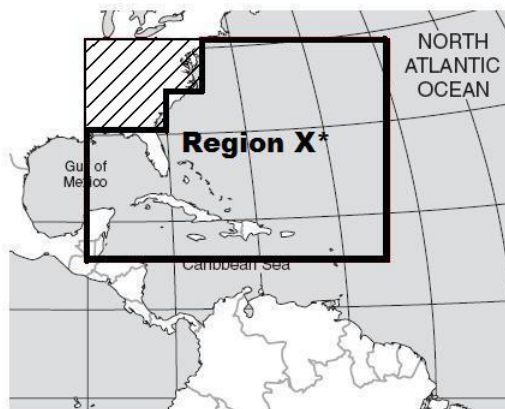


Figure 14: The thick black line outlines Region X*. The diagonal hatching identifies the area that is in Region X but not in Region X*. Adapted from <http://www.eduplace.com/ss/maps/pdf/americas.pdf>.

From the perspective of a decision maker in the U.S. who is preparing for a possible strike, it seems reasonable to disregard locations outside Region X* when considering which position states to include in Ω . This reduces the region of interest to a smaller box that still “captures” most of the cyclones that eventually made landfall.

This region analysis and the results of the regressions discussed in Section 4 are depicted graphically in Figure 15. The graph shows which factors in which areas of Region X* were found to be significant. The color of the shading indicates the predictor variables that had a statistically significant influence on the probability of landfall. For example, the black shading in the region bounded by 25°N , 30°N , 65°W , and 70°W illustrates that the direction of travel, wind speed, and speed of forward motion all had a significant effect on whether or not a cyclone in this area eventually made landfall in the U.S.

These results can be used to identify a revised state space, Ω_r , and the size of Ω_r can be determined. In the following discussion, the term “characteristics” refers to direction of travel, wind speed, and speed of forward motion. Recall that each characteristic can take on one of three values: “north,” “west,” or “other” for direction of travel; “depression-force,” “storm-force,” or “hurricane-force” for wind speed; and “slow,” “medium,” or “fast” for speed of forward motion.

For each of the 7° by 5° boxes in Region X* in which none of the three characteristics was found to be significant, there are 25 possible states. This is because a cyclone in one of these boxes could have its center located in any of the 25° by 1° cells.

For each of the 6° by 5° boxes in Region X* in which only one of the three characteristics was significant, there are 75 possible states (3 possible values of the characteristic in each of the 25° by 1° cells). Analogous reasoning shows that there are $225 = 3 \times 3 \times 25$ states in each of the 10° by 5° boxes with exactly two characteristics of significance and

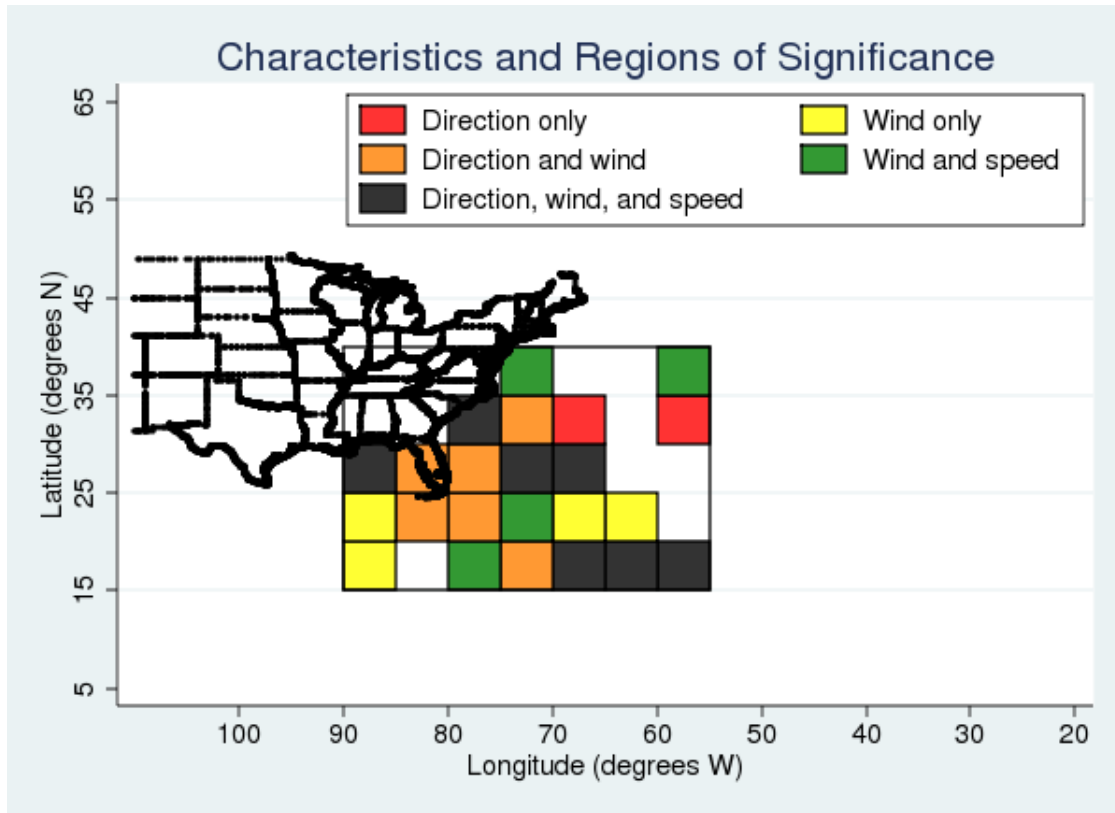


Figure 15: Map displaying the characteristics of significance within Region X*.

$675 = 3 \times 3 \times 3 \times 25$ states in each of the 7 5° by 5° boxes with three significant characteristics. Recall that the state space contains 1 cell denoting cyclone termination. Then, using Region X* and the regression results depicted in Figure 15, the size of the revised state space Ω_r can be calculated as

$$\begin{aligned}
 |\Omega_r| &= (25 \times 7) + (75 \times 6) + (225 \times 10) + (675 \times 7) + 1 \\
 &= 7601
 \end{aligned}
 \tag{10}$$

Note that if the full $0^\circ - 70^\circ\text{N}$ and $0^\circ - 100^\circ\text{W}$ region were used and all three characteristics were included as state variables in each of the 7000 1° by 1° cells, then the state space would contain 189,001 states ($7000 \times 3 \times 3 \times 3 + 1$).

4.7 Improved Decision Making with a Revised State Space

The following example illustrates how a decision maker can benefit from a more sophisticated state space that includes additional state variables. Suppose that two cyclones, A and B, are in the same 1° by 1° region bounded by 25°N , 26°N , 65°W , and 66°W . In the Markov model used by RH, the state of a cyclone in this region is defined only by position, since this cell lies outside Region RH, in which direction is included as a state variable. However, if our revised state space, Ω_r , is used, the state of a cyclone in this cell includes its position, direction of travel, wind speed, and speed of forward motion.

Suppose that cyclone A is traveling in the “west” or “other” direction, contains hurricane-force winds, and is moving at less than 30 km per hr. Also suppose that cyclone B is traveling northwards with depression- or storm-force winds and is moving at a speed greater than 30 km per hr. The results of the logistic regression analysis suggest that storm A is more likely than storm B to make landfall in the U.S.

However, for a particular target location, the strike probability p_j is the same for both cyclones when using the RH Markov model. This is because both storms are in the same 1° by 1° cell and therefore are in the same state j . Recall that, for each state j , the dynamic policy π_d specifies the action a decision maker will take when the cyclone is in that state. Therefore, the optimal action a_j must necessarily be equal for both cyclones.

However, when using the revised state space, cyclones A and B are in different states (call them k and m , respectively). The probability that cyclone A will strike a certain target location on the U.S. coastline is higher than the probability that cyclone B will hit this same target. That is, $p_k > p_m$. Therefore, perhaps action a_k equals one in this scenario, while action a_m equals zero, suggesting that a decision maker with assets at the target location should begin preparations immediately if the cyclone is in state k and wait for an updated forecast if the cyclone is in state m . The refined Markov state space could reduce the occurrence of strikes on unprepared targets and also prevent a decision maker from making costly preparations that later turn out to be unnecessary.

5 Conclusion

In this study, we used logistic regression to identify storm characteristics that affected the probability a tropical cyclone would strike the U.S. coastline. Three of the five factors examined – direction of travel, wind speed, and speed of forward motion – were found to be significant. In formulating a revised state space Ω_r for the Markov model of cyclone motion, we considered a smaller region of the Atlantic and added additional state variables in areas within the region in which these variables had a significant effect on landfall. Decreasing the size of the region of interest while including additional state variables keeps the size of the state space manageable while improving the quality of the Markov model. To extend this research, we would like to examine additional storm characteristics that could be added as state variables in the Markov state space. Two features of interest are the central pressure and the number of observations taken to reach maximum wind speed.

A more sophisticated model of tropical cyclone motion, when coupled with RH’s dynamic decision model, will allow decision makers to make more informed choices when faced with an oncoming storm.

References

- [1] Frederick K. Lutgens and Edward J. Tarbuck. *The Atmosphere : An Introduction to Meteorology*. Prentice Hall, eighth edition, 2001.
- [2] John C. Whitehead. One million dollars per mile? The opportunity costs of hurricane evacuation. *Ocean and Coastal Management*, 46(11–12):1069–1083, Dec 2003.
- [3] Eva Regnier and Patrick A. Harr. A dynamic decision model applied to hurricane landfall. *Weather and Forecasting*, 21(5):764–780, Oct 2006.

- [4] Gregory K. Miller. *Probability: Modeling & Applications to Random Processes*. Wiley-Interscience, 2006.
- [5] Eva Regnier. Public evacuation decisions and hurricane track uncertainty. *Management Science*, 54(1):16–28, Jan 2008.
- [6] Anthony P. Kirvan. Latitude/Longitude, NCGIA Core Curriculum in GIScience.
- [7] National Oceanic and Atmospheric Administration. What is an El Niño?
- [8] National Oceanic and Atmospheric Administration. Answers to La Niña frequently asked questions.