The Development and Evaluation of a Tropical Cyclone Probabilistic Landfall Forecast Product

BENJAMIN C. TRABING,[a] KATE D. MUSGRAVE,[a] MARK DEMARIA,[a] BRIAN C. ZACHRY,[b] MICHAEL J. BRENNAN,[c] AND EDWARD N. RAPPAPORT[c]

[a] *CIRA/CSU, Fort Collins, Colorado*
[b] *Travelers Indemnity Company, Hartford, Connecticut*
[c] *NOAA/NWS/National Hurricane Center, Miami, Florida*

(Manuscript received 14 November 2022, in final form 4 May 2023, accepted 11 May 2023)

ABSTRACT: Improving estimates of tropical cyclone forecast uncertainty remains an important goal of the Hurricane Forecast Improvement Project (HFIP). Intensity forecast uncertainty near landfall is especially complicated because intensity forecasts depend on track forecasts. Ensembles can be difficult to interpret near land due to differences in both spatial and temporal resolution and differences in landfall timing (if at all) and location. The Monte Carlo Wind Speed Probability (WSP) model is a statistical ensemble based on the error characteristics of forecasts by the National Hurricane Center (NHC) and the spread of several track forecast models. The landfall distribution product (LDP) introduced in this paper was developed to use the statistical ensemble of forecasts from the WSP model to estimate both the track and intensity forecast uncertainty associated with potential landfalls. The LDP includes probabilistic intensity estimates as well as estimates of the most likely and reasonable strongest intensity at landfall. These products could communicate concise intensity uncertainty information to users at risk for tropical cyclone impacts. Demonstration on a retrospective dataset from 2010 to 2018 and evaluation of the LDP on the 2020–21 Atlantic hurricane seasons shows that the probability of landfall and the landfall intensity probabilities generated by the WSP model are reliable and potentially useful for preparedness decision-making. A case study of Hurricane Ida (2021) highlights how the LDP can be implemented to communicate landfall uncertainty to a broad range of users.

SIGNIFICANCE STATEMENT: With the new landfall distribution product (LDP), the National Hurricane Center can provide both track and intensity forecast uncertainty surrounding the landfall of hurricanes. The issuance of a reasonable worst case scenario for the strongest winds that could impact a region could amplify messaging to encourage people to take appropriate action prior to a landfall.

KEYWORDS: Tropical cyclones; Operational forecasting; Probability forecasts/models/distribution

1. Introduction

Improving estimates of tropical cyclone forecast uncertainty remains an important goal of the Hurricane Forecast Improvement Project (HFIP; Gall et al. 2013). The National Hurricane Center (NHC) provides track forecast uncertainty in the form of a "cone of uncertainty"; however, no such uncertainty information is provided for intensity forecasts. Providing estimates of both track and intensity forecast uncertainty is essential to communicate the risks associated with tropical cyclones (Marks et al. 2019). Although the NHC currently provides wind speed probabilities for 34-, 50-, and 64-kt winds (1 kt $\approx$ 0.51 m s^{-1}), a product highlighting the possible range of intensities or the strongest reasonable intensity is not available.

The Monte Carlo Wind Speed Probability (WSP) model, which is able to capture some aspects of forecast uncertainty for specified wind speed thresholds, uses the NHC forecast and climatological error distribution to estimate the probability that sustained 34-, 50-, and 64-kt winds will occur at a given location (DeMaria et al. 2013). Additional tools to estimate intensity forecast uncertainty have been developed such as the Goerss predicted consensus error (GPCE; Goerss 2007; Goerss and Sampson 2014) and the Prediction of Intensity Model Error (PRIME) model (Bhatia and Nolan 2015). However, these uncertainty estimates were developed as forecaster guidance and are not provided to the public because they do not directly estimate the uncertainty of the NHC official forecast. The intensity bias and uncertainty scheme (IBUS), which provides a climatological uncertainty estimate based on NHC's intensity change error distribution, has been shown to provide intensity forecast uncertainty similar to the "cone of uncertainty" (Trabing et al. 2022). However, since the IBUS is a climatological intensity uncertainty estimate, it is not able to incorporate track uncertainty into the intensity uncertainty which is essential near landfall.

The NHC had forecast products in the past that provided intensity uncertainty to the public. A table was provided with the probability that a tropical cyclone would be within each of the Saffir–Simpson hurricane wind speed scale (Schott et al. 2012) categories at verifying forecast time periods (Stewart 2012). The NHC also provided a plot, based on a similar

Denotes content that is immediately available upon publication as open access.

Corresponding author: Benjamin C. Trabing, btrabing@colostate.edu

DOI: 10.1175/WAF-D-22-0199.1

framework, that showed as a function of time the forecast maximum intensity and the intensities (higher and lower) that could be expected to occur 10% and 20% of the time. Because the intensity category probabilities were valid only at fixed verifying forecast times ranging from 12 to 120 h with 24-h resolution after 48 h, timing uncertainty owing to tropical cyclone landfalls was poorly resolved leading to low-biased probabilities relative to the actual prelandfall intensity. This shortcoming for essential landfall forecasts led to the products' discontinuation after the 2013 Atlantic hurricane season.

Limitations of previous products suggest that future products need to account for uncertainty near tropical cyclone landfalls. Forecasting tropical cyclone intensity near landfall is challenging because small differences in the landfall timing or location can lead to large intensity forecast errors. For example, the NHC had above average skill in forecasting the intensity for Hurricane Ida through 72 h; however, the 96-h forecasts were considered unskillful relative to climatology and persistence because Hurricane Ida meandered near the coast of Louisiana before moving inland and weakening (Beven et al. 2022). Since tropical cyclone intensity near land is often track dependent, an intensity product should focus explicitly on regions where landfall may occur in order to highlight track dependent intensity solutions.

In this study, we will explore the use of landfall probabilities to estimate both track and intensity uncertainty on daily time scales. The probability of a tropical cyclone landfall is already available to the public as a climatological value through various sources. Brettschneider (2008) developed probability of landfall forecasts using climatological data as a function of current tropical cyclone position. In addition, landfall probabilities are frequently provided with seasonal tropical cyclone forecasts to communicate a population's risk of tropical cyclone impacts prior to the start of the hurricane season (Klotzbach et al. 2022). A probability of landfall framework could be a simple way to communicate tropical cyclone risk on both short and long time scales.

Here we will document, evaluate, and discuss a new tool to communicate forecast uncertainty to the public. The new landfall distribution product (LDP) converts the NHC's deterministic track and intensity forecasts to probabilistic estimates of track, intensity, and timing uncertainty around potential landfalls within specified regions. The LDP will allow us to answer questions such as "What is the probability that Hurricane X will make landfall in Mississippi?" and "If Hurricane X makes landfall in Mississippi, what is the probability it will be a major hurricane?" Section 2 will detail the data and methods. The testing of the new tool on retrospective forecasts will be shown in section 3. An evaluation of the LDP during the 2020–21 Atlantic hurricane seasons will be shown in section 4 with a case evaluation on Hurricane Ida (2021) in section 5. Section 6 will summarize the LDP and offer conclusions.

2. Development of the landfall product

a. Data

The new landfall product is created using the Monte Carlo WSP model (DeMaria et al. 2009, 2013). The WSP model creates a statistical ensemble with 1000 members centered on the NHC forecast. The spread of the ensemble is based on a climatology of track and intensity errors, where the track spread is additionally a function of the track spread of commonly used forecast models. Each of the statistical realizations in the WSP model are a reasonable forecast that includes an intensity decay correction when the forecasted tropical cyclone moves over land. The WSP model is employed here because the model is already operationally used by NHC to communicate the probability that sustained winds at specific locations will reach 34, 50, and 64 kt. Using the same model ensures consistency between the new landfall product and existing NHC products.

A set of retrospective forecasts from the WSP model is used to develop the landfall distribution product because the version of the WSP model can vary and the spread of the forecasts changes from year to year based on the NHC's 5-yr error distribution. The retrospective forecasts used here are from the 2020 version of the model and include 450 forecast cycles from tropical cyclones that made landfall from 2010 to 2018. The NHC best track dataset (HURDAT2) is used to verify the new product, which includes verifying landfall locations, times, and intensities from 1851 to 2021 (Landsea and Franklin 2013). Within the best track and the retrospective sample, there are 84 verifiable landfalls within the Atlantic basin that do not include landfalls on small islands which are not part of the LDP. The LDP was run internally at the NHC during the 2020–21 Atlantic hurricane seasons, which had 39 verifying landfalls within the regions defined in the LDP.

Landfalls by definition require a tropical cyclone center to cross a coastline. Landfalls are objectively defined using the distance to land field from the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria and Kaplan 1994, 1999; DeMaria et al. 2005). The distance to land field includes major landmasses but does not include small islands that will have little impact on tropical cyclone intensity. Islands including Bermuda, the Bahamas, the Florida Keys, or the lesser Antilles are not included in the distance to land field and will not be included in this version of the LDP. Small islands are not included in the LDP because the probability of landfall will be low due to its relationship with the coastline size. In addition, landfall is less important for small islands where the maximum winds can impact the island while the center remains offshore.

The LDP will provide probabilistic forecasts for specific regions within the Atlantic and east Pacific. The regions that are included in the LDP are shown in Fig. 1 for reference. The regions employed in the product include U.S. and Mexico state boundaries along the Atlantic and east Pacific coastlines, as well as countries in Central America. Larger islands, such as Cuba, Hispaniola, Puerto Rico, and Jamaica are included in the LDP. Regions with large coastlines such as Texas, Louisiana, Cuba, and Florida were split into two or three equidistant smaller regions to reduce false alarms (Fig. 1). The inclusion of a finer resolution version of this product is the focus of future work.

b. Methods

The first step in creating the LDP is to identify all the landfalls within the statistical ensemble. A landfall is defined as

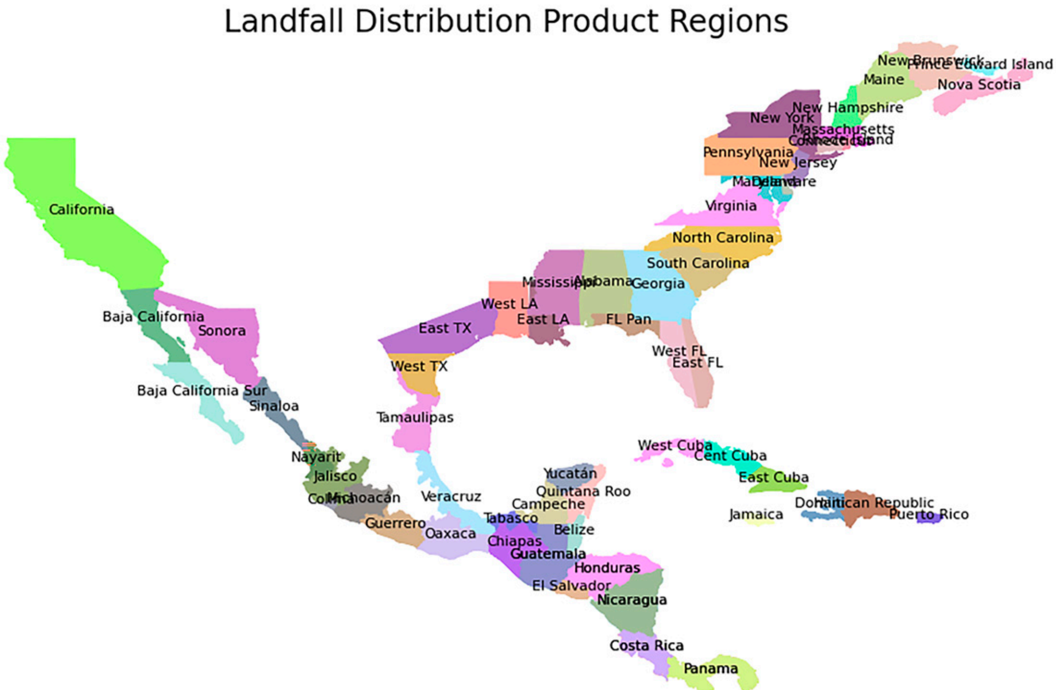


FIG. 1. The regions in the North Atlantic and east Pacific basins that are included in the LDP. Small islands are not included in the landfall product.

when the center of the tropical cyclone moves over land and is objectively calculated for each statistical realization in the ensemble. To do this, the track forecasts are interpolated from hourly to 10-min resolution and the distance to land is calculated along the track of each ensemble member. The landfalls are objectively defined by identifying any time where the distance to land parameter was both decreasing and changed sign. Once the landfalls are objectively found, the landfall intensity is defined as the maximum intensity within the 1 h prior to landfall to reduce any potential weakening in the model that has already begun due to the ongoing landfall. Statistical realizations that made landfall below 15 kt are considered dissipated and not included in the landfall probabilities.

Once all the landfalls are objectively identified within every statistical realization in the WSP model, the landfalls are then classified by region. The North Atlantic and east Pacific regions are shown in Fig. 1. To improve hazard communication when landfalls occur near boundaries and the heterogeneity in region size, an individual landfall is allowed to be classified into more than one region. The region of landfall includes all regions within the radius of maximum wind (RMW) at the landfall location. The RMW is statistically estimated by the WSP based on climatological values because the NHC does not provide forecasts for the RMW. Allowing landfalls in multiple regions within the RMW substantially improves the landfall probabilities along the northeastern U.S. coastline where the probability of landfall would otherwise be too sensitive to coastline size due to the small sizes of the regions. To prevent the probabilities from exceeding 100%, a statistical realization is only allowed to make landfall in a specific region one time over the 5-day forecast. If a statistical realization does make landfall in a region twice, for instance because of a loop in the track, the earliest landfall is used to calculate the intensity probabilities.

After all the landfalls have been classified into regions (Fig. 1), the intensity probabilities are calculated for every region where a landfall has occurred. Because the intensity probabilities are conditioned based on landfall occurring, we set a 5% lower bound on the probability of landfall over 5 days. Intensity probabilities are not calculated within regions where there is a <5% probability of landfall to ensure a large enough sample to calculate the conditional intensity probabilities and to not communicate a landfall intensity when the likelihood of landfall is too low. To concisely communicate the landfall intensities to the public, intensities at landfall are categorized by the Saffir–Simpson hurricane wind speed scale (Schott et al. 2012). Deterministic estimates of the most likely landfall intensity, the strongest reasonable landfall intensity, the most likely landfall time, and the earliest reasonable landfall time are also calculated for each region. The most likely landfall times and landfall intensity are defined as the median of the distributions to limit distribution biases. The strongest reasonable landfall intensity is defined as the intensity with a 10% probability of exceedance and the earliest reasonable landfall time is estimated as the time in which 90% of all landfalls occur after. The definition for the strongest reasonable landfall intensity provides an intensity that could possibly be reached and is consistent with the products used for communicating storm surge risk (Morrow et al. 2015).

After calculating the conditional intensity probabilities and the intensity forecast estimates for each region, a check is

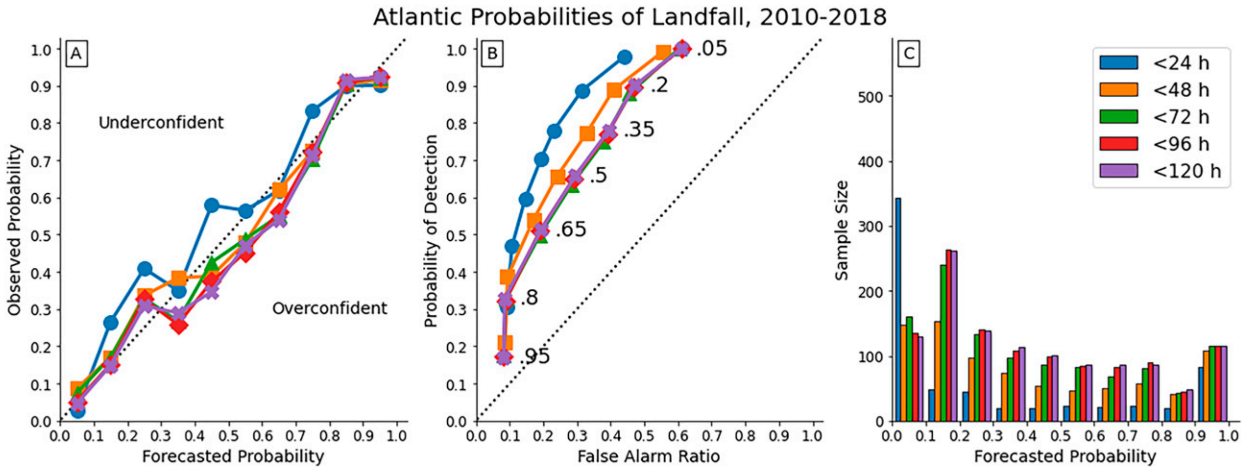


FIG. 2. (a) Reliability diagram of the forecast landfall probabilities vs observed probability of landfall using 0.1 bins. (b) ROC diagram of the probability of landfall with a 0.15 threshold interval. (c) Histogram of the sample size within the retrospective sample using 0.1 bins. Colored lines denote the probabilities within daily intervals through 5 days.

added to the LDP to ensure that the most likely intensity at landfall matches NHC's forecast. The WSP model skews the intensity distributions for higher intensity forecasts based on NHC's error forecast characteristics, because there are intrinsic limits on how strong a tropical cyclone can become (DeMaria et al. 2013). As a result, the most likely intensity will tend to be biased too low relative to the NHC official forecast for strong hurricanes because of the skewed distribution. To impose consistency with the NHC official forecast (OFCL), the deviation from the most likely intensity and the intensity from the most recent OFCL forecast is calculated. The OFCL offset is calculated from a 12-h linearly interpolated OFCL forecast and applied based on the most likely time of landfall for each region. If landfall occurs between the interpolated forecast hours, the intensity is held constant from the previous 12-h forecast point through the landfall time. The OFCL offset is added to each of the statistical realizations to center the entire distribution of intensities around NHC's forecast that leads to a slight high bias in forecast intensities relative to the WSP model for major hurricanes. The errors for the most likely landfall intensity for each region will therefore exactly match NHC's intensity forecast error distribution.

3. Analysis on retrospectives

An evaluation of 40 tropical cyclones from 2010 to 2018 in the North Atlantic is conducted using the retrospectives of the 2020 version of the WSP model. The probability of landfall within a given region is verified using HURDAT2 (Landsea and Franklin 2013). The region of landfall is derived from the best track dataset following the same methodology used to identify the region(s) of landfall in the LDP. Since forecasts can have multiple landfalls, we set a ±24-h buffer window between the forecast and verifying landfall times for categorizing false alarms. The intensity probabilities are verified against the observed landfall intensities found in the best track dataset in validating regions. Note that because the retrospectives were not run on every tropical cyclone, we only verify times when the model was run, which could lead to a reduced number of missed forecast cases.

First, we will examine how well the LDP is able to estimate the probability of landfall, which is a measure of the track forecasting skill. Figure 2 shows a reliability diagram, receiver operating characteristic (ROC) curve, and the histogram of landfall probabilities for the retrospective sample as a function of time period. The landfall probabilities are valid for landfalls that occur within the set time windows. This means that a region could have landfall probabilities of 90% within 120 h if the TC is expected to make landfall in 24 h. The reliability diagram in Fig. 2a shows how frequently landfalls occur in the sample relative to the forecast frequency of landfall. The reliability curve shows that overall, the landfall probabilities are fairly well calibrated. The probability that landfall will occur within 24 h shows a 10%–15% low bias for low probability landfall events and overall tends to be underconfident. There is a 5%–10% overforecast bias for 48–120-h forecasts for moderate probability landfall events. It should be noted that there is very little distinction between the 72-, 96-, and 120-h forecast probabilities because of the small sample size of retrospective forecasts with verifying landfalls beyond 72 h (Fig. 2c).

To show the potential utility of the landfall probabilities, the ROC curve is examined (Fig. 2b). ROC curves show how the probability of detection and false alarm ratio change when different probability thresholds are chosen for action in all forecasts. The probability of detection is defined as the fraction of cases where landfall is correctly forecast and the false alarm ratio is the fraction of cases where landfall is forecast but do not verify. The equations to calculate the contingency table can be found in DeMaria et al. (2021). The ROC curve shows that at all times the curves are above the one-to-one dotted line which indicates that the probabilities have utility in discriminating between landfall and nonlandfall events. The

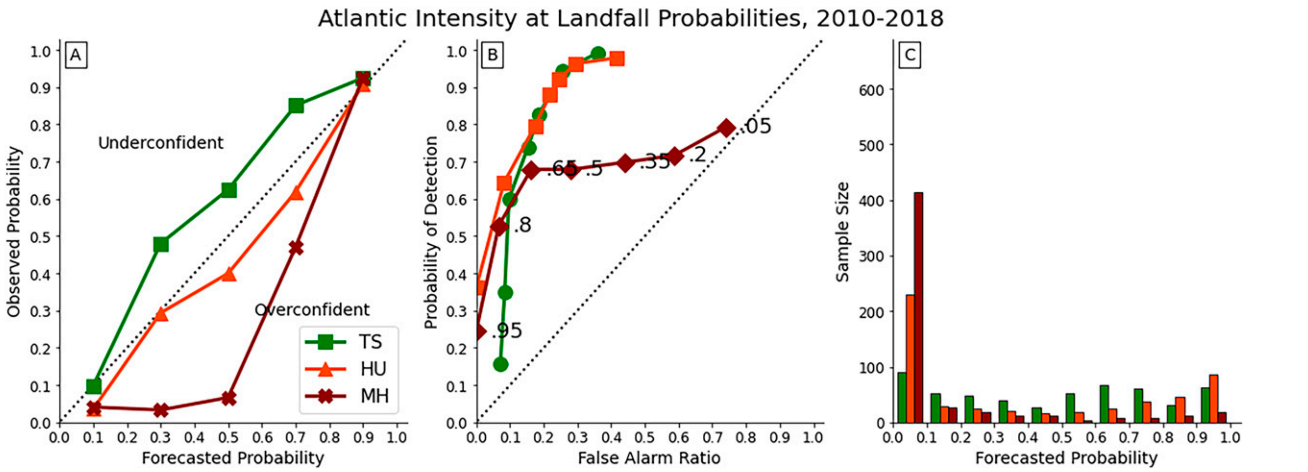


FIG. 3. (a) Reliability diagram of the forecast vs observed conditional probability that the tropical cyclone would be a tropical storm, hurricane, or major hurricane using 0.2 bins. (b) ROC diagram of the intensity probabilities with a 0.15 threshold interval. (c) Histogram of the sample size within the retrospective sample using 0.1 bins. Colored lines denote the conditional probabilities of a tropical storm (green), hurricane (orange), and major hurricane (maroon) through 5 days.

lowest probability of detection and false alarm ratio occur with a high threshold of 0.95. There are few probabilities that reach a 0.95 threshold which leads to a low probability of detection; however, a large proportion of these forecasts verify leading to a low false alarm ratio (Fig. 2a). As expected, the area under the curve (AUC), which is a measure of the utility, is lead time dependent. The utility of the LDP is greatest for 24-h forecasts with decreasing AUC for 48- and 72-h forecasts. Because of the low sample size of landfalls beyond 72 h in the retrospectives, the 96- and 120-h landfall probability discrimination is nearly identical with minor differences for low discrimination thresholds.

Next, we examine how well the conditional intensity probabilities in the retrospectives perform. Figure 3 shows the reliability, ROC curve, and histogram for the conditional probabilities that the tropical cyclone will be a tropical storm, hurricane, or major hurricane if landfall occurs. Only three category ranges are shown because of a low sample size for each of the Saffir–Simpson wind scale categories and to correspond with the convention used by the NHC. Because the intensity probabilities are conditional, a corresponding landfall in HURDAT2 is required to be included. The reliability diagram (Fig. 3a) shows that the probability that a tropical cyclone will be a hurricane is the most reliable but with a slight overforecast bias. The reliability for tropical storms shows a 15%–20% underforecast in the moderate probabilities, while the major hurricane reliability has a large overforecast in the same range. The sample size is partially to blame for the major hurricane probability bias because there are very few nonzero major hurricane probabilities in the retrospective sample (Fig. 3c). The applied offset to match NHC's forecasted intensity also contributes to the high bias in major hurricane probabilities; however, there is still a high bias at low probability values if this offset is removed for the retrospective sample (not shown).

An examination of the ROC curve shows that the LDP has utility in discriminating between the intensities at landfall (Fig. 3b). The conditional probabilities are particularly useful when predicting a tropical storm or hurricane. The probability of detection for major hurricanes changes very little for thresholds below 0.8 because of a lack of an adequate sample size in those ranges. The probability of detection for major hurricanes also maximizes near 80% as there are major hurricane landfalls that occurred when the conditional intensity forecast had a <2% probability. The lower probability of detection for major hurricanes is because some of the storms within the retrospective sample underwent rapid intensification (RI), which is defined as a 30 kt or greater increase in the maximum winds within 24 h (Kaplan et al. 2010), prior to landfall. A low bias in forecasted intensity during RI is persistent in NHC's climatological forecast error distributions (Trabing and Bell 2020). RI and rapid weakening (RW), which is defined as a 30 kt or greater decrease in the maximum winds within 24 h (Wood and Ritchie 2015), remains a difficult forecast challenge (Cangialosi et al. 2020).

Although there are limitations in the sample size within the retrospective sample, the LDP is able to reliably provide probabilities that landfall will occur within regions in the Atlantic basin. The intensity category probabilities conditioned on landfall show promise; however, additional testing with a larger sample size is required for major hurricanes. To better understand the benefits and limitations of the landfall product, we will next examine a verification of the forecasts during the 2020 and 2021 Atlantic hurricane seasons.

4. Evaluation of the landfall product for the 2020–21 Atlantic seasons

The LDP was run internally at the NHC during the 2020 and 2021 Atlantic and east Pacific hurricane seasons. Here we will conduct a verification of the probability of landfall, the conditional intensity probabilities, and the deterministic intensity estimates on the combined 2-yr dataset in the Atlantic

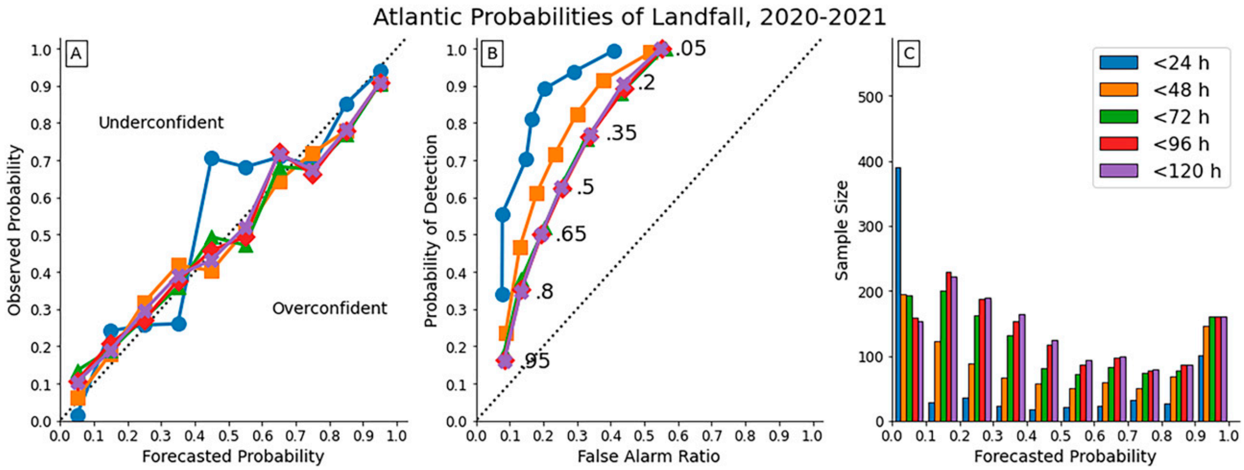


FIG. 4. As in Fig. 2, but for Atlantic tropical cyclones from 2020 to 2021.

basin only. A verification is not conducted for the east Pacific because of a small number of landfalling tropical cyclone forecasts. The 2020 and 2021 Atlantic hurricane seasons had a large number of landfalling tropical cyclones with several major hurricane landfalls that provide a robust analysis of the LDP.

First, we will examine the performance of the landfall probabilities from the LDP during the 2020–21 Atlantic seasons. Figure 4 shows the reliability diagram, ROC curve, and sample size histogram. The reliability diagram (Fig. 4a) shows that the probability of landfall is a slight improvement over the results from the retrospective analysis at most forecast hours (Fig. 2a). The probability of landfall forecast shows a reduced overforecast bias as a function of the time until landfall and a reduced bias for moderate probability events. The probability of landfall within 24 h shows a jump from an overconfident to underconfident forecast which is partially due to the lower sample size of probabilities in the 30%–40% range (Fig. 4c). Overall, the probability of landfall forecasts are reliable and well calibrated owing to the skill of NHC forecasters.

The ROC curve in Fig. 4b shows the usefulness of the landfall probabilities in discriminating between a landfall and nonlandfall event during the 2020–21 Atlantic seasons. The ROC curves are similar to the retrospective analysis for all forecast lengths which indicates that the utility of the LDP is not sensitive to the sample of cases (Fig. 2b). Regardless of the value used to discriminate action, landfall probabilities within all times are able to capture more landfall events than false alarms with a positive AUC. Beyond 72 h there is little difference in the ROC curves suggesting that the most utility in the landfall probabilities is coming from forecasts within 72 h. The ROC curve shows that a ~30%–40% cutoff probability would maximize the probability of detection and minimize the false alarm ratio for forecasting landfall regions using the LDP.

The conditional intensity probabilities in Fig. 5 show consistent reliability and ROC curves for tropical storms and hurricanes between the 2020–21 seasons and the retrospective sample (Fig. 3). The tropical storm and hurricane conditional probabilities are overall reliable and make up the largest

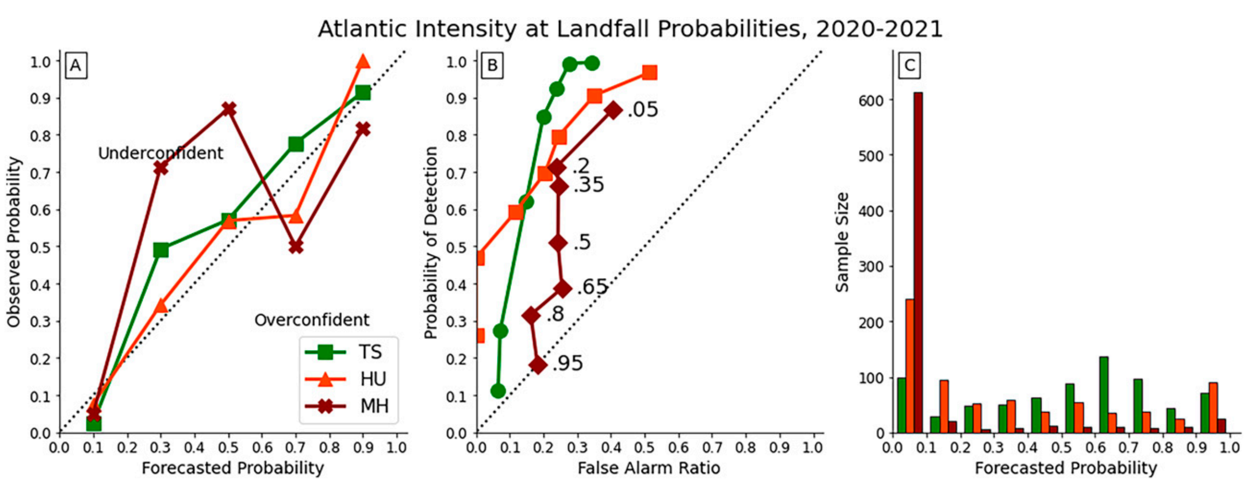


FIG. 5. As in Fig. 3, but for Atlantic tropical cyclones from 2020 to 2021.

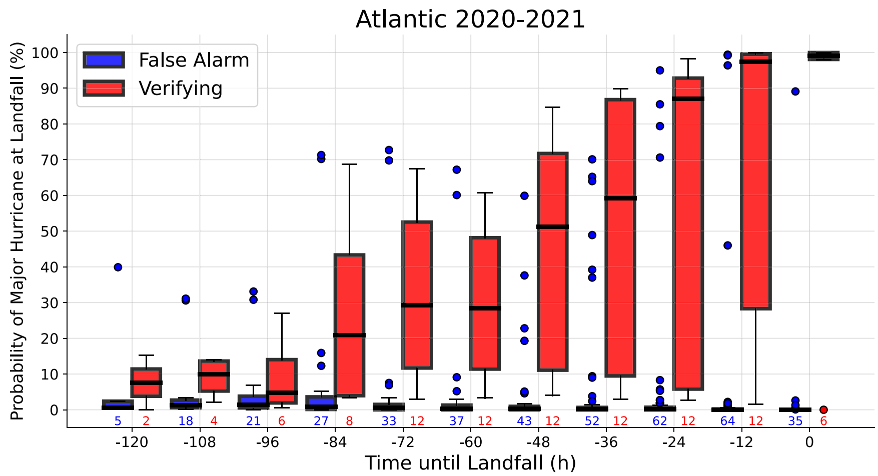


FIG. 6. Box-and-whisker plot of the probabilities that the tropical cyclone will be a major hurricane if landfall occurs as a function of lead time for the 2020–21 Atlantic cases. Red boxes (verifying) are the probabilities for verifying major hurricane landfalls, and blue boxes (false alarms) are for landfalls where the intensity was below major hurricane strength. The red and blue numbers indicate the sample size of verifying and false alarm cases, respectively.

component of nonzero sample size probabilities (Fig. 5c). The major hurricane probabilities have the lowest sample size of nonzero forecasts and shows a mixed signal in its reliability with an underforecast of low probabilities and overforecast of high probabilities. The ROC curve shows that despite the reliability biases of the major hurricane probabilities, the intensity probabilities provide useful forecasts given that the probability of detection is larger than the false alarm ratio for nearly all of the thresholds.

To better understand the major hurricane probabilities, we can examine how the conditional major hurricane probabilities evolve as a function of time until landfall. Figure 6 shows how the probability that the tropical cyclone would be a major hurricane evolves for verifying landfalls during 2020–21. False alarms are also shown in Fig. 6 which includes forecasts for tropical cyclones that made landfall below major hurricane strength. For verifying major hurricane landfalls, the intensity probabilities remain low at lead times beyond 84 h but steadily increase at shorter lead times. The opposite is true for most false alarm forecasts where the probability of a major hurricane tends to decrease until the time of landfall. Within 84 h there are larger mean probabilities of a major hurricane landfall that are statistically significant at the 95% confidence interval from the false alarm distribution. There are a couple of notable exceptions in the false alarm cases which are considered outliers from the rest of the false alarm distribution. The largest false alarms came from Hurricane Delta (2020) which had two landfalls just under major hurricane strength as a category-2 hurricane. NHC intensity forecast errors for Hurricane Delta were fairly large with an unexpected RI in the Caribbean followed quickly by RW (Cangialosi and Berg 2021). A second intensification followed by weakening occurred in the Gulf of Mexico leading to additional major hurricane false alarms by the LDP. RI and RW forecasts remain a difficult forecast challenge.

In addition to providing new estimates of the probability of landfall, the LDP also provides a most likely intensity at landfall, strongest reasonable estimate of the landfall intensity, earliest reasonable landfall time, and most likely landfall time. Figure 7 shows a joint histogram of the forecast versus observed for each deterministic quantity. The red line on each plot shows the reliability between the forecast and observed quantity. The most likely landfall time is close to the one-to-one line through 60 h with a slight bias toward landfall occurring later than expected after 60 h (Fig. 7a). The earliest reasonable landfall time in Fig. 7b shows a consistent early bias which is expected given its definition as the 10th percentile of landfall times. The uncertainty associated with the earliest reasonable landfall time is roughly 6 h for 12-h forecasts and increases to ~12-h of uncertainty by 36 h. By day 4, there is ~24 h of uncertainty associated with the landfall time provided by the LDP. The earliest reasonable landfall time appears to overestimate the landfall timing uncertainty beyond 72 h compounding with NHC's forecast bias for the most likely landfall time. No landfalls occurred prior to the earliest reasonable beyond 72 h, which would be expected as the 10th percentile.

The most likely landfall intensity shown in Fig. 7c has much larger biases compared to the landfall timing forecast, which is to be expected as intensity forecasts are generally more uncertain than track forecasts. The most likely landfall intensity shows a negative (weak intensity) bias for tropical cyclone landfalls of storms weaker than 100 kt and a slight high (strong intensity) bias, mostly coming from Hurricane Delta, for tropical cyclones stronger than 100 kt. The most likely landfall intensity provided by the LDP is bias corrected to match NHC's official forecast, so this bias is expected given the challenging landfall forecasts in 2020–21. Figure 8 shows the distribution of all 2020–21 Atlantic landfall intensities and the 24-h over-ocean intensity change prior to landfall

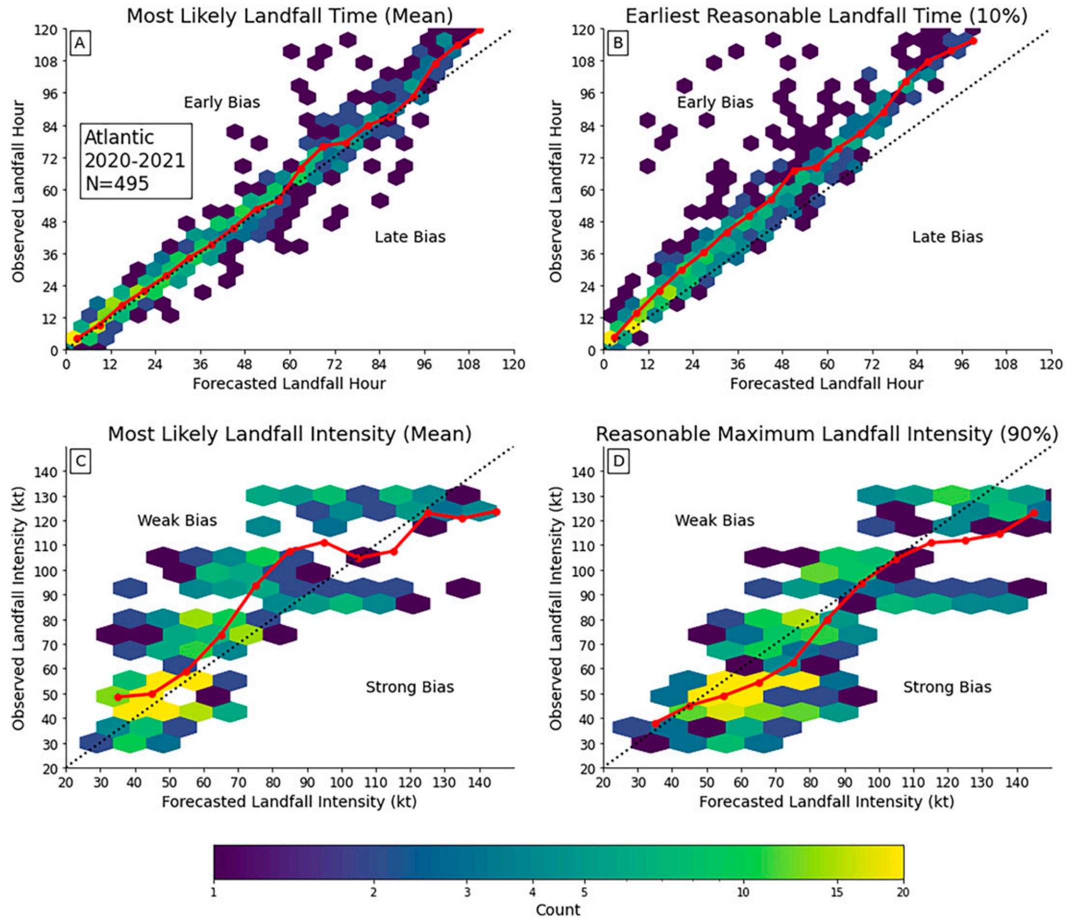


FIG. 7. Histograms of the forecasted vs observed (a) most likely landfall time, (b) earliest reasonable landfall time, (c) most likely landfall intensity, and (d) reasonable maximum landfall intensity for the verifiable landfalls in the Atlantic basin from 2020 to 2021. The red line shows the mean observed value calculated in 6-hourly bins for the landfall times and 10-kt bins for the landfall intensities. The black dotted line is the one-to-one line for an unbiased forecast.

from HURDAT2. Intensity errors are well correlated with the intensity change prior to landfall, of which there were a large number of intensification events through landfall in the 2020–21 seasons (Fig. 8b). From 2020 to 2021 the average intensity trend for tropical cyclones making landfall was an increase of 12 kt in 24 h.

The reasonable maximum landfall intensity, which we defined as the 90th percentile of the landfall intensity distribution, shows a positive (strong intensity) forecast bias for most tropical cyclone intensities (Fig. 7d). A high intensity bias is expected for a perfect intensity forecast given that it will always be larger than the most likely landfall intensity. During the 2020–21 Atlantic hurricane seasons, the reasonable maximum landfall intensity was closer to the one-to-one line than the most likely landfall intensity for most intensity ranges. The reduced expected bias from the reasonable maximum landfall intensity is due to the large number of intensification events just prior to landfall that are difficult to accurately predict (Fig. 8b). The reasonable maximum landfall intensity accounts for the potential for additional intensification to occur before landfall, and the distribution of intensity change over recent years justifies its use for the communication of landfall intensity uncertainty. The reasonable maximum landfall intensity could be used to provide landfall intensity uncertainty in a more concise way to the public than conditional landfall intensity probabilities which will be illustrated in the following section.

Here we have shown that the landfall distribution product is capable of providing reliable and useful landfall probabilities and conditional intensity probabilities at landfall. The LDP is also able to provide estimates to gauge the uncertainty in the landfall timing and intensity. Emergency managers consider the earliest onset times for hazardous winds instead of the most likely time of maximum winds at landfall presented here. However, the intensity uncertainty information could be valuable to emergency managers and decision-makers at risk of a tropical cyclone landfall.

5. Case example for Hurricane Ida (2021)

We will now show a case study of how the LDP performed for Hurricane Ida (2021). The first intensity forecast of Tropical Depression Nine (later Hurricane Ida) by NHC was unique

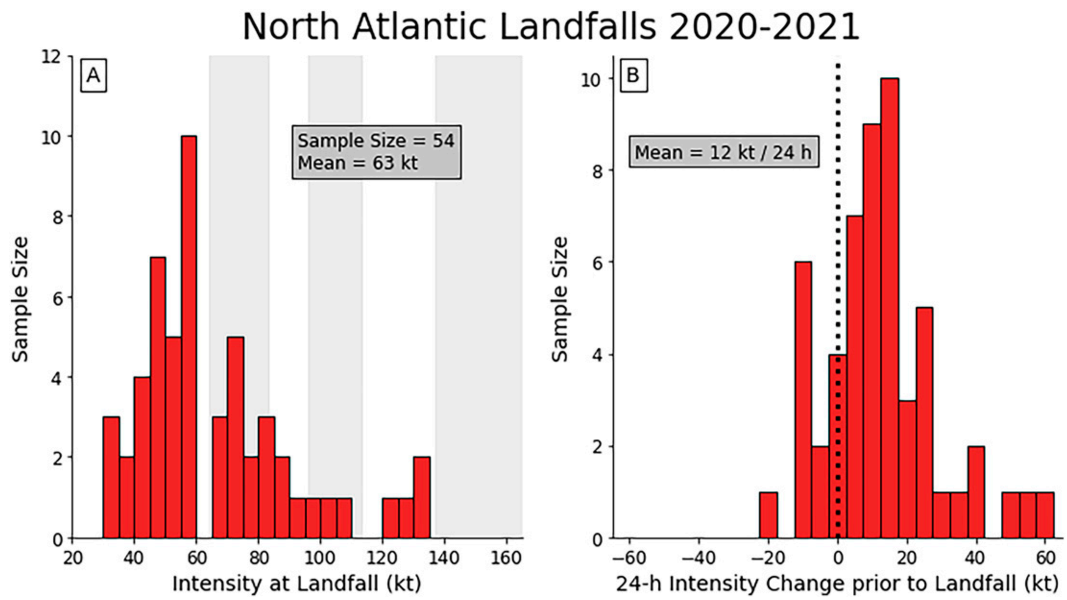
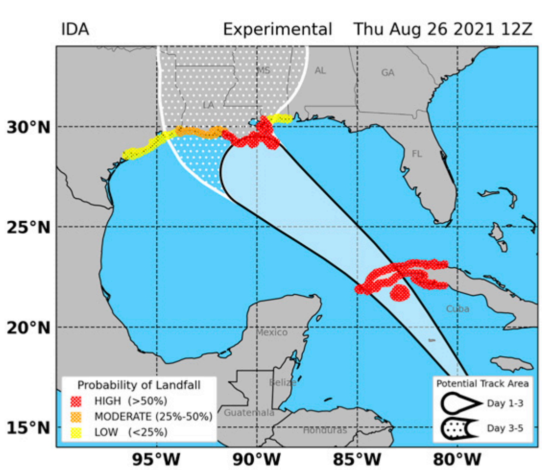


FIG. 8. The distribution of all landfalls in the Atlantic in 2020–21 from the NHC best track dataset. The distribution of (a) landfall intensities and (b) 24-h intensity change prior to landfall. The 24-h intensity change is calculated relative to the synoptic hour prior to landfall.

because it was one of the highest intensity change forecasts ever made by NHC for a tropical depression. Figure 9 shows an example graphic with select probabilistic and deterministic output for the first forecast of what became Hurricane Ida. The graphic shows the communication potential for the probability of landfall with conditional probabilities for the expected intensity at landfall. A strongest reasonable landfall intensity is provided which can be used to communicate an upper bound for the intensity forecast with a 10% chance of exceedance. On 26 August at 1200 UTC, Ida was a 30-kt tropical depression and forecasted to become a 95-kt hurricane in 72 h. From the NHC's forecast, the LDP estimated that the strongest reasonable landfall intensity for Louisiana would be a category-4 hurricane with maximum sustained wind speeds reaching 120 kt. The verifying landfall intensity of Hurricane Ida was a category-4 hurricane with 130 kt. At this time, NHC does not provide any estimates of intensity forecast uncertainty, but the strongest reasonable landfall intensity could fill that role in a public product.

The evolution of the probability of landfall and conditional intensity probabilities for Hurricane Ida is shown in Fig. 10. The genesis of Hurricane Ida occurred in the Caribbean just over 72 h prior to landfall in eastern Louisiana. The probability that landfall would occur along the eastern Louisiana



Tropical Cyclone Intensity Probabilities at Landfall

Region	Probability of Landfall	Tropical Depression or Weaker (<39 mph)	Tropical Storm (39-73 mph)	Hurricane Cat 1-5 (>74 mph)	Major Hurricane Cat 3-5 (>110 mph)	Strongest Reasonable Landfall Intensity
W CUBA	82%	7%	75%	18%	<2%	Cat 1 (80 mph)
East TX	19%	<2%	16%	83%	41%	Cat 4 (130 mph)
West LA	48%	<2%	9%	91%	50%	Cat 4 (140 mph)
East LA	50%	<2%	7%	93%	54%	Cat 4 (145 mph)
MS	12%	<2%	13%	86%	51%	Cat 4 (140 mph)

FIG. 9. An example graphic for the first forecast of Hurricane Ida (2021) at 1200 UTC 26 Aug. (left) The 5-day cone of uncertainty with hatched regions colored by the probability of landfall occurring in 5 days. (right) Table showing the probabilities of landfall, the conditional intensity probabilities if landfall occurs in each region, and the strongest reasonable landfall intensity for each region.

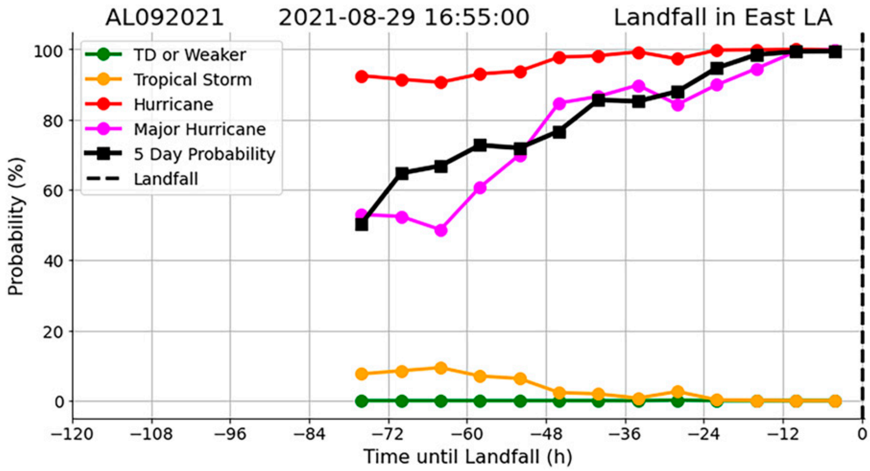


FIG. 10. The evolution of the intensity and landfall probabilities with time leading up until the landfall of Hurricane Ida in eastern Louisiana on 29 Aug. The black line with squares shows the probability that landfall will occur in 5 days. The colored lines show the conditional probabilities if landfall does occur in eastern Louisiana.

coastline began at ~50% just over 72 h out and steadily increased as the time until landfall decreased. Just 36 h prior to landfall, the probability of landfall in eastern Louisiana eclipsed 85% suggesting the strong likelihood that landfall would occur there. The conditional intensity probabilities showed that the likelihood that Ida would be making landfall as a tropical storm, a tropical depression, or weaker was less than 10% and there was a high confidence that it would be a hurricane at landfall for every forecast. The probability that Ida would be a major hurricane generally followed a similar evolution as the total probability of landfall. The probability of a major hurricane began ~50% at −72 h and climbed to above 80% 48 h prior to landfall. The already high probabilities of a major hurricane making landfall in Louisiana reached the maximum of 98% within 12 h of landfall.

The evolution of the most likely and reasonable maximum landfall intensity for Hurricane Ida is shown in Fig. 11. The forecasts for the most likely maximum wind speeds at landfall show a persistent low bias matching the NHC forecasted intensity at landfall. Although NHC explicitly forecasted RI of Hurricane Ida, the intensification rate was still underestimated owing to the challenges of forecasting RI in the limited time prior to landfall. Despite the initial low bias in forecasted intensity 72 h prior to landfall, the most likely intensity gradually approached the true intensity and was consistent with time. The reasonable maximum intensity at landfall in eastern Louisiana was initially lower than the true landfall intensity. Approximately 48 h prior to

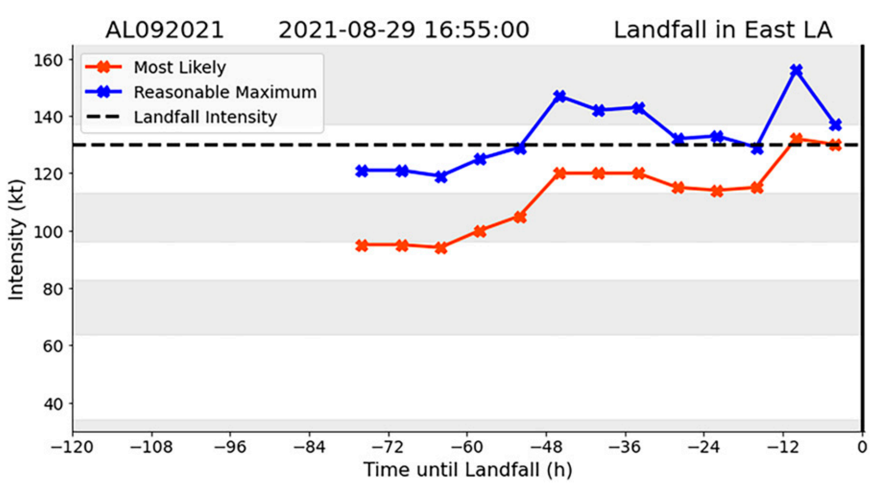


FIG. 11. The evolution of the most likely and reasonable maximum intensity estimates as a function of lead time until the landfall of Hurricane Ida in eastern Louisiana on 29 Aug. The black dashed line shows the verifying landfall intensity. The alternating gray and white shading corresponds to the Saffir–Simpson Hurricane wind speed scale.

landfall, the strongest reasonable intensity became larger than the true landfall intensity. The difference between the most likely and reasonable maximum decreases as the landfall time approaches corresponding to NHC's forecast error characteristics. A higher amount of uncertainty was suggested between the forecasts 18 and 12 h prior to landfall where the forecasted landfall intensity jumped by 15 kt.

The performance of the LDP during Hurricane Ida shows how NHC's forecasts are generally consistent with time. The continuity between forecasts is a critical component to the product, which can improve the users' trust in the forecasts. The LDP showed that 72 h prior to landfall of Hurricane Ida there was a 50% chance that Ida would make landfall in eastern Louisiana and a 50% chance that Ida would be a major hurricane if it did make landfall. The probabilities continued to increase with time as the landfall got closer, which allows for the public to continually be updated with the wind risks associated with the approaching hurricane. When provided in real time, the LDP could be a valuable tool to provide landfall intensity probabilities to decision makers with a consistent and continuous measure of uncertainty.

6. Summary and conclusions

A new landfall distribution product (LDP) has been developed and tested for use at the National Hurricane Center to provide probabilistic estimates of maximum winds during tropical cyclone landfalls. The new product uses the Monte Carlo Wind Speed Probability model which allows for consistent messaging with other NHC products and uses skill from the NHC forecasters. The product was tested during the 2020–21 Atlantic hurricane seasons and has the ability to provide reliable and useful probabilities of landfall through five days. The LDP also shows the ability to discriminate between the landfall potential for tropical storms, hurricanes, and major hurricanes through 5 days with the best results for landfalls within 72 h.

The LDP provides estimates of most likely landfall intensity, the reasonable maximum landfall intensity, the most likely landfall time, and the earliest reasonable landfall time. The reasonable maximum landfall intensity is the maximum intensity that could reasonably be reached given the forecast and is defined as the intensity with a 10% probability of exceedance. This definition of a reasonable maximum is consistent with probabilistic storm surge forecasts. Similarly the earliest reasonable landfall time is estimated as the time after which 90% of the landfalls occur. The time of landfall product is well calibrated, particularly for landfall forecasts within 72 h with a small tendency for an early bias beyond 72 h. The deterministic landfall intensity estimates tended to be biased too low compared to the observations, with the largest biases occurring for forecasts beyond 72 h. The underforecast intensity bias at landfall is directly attributed to NHC forecast errors and the difficult challenge of forecasting RI up to landfall, which has occurred frequently in recent years.

There are still several limitations of the landfall distribution product presented here that should be discussed. First, the probabilistic intensity forecasts do not inherently capture track-dependent intensity forecasts, although the LDP does account for decay due to potential land interactions. For example, if a cross-track bias would lead to a vastly improved thermodynamic environment or the movement of the tropical cyclone over a warm ocean eddy, the potential for intensification in that forecast solution is not present. In addition, the LDP uses statistically derived RMWs for each statistical realization in the WSP model to categorize landfall regions and not explicit forecasts for the RMW. Although NHC does not forecast RMW, a RMW forecast and error distribution could be derived from the maximum intensity and wind radii (Chavas and Knaff 2022). Future work will help to address these limitations.

Another area of potential bias is that the WSP does not capture any trends in intensity between the final over-ocean official forecast time and landfall. For example, when landfall is forecast to occur at hour 94, which is 2 h prior to the 96-h forecast and 22 h after the 72-h forecast, there are still 22 h of potential intensity change that is not captured. A fixed intensity approach prior to landfall was used to avoid creating artificial peaks in the intensity forecast prior to landfall that do not directly represent NHC's official forecast. A solution to this issue would be for NHC to explicitly forecast a landfall intensity that could be directly included in the WSP model.

While the methods here have described how the landfall distribution product can be used as a communication tool for NHC, the same framework could be applied to the global ensembles. A probability of landfall could easily be calculated from a set of global ensembles; however, numerical techniques to improve the intensity evolution and intensity spread of the coarse resolution global models is still needed. A probability of landfall framework could potentially be more useful at extended forecast ranges to concisely communicate changing risks from tropical cyclone winds. The expansion of the methods used here to multimodel global ensemble data to use as a forecast tool is ongoing.

Acknowledgments. This work has been funded by HFIP Award NA19OAR4320073. The authors would like to acknowledge Dan Brown, Eric Blake, and Robbie Berg for their helpful comments in developing the landfall distribution product graphics. This work has been improved by comments from Alan Brammer, Kathy Haynes, and five anonymous reviewers.

Data availability statement. The source code for the landfall distribution product can be found in an online repository https://gitlab.com/Btrabing/ldp.

REFERENCES

Beven, J. L., A. Hagen, and R. Berg, 2022: National Hurricane Center Forecast Tropical cyclone report: Hurricane Ida (26 August–1 September 2021). NOAA/NHC Tech. Rep. AL092021, 163 pp., https://www.nhc.noaa.gov/data/tcr/AL092021_Ida.pdf.

Bhatia, K. T., and D. S. Nolan, 2015: Prediction of Intensity Model Error (PRIME) for Atlantic basin tropical cyclones.

Wea. Forecasting, **30**, 1845–1865, https://doi.org/10.1175/WAF-D-15-0064.1.

Brettschneider, B., 2008: Climatological hurricane landfall probability for the United States. *J. Appl. Meteor. Climatol.*, **47**, 704–716, https://doi.org/10.1175/2007JAMC1711.1.

Cangialosi, J. P., and R. Berg, 2021: National Hurricane Center forecast tropical cyclone report: Hurricane Delta (4–10 October 2020). NOAA/NHC Tech. Rep. AL262020, 46 pp., https://www.nhc.noaa.gov/data/tcr/AL262020_Delta.pdf.

——, E. Blake, M. DeMaria, A. Penny, A. Latto, E. Rappaport, and V. Tallapragada, 2020: Recent progress in tropical cyclone intensity forecasting at the National Hurricane Center. *Wea. Forecasting*, **35**, 1913–1922, https://doi.org/10.1175/WAF-D-20-0059.1.

Chavas, D. R., and J. A. Knaff, 2022: A simple model for predicting the tropical cyclone radius of maximum wind from outer size. *Wea. Forecasting*, **37**, 563–579, https://doi.org/10.1175/WAF-D-21-0103.1.

DeMaria, M., and J. Kaplan, 1994: A Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic basin. *Wea. Forecasting*, **9**, 209–220, https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2.

——, and ——, 1999: An updated Statistical Hurricane Intensity Prediction Scheme (SHIPS) for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **14**, 326–337, https://doi.org/10.1175/1520-0434(1999)014<0326:AUSHIP>2.0.CO;2.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, https://doi.org/10.1175/WAF862.1.

——, 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. *Wea. Forecasting*, **24**, 1573–1591, https://doi.org/10.1175/2009WAF2222286.1.

——, and Coauthors, 2013: Improvements to the operational tropical cyclone wind speed probability model. *Wea. Forecasting*, **28**, 586–602, https://doi.org/10.1175/WAF-D-12-00116.1.

——, J. L. Franklin, M. J. Onderlinde, and J. Kaplan, 2021: Operational forecasting of tropical cyclone rapid intensification at the National Hurricane Center. *Atmosphere*, **12**, 683, https://doi.org/10.3390/atmos12060683.

Gall, R., J. Franklin, F. Marks, E. N. Rappaport, and F. Toepfer, 2013: The hurricane forecast improvement project. *Bull. Amer. Meteor. Soc.*, **94**, 329–343, https://doi.org/10.1175/BAMS-D-12-00071.1.

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.*, **135**, 1985–1993, https://doi.org/10.1175/MWR3390.1.

——, and C. R. Sampson, 2014: Prediction of consensus tropical cyclone intensity forecast error. *Wea. Forecasting*, **29**, 750–762, https://doi.org/10.1175/WAF-D-13-00058.1.

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. *Wea. Forecasting*, **25**, 220–241, https://doi.org/10.1175/2009WAF2222280.1.

Klotzbach, P. J., M. M. Bell, and A. J. DesRosiers, 2022: Forecast of Atlantic seasonal hurricane activity and landfall strike probability for 2022. Colorado State University, 43 pp., https://tropical.colostate.edu/Forecast/2022-07.pdf.

Landsea, C. W., and J. L. Franklin, 2013: Atlantic hurricane database uncertainty and presentation of a new database format. *Mon. Wea. Rev.*, **141**, 3576–3592, https://doi.org/10.1175/MWR-D-12-00254.1.

Marks, F. J., N. Kurkowski, M. DeMaria, and M. Brennan, 2019: Hurricane forecast improvement project years ten to fifteen strategic plan. NOAA, 83 pp., https://hfip.org/sites/default/files/documents/hfip-strategic-plan-20190625.pdf.

Morrow, B. H., J. K. Lazo, J. Rhome, and J. Feyen, 2015: Improving storm surge risk communication: Stakeholder perspectives. *Bull. Amer. Meteor. Soc.*, **96**, 35–48, https://doi.org/10.1175/BAMS-D-13-00197.1.

Schott, T., and Coauthors, 2012: The Saffir-Simpson hurricane wind scale. NOAA/NHC Tech. Rep., 4 pp., https://www.nhc.noaa.gov/pdf/sshws.pdf.

Stewart, S., 2012: Tropical Storm Isaac wind speed probabilities. NOAA/NHC, accessed 16 August 2022, https://www.nhc.noaa.gov/archive/2012/al09/al092012.wndprb.038.shtml?.

Trabing, B. C., and M. M. Bell, 2020: Understanding error distributions of hurricane intensity forecasts during rapid intensity changes. *Wea. Forecasting*, **35**, 2219–2234, https://doi.org/10.1175/WAF-D-19-0253.1.

——, K. D. Musgrave, M. DeMaria, and E. Blake, 2022: A simple bias and uncertainty scheme for tropical cyclone intensity change forecasts. *Wea. Forecasting*, **37**, 1957–1972, https://doi.org/10.1175/WAF-D-22-0074.1.

Wood, K. M., and E. A. Ritchie, 2015: A definition for rapid weakening of North Atlantic and eastern North Pacific tropical cyclones. *Geophys. Res. Lett.*, **42**, 10 091–10 097, https://doi.org/10.1002/2015GL066697.