

OPERATIONAL PREDICTION SYSTEM NOTES (OPS NOTES)

Lightning-Based Tropical Cyclone Rapid Intensification Guidance^①

CHRISTOPHER J. SLOCUM^②,^a JOHN A. KNAFF,^a AND STEPHANIE N. STEVENSON^b

^a NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado

^b NOAA/National Hurricane Center, Miami, Florida

(Manuscript received 30 August 2022, in final form 20 March 2023, accepted 15 April 2023)

ABSTRACT: With several seasons of Geostationary Lightning Mapper (GLM) data, this work revisits incorporating lightning observations into operational tropical cyclone rapid intensification guidance. GLM provides freely available, real-time lightning data over the central and eastern North Pacific and North Atlantic Oceans. A long-term lightning dataset is needed to use GLM in a statistical–dynamical operational application to capture the relationship between lightning and the rare occurrence of rapid intensification. This work uses the World Wide Lightning Location Network (WWLLN) dataset from 2005 to 2017 to develop lightning-based predictors for rapid intensification guidance models. The models mimic the operational Statistical Hurricane Intensity Prediction Scheme Rapid Intensification Index and Rapid Intensification Prediction Aid frameworks. The frameworks are averaged to form a consensus as a means to isolate the impact of the lightning predictors. Two configurations for lightning predictors are assessed: a spatial configuration with 0–100-km inner core and 200–300-km rainband area for the preceding 6-h predictors and a temporal configuration with an inner core only for the preceding 0–1, 0–6, and 6–12 h. When tested on the 2018–21 seasons, the temporal configuration adds skill primarily to the 12–48-h forecasts when compared to the no-lightning version and rapid intensification operational consensus. When WWLLN is replaced with GLM, minor changes to the prediction are observed suggesting that this approach is suitable for operational applications and provides a new baseline for tropical cyclone lightning-based rapid intensification aids.

SIGNIFICANCE STATEMENT: The forecasting of rare, yet critical, tropical cyclone rapid intensification events continues to be challenging. The current operational tools to anticipate rapid intensity changes use a combination of numerical weather prediction–derived environmental conditions and satellite-based cloud top temperature variations of deep convection. Here, we use freely available Geostationary Lightning Mapper data, which provide independent information about convection, in similar intensity guidance frameworks using temporal and spatial aspects of lightning variability. Our results show an improvement in short-term (12–48 h) rapid intensification forecasts by using temporal lightning information, and our investigation highlights that users of Geostationary Lightning Mapper lightning information should be cognizant of the influence and impact of land on these observations.

KEYWORDS: Lightning; Tropical cyclones; Forecast verification/skill; Operational forecasting; Intensification; Machine learning

1. Introduction

Lightning observations provide valuable information over the open ocean. In oceanic regions, lightning observations are one of the few direct measurements of convective activity (Mauchly 1923; MacGorman and Rust 1998). Lightning supplies unique, indirect information about the dynamics and microphysics of small-scale convection (for examples of this in data assimilation, see Mansell et al. 2007; Fierro et al. 2019).

In aggregate, lightning flashes highlight a response to environmental conditions. For tropical cyclones, large- and convective-scale lightning activity measures might be useful for diagnosing structural and intensity changes. While intensity forecasts are improving, the tropical cyclone community shows that rapid intensification remains a forecast challenge (DeMaria et al. 2014, 2021) and continues to focus on forecast improvement efforts (Gall et al. 2013; Marks et al. 2018). While the definition for rapid intensification is arbitrary, the community applies this term to the tail of the intensity change distribution within a basin, the 95th percentile of intensification rates—30 kt in 24 h (15.4 m s^{-1} ; $1 \text{ kt} = 0.514 \text{ m s}^{-1}$) in the North Atlantic Ocean basin (Kaplan and DeMaria 2003). To improve critical rapid intensification forecasts through incorporating measures of convection, the tropical cyclone community is exploring applications of lightning data. Following Alexander et al. (1999), Apodaca et al. (2014) use and assess the impact of lightning observations

^① Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/WAF-D-22-0157.s1>.

Corresponding author: Chris Slocum, Christopher.Slocum@noaa.gov

in data assimilation systems for mesoscale modeling in improving the representation of deep convection.

In preparation for the Geostationary Lightning Mapper (GLM) instrument on the Geostationary Operational Environmental Satellite (GOES) R-series satellites [for details about the instrument, see [Goodman et al. \(2013\)](#)], [DeMaria et al. \(2012\)](#) test lightning-based predictors in the Statistical Hurricane Intensity Prediction Scheme (SHIPS) Rapid Intensification Index (SHIPS-RII; [Kaplan and DeMaria 2003](#)). The authors use land-based lightning detection network data from the World Wide Lightning Location Network (WWLLN). With data collection beginning in the early 2000s, WWLLN is one of the longest, continually operated ground-based lightning network datasets available (e.g., [Lay et al. 2004](#); [Rodger et al. 2005, 2006](#)). Through a prolonged period of data collection, WWLLN captures convective information for rare events like tropical cyclone rapid intensification. In the [DeMaria et al. \(2012\)](#) work, the authors show that inner-core lightning activity (defined as the preceding 6-h density of WWLLN strokes inside a radius of 100 km of the storm center) correlates with rapid weakening for North Atlantic hurricanes and rainband area lightning activity (defined as the density in a 200–300-km annulus around the storm center) correlates with average- and rapid-intensity increases. Since lightning is associated with deep convection, their findings for the rainband area are perplexing because these findings do not align with theoretical ideas (e.g., [Schubert and Hack 1982](#)) nor convective-based observational studies (e.g., [Rogers et al. 2012, 2013](#)) for rapid intensification. However, their work suggests that lightning predictors capture both the large-scale environmental influences in addition to convective indicators of rapid intensification in the inner core. For the inner core, the [DeMaria et al. \(2012\)](#) conclusions are consistent with early work on tropical cyclone lightning ([Molinari et al. 1994, 1999](#); [Black and Hallett 1999](#)) indicating that tropical cyclone lightning activity rarely occurs within 100 km of the center of a tropical cyclone. The rarity of lightning in the tropical cyclone eyewall is due to the rarity of intense updrafts with vertical velocities exceeding 10 m s^{-1} in the presence of supercooled water ([Black and Hallett 1999](#)). Also, the [DeMaria et al. \(2012\)](#) results are comparable to studies (e.g., [Jiang and Ramirez 2013](#); [Stevenson et al. 2016](#); [Xu et al. 2017](#)) that show a spike in lightning in the inner core could signify weakening. But, the [DeMaria et al. \(2012\)](#) results are inconsistent with some studies that suggest an outbreak of lightning in the inner-core signals rapid intensification ([Molinari et al. 1994](#); [Price et al. 2009](#); [Pan et al. 2010, 2014](#); [Abarca et al. 2011](#); [Stevenson et al. 2014](#); [Zhang et al. 2015](#); [Zawislak et al. 2016](#); [Vagasky 2017](#); [Fierro et al. 2018](#)). And the inner-core lightning relationship to intensity change is sensitive to the placement of convective updrafts and lightning bursts relative to the radius of maximum wind ([Rogers et al. 2012, 2013, 2016](#); [Stevenson et al. 2018](#)).

The contradictions between lightning studies are problematic, especially when trying to develop and transition a stable rapid intensification aid to operations. Despite the conflicted conceptual model of how lightning relates to tropical cyclone intensity changes, we believe that the selection of dataset,

stratification of data, and criteria for case inclusion and duration could explain many of the discrepancies. Here, we do not try to unearth a unifying theory for lightning and tropical cyclones. However, now that we have more lightning data, we want to understand how area lightning density predictors capture changes in lightning activity in the tropical cyclone inner-core area and in the tropical cyclone environment during rapid intensification.

We cover the datasets used in our implementation of lightning-based rapid intensification guidance in [section 2](#). We revisit the operational approach proposed by [DeMaria et al. \(2012\)](#) by calculating inner-core and rainband area lightning predictors as well as calculating new time-lagged metrics in [section 3](#). In calculating both the metric formulation proposed by [DeMaria et al. \(2012\)](#) and the time-lagged formulation, we assess the role of the rainband predictors and their relationship to land. In [section 4](#), we outline the rapid intensification guidance frameworks and training procedures. Here, we use the WWLLN dataset for training because the period of record from *GOES-16/GOES-17* GLM currently is not long enough to be statistically robust. In [section 5](#), we evaluate the performance using the 2018–21 training dataset using the WWLLN data and discuss the differences between the [DeMaria et al. \(2012\)](#) spatial configuration that uses inner-core and rainband lightning density and the temporal configuration proposed here. Because the guidance will use GLM in real time for operations, we compare WWLLN and *GOES-16* GLM-based predictors and evaluate these predictors during the 2018–21 seasons in [section 6](#). This comparison aids in understanding the merit for transitioning a rapid intensification aid with GLM lightning-based predictors to operations. [Section 6](#) also discusses the independent performance of the temporal configuration of the lightning-based guidance that ran experimentally at the National Hurricane Center during the 2022 North Atlantic and eastern North Pacific hurricane seasons. We believe that this work highlights the value that *GOES-16/GOES-17* GLM lightning predictors provide in an updated baseline for the operational lightning applications for tropical cyclones, and we discuss potential future extensions in our discussion and conclusions, [section 7](#).

2. Data

We focus on tropical systems that are within the eastern and central North Pacific and North Atlantic. To represent tropical cyclone characteristics and lightning, we employ four datasets: 1) Automated Tropical Cyclone Forecast (ATCF) system best track and probability-aids (i.e., “e-deck”) database, 2) operational and developmental SHIPS large-scale environmental diagnostic dataset, 3) WWLLN data, and 4) GOES GLM data.

a. Tropical cyclone characteristics data

1) AUTOMATED TROPICAL CYCLONE FORECAST DATABASE

The ATCF system ([Sampson and Schrader 2000](#); [NRL 2020](#)) combines forecaster information generated by the National

Hurricane Center and Central Pacific Hurricane Center to supply estimates of storm location and intensity. Here, we extract values from the best track and probability-aid databases. From the best track database, we use the position, intensity (i.e., maximum sustained 1-min averaged wind), and 12-h intensity change (persistence). The top third of Table 1 has the storm characteristics taken from the ATCF database, variable units, and a description. The ATCF best track database intensity estimates have uncertainties (Landsea and Franklin 2013; Torn and Snyder 2012; Combet et al. 2020) that have implications for machine learning applications (Bonfanti et al. 2018; Kumler-Bonfanti et al. 2020). In addition to the best track database, we use the probability-aid database to extract operational rapid intensification consensus (RIOCI; Kaplan et al. 2015; DeMaria et al. 2021) forecasts during the 2018–21 seasons to use as a reference or baseline model. RIOCI averages the current operational SHIPS-RII linear discriminant analysis model, a logistic growth equation model, and the Rozoff and Kossin (2011) naïve Bayesian model.

2) SHIPS LARGE-SCALE DIAGNOSTIC DATASET

The large-scale diagnostic files provide the SHIPS-RII guidance with storm environment and convective context. The large-scale diagnostic algorithm calculates environmental quantities from global numerical weather prediction models, mainly the Global Forecast System, and convective parameters from operational geostationary satellites.

The diagnostic algorithm azimuthally averages quantities using storm-centric, area-weighted annuli, the bounding radii of which change depending on the parameter and the perceived physical impact to storm intensity. DeMaria and Kaplan (1994) outline how the algorithm formulates calculations for each variable, which is like the current operational SHIPS model with additions and changes based on advances by the research community and parameter performance in operations. To give the community a consistent dataset, the SHIPS Developmental Dataset contains the current suite of environmental and convective predictors from 1982 to present (RAMMB 2021). While consistent in form, the SHIPS Developmental Dataset contains inhomogeneities associated with upgrades to the operational Global Forecast System. Slocum et al. (2022) explore and quantify these inhomogeneities by using the SHIPS Developmental Dataset and model diagnostics calculated from the European Centre for Medium-Range Weather Forecasts fifth-generation atmospheric reanalysis (Hersbach et al. 2020). The authors show that the mean and median deviations of the developmental dataset quantities are relatively stable over the period used here with the standard deviation decreasing over time. Slocum et al. (2022) also highlight how tropical cyclone position errors impact diagnostic quantities like generalized wind shear, where 0.1° displacement can result in 0.5-kt shear differences. In addition to the developmental dataset, we use the operational North Atlantic and eastern and central North Pacific large-scale SHIPS diagnostic files during the 2018–21 seasons. Because additional variability arises from using the operational track and intensity analysis errors in the working best track data, we use the operational SHIPS diagnostics files to assess real-time performance and capture algorithm behavior in operations. We

note that verification uses the postseason final best track intensity values.

b. Lightning observations

Here, we want to leverage GLM because of its availability for operations. But, because GLM is a recent addition and we need a long training dataset to develop a statistical–dynamical rapid intensification product, we combine GLM with WWLLN.

1) WORLD WIDE LIGHTNING LOCATION NETWORK

The WWLLN (Lay et al. 2004) detects cloud-to-ground and intracloud lightning and is one of the longest-operating ground-based lightning networks. The network serves as a baseline for lightning products as a result. Like other ground-based lightning networks, the WWLLN lightning detection efficiency decreases over the ocean. DeMaria et al. (2012) find that detection improves as the number of stations increased from the early to late 2000s. Using the Lightning Imaging Sensor on NASA's Tropical Rainfall Measuring Mission climatology from Cecil (2001), the authors create a seasonal correction for the WWLLN dataset to account for the decreased detection efficiency. Here, we apply the bias correction approach outlined by Stevenson et al. (2018) and the seasonal correction adjustment factors from 2005 to 2020 for the North Atlantic Ocean, eastern and western North Pacific Ocean basins (Fig. 1). Since the eastern and central North Pacific basins rapid intensification algorithm is the same, we assume that the eastern North Pacific adjustment factors are characteristic of the central North Pacific while the western North Pacific might be more endemic—as global geostationary lightning measurements become available, the assumption for how lightning needs to be corrected to align with seasonal climatology can be revisited. This approach produces different, but similar values to those in Table 1 of DeMaria et al. (2012). In recent years, the basins show that the WWLLN adjustment factors level off to a value between two and four.

2) GOES-R GEOSTATIONARY LIGHTNING MAPPER

The GLM is an optical sensor on the R-series of the GOES intended to detect total lightning (i.e., cloud-to-ground and intracloud). The GLM instrument measures reflectance in the oxygen band with a central wavelength of 0.7774 μm at a 2-ms frame rate. To detect lightning, the processing algorithm calculates the difference between pairs of images and compares pixel values to background climatology. The resulting product denotes valid pixels that exceed the climatological background threshold as an “event.” The processing algorithm links events in a single image together into a mass-weighted centroid in space and labeled as a “group.” Then, the algorithm connects the groups to a mass weighted centroid in time and space and labeled as a “flash” (Goodman et al. 2013). The Level 2 GLM product supplies the data for the events, groups, and flashes. In this work, we use the latitude and longitude values for the flash data. The Level 2 GLM product flash data has parallax-induced geolocation error (Virts and Koshak 2020) that we do not correct since the displacement error is small with respect to the areal coverage used in our application.

TABLE 1. The top third of the table provides a list of storm characteristics from the Automated Tropical Cyclone Forecast system best track database (Sampson and Schrader 2000), the units for the characteristics, and a description. The middle and bottom thirds of the table list large-scale storm environment and convective parameters taken from the developmental and operational SHIPS datasets (e.g., DeMaria and Kaplan 1994; RAMMB 2021).

| Quantity | Abbreviation | Units | Description |
|----------------------------|--------------|------------------------|---|
| Storm characteristics | | | |
| Latitude | LAT | ° | The latitude of the center of the tropical cyclone |
| Longitude | LON | ° | The longitude of the center of the tropical cyclone |
| Intensity | V_{\max} | kt | The 1-min maximum wind speed |
| Persistence | PER | kt | Intensity change in the preceding 12-h period |
| Environment parameters | | | |
| Potential intensification | POT | kt | Remainder of the empirical potential intensity calculated from the sea surface temperature and intensity |
| Vertical wind shear | SHDS | kt | Vertical wind shear magnitude calculated within a radius of 500 km where <i>S</i> is “C” for an 850–200-hPa deep layer or “G” for generalized column shear |
| Relative humidity | RHLL | % | Relative humidity averaged within a 200–800-km annulus around the storm center; <i>LL</i> is the level with “LO” for 850–700 hPa and “MD” for 700–500 hPa |
| 200-hPa divergence | DDDD | $s^{-1} \times 10^7$ | 200-hPa divergence from 0 to 1000 km; <i>DDD</i> represents center with “200” for the surface center and “IVC” for the 850-hPa vortex center |
| Tangential wind tendency | ATWT | $m s^{-2}$ | Tendency of the model average symmetric tangential wind from 0 to 600 km |
| Temperature advection | TADV | $K s^{-1} \times 10^6$ | Temperature advection between 850 and 700 hPa within a radius of 500 km |
| Eddy momentum flux | REFC | $m s^{-1} day^{-1}$ | 200-hPa average relative eddy momentum flux convergence in a 100–600-km annulus |
| Ocean heat content | OHC | $kJ cm^{-2}$ | Ocean heat content relative to the 26°C isotherm derived from the Navy Coupled Ocean Data Assimilation system analyses |
| Convective parameters | | | |
| GOES IR | PCTB | % | Percent area of GOES brightness temperatures < <i>TB</i> within a 50–200-km annulus, where <i>TB</i> is either -30° , -50° , or $-60^\circ C$ (denoted as “30,” “50,” “60”) |
| GOES IR standard deviation | TBSD | % | Standard deviation of GOES brightness temperature within a radius of 200 km |
| GOES IR minimum radius | RMNT | km | Radius of minimum GOES brightness temperature between 0 and 150 km |
| Size metric | fR5 | % | GOES IR fractional deviation from the climatological infrared tropical cyclone size metric |
| Inner-core lightning | LHHI | $km^{-2} yr^{-1}$ | Inner-core region lightning density between 0 and 100 km in scaled radius using fR5; <i>HH</i> denotes the 0–1-h (“01”), 0–6-h (“06”), and 6–12-h (“12”) period |
| Rainband lightning | LHHR | $km^{-2} yr^{-1}$ | Rainband region lightning density between 200 and 300 km in scaled radius using fR5; <i>HH</i> denotes the 0–6-h (“06”) period |

Here, we use *GOES-16* GLM flashes from 2018 to 2021 in the North Atlantic. While preliminary, nonoperational *GOES-16* GLM data from 2017 do exist, we restrict data to after the checkout period for the instrument once the *GOES-16* satellite reached its operational orbit position at $75.2^\circ W$. And because the *GOES-17* GLM did not become operational until after the 2018 hurricane season ended and the *GOES-16* GLM domain extends into the western part of the eastern North Pacific, we only use *GOES-16* to assess the impact of lightning observations even though we run the model using both GLM instruments. We note that changes to GLM processing to flag and correct erroneous GLM output are ongoing [i.e., the so-called “Bermuda bars” documented by

Rudlosky et al. (2019)]. To date, the GLM data are not being reprocessed with a consistent version of the algorithm to produce a homogeneous Level 2 product. We will highlight an example of erroneous GLM output and discuss the impact on applications for tropical cyclones in section 6.

3. Tropical cyclone lightning density

Tropical cyclones exhibit a bimodal distribution of lightning in radius. The peaks occur near the eyewall and in the outer rainband area (e.g., Molinari et al. 1999; Cecil et al. 2002; Abarca et al. 2011; Stevenson et al. 2016). To exploit this, DeMaria et al. (2012) define two areas for calculating

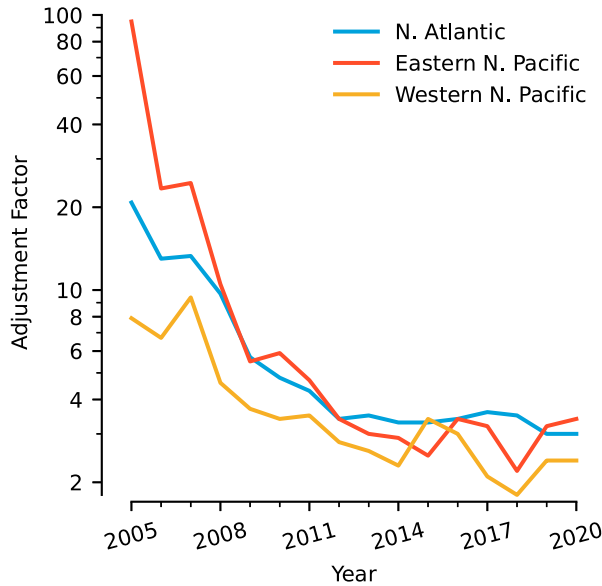


FIG. 1. Adjustment factors using the methods from [Stevenson et al. \(2018\)](#) applied to the WWLLN dataset from 2005 to 2020 for the eastern (red curve) and western (yellow curve) North Pacific and North Atlantic (blue curve) ocean basins.

lightning predictors: inner core and rainband. They set the inner-core area from 0 to 100 km and the rainband area from 200 to 300 km with a gap between the two areas representing the climatological minimum in tropical cyclone lightning [for an updated radial climatology, see Fig. 7 in [Stevenson et al. \(2016\)](#)]. Through later experimentation, [DeMaria et al. \(2012\)](#) find that the variance explained increases by selecting different ranges for each Northern Hemisphere tropical cyclone basin. These ad hoc adjustments to the bounding radii relate to the typical storm size characteristic of each basin ([Chavas and Emanuel 2010](#); [Knaff et al. 2014b](#)). In general, the eastern North Pacific basin has smaller storms than other basins on average, the western North Pacific basin has the largest storms, and the North Atlantic basin is more variable regardless of the size metric (e.g., [Knaff et al. 2007, 2014b](#)).

Here, we calculate an axisymmetric, average lightning density for the two areas described in [DeMaria et al. \(2012\)](#) and defined as

$$LD = \frac{t_y \sum_n f_n; r_j \leq r_n < r_{j+1}}{t_h \frac{1}{2}(r_{j+1}^2 - r_j^2)}, \quad (1)$$

where n is a single flash, N is the total number of flashes, t_y is the number of hours in a year (i.e., 8766 h), t_h is the period (in this work, 1 or 6 h), r_j and r_{j+1} are the radial bin bounds, and $f_{r_j \leq r_n < r_{j+1}}$ are the flashes within the bin. Note that (1) has units of flashes per square kilometer per year. Before radially binning the lightning flashes, we calculate a scaled radius r by dividing the physical radius of the flash from the storm center, r , by the fractional radius of 5-kt wind (fR5). [Knaff et al. \(2017\)](#) define fR5 as the storm size relative to climatology using the geostationary satellite longwave infrared imagery—fR5 is available

in real time in the SHIPS diagnostic files. [Figure 2](#) shows how our scaling impacts what lightning observations we use in calculating the lightning density for Hurricane Dorian at 0600 UTC 1 September 2019, a smaller than average storm with an fR5 of 75% of normal and intense inner-core lightning activity, and Hurricane Laura at 1800 UTC 26 August 2020, a larger than average storm with an fR5 of 122% of normal and minimal inner-core lightning activity. While the differences in physical and scaled radius are minor as the figure shows, we find that this scaling creates a clearer distinction between inner-core and rainband lightning by reducing the variance of lightning density values near 150 km in scaled radius space (e.g., [Knaff et al. 2014a](#)).

For tropical cyclone centric lightning observations, the frequency distribution of lightning density in (1) is skewed with values shifted toward zero—characteristic of a lognormal distribution. Since the operational machine learning technique assumes that predictors are normally distributed, we want to normalize the quantity. [DeMaria et al. \(2012\)](#) take the square root of the lightning density. While their approach reduces extreme values and shifts the distribution toward normal, this step does not sufficiently normalize the data. From (1), we transform the data with the natural logarithm of one plus the square root of lightning density:

$$\mathcal{L} = \ln(1 + LD^{1/2}). \quad (2)$$

In [section 6](#), we will evaluate how the transformed lightning density values compare between WWLLN and GLM as well as show the impact of the transform defined in (2).

4. Lightning-based rapid intensification guidance

To assess the contribution of the lightning predictors, we leverage two different operational rapid intensification guidance frameworks. With the two frameworks, we create control runs with no lightning and experimental runs with two lightning-based convective predictor configurations. Because these two rapid intensification guidance frameworks exhibit a degree of independence, we calculate an equally weighted average or consensus between the two frameworks to understand if the change in skill contributed by the lightning-based predictors is relevant regardless of the framework. Also, note that we only show results in the context of the consensus based on a simple average to create a framework-agnostic approach as a means to isolate the influence of the lightning predictors.

a. Guidance framework setup

For the first framework, we construct a basin-specific baseline model used operationally that is similar to the [DeMaria et al. \(2012\)](#) SHIPS-RII setup. The SHIPS-RII calculates predictor coefficients with a linear discriminant analysis ([Kaplan and DeMaria 2003](#)). In the operational SHIPS-RII model, the large-scale diagnostic predictors include 0–500-km deep-layer vertical wind shear (SHDC), 200–800-km lower-tropospheric relative humidity (RHLO), 0–1000-km 200-hPa divergence (D200), and the tendency of the model average symmetric tangential wind from 0 to 600 km (ATWT). For ocean parameters, the model uses ocean-heat content (OHC) and sea surface temperature coupled with V_{\max} to calculate potential

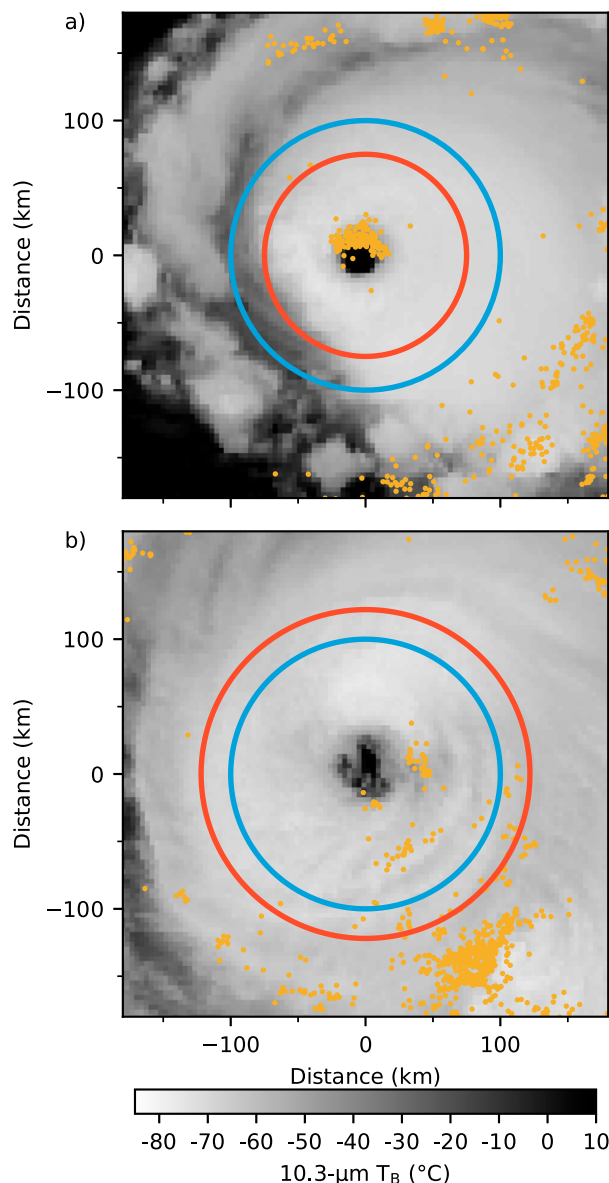


FIG. 2. The 10.3- μm longwave infrared imagery from *GOES-16* overlaid with lightning strokes from WWLLN (yellow points) and effective area for calculating of the lightning density from 100 km in physical space (blue curve) and scaled space (red curve) for (a) Hurricane Dorian at 0600 UTC 1 Sep 2019 with an fR5 of 75% of normal and (b) Hurricane Laura at 1800 UTC 26 Aug 2020 with an fR5 of 122% of normal.

intensification (POT). For convective parameters, the model uses a percent area with brightness temperature less than -30°C (PC30) and the standard deviation of the brightness temperature within 200 km of the storm (TBSD). From the working best track database files, the model uses the maximum sustained wind (V_{max}), the square of V_{max} , and 12-h intensity change (e.g., persistence; PER).

For the second framework, we mimic the linear discriminant analysis component of the Rapid Intensification Prediction

Aid (RIPA) model (Knaff et al. 2018, 2020)—RIPA also has a logistic regression component that we do not use here. With being developed for the Joint Typhoon Warning Center's area of responsibility (i.e., western North Pacific, Indian Ocean, Southern Hemisphere), RIPA has a separate set of model predictors from the SHIPS-RII, which has an eastern North Pacific and North Atlantic operational focus. But this configuration is skillful in the eastern North Pacific and North Atlantic (Franklin 2021). For model predictors, RIPA uses the 0–500-km generalized vertical wind shear (SHDG; Knaff et al. 2005), 200–800-km middle-tropospheric relative humidity (RHMD), 0–1000-km 200-hPa divergence using the 850-hPa vortex location (DIVC), average relative eddy momentum flux convergence in a 100–600-km annulus (REFC), and temperature advection between 850 and 700 hPa from 0 to 500 km (TADV). For ocean predictors, RIPA also uses OHC and sea surface temperature for the POT calculation. For the infrared-based convective parameters, RIPA also uses TBSD. But RIPA uses the percent area with brightness temperature less than -50° and -60°C (PC50 and PC60, respectively), radius of minimum brightness temperature between 0 and 150 km (RMNT), and the fractional deviation from the climatological infrared tropical cyclone size metric (fR5). RIPA uses V_{max} and PER from the best track data. Unlike prior rapid intensification guidance aids, RIPA includes caps on V_{max} and PER. By using the caps, RIPA has more independence between these predictors (Knaff et al. 2020).

For the lightning predictors, we define the inner-core 0–100-km scaled-radius area and rainband 200–300-km scaled-radius area lightning density predictors. For the first configuration (the spatial configuration), we mimic DeMaria et al. (2012) but take the natural logarithm of one plus the square root of the lightning density, defined by (2), for the two areas for the 0–6-h period. For the second configuration (the temporal configuration), we calculate the normalized lightning density with (2) from 0 to 1 h, from 0 to 6 h, and from 6 to 12 h for the inner-core area only.

We list the parameters for both configurations in the bottom third of Table 1. And in being consistent with other rapid intensification guidance (DeMaria and Kaplan 1994; DeMaria et al. 2012, 2014; Kaplan and DeMaria 2003; Kaplan et al. 2015; Knaff et al. 2018, 2020), we note that the vertical wind shear, relative humidity, 200-hPa divergence, average symmetric tangential wind tendency, temperature advection, and eddy momentum flux convergence are the average value over the forecast period (i.e., from $t = 0$ to $t = 12, 24, 36, 48, 72$ h) while the current intensity is at the analysis time $t = 0$ h and potential intensification and ocean heat content follow the forecast track on fixed ocean state fields.

Consensus techniques for tropical cyclone forecasting are common and improve tropical cyclone guidance skill (Goerss et al. 2004; Sampson et al. 2008). For the consensus, we equally weigh the SHIPS-RII and RIPA probabilities for the no-lightning predictor control and the two lightning predictor configurations (i.e., spatial and temporal). If the consensus with the lightning information is different from our skillful baselines, then we can argue that lightning-based predictors adds skill for the operational model.

TABLE 2. The rapid intensification guidance thresholds with the change in intensity ΔV_{\max} over a period Δt (first column). The frequency of the threshold occurring (i.e., sample base rate) in the North Atlantic and eastern and central North Pacific for the 2005–17 training sample and 2018–21 testing sample in parentheses (second column). The preceding 0–6-h mean inner-core lightning density (L06I) rapid intensification (RI) threshold and non-RI cases for the North Atlantic and eastern and central North Pacific in parentheses (third column). The threshold probability or trigger used to assume “yes” rapid intensification for validating the linear discriminant analysis models (last column).

| Threshold | | Lightning density ($\text{km}^{-2} \text{yr}^{-1}$) | | | |
|------------------------|----------------|---|-----------|-----------|-------------|
| ΔV_{\max} (kt) | Δt (h) | Frequency (%) | RI | No RI | Trigger (%) |
| 20 | 12 | 6.2 (7.2) | 3.2 (0.6) | 1.8 (0.5) | 20.0 |
| 25 | 24 | 12.8 (13.6) | 2.2 (0.4) | 1.7 (0.5) | 20.0 |
| 30 | 24 | 8.8 (9.3) | 2.3 (0.4) | 1.7 (0.5) | 22.5 |
| 35 | 24 | 5.7 (6.3) | 2.5 (0.5) | 1.7 (0.5) | 22.5 |
| 40 | 24 | 3.8 (4.0) | 3.7 (0.4) | 1.7 (0.5) | 17.5 |
| 45 | 36 | 6.7 (7.0) | 2.0 (0.3) | 1.6 (0.5) | 17.5 |
| 55 | 48 | 6.6 (6.0) | 1.5 (0.2) | 1.5 (0.5) | 12.5 |
| 65 | 72 | 6.9 (5.9) | 0.8 (0.2) | 1.5 (0.5) | 15.0 |

b. Guidance training

For the linear discriminant analysis, we use the SHIPS Developmental Dataset from 2005 to 2017 and train the SHIPS-RII and RIPA framework with the basins independently. Note that for RIPA, Knaff et al. (2018, 2020) train the operational RIPA model with the western North Pacific, Indian Ocean, and Southern Hemisphere basins. While the rapid intensification threshold is 30 kt in 24 h in the North Atlantic Ocean, operational rapid intensification aids predict multiple rapid intensification thresholds (Table 2). Each rapid intensification threshold has an independent set of coefficients calculated from the linear discriminant analysis. For linear discriminant analysis, each input feature (i.e., predictor) should have a Gaussian distribution and should exhibit some degree of independence from other candidate features. While not shown here, the variables in RIPA are more independent than those used with SHIPS-RII, which lends to the independence between the two linear discriminant analysis models and why we choose to assess the value of lightning-based predictors using these two guidance frameworks.

Table 2 shows the mean 0–6-h inner-core lightning density between the rapid intensification and no rapid intensification groups. Table 3 shows the lightning predictor coefficient values for the linear discriminant analysis for the 30 kt in 24 h

threshold for the SHIPS-RII and RIPA configurations for the eastern and central North Pacific and North Atlantic basins (all model predictor coefficient values are in Table S1 in the online supplemental material). For the spatial lightning predictors for the 30 kt in 24 h threshold, we see that the coefficient is negative for the inner-core area (L06I) for the North Atlantic and eastern and central North Pacific. Negative lightning predictor coefficients indicate that lightning activity is decreased for rapid intensification events in comparison to nonrapid intensification events. The sign of the coefficients is consistent with DeMaria et al. (2012). However, in DeMaria et al. (2012), the rainband areas (L06R) had a positive and neutral contribution, respectively. But, here, we see that the rainband area lightning density contribution for the North Atlantic is neutral to slightly negative. For the temporal configuration outlined here, the inner core predictors for 0–1 h (L01I), 0–6 h (L06I), and 6–12 h (L12I) are negative for the North Atlantic and eastern and central North Pacific. The basins differ on the importance of the 6–12-h predictor (L12I) with the eastern and central North Pacific showing a larger, negative contribution.

c. Guidance independence

In all rapid intensification guidance, environmental and storm metadata metrics such as intensity, persistence, potential intensification, sea surface temperature or ocean heat content,

TABLE 3. The linear discriminant analysis standardized coefficients for the lightning density predictors (i.e., L01I, L06I, L12I, L06R; see Table 1 for descriptions) for the 30 kt in 24 h rapid intensification threshold (for the coefficients for all predictors, see Table S1). The table is split in two for the North Atlantic and eastern and central North Pacific. Within the basins, the two configurations for spatial and temporal are given. Within each lightning predictor configuration, the values for the SHIPS and RIPA frameworks are provided.

| Predictor | North Atlantic | | | | Eastern and central North Pacific | | | |
|-----------|----------------|--------|----------|--------|-----------------------------------|--------|----------|--------|
| | Spatial | | Temporal | | Spatial | | Temporal | |
| | SHIPS | RIPA | SHIPS | RIPA | SHIPS | RIPA | SHIPS | RIPA |
| L01I | — | — | −0.115 | −0.249 | — | — | −0.162 | −0.204 |
| L06I | −0.140 | −0.284 | −0.022 | −0.073 | −0.316 | −0.498 | −0.052 | −0.068 |
| L12I | — | — | −0.094 | −0.079 | — | — | −0.370 | −0.385 |
| L06R | −0.096 | −0.074 | — | — | 0.059 | 0.125 | — | — |

convective predictors, and vertical wind shear dominate as the most important predictors (see Table S1; Kaplan and DeMaria 2003; Kaplan et al. 2015; Knaff et al. 2018, 2020). As a consequence, other predictors assist at the margin of the probability threshold. To assess the differences in model probabilistic output, we use the 2005–17 test dataset seasons to evaluate the model independence by calculating a Spearman's rank correlation r_s and conducting paired Student's t test with an $\alpha = 0.05$. As expected, the model output probabilities between the no lightning and the two consensus lightning configurations are all correlated with r_s values between 0.83 and 0.98 across the eight rapid intensification thresholds. But, from the Student's t test, the p values between no lightning and the two consensus lightning configurations range from 7.3×10^{-5} to 3.5×10^{-37} allowing us to reject the null hypothesis that models with lightning predictors do not produce a different probabilistic forecast output. This suggests that the addition of the lightning predictors are indeed shifting the output probability distribution. Figure 3 shows the p values from the paired Student's t test between the temporal and spatial consensus configurations. We can reject the null hypothesis that there is no difference in the probabilistic model output for the 20 kt in 12 h, 35 and 40 kt in 24 h, and 65 kt in 72 h. However, for the more common rapid intensification thresholds of 25 and 30 kt in 24 h, 45 kt in 36 h, and 55 kt in 48 h, we cannot reject the null hypothesis. The 30 kt in 24 h appears to be consistent with the linear discriminant analysis coefficients for the rainband region lightning in our spatial configuration are relatively small (Table 3) meaning that the inner-core lightning is dominating. While these four rapid intensification threshold output probabilities are not significantly different, this finding may also point to lightning activity being influenced by proximity to land. The implications of a physical difference in lightning related to land are discussed in section 5b.

5. Lightning predictor evaluation

To assess the value added by the lightning density predictors, we evaluate the consensus forecast created from the SHIPS-RII and RIPA models with the North Atlantic and eastern North Pacific operational large-scale diagnostic files for 2018–21. We limit our evaluation to individual analysis times when the tropical system has an initial intensity greater than 34 kt and does not make landfall during the rapid intensification forecast period. For the evaluation metrics, we use hit rate (i.e., probability of detection) defined as $H = a/(a + c)$ where a , b , c , and d are values in a contingency table that stand for true positive, false positive, false negative, and true negative, respectively, and success ratio (i.e., one minus the false alarm ratio) defined as

$$SR = \frac{a}{a + b}, \quad (3)$$

which provides the fraction of the predicted rapid intensification events that verify [for a more thorough discussion of scalar contingency table and the metric formulation, see Wilks (2019)].

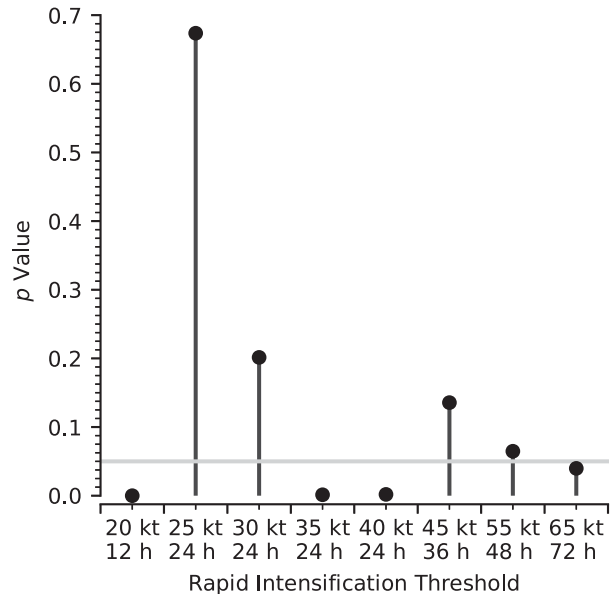


FIG. 3. A bar chart of the p values from a paired Student's t test for each of the rapid intensification thresholds. The gray horizontal line denotes $\alpha = 0.05$, which is used in assessing statistical significance.

With any metric computed from the contingency table, we can create a skill score. Skill scores compare model performance to a reference defined as

$$S = \frac{S - S_{\text{Ref}}}{S_{\text{Perf}} - S_{\text{Ref}}}, \quad (4)$$

where S is the value of the skill metric for the candidate model, S_{Ref} is the reference model score, and S_{Perf} is a perfect score in the metric that is either zero or one. The reference in meteorological applications is often climatology, persistence, or other skillful, baseline model. However, climatology is ill-defined for rapid intensification. Often, developers use frequency across all forecast periods as the reference state (e.g., Kaplan et al. 2015). However, this approach neglects the regional and environmental variability that plays vital roles in intensification, so the frequency produces an erroneous skill representation (Hamill and Juras 2006). In this context, persistence would also not resolve rapid intensification. Instead of comparing skill to climatology and persistence, we use the no-lightning version and the operational rapid intensification consensus (RIOCI) as our reference or baseline models.

For the rapid intensification, the operational guidance developers customarily define a threshold or trigger probability value for which a forecast exceeding this value is “yes” rapid intensification [for information on converting probabilistic forecasts to non-probabilistic forecasts, see the discussion in Wilks (2019)]. To create a skillful baseline, we define the threshold probability using the no-lightning consensus model on the training dataset to optimize the Heidke skill score and frequency bias ratio. The Heidke skill score (HSS) is defined as

$$\text{HSS} = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}, \quad (5)$$

where $\text{HSS} = 1$ is a perfect score and where the skill score attempts to quantify the accuracy relative to random chance. For rare events, the Heidke skill score approaches $2a/(2a + b + c)$ (referred to as the F_1 score; see the [appendix](#) for more detail). The frequency bias ratio B is defined as

$$B = \frac{a + b}{a + c}, \quad (6)$$

where $B = 1$ is an unbiased forecast model, $B < 1$ is underforecasts, and $B > 1$ is overforecasts. Since individual predictions can skew skill metrics, we perform a bootstrap resample with 100 000 permutations over the 2005–17 WWLLN training dataset. From the bootstrap resampling of the 2005–17 training dataset in [Fig. 4](#), the HSS for all thresholds levels off between threshold probability values of 15% and 40%. Within that range, we then select a threshold probability to the nearest 2.5% where the frequency bias ratio B is greater than one and HSS is maximized. For the 24-h thresholds, we choose 20%, 22.5%, 22.5%, and 17.5% for 25, 30, 35, and 40 kt, respectively. And we select 20%, 17.5%, 12.5%, and 15% for the 12-, 36-, 48-, and 72-h thresholds (also, see values in [Table 2](#)). Based on the selected threshold probabilities, [Fig. 5](#) shows that the spatial and temporal lightning configurations generally increase the success ratio SR and decrease the hit rate H while maintaining a relatively constant threat score, $\text{TS} = a/(a + b + c)$, for the 2005–17 training dataset. For the 20 kt in 12 h threshold (shades of blue in [Fig. 5b](#)), neither the spatial nor temporal configurations show changes to the skill over the no-lightning baseline consensus. For the 25 kt in 24 h thresholds (shades of blue in [Fig. 5a](#)), the spatial and temporal configurations have a similar threat score to and a frequency bias ratio closer to one than the no-lightning baseline. The 30 kt in 24 h has no discernible difference. For the 35 and 40 kt in 24 h, the spatial configuration appears to degrade the skill but is within the bounds provided by the standard deviation. For the 45 kt in 36 h threshold (shades of green in [Fig. 5b](#)), the lightning predictor configurations lend to underforecasting. With rare thresholds like at longer lead times, the standard deviation tends to increase highlighting the sensitivity to the small training sample.

a. Testing dataset assessment

For the 2018–21 eastern and central North Pacific and North Atlantic hurricane seasons, we test the rapid intensification guidance run on the operational large-scale diagnostic files. Unlike the SHIPS Development Dataset used for the 2005–17 training dataset, these operational large-scale diagnostic files contain errors associated with the operational estimates of storm position and current intensity as well as errors in the forecast fields. Given these sources of additional error, we repeat our bootstrap resampling approach to reflect some of this inherent uncertainty and used these values in

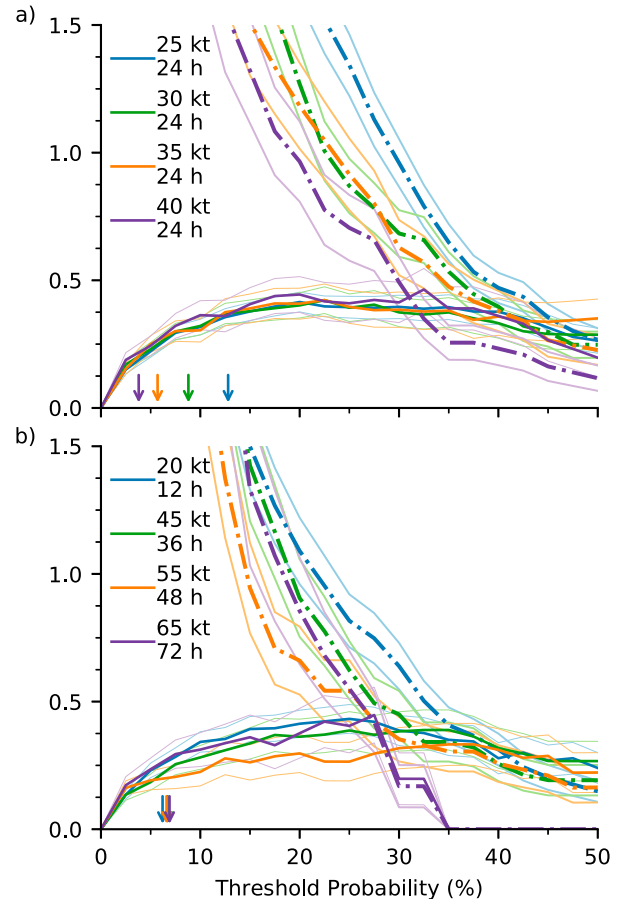


FIG. 4. The B (dashed curves) and HSS (solid curves) values for (a) the 24-h rapid intensification thresholds (i.e., 25, 30, 35, 40 kt in 24 h) and (b) the 12-, 36-, 48-, and 72-h rapid intensification thresholds where the mean (thick curves) and standard deviation (thin curves) are calculated from 0% to 50% at 2.5% increments calculated from a bootstrap resample with 100 000 permutations. The arrows denote the frequency at which rapid intensification threshold occurs in the 2005–17 training dataset (also listed in [Table 2](#)).

calculating the hit rate and success ratio shown in [Fig. 6](#) for the 2018–21 testing dataset. For the 20 kt in 12 h threshold, the spatial and temporal configuration do not improve the hit rate nor success ratio over the no-lightning baseline consensus or RIOC. For the 24-h thresholds, the spatial configuration shows mixed performance over the no-lightning baseline. But the temporal configuration shows improvement over the no-lightning and RIOC. For the 35 and 40 kt, the improvement across metrics is $\sim 10\%$ – 25% relative to the no-lightning model and $\sim 5\%$ – 10% relative to RIOC. For the 45 kt in 36 h threshold, the temporal lightning predictor configuration greatly outperforms the no-lightning configuration RIOC baselines. However, the spatial configuration for the 45 kt in 36 h threshold appears slightly better. Convective predictors such as lightning density appear to influence intensity forecasts in the short-term and generally do not aid in improving the 72-h lead times when compared to the no-lightning

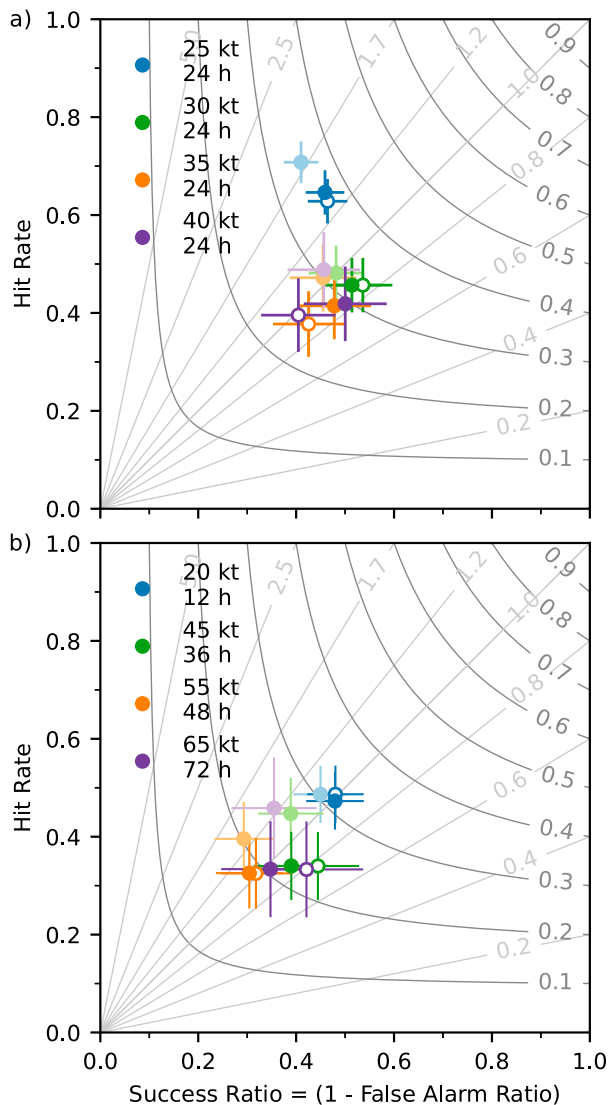


FIG. 5. A categorical performance diagram for the 2005–17 training dataset seasons with the SR on the abscissa, the hit rate H on the ordinate, frequency bias ratio B in the light gray lines, and threat score in the dark gray curves (Roebber 2009). The circles are the mean and error bars are the standard deviation for a fixed threshold probability value per rapid intensity threshold (see Table 2) calculated from a bootstrap resample with 100 000 permutations. The (a) 24-h and (b) 12-, 36-, 48-, and 72-h rapid intensification thresholds SR and H values are shown for the no-lightning (light, closed circle), spatial (dark, open circle), and temporal (dark, closed circle).

baseline. Also, given that long-lead thresholds are rare, the no-lightning consensus also shows a degradation compared to RIOCI, which points to our 13-yr training dataset being insufficient at discriminating between the rapid intensification events. Despite this, the verification of the 2018–21 seasons in the testing dataset shows that adding lightning observations in the form of the temporal configuration increases skill in the 24–36-h lead times with respect to RIOCI.

b. Distance to land

To address whether lightning predictors add value to rapid intensification forecasts, we must attempt to untangle the discrepancy between the spatial and temporal predictor configurations. As previously noted, DeMaria et al. (2012) show a perplexing relationship between rapid intensification and rainband lightning in the North Atlantic. And the authors indicate that this rainband area relationship does not exist in the eastern North Pacific—in fact, the rainband area has no correlation to intensification there. Considering that our longer training dataset yielded a different relationship, we believe that the rainband area predictor is contaminated by convection over land, impacted by unstable land-originating air masses (MacGorman and Rust 1998), and influenced by aerosols (Stolz et al. 2015), and that this is a plausible explanation for why the rainband area had a positive contribution in the DeMaria et al. (2012) and a neutral to slightly negative contribution here. In the North Atlantic, tropical cyclones tend to rapidly intensify in portions of the basin close to land (Kaplan and DeMaria 2003) causing the rainband area predictor to capture noncausal spikes in lightning activity. While the detection efficiency of land-based lightning detection networks is a function of land, we use the SHIPS distance-to-land database, which includes continents and major, storm-influencing islands to understand the possible relationship between land and our lightning predictors. We plot the 2005–21 North Atlantic and eastern and central North Pacific 30 kt in 24-h rapid intensification threshold events with respect to distance to land in Figs. 7a and 7d. From Figs. 7b, 7c, 7e, and 7f, lightning density increases as storms approach land with the rainband area showing a clearer relationship likely due to rainbands being over land in regions like the Caribbean, Gulf of Mexico, or Pacific coast of Mexico. With this relationship to land, we believe that any correlation between rapid intensification and lightning in the rainband area is not causal but is endemic to the location and frequency of which rapid intensification occurs in the basin in the training sample.

Returning to Fig. 6, the spatial and temporal lightning configurations have similar performance for the 30 kt in 24 h rapid intensification threshold. From the Heidke skill score, the accuracy difference is $\sim 0.5\%$ (0.4184 and 0.4229, respectively). To understand the impact of land on the spatial and temporal lightning configurations, we separate the forecasts into two groups: tropical cyclone center positions with a distance from land that is 1) ≥ 800 km (1817 forecasts) and 2) ≤ 600 km (443 forecasts). For the ≥ 800 km from land, the difference is less than $\sim 0.9\%$ (0.3682 and 0.3591, respectively) in Heidke skill score between the spatial and temporal configuration, which is consistent for our all storms metrics. When storms are within 600 km of land, the spatial configuration is $\sim 4.5\%$ worse than the temporal (0.5265 and 0.5716, respectively).

From evaluating the relationship between lightning density and distance to land, we opt for the temporal configuration in our operational implementation of the rapid intensification guidance. By doing so, we largely avoid the spurious correlation between rainband lightning density and rapid intensification

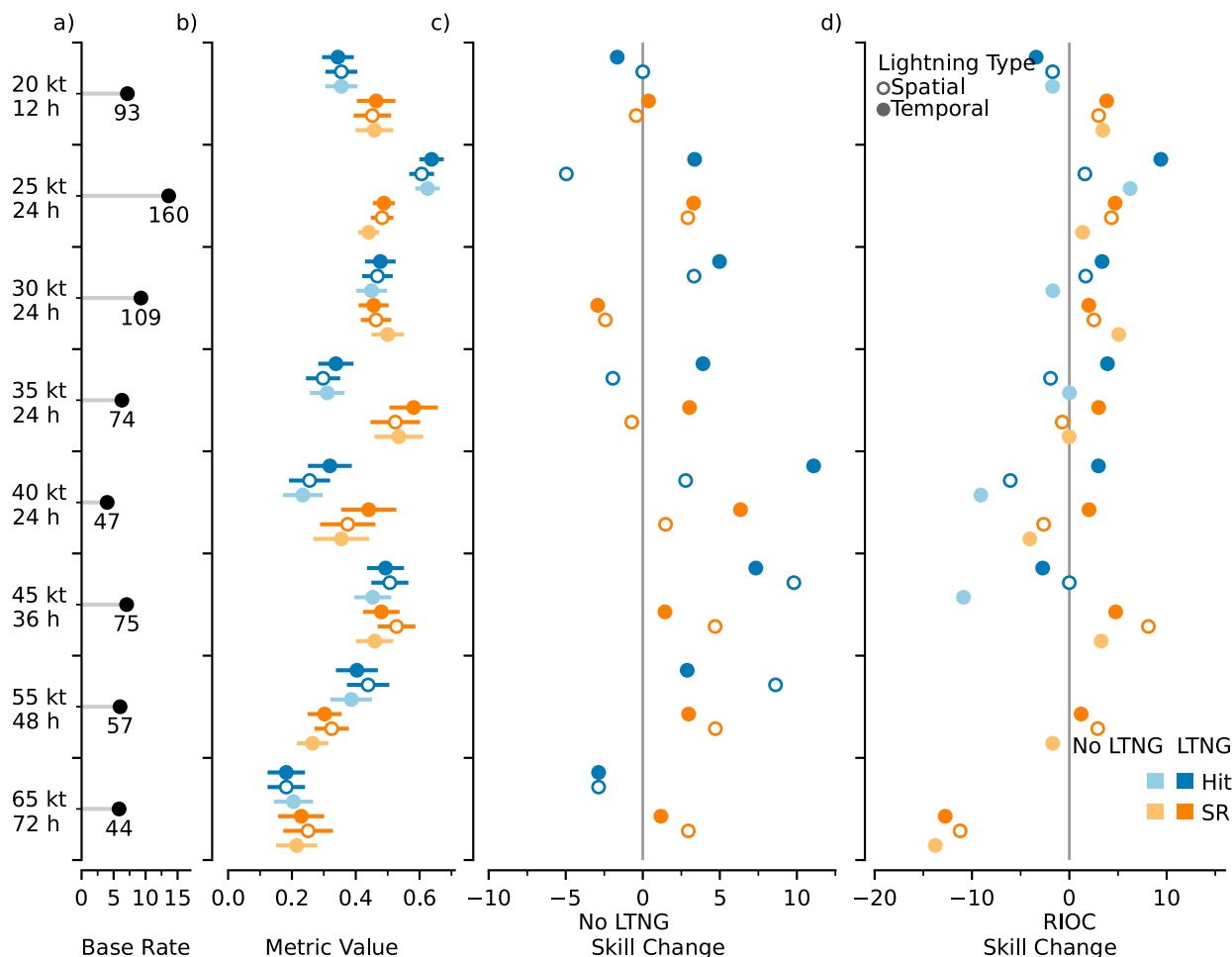


FIG. 6. For each of the rapid intensification thresholds along the ordinate, (a) the 2018–21 sample base rate (i.e., percent of forecast periods) for the North Atlantic and eastern and central North Pacific that reach the threshold along abscissa with the number of cases labeled by the circle. (b)–(d) The hit rate (blue) and success ratio (SR; orange) are shown for the spatial (dark, open circle) and temporal (dark, closed circle) lightning predictor consensus and consensus without lightning (No LTNG; light, closed circle) for the scalar metric values in (b), the skill change (or skill) relative to the consensus without lightning (No LTNG) in (c), and the skill change (or skill) relative to the rapid intensification operational consensus (RIOC) in (d). The horizontal bars in (b) represent the standard deviation from a bootstrap resample with 100 000 permutations. Note that we do not show some large negative values in (c) and (d) for longer lead times.

noted in DeMaria et al. (2012) for the North Atlantic. And we believe that this comparison assists in explaining the discrepancy between the North Atlantic and eastern North Pacific rainband lightning density coefficients (DeMaria et al. 2012). The influence of land is reduced for the inner-core lightning predictors. Considering that we already add seasonal climatology adjustments and regress GLM onto the WWLLN observations as will be highlighted in the next section, we do not want to add another adjustment for land and the inner-core lightning. A longer record from GLM might allow for a land-based adjustment.

6. Using GLM in operations

In DeMaria et al. (2012), the authors aspire to use GLM lightning data in the operational SHIPS-RII. The benefit of

using GLM lightning data is that they are available in real time for operations. However, with rapid intensification being infrequent, developers need to train the model with a sufficiently large training dataset to ensure a robust stable statistical-dynamical model—a problem for GLM data. As noted earlier, the goal of this work is to understand if we can create a longer, larger training dataset to allow us to still use GLM in an operational setting. Figure 8 shows that for a single storm (Hurricane Florence 2018) the two lightning data sources appear similar. However, an outlier at 0000 UTC 11 September is apparent. Initially, it appears that GLM captured a spike in lightning activity as intensification ended. However, at this time, Hurricane Florence passed through the imager-induced erroneous region of false lightning detection that appears in the GLM L2 product events, groups, and flashes referred to as the “Bahama bar” (Rudlosky et al. 2019).

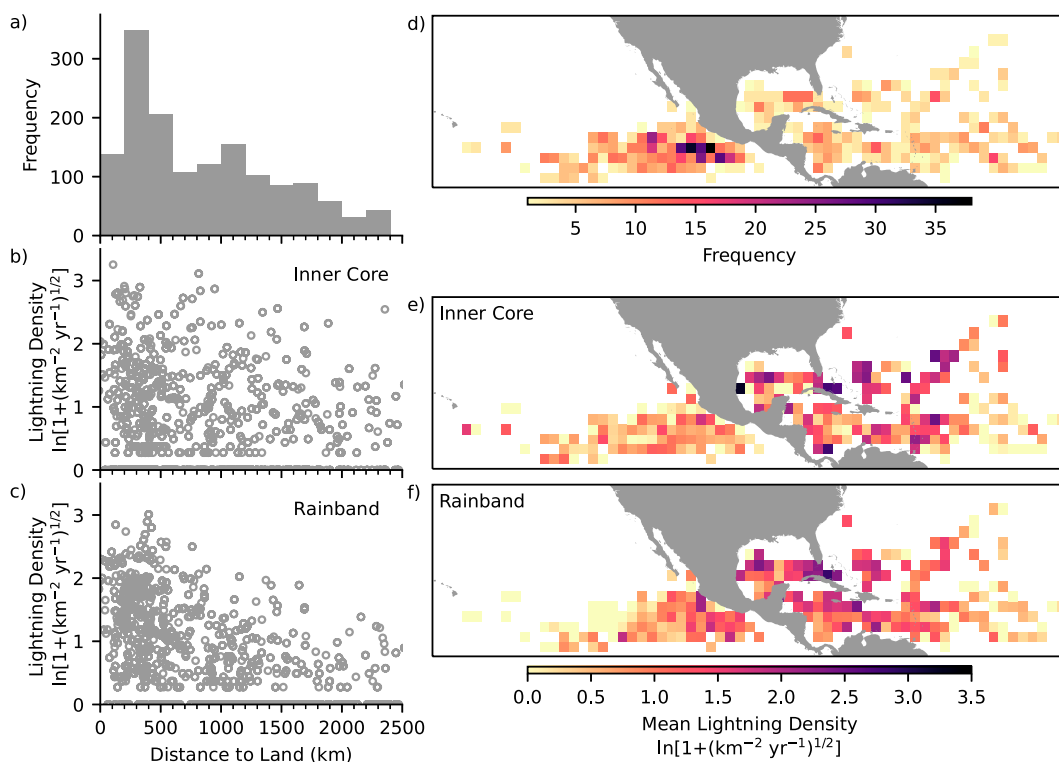


FIG. 7. The 2005–21 North Atlantic and eastern and central North Pacific 30 kt in 24 h rapid intensification threshold events with respect to distance to land for the (a) frequency binned by 200 km, (b) 0–6-h inner-core lightning density \mathcal{L} , (c) 0–6-h rainband lightning density, and $2^\circ \times 2^\circ$ binned (d) rapid intensification frequency, (e) mean inner-core lightning density, and (f) mean rainband lightning density. Note that storms that pass over land or make landfall in the 24 h from the initial time are removed.

While outliers caused by the Bahama bar and reverse Bahama bar (data devoid) regions exist, we can substitute the WWLLN-based predictors with the GLM-based predictors assuming that the WWLLN strokes and GLM flashes correlated and can be biased corrected. We bin, plot, and correlate lightning density values for GLM and WWLLN (Fig. 9). From the figure, the impact of using the natural logarithm of the square root of the lightning density plus one following (2) more realistically captures the lightning data and improves the representation of the high end of the lightning density distribution. To understand the relationship between WWLLN and GLM-based inner-core predictors, we calculate the Spearman's rank correlation coefficient r_s between GLM and WWLLN 6-hourly inner-core lightning density. The r_s values for a tropical cyclone with maximum sustained winds exceeding 95 kt (i.e., major hurricane), exceeding 64 kt (i.e., hurricane-force wind), between 34 kt (tropical storm force) and 64 kt, and less than 34 kt are 0.91 (188 cases), 0.90 (506), 0.93 (1612), and 0.94 (518), respectively. Given that the correlation between the WWLLN stroke-based and GLM flash-based lightning density values is statistically significant across storm intensities, we opt for a simple piecewise linear regression. For the 0–100-km inner-core area bin, the piecewise linear regression between the GLM and WWLLN \mathcal{L} values calculated with (2) has a slope m of 1.26 and intercept b of 0.20 for systems weaker

than 64 kt and $m = 1.17$ and $b = 0.09$ for systems stronger than 64 kt. Using these linear regressions, we run the lightning-based models trained with the 2005–17 WWLLN predictors with the corrected GLM predictors.

While the lightning densities are highly correlated, we want to understand the practical impact. Figure 10 shows a case

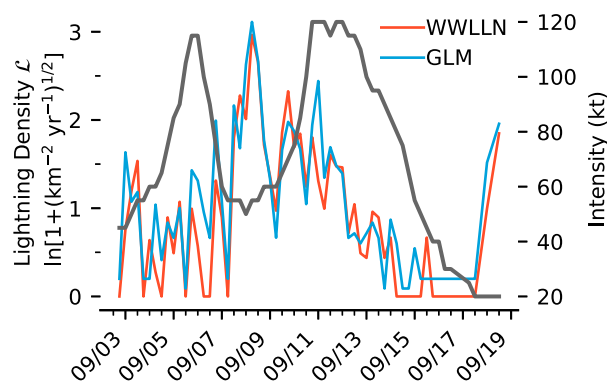


FIG. 8. Hurricane Florence (2018) lightning density \mathcal{L} from 2 to 19 Sep for the 0–6-h 0–100-km inner core for WWLLN (red curve) and GLM (blue curve). The intensity (gray curve) is read using the right-hand-side ordinate. Note that the spike in lightning density at 0000 UTC 11 Sep 2018 is from a Bahama bar.

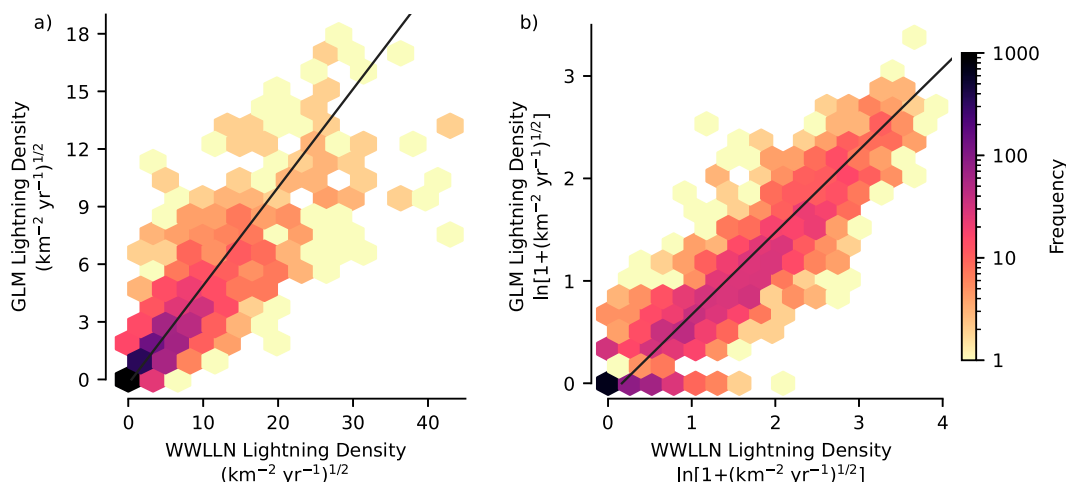


FIG. 9. The joint distribution of the inner-core lightning density for 2018–21 in the North Atlantic and eastern North Pacific for WWLLN on the abscissa and *GOES-16* GLM on the ordinate. (a) The square root of the lightning density calculated using (1) and (b) the values after calculating \mathcal{L} , the natural logarithm of the square root of the lightning density plus one using (2). The black line is the linear regression between GLM and WWLLN for all tropical cyclone intensities.

where 1) WWLLN misses activity observed by GLM in Hurricane Florence on 0600 UTC 6 September 2018 (Fig. 10a) and 2) GLM misses activity observed by WWLLN in Hurricane Delta on 0000 UTC 6 October 2020 (Fig. 10b). We note that these discrepancies between the instruments do not have a source as clear as the Bahama bar. Errors may arise from distance-to-land detection efficiency issues for WWLLN or high cloud optical thickness for GLM (Rutledge et al. 2020). Using the testing dataset for 2018–21, we calculate the probabilities of rapid intensification thresholds for the temporal lightning predictor configuration and *GOES-16* GLM. Then, we calculate the percentage of cases that would change from rapidly intensifying (true and false positives) to nonrapidly intensifying using the temporal lightning predictors. On average, we see four fewer forecasts of rapid intensification per season per threshold based on the sample base rates for the 2018–21 seasons. For true positives, we see on average one less correct forecast of rapid intensification per season per threshold. The occurrence of these differences in true positives between using GLM data in the models trained using WWLLN data is infrequent relative to the total number of forecasts per season.

During the 2022 North Atlantic and eastern North Pacific hurricane seasons, the temporal lightning density predictor configuration of the SHIPS-RII and RIPA models ran as a real-time experimental demonstration at the National Hurricane Center. Using the preliminary, working best track data—the operational assessment of tropical cyclone intensity that has not gone through a postseason analysis, we assess the real-time performance of the consensus output in comparison to RIOC. Figure 11 shows the success ratio and hit rate for each threshold. From Fig. 11a, the lightning-based guidance performed better than RIOC for the 25, 30, and 35 kt in 24 h thresholds. But, underperformed for the 40 kt in 24 h threshold. For the other thresholds shown in Fig. 11b, the lightning guidance demonstrated

similar performance with 45 kt in 36 h and 55 kt in 48 h had better threat scores but 20 kt in 12 h and 65 kt in 72 h had worse threat scores. While some thresholds performed worse, this degradation for these thresholds is not outside the characteristic behavior of other rapid intensification guidance to raise concerns about performance in future seasons if the lightning-based predictors transition into operations.

7. Discussion and conclusions

Here, we revisit and document using lightning data as a predictor in the rapid intensification guidance motivated by DeMaria et al. (2012). Now that GLM is in orbit, we want to leverage the additional decade of WWLLN data to understand the best predictors and to understand how we can use the GLM instrument. We use ground-based lightning observations from 2005 to 2017 from WWLLN to develop predictors. And we introduce new approaches for preprocessing the data to consider storm size using fR5 and to normalize the lightning density values. To isolate the impact of the lightning predictors, we use two guidance frameworks, SHIPS-RII and RIPA, to see if the lightning predictors add value in a consensus. Next, we evaluate the model performance for the 2018 through 2021 seasons using the operational input. From our approach, we show that taking advantage of how lightning varies over the 12-h prior to rapid intensification yields improved guidance in the short term (e.g., 12–48 h) when compared to the no-lightning baseline and operational consensus (RIOC). We also show that the original inner-core and rainband area predictors proposed by DeMaria et al. (2012) do not exploit lightning data fully given performance relative to the temporal predictors outlined here. Here, we also show that the rainband area lightning degrades performance during the 2018–21 seasons. While lightning in the rainband area near land appears to degrade performance in aggregate for rapid

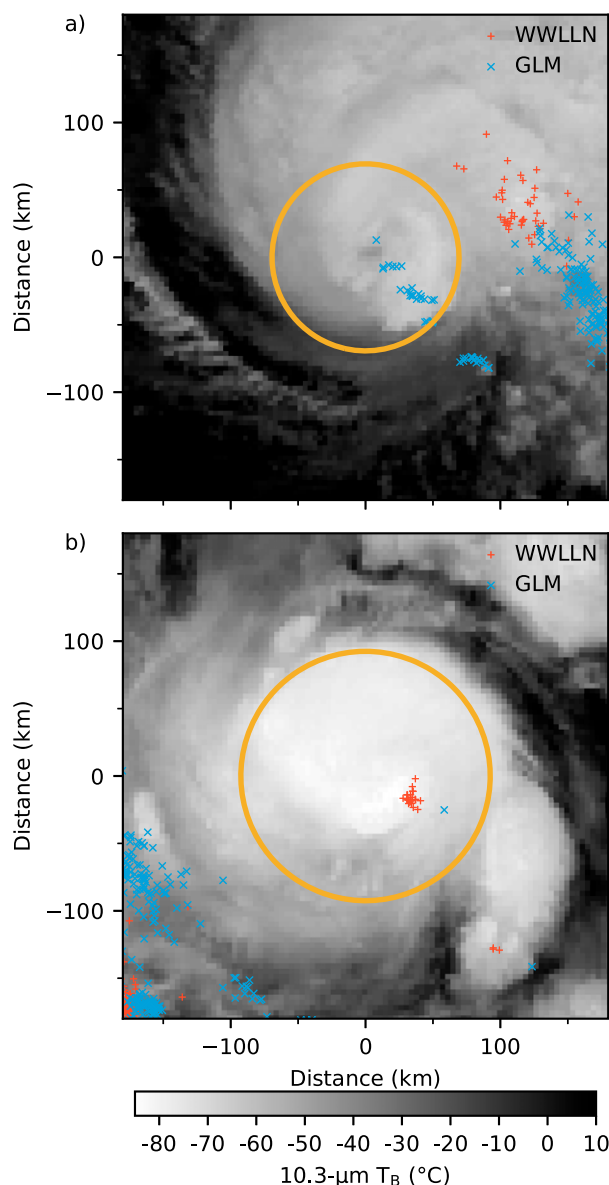


FIG. 10. Cases where lightning observations are missing in the 0–6-h fR5-scaled inner-core area (yellow circle) for (a) Hurricane Florence at 0600 UTC 6 Sep 2018 for WWLLN (red “+” symbols) compared with GLM (blue “x” symbols) and (b) Hurricane Delta at 0000 UTC 6 Oct 2020 for GLM compared with WWLLN. The lightning data are overlaid on *GOES-16* 10.3- μm longwave infrared imagery.

intensification guidance, rainband lightning may have utility on a case-by-case basis and with respect to vertical wind shear (Stevenson et al. 2016).

While we train the model using WWLLN since statistical-dynamical rapid intensification guidance requires longer data records than what is currently available solely from GLM, we also discuss the impact of using GLM in place of WWLLN in operations. In converting the WWLLN strokes and GLM flashes into lightning density, we find that both datasets are

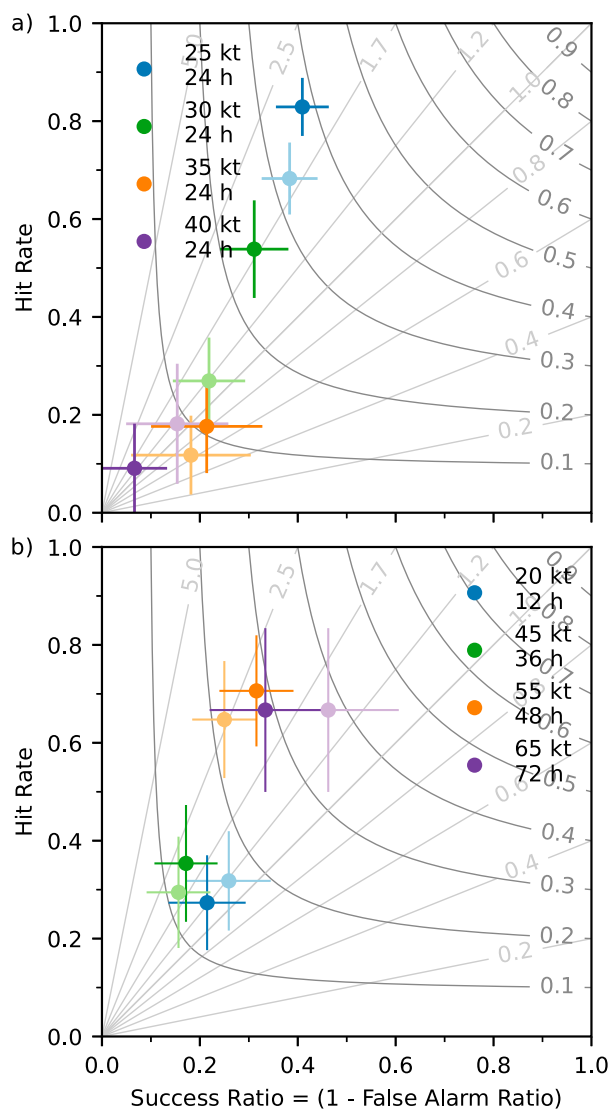


FIG. 11. A categorical performance diagram for the 2022 North Atlantic and eastern North Pacific hurricane season using the preliminary, working best track data with the SR on the abscissa, the hit rate H on the ordinate, frequency bias ratio B in the light gray lines, and threat score in the dark gray curves (Roebber 2009). The circles are the mean and error bars are the standard deviation for a fixed threshold probability value per rapid intensity threshold (see Table 2) calculated from a bootstrap resample with 100 000 permutations. The (a) 24-h and (b) 12-, 36-, 48-, and 72-h rapid intensification thresholds SR and H values are shown for the RIOCI (light circle) and temporal (dark circle).

highly correlated. We show that discrepancies between running the trained WWLLN-based models with either the WWLLN or GLM data are trivial and should have a minor impact on operational implementation of lightning-based guidance. Further into the GLM mission, the community can explore other approaches for blending the WWLLN and GLM datasets to create a robust, long-term tropical cyclone centric lightning dataset.

Throughout this work, we suggest that to fully understand lightning with respect to intensity change, we need to continue to understand changes in time at multiple time scales. From GOES-R GLM imagery, storms clearly exhibit short-lived bursts of lightning activity that are lost in the 1- and 6-h predictor quantities. However, this work does add to the discussion of the role of lightning in tropical cyclone intensity trends by supporting the previous finding presented in our introduction. But, obvious questions, whether the lightning density metric extracts the most value from the lightning observations and whether linear discriminate analysis can extract this value, remain. And we believe that there is likely utility in other GLM lightning metrics (e.g., lightning extent density and area) and in using higher temporal variability information rather than the preceding 0–1-, 0–6-, and 6–12-h symmetric lightning density quantities so that more abrupt changes to the storm can be captured. Work of this type is also limited by the rarity of rapid intensification events, and uncertainty and frequency of intensity estimates in both operations and best tracks. Each of these issues and questions can and should be addressed in the future. That being said, GLM information, as used here, assisted in improving our ability to anticipate rapid intensification events as shown by the evaluation of the 2022 North Atlantic and eastern North Pacific hurricane season experimental demonstration, and these methods represent an incremental improvement in our abilities.

Finally, we want to reiterate that we believe that these models developed here serve as a baseline for the community. These baseline models can be extended as additional environmental satellite agencies launch lightning mappers on geostationary satellite platforms. For example, the approach outlined for the GOES-R series can be extended to lightning observations from the Japan Meteorological Agency *Himawari-8/Himawari-9* follow-on program by creating the appropriate lightning density values and using the western Pacific Ocean basin seasonal correction adjustment factors, Fig. 1 (Bessho 2019). In this work, we scaled using the *fr5* metric. But physical and machine learning models are improving our operational estimates of the radius of maximum wind allowing tropical cyclone warning centers to create postseason analyses of these metrics (Knaff et al. 2021; Landsea 2022a,b; Chavas and Knaff 2022). This information along with gridded lightning products (Duran et al. 2021) may allow us to relate lightning to intensity change based on our physical (Schubert and Hack 1982) and observational (Rogers et al. 2013, 2016) understanding. These baseline models will allow the community to measure skill of future work as the community learns to exploit the always-available lightning data provided by the GLM instruments.

Acknowledgments. The authors wish to thank Naufal Razin and Ben Trabling for their comments on an early version of this manuscript as well as three anonymous reviewers. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce. The authors wish to thank the World Wide Lightning Location

Network (<http://wwln.net>), a collaboration among over 50 universities and institutions, for providing the lightning location data used in this paper.

Data availability statement. The *GOES-16/GOES-17* Advanced Baseline Imager longwave infrared and Geostationary Lightning Mapper Level 2 data are available from NOAA's Comprehensive Large Array-data Stewardship System (CLASS). Information about requesting WWLLN data are available from the University of Washington at <https://wwln.net/>. The operational and final Automated Tropical Cyclone Forecast (ATCF) best track database (Sampson and Schrader 2000) and the operational large-scale diagnostic files are available from the National Hurricane Center at <https://ftp.nhc.noaa.gov/atcf/>. The SHIPS Developmental Dataset predictor files is available from the Cooperative Institute for Research in the Atmosphere at http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp.

APPENDIX

Heidke Skill Score, F_1 Score Metric, and Rare Events

Skill assessment is a necessary element for developing, testing, and evaluating models either for insight into a phenomenon or transition to operations. However, selecting the best or appropriate skill metric is unclear for rare events like tropical cyclone rapid intensification. Not only does the rarity present problems, but the labeled datasets contain uncertainties in addition to small sample sizes and sample imbalances. Therefore, developing models to predict rare events takes special care to avoid both large biases and overfitting while producing a tool that results in useful guidance. The concept of “usefulness” is somewhat subjective, but entails maximizing the hit rate H , while minimizing the potential for false alarms. And depending on the event of interest like rapid intensification, forecasters or decision makers may accept a small degree of forecast bias, particularly over forecasting. Here, we examine two skill metrics, the Heidke skill score (HSS) and the F_1 score metric—widely used in classification problems—for their behavior and strengths in assessing model skills associated with rare events.

The Heidke skill score provides a measure of how well a forecast does with respect to random chance. As mentioned in section 5, the Heidke skill score, (5), approaches $2a/(2a + b + c)$ for rare events. While not apparent in this form, $2a/(2a + b + c)$ is the F_1 score metric, a special case of the Sørensen–Dice coefficient. Typically, the F_1 score metric is written as the harmonic mean between success ratio SR (i.e., precision) and hit rate H (i.e., recall). To explore the relationship between the Heidke skill score and F_1 score metric, we use the Hogan and Mason (2011) framework for comparing classification evaluation metrics with respect to the base rate s , hit rate H , and false alarm rate F (i.e., probability of false detection). Here, base rate, $s = (a + c)/n$, is the sample event frequency or event climatology, where n is the summation of a , b , c , and d . As a reminder, hit rate H (i.e., probability of detection) is given

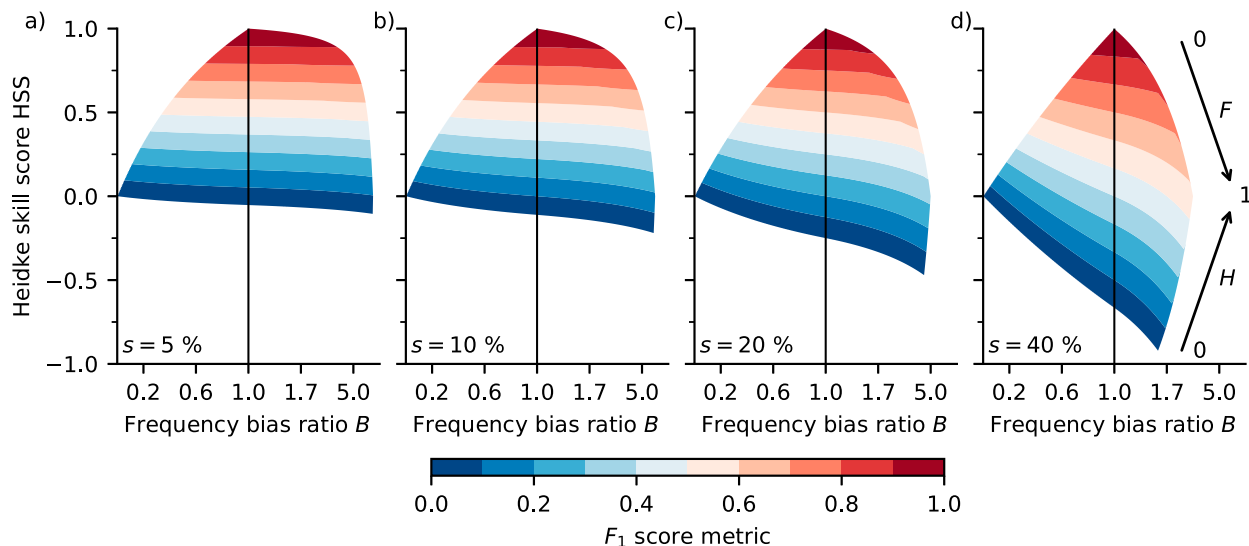


FIG. A1. Skill-bias diagrams with the Heidke skill score [HSS; (A2)], on the ordinate and frequency bias ratio [B; (A1)] on the abscissa, and contoured values of the F_1 score metric (A4). The contours for the F_1 score metric are for when the base rate s is (a) 5%, (b) 10%, (c) 20%, and (d) 40%. Annotations in (d) show the relationship between the skill-bias diagram and the hit rate H and false alarm rate F . Note that the abscissa is scaled for $B > 1$ so that the degree of over-/underforecasting is symmetric around $B = 1$.

by $a/(a + c)$. The false alarm rate F is defined as $b/(b + d)$. Using s , H , and F , we can write $a/n = sH$, $b/n = (1 - s)F$, $c/n = s(1 - H)$, and $d/n = (1 - s)(1 - F)$.

With these definitions, we can write the frequency bias ratio [B; (6)] as

$$B = \frac{a + b}{a + c} = \frac{(1 - s)}{s} F + H, \quad (\text{A1})$$

the Heidke skill score [HSS; (5)] as

$$\begin{aligned} \text{HSS} &= \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)} \\ &= \frac{2s(1 - s)(H - F)}{s + s(1 - 2s)H + (1 - s)(1 - 2s)F}, \end{aligned} \quad (\text{A2})$$

and the success ratio [SR; (3)] as

$$\begin{aligned} \text{SR} &= \frac{a}{a + b} \\ &= \frac{sH}{(1 - s)F + sH}. \end{aligned} \quad (\text{A3})$$

As mentioned, we can write the F_1 score metric as

$$\begin{aligned} F_1 &= \frac{2a}{2a + b + c} \\ &= 2 \frac{\text{SR} \times H}{\text{SR} + H} \\ &= \frac{2sH}{s(1 + H)(1 - s)F}, \end{aligned} \quad (\text{A4})$$

where the top line of the equation is in terms of a , b , c , d , the middle line is the harmonic mean, and the bottom line is in terms of s , H , F .

From (A1)–(A4), we can see that each metric is dependent on the base rate s , which suggests that the scoring metrics will have different meanings for balanced and imbalanced classification problems. To assess the impact of s , we employ a skill-bias diagram with HSS on the ordinate and frequency bias ratio B on the abscissa. Figure A1 shows contours of F_1 in the skill-bias space for base rates, $s = 5\%$, 10% , 20% , 40% . Note that the distorted quadrilateral-like shape can be interpreted in terms of hit rate H and false alarm rate F , with H along the upper-left and lower-right edges and F along the lower-left and upper-right edges. As annotated in Fig. A1d, H increases from 0 to 1 from the lower left to the upper right, and F increases from 0 to 1 from the upper left to lower right. For $s = 5\%$ (Fig. A1a), we see that for both the HSS and F_1 score are relatively independent of the frequency bias ratio B . While Ebert and Milne (2022) see this as a shortcoming, this can be a strength when used in conjunction with the frequency bias ratio B if the application desires a specific tendency to either over- or underforecast. As HSS approaches F_1 for low values of s , the range of HSS decreases from -1 to 1 to ~ 0 to 1 , where $\text{HSS} \leq 0$ represents a model with no forecast skill. For the F_1 score metric, we see that F_1 can be artificially increased by overforecasting (i.e., $B > 1$) as the base rate s increases from 5% to 40%. For example, when $F_1 = 0.5$ for $s = 40\%$ —a relatively balanced classification problem, the meaning as defined by the Heidke skill score changes from skillful to no skill as B approaches 2. As shown, the HSS limits the potential of a model maximizing skill by dramatically overforecasting infrequent events, and appears more appropriate as a metric for skill assessment for a wider range of forecast applications.

Rapid intensification is a rare event (i.e., low base rate s) that would allow us to assume that HSS and F_1 are equivalent for most rapid intensification thresholds. However, we opt to optimize our probability threshold value in section 5 with the HSS and $B > 1$ as shown in Fig. 4 since for some rapid intensification thresholds, we have a local maximum in our HSS values on either side of $B = 1$. In doing so, we avoid artificially increasing our chosen score metric through overforecasting and selecting a probability threshold that yields a higher frequency bias ratio B .

REFERENCES

- Abarca, S. F., K. L. Corbosiero, and D. Vollaro, 2011: The World Wide Lightning Location Network and convective activity in tropical cyclones. *Mon. Wea. Rev.*, **139**, 175–191, <https://doi.org/10.1175/2010MWR3383.1>.
- Alexander, G. D., J. A. Weinman, V. M. Karyampudi, W. S. Olson, and A. C. L. Lee, 1999: The effect of assimilating rain rates derived from satellites and lightning on forecasts of the 1993 superstorm. *Mon. Wea. Rev.*, **127**, 1433–1457, [https://doi.org/10.1175/1520-0493\(1999\)127<1433:TEOARR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<1433:TEOARR>2.0.CO;2).
- Apodaca, K., M. Zupanski, M. DeMaria, J. A. Knaff, and L. D. Grasso, 2014: Development of a hybrid variational-ensemble data assimilation technique for observed lightning tested in a mesoscale model. *Nonlinear Processes Geophys.*, **21**, 1027–1041, <https://doi.org/10.5194/npg-21-1027-2014>.
- Bessho, K., 2019: Status of Himawari-8/9 and their follow-on satellites program. *2019 Joint Satellite Conf.*, Boston, MA, Amer. Meteor. Soc., 8A.2A, <https://ams.confex.com/ams/JOINTSATMET/meetingapp.cgi/Paper/371501>.
- Black, R. A., and J. Hallett, 1999: Electrification of the hurricane. *J. Atmos. Sci.*, **56**, 2004–2028, [https://doi.org/10.1175/1520-0469\(1999\)056<2004:EOTH>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<2004:EOTH>2.0.CO;2).
- Bonfanti, C., L. Trailovic, J. Stewart, and M. Govett, 2018: Machine learning: Defining worldwide cyclone labels for training. *Proc. IEEE 21st Int. Conf. on Information Fusion (FUSION)*, Cambridge, United Kingdom, Institute of Electrical and Electronics Engineers, 753–760, <https://doi.org/10.23919/ICIF.2018.8455276>.
- Cecil, D. J., 2001: LIS/OTD 0.5 degree high resolution full climatology (HRFC) V2.3.2015. NASA, accessed 13 July 2015, <https://doi.org/10.5067/LIS/LIS-OTD/DATA302>.
- , E. J. Zipser, and S. W. Nesbitt, 2002: Reflectivity, ice scattering, and lightning characteristics of hurricane eyewalls and rainbands. Part I: Quantitative description. *Mon. Wea. Rev.*, **130**, 769–784, [https://doi.org/10.1175/1520-0493\(2002\)130<0769:RISALC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<0769:RISALC>2.0.CO;2).
- Chavas, D. R., and K. A. Emanuel, 2010: A QuikSCAT climatology of tropical cyclone size. *Geophys. Res. Lett.*, **37**, L18816, <https://doi.org/10.1029/2010GL044558>.
- , 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>.
- Comboto, C., A. Mouche, J. Knaff, Y. Zhao, Y. Zhao, L. Vinour, Y. Quilfen, and B. Chapron, 2020: Extensive high-resolution Synthetic Aperture Radar (SAR) data analysis of tropical cyclones: Comparisons with SFMR flights and best track. *Mon. Wea. Rev.*, **148**, 4545–4563, <https://doi.org/10.1175/MWR-D-20-0005.1>.
- 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](https://doi.org/10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2).
- , R. T. DeMaria, J. A. Knaff, and D. Molnar, 2012: Tropical cyclone lightning and rapid intensity change. *Mon. Wea. Rev.*, **140**, 1828–1842, <https://doi.org/10.1175/MWR-D-11-00236.1>.
- , C. R. Sampson, J. A. Knaff, and K. D. Musgrave, 2014: Is tropical cyclone intensity guidance improving? *Bull. Amer. Meteor. Soc.*, **95**, 387–398, <https://doi.org/10.1175/BAMS-D-12-00240.1>.
- , 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>.
- Duran, P., and Coauthors, 2021: The evolution of lightning flash density, flash size, and flash energy during Hurricane Dorian's (2019) intensification and weakening. *Geophys. Res. Lett.*, **48**, e2020GL092067, <https://doi.org/10.1029/2020GL092067>.
- Ebert, P. A., and P. Milne, 2022: Methodological and conceptual challenges in rare and severe event forecast verification. *Nat. Hazards Earth Syst. Sci.*, **22**, 539–557, <https://doi.org/10.5194/nhess-22-539-2022>.
- Fierro, A. O., S. N. Stevenson, and R. M. Rabin, 2018: Evolution of GLM-observed total lightning in Hurricane Maria (2017) during the period of maximum intensity. *Mon. Wea. Rev.*, **146**, 1641–1666, <https://doi.org/10.1175/MWR-D-18-0066.1>.
- , Y. Wang, J. Gao, and E. R. Mansell, 2019: Variational assimilation of radar data and GLM lightning-derived water vapor for the short-term forecasts of high-impact convective events. *Mon. Wea. Rev.*, **147**, 4045–4069, <https://doi.org/10.1175/MWR-D-18-0421.1>.
- Franklin, J. L., and M. Onderlinde, 2021: Verification of RI measures in 2021. HFIP Annual Meeting, NOAA, 21 pp., <https://hfip.org/sites/default/files/events/269/330-franklin-verification-riptxt.pdf>.
- 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., C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638, [https://doi.org/10.1175/1520-0434\(2004\)019<0633:AHOWNP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2).
- Goodman, S. J., and Coauthors, 2013: The GOES-R Geostationary Lightning Mapper (GLM). *Atmos. Res.*, **125–126**, 33–49, <https://doi.org/10.1016/j.atmosres.2013.01.006>.
- Hamill, T. M., and J. Juras, 2006: Measuring forecast skill: Is it real skill or is it the varying climatology? *Quart. J. Roy. Meteor. Soc.*, **132**, 2905–2923, <https://doi.org/10.1256/qj.06.25>.
- Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, <https://doi.org/10.1002/qj.3803>.
- Hogan, R. J., and I. B. Mason, 2011: Deterministic forecasts of binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd ed. I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 31–59.
- Jiang, H., and E. M. Ramirez, 2013: Necessary conditions for tropical cyclone rapid intensification as derived from 11 years of TRMM data. *J. Climate*, **26**, 6459–6470, <https://doi.org/10.1175/JCLI-D-12-00432.1>.
- Kaplan, J., and M. DeMaria, 2003: Large-scale characteristics of rapidly intensifying tropical cyclones in the North Atlantic

- basin. *Wea. Forecasting*, **18**, 1093–1108, [https://doi.org/10.1175/1520-0434\(2003\)018<1093:LCORIT>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<1093:LCORIT>2.0.CO;2).
- , and Coauthors, 2015: Evaluating environmental impacts on tropical cyclone rapid intensification predictability utilizing statistical models. *Wea. Forecasting*, **30**, 1374–1396, <https://doi.org/10.1175/WAF-D-15-0032.1>.
- Knaff, J. A., C. R. Sampson, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688–699, <https://doi.org/10.1175/WAF863.1>.
- , —, —, T. P. Marchok, J. M. Gross, and C. J. McAdie, 2007: Statistical tropical cyclone wind radii prediction using climatology and persistence. *Wea. Forecasting*, **22**, 781–791, <https://doi.org/10.1175/WAF1026.1>.
- , M. DeMaria, S. Longmore, and R. DeMaria, 2014a: Improving tropical cyclone guidance tools by accounting for variations in size. *31th Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 51, <https://ams.confex.com/ams/31Hurr/webprogram/Paper244165.html>.
- , S. P. Longmore, and D. A. Molenaar, 2014b: An objective satellite-based tropical cyclone size climatology. *J. Climate*, **27**, 455–476, <https://doi.org/10.1175/JCLI-D-13-00096.1>.
- , C. R. Sampson, and G. Chirokova, 2017: A global statistical–dynamical tropical cyclone wind radii forecast scheme. *Wea. Forecasting*, **32**, 629–644, <https://doi.org/10.1175/WAF-D-16-0168.1>.
- , —, and K. D. Musgrave, 2018: An operational rapid intensification prediction aid for the western North Pacific. *Wea. Forecasting*, **33**, 799–811, <https://doi.org/10.1175/WAF-D-18-0012.1>.
- , —, and B. R. Strahl, 2020: A tropical cyclone rapid intensification prediction aid for the Joint Typhoon Warning Center's areas of responsibility. *Wea. Forecasting*, **35**, 1173–1185, <https://doi.org/10.1175/WAF-D-19-0228.1>.
- , and Coauthors, 2021: Estimating tropical cyclone surface winds: Current status, emerging technologies, historical evolution, and a look to the future. *Trop. Cyclone Res. Rev.*, **10**, 125–150, <https://doi.org/10.1016/j.tcr.2021.09.002>.
- Kumler-Bonfanti, C., J. Stewart, D. Hall, and M. Govett, 2020: Tropical and extratropical cyclone detection using deep learning. *J. Appl. Meteor. Climatol.*, **59**, 1971–1985, <https://doi.org/10.1175/JAMC-D-20-0117.1>.
- Landsea, C. W., 2022a: Updated Northeast/North Central HURDAT2 for 2021 ... with RMW. Tropical Communities, accessed 7 April 2022, <http://tstorms.org/>.
- , 2022b: The revised Atlantic hurricane database (HURDAT2). NOAA, 8 pp., <https://www.nhc.noaa.gov/data/hurdat/hurdat2-format-atl-1851-2021.pdf>.
- , 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>.
- Lay, E. H., R. H. Holzworth, C. J. Rodger, J. N. Thomas, O. Pinto Jr., and R. L. Dowden, 2004: WWLL global lightning detection system: Regional validation study in Brazil. *Geophys. Res. Lett.*, **31**, L03102, <https://doi.org/10.1029/2003GL018882>.
- MacGorman, D. R., and W. D. Rust, 1998: *The Electrical Nature of Storms*. Oxford University Press, 422 pp.
- Mansell, E. R., C. L. Ziegler, and D. R. MacGorman, 2007: A lightning data assimilation technique for mesoscale forecast models. *Mon. Wea. Rev.*, **135**, 1732–1748, <https://doi.org/10.1175/MWR3387.1>.
- Marks, F., N. Kurkowski, M. DeMaria, and M. Brennan, 2018: NOAA Hurricane Forecast Improvement Project: Years ten to fifteen strategic plan. NOAA, 83 pp., hftp.org/sites/default/files/documents/hftp-strategic-plan-20190625.pdf.
- Mauchly, S. J., 1923: On the diurnal variation of the potential gradient of atmospheric electricity. *Terr. Magn. Atmos. Electr.*, **28**, 61–81, <https://doi.org/10.1029/TE028i003p00061>.
- Molinari, J., P. K. Moore, V. P. Idone, R. W. Henderson, and A. B. Saljoughy, 1994: Cloud-to-ground lightning in Hurricane Andrew. *J. Geophys. Res.*, **99**, 16 665–16 676, <https://doi.org/10.1029/94JD00722>.
- , —, and —, 1999: Convective structure of hurricanes as revealed by lightning locations. *Mon. Wea. Rev.*, **127**, 520–534, [https://doi.org/10.1175/1520-0493\(1999\)127<0520:CSOHAR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<0520:CSOHAR>2.0.CO;2).
- NRL, 2020: ATCF probability format. NRL, accessed 16 August 2022, https://www.nrlmry.navy.mil/atcf_web/docs/database/new/edeck.txt.
- Pan, L., X. Qie, D. Liu, D. Wang, and J. Yang, 2010: The lightning activities in super typhoons over the northwest Pacific. *Sci. China Earth Sci.*, **53**, 1241–1248, <https://doi.org/10.1007/s11430-010-3034-z>.
- , —, and D. Wang, 2014: Lightning activity and its relation to the intensity of typhoons over the northwest Pacific Ocean. *Adv. Atmos. Sci.*, **31**, 581–592, <https://doi.org/10.1007/s00376-013-3115-y>.
- Price, C., M. Asfur, and Y. Yair, 2009: Maximum hurricane intensity preceded by increase in lightning frequency. *Nat. Geosci.*, **2**, 329–332, <https://doi.org/10.1038/ngeo477>.
- RAMMB, 2021: SHIPS developmental data. Regional and Mesoscale Meteorology Branch, Cooperative Institute for Research in the Atmosphere, Colorado State University, accessed 22 April 2021, http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/developmental_data.asp.
- Rodger, C. J., J. B. Brundell, and R. L. Dowden, 2005: Location accuracy of VLF World Wide Lightning Location (WWLL) network: Post-algorithm update. *Ann. Geophys.*, **23**, 277–290, <https://doi.org/10.5194/angeo-23-277-2005>.
- , S. Werner, J. B. Brundell, E. H. Lay, N. R. Thompson, R. H. Holzworth, and R. L. Dowden, 2006: Detection efficiency of VLF World-Wide Lightning Location Network (WWLLN): Initial case study. *Ann. Geophys.*, **24**, 3197–3214, <https://doi.org/10.5194/angeo-24-3197-2006>.
- Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24**, 601–608, <https://doi.org/10.1175/2008WAF2222159.1>.
- Rogers, R. F., S. Lorsolo, P. Reasor, J. Gamache, and F. Marks, 2012: Multiscale analysis of tropical cyclone kinematic structure from airborne Doppler radar composites. *Mon. Wea. Rev.*, **140**, 77–99, <https://doi.org/10.1175/MWR-D-10-05075.1>.
- , P. Reasor, and S. Lorsolo, 2013: Airborne Doppler observations of the inner-core structural differences between intensifying and steady-state tropical cyclones. *Mon. Wea. Rev.*, **141**, 2970–2991, <https://doi.org/10.1175/MWR-D-12-00357.1>.
- , J. A. Zhang, J. Zawislak, H. Jiang, G. R. Alvey III, E. J. Zipser, and S. N. Stevenson, 2016: Observations of the structure and evolution of Hurricane Edouard (2014) during intensity change. Part II: Kinematic structure and the distribution of deep convection. *Mon. Wea. Rev.*, **144**, 3355–3376, <https://doi.org/10.1175/MWR-D-16-0017.1>.
- Rozoff, C. M., and J. P. Kossin, 2011: New probabilistic forecast models for the prediction of tropical cyclone rapid intensification. *Wea. Forecasting*, **26**, 677–689, <https://doi.org/10.1175/WAF-D-10-05059.1>.

- Rudlosky, S. D., S. J. Goodman, K. S. Virts, and E. C. Bruning, 2019: Initial geostationary lightning mapper observations. *Geophys. Res. Lett.*, **46**, 1097–1104, <https://doi.org/10.1029/2018GL081052>.
- Rutledge, S. A., K. A. Hilburn, A. Clayton, B. Fuchs, and S. D. Miller, 2020: Evaluating geostationary lightning mapper flash rates within intense convective storms. *J. Geophys. Res. Atmos.*, **125**, e2020JD032827, <https://doi.org/10.1029/2020JD032827>.
- Sampson, C. R., and A. J. Schrader, 2000: The Automated Tropical Cyclone Forecasting System (version 3.2). *Bull. Amer. Meteor. Soc.*, **81**, 1231–1240, [https://doi.org/10.1175/1520-0477\(2000\)081<1231:TATCFS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1231:TATCFS>2.3.CO;2).
- , J. L. Franklin, J. A. Knaff, and M. DeMaria, 2008: Experiments with a simple tropical cyclone intensity consensus. *Wea. Forecasting*, **23**, 304–312, <https://doi.org/10.1175/2007WAF2007028.1>.
- Schubert, W. H., and J. J. Hack, 1982: Inertial stability and tropical cyclone development. *J. Atmos. Sci.*, **39**, 1687–1697, [https://doi.org/10.1175/1520-0469\(1982\)039<1687:ISATCD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1982)039<1687:ISATCD>2.0.CO;2).
- Slocum, C. J., M. N. Razin, J. A. Knaff, and J. P. Stow, 2022: Does ERA5 mark a new era for resolving the tropical cyclone environment? *J. Climate*, **35**, 7147–7164, <https://doi.org/10.1175/JCLI-D-22-0127.1>.
- Stevenson, S. N., K. L. Corbosiero, and J. Molinari, 2014: The convective evolution and rapid intensification of Hurricane Earl (2010). *Mon. Wea. Rev.*, **142**, 4364–4380, <https://doi.org/10.1175/MWR-D-14-00078.1>.
- , —, and S. F. Abarca, 2016: Lightning in eastern North Pacific tropical cyclones: A comparison to the North Atlantic. *Mon. Wea. Rev.*, **144**, 225–239, <https://doi.org/10.1175/MWR-D-15-0276.1>.
- , —, M. DeMaria, and J. L. Vigh, 2018: A 10-year survey of tropical cyclone inner-core lightning bursts and their relationship to intensity change. *Wea. Forecasting*, **33**, 23–36, <https://doi.org/10.1175/WAF-D-17-0096.1>.
- Stolz, D. C., S. A. Rutledge, and J. R. Pierce, 2015: Simultaneous influences of thermodynamics and aerosols on deep convection and lightning in the tropics. *J. Geophys. Res. Atmos.*, **120**, 6207–6231, <https://doi.org/10.1002/2014JD023033>.
- Torn, R. D., and C. Snyder, 2012: Uncertainty of tropical cyclone best-track information. *Wea. Forecasting*, **27**, 715–729, <https://doi.org/10.1175/WAF-D-11-00085.1>.
- Vagasky, C., 2017: Enveloped eyewall lightning: The EEL signature in tropical cyclones. *J. Oper. Meteor.*, **5**, 171–179, <https://doi.org/10.1519/nwajom.2017.0514>.
- Virts, K. S., and W. J. Koshak, 2020: Mitigation of geostationary lightning mapper geolocation errors. *J. Atmos. Oceanic Technol.*, **37**, 1725–1736, <https://doi.org/10.1175/JTECH-D-19-0100.1>.
- Wilks, D. S., 2019: *Statistical Methods in the Atmospheric Sciences*. 4th ed. Elsevier, 840 pp.
- Xu, W., S. A. Rutledge, and W. Zhang, 2017: Relationships between total lightning, deep convection, and tropical cyclone intensity change. *J. Geophys. Res. Atmos.*, **122**, 7047–7063, <https://doi.org/10.1002/2017JD027072>.
- Zawislak, J., H. Jiang, G. R. Alvey III, E. J. Zipser, R. F. Rogers, J. A. Zhang, and S. N. Stevenson, 2016: Observations of the structure and evolution of Hurricane Eduoard (2014) during intensity change. Part I: Relationship between the thermodynamic structure and precipitation. *Mon. Wea. Rev.*, **144**, 3333–3354, <https://doi.org/10.1175/MWR-D-16-0018.1>.
- Zhang, W., Y. Zhang, D. Zheng, and L. Xu, 2015: Relationship between lightning activity and tropical cyclone intensity over the northwest Pacific. *J. Geophys. Res. Atmos.*, **120**, 4072–4089, <https://doi.org/10.1002/2014JD022334>.