A Simple Model for Predicting the Tropical Cyclone Radius of Maximum Wind from Outer Size⌖

DANIEL R. CHAVAS[a] AND JOHN A. KNAFF[b]

[a] *Purdue University, Department of Earth, Atmospheric, and Planetary Sciences, West Lafayette, Indiana*
[b] *NOAA/Center for Satellite Applications and Research, Fort Collins, Colorado*

ABSTRACT: The radius of maximum wind (R_{max}) in a tropical cyclone governs the footprint of hazards, including damaging wind, surge, and rainfall. However, R_{max} is an inconstant quantity that is difficult to observe directly and is poorly resolved in reanalyses and climate models. In contrast, outer wind radii are much less sensitive to such issues. Here we present a simple empirical model for predicting R_{max} from the radius of 34-kt (1 kt $\approx$ 0.51 m s^{-1}) wind ($R_{17.5\,ms}$). The model only requires as input quantities that are routinely estimated operationally: maximum wind speed, $R_{17.5\,ms}$, and latitude. The form of the empirical model takes advantage of our physical understanding of tropical cyclone radial structure and is trained on the Extended Best Track database from the North Atlantic 2004–20. Results are similar for the TC-OBS database. The physics reduces the relationship between the two radii to a dependence on two physical parameters, while the observational data enables an optimal estimate of the quantitative dependence on those parameters. The model performs substantially better than existing operational methods for estimating R_{max}. The model reproduces the observed statistical increase in R_{max} with latitude and demonstrates that this increase is driven by the increase in $R_{17.5\,ms}$ with latitude. Overall, the model offers a simple and fast first-order prediction of R_{max} that can be used operationally and in risk models.

SIGNIFICANCE STATEMENT: If we can better predict the area of strong winds in a tropical cyclone, we can better prepare for its potential impacts. This work develops a simple model to predict the radius where the strongest winds in a tropical cyclone are located. The model is simple and fast and more accurate than existing models, and it also helps us to understand what causes this radius to vary in time, from storm to storm, and at different latitudes. It can be used in both operational forecasting and models of tropical cyclone hazard risk.

KEYWORDS: Tropical cyclones; Hurricanes/typhoons; Statistical techniques; Risk assessment

1. Introduction

The radius of maximum wind (R_{max}) in a tropical cyclone (TC) principally determines the location and areal extent of storm hazards, including extreme wind and coastal rainfall (Lonfat et al. 2007; Lu et al. 2018; Xi et al. 2020) and storm surge (Penny et al. 2021; Irish et al. 2008; Irish and Resio 2010). Hence, a simple first-order prediction of R_{max} has significant value for both operational forecasting and risk assessment. However, R_{max} is a noisy quantity and difficult to observe directly in real storms due to the turbulent nature of the moist convective inner-core boundary layer (Shea and Gray 1973; Kossin et al. 2007; Sitkowski et al. 2011; Kepert 2017; Stern et al. 2020). The R_{max} is also poorly resolved by most numerical weather prediction models, especially climate models, as low horizontal resolution acts to smooth the inner-core structure radially outward (Reed and Jablonowski 2011; Gentry and Lackmann 2010; Rotunno and Bryan 2012). In contrast, the outer storm circulation is relatively quiescent and hence less variable in space and time (Frank 1977; Cocks and Gray 2002; Chavas and Lin 2016). As a result, a measure of the size of the broad outer circulation ("outer

size"), such as the radius of 34-kt (1 kt $\approx$ 0.51 m s^{-1}) wind, is much easier to resolve, less sensitive to turbulence, and is predictable operationally (Knaff and Sampson 2015). Could we use information about outer size to help predict R_{max}?

To do so, we use storm structure to link outer size to R_{max}. Fundamentally, R_{max} is strongly dependent on the outer circulation because the source of angular momentum at R_{max} is inward advection from larger radii (Palmén and Riehl 1957). Moreover, because surface friction removes angular momentum from air parcels as they spiral inward toward R_{max}, angular momentum must gradually decrease moving inward toward R_{max}. Chavas et al. (2015) developed a physical model for the complete radial structure of the low-level angular momentum distribution and, in turn, the cyclonic wind field (we refer to this model hereafter as "C15"). Chavas and Lin (2016) demonstrated that the C15 model predicts the qualitative dependence of R_{max} on key parameters that are found in observations: R_{max} tends to be larger for storms that are weaker (smaller V_{max}), whose outer circulation is larger, or that are farther poleward. This result complements preceding empirical work showing similar dependencies on intensity and latitude (Kossin et al. 2007; Knaff et al. 2015). Hence, the C15 model

---

Publisher's Note: This article was revised on 10 May 2022 to correct Fig. 2, which included an incorrect panel (a) when originally published. The first sentence of the Data Availability Statement has also been revised to include an updated DOI link.

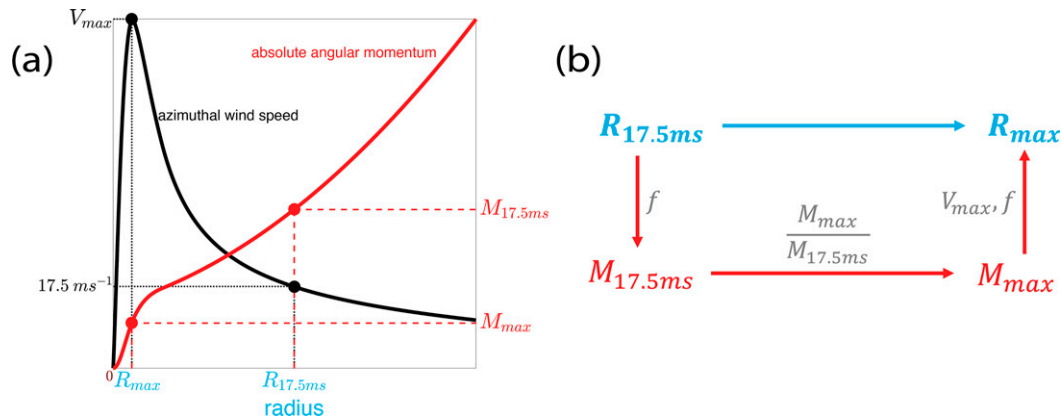---

Corresponding author: Daniel R. Chavas, drchavas@gmail.com

Fig. 1. (a) Example radial structure of azimuthal wind speed (black) and absolute angular momentum (red), with $R_{17.5\,ms}$ and R_{max} denoted in blue. (b) Conceptual framework of our model for predicting R_{max} from $R_{17.5\,ms}$ via the radial structure of absolute angular momentum [Eqs. (2)–(4)].

appears to capture the fundamental first-order physics linking outer size to R_{max}. The structural model also has been used to explain the relationship between the TC minimum central pressure and maximum wind speed (Knaff and Zehr 2007; Courtney and Knaff 2009; Chavas et al. 2017). Moreover, the C15 model explicitly identifies a small number of physical parameters that govern how R_{max} depends on V_{max}, latitude, and outer size.

In principle, the C15 wind structure model could be used to directly predict R_{max} from outer size (e.g., Davis 2018). However, the C15 model is idealized and thus has biases relative to observations in representing the relationships between radii (Chavas et al. 2015; Chavas and Lin 2016). In contrast, one may predict R_{max} using purely empirical regression models based on observed quantities such as storm intensity or satellite-derived quantities based on visible, infrared, or microwave imagery (Mueller et al. 2006; Willoughby et al. 2006; Kossin et al. 2007; Vickery and Wadhera 2008; Knaff et al. 2015). However, a purely empirical approach ignores valuable information encoded in the physics of storm structure described in the C15 model that can help relate R_{max} to other, larger wind radii.

An ideal alternative is a hybrid model that takes advantage of both observations and physical theory simultaneously. One way to do this is to exploit the theoretical basis for the C15 structure model, rather than the model itself, in the design of an empirical regression model fit to observations. This approach would use C15 physical theory to dictate model predictors and predictands but then use observations to define their true relationship in nature. Moreover, such a model may potentially be applied to predict how R_{max} may change in a future climate state, where no observational data are available.

Given the above knowledge gap, the principal objectives of this work are:

1) To develop a simple predictive model for R_{max} from outer size (here the radius of 34-kt wind) that exploits the benefits of both observations and physical theory,
2) To estimate model parameters from historical tropical cyclone data in the North Atlantic basin,

3) To evaluate the model against the observational record, and
4) To compare model performance against existing predictive models for R_{max}.

In pursuing these objectives, our primary goal is to develop a simple model that can both predict R_{max} for TC applications across research, forecasting, and risk analysis, and also help explain the underlying physics linking R_{max} to outer size (Shmueli et al. 2010).

Section 2 describes the methodology. Section 3 presents the results of our model and comparison with other existing predictive models. Finally, section 4 summarizes our findings and discusses avenues for future work.

2. Methodology

We begin with a conceptual overview of our model design. We then describe the theory for the physical component of the model. Finally, we describe the datasets and regression methodology used to define the final predictive model.

For our purposes, we focus in this paper exclusively on the maximum extent of the radius of 34-kt wind (17.5 m s^{-1}; hereafter $R_{17.5\,ms}$) as our measure of outer size. The $R_{17.5\,ms}$ along with the intensity (V_{max}), the maximum sustained 1-min wind, are routinely estimated in operations (OFCM 2020). The $R_{17.5\,ms}$ is also typically located in the outer circulation where convection is minimal (with the exception of very weak storms) and thus it tends to covary minimally with V_{max} (Merrill 1984; Chavas and Lin 2016). However, this approach can in principle be applied to any wind radius if given appropriate data (e.g., $R_{8\,ms}$; Schenkel et al. 2017).

a. Conceptual overview

A conceptual diagram of our model is shown in Fig. 1. A TC is an overturning circulation in which boundary layer air parcels at larger radii flow radially inward toward the center in the presence of background rotation (Riehl 1950; Wing et al. 2016). The absolute angular momentum (hereafter simply "angular momentum") of an air parcel is given by

$$M = rV + \frac{1}{2}fr^2, \quad (1)$$

where r is radius, V is the azimuthal wind speed, $f = 2\Omega\sin\phi$ is the Coriolis parameter, $\Omega = 7.292 \times 10^{-5}$ s^{-1} is Earth's rotation rate, and ϕ is the latitude of the storm center. The first term is the relative angular momentum associated with the storm circulation. The second term is the planetary angular momentum associated with the projection of Earth's rotation onto the TC's axis of rotation (vertical) at the latitude of the storm center[1]. Given the value of the Coriolis parameter f, the radial structure of angular momentum can be directly translated to the radial structure of the azimuthal wind via Eq. (1) (Fig. 1a).

As air parcels spiral inward from larger radii toward R_{max}, they gradually lose angular momentum due to surface friction (Fig. 1a). Hence M decreases moving radially inward (this is also necessary in order to be inertially stable). Thus, we may understand the relationship between R_{max} and $R_{17.5ms}$ in terms of the fraction of absolute angular momentum that has been *lost* between $R_{17.5ms}$ and R_{max}: $M_{max}/M_{17.5ms}$. This quantity will always be less than or equal to one. Given values for $R_{17.5ms}$ and ϕ (with $V_{17.5ms} = 17.5$ m s^{-1}), we can define $M_{17.5ms}$. Given a value of $M_{max}/M_{17.5ms}$, we can calculate M_{max}. Finally, given a value of V_{max} and ϕ we can solve for R_{max}. The quantities V_{max}, $R_{17.5ms}$, and ϕ are already estimated in operations every 6 h (OFCM 2020). These steps are summarized conceptually in Fig. 1b.

Mathematically, our model is defined as follows:

1) Calculate $M_{17.5ms}$: (Inputs: $R_{17.5ms}$ and f)

$$M_{17.5ms} = R_{17.5ms} \times (17.5 \text{ m s}^{-1}) + \frac{1}{2}fR_{17.5ms}^2. \quad (2)$$

2) Calculate M_{max}: (Inputs: $M_{17.5ms}$)

$$M_{max} = \left(\frac{M_{max}}{M_{17.5ms}}\right)M_{17.5ms}. \quad (3)$$

3) Solve for R_{max} from M_{max}: (Inputs: V_{max} and f)

$$R_{max} = \frac{V_{max}}{f}\left(\sqrt{1 + \frac{2fM_{max}}{V_{max}^2}} - 1\right). \quad (4)$$

The only missing information in the model is an estimate of $M_{max}/M_{17.5ms}$ in Eq. (3), which we describe in section 2b.

Note that it is important to average across quadrants rather than considering individual quadrants, for two reasons. First, the air at R_{max} is expected to be a mix of inflowing air parcels from all azimuths. This is because air at R_{max} and $R_{17.5ms}$ within the same quadrant are not directly linked: air parcels

complete one rotation around the storm center very quickly at R_{max} [~1 h for $(R_{max}, V_{max}) = (30$ km, 50 m s$^{-1})$] and much more slowly at $R_{17.5ms}$ (~1 day for $R_{17.5ms} = 300$ km). Second, averaging quadrants reduces noise associated with azimuthal asymmetries in storm structure (e.g., Uhlhorn et al. 2014). More practically, operational estimates of R_{max} are currently not quadrant specific. Quadrant-specific data could in principle be used to model the relationship, but this would require detailed information about the provenance of air parcels at R_{max}, which is quite complex to ascertain and not readily available.

As an aside, the above Lagrangian perspective does not imply that the radius of maximum wind itself follows a material M-surface. Indeed, radial inflow remains strong at R_{max} itself, and the precise location of R_{max} is not easily ascertained but rather emerges from the interplay between the radial distributions of the inward flux of angular momentum and the loss of angular momentum due to surface friction (Chen and Chavas 2020).

b. Model for $M_{max}/M_{17.5ms}$

The C15 model provides a physical model for the complete low-level radial structure of M that combines the solution of Emanuel and Rotunno (2011) for the inner deep-convecting region and the solution of Emanuel (2004) for the outer nonconvecting region (see Fig. 1a for an example solution). The C15 model does not have a true analytic solution, as it requires two fast numerical integrations to solve for the outer solution and to match it to the inner solution. As noted earlier, Chavas and Lin (2016) showed that this model can capture the characteristic modes of variability in the wind field found in nature.

The C15 model predicts that $M_{max}/M_{17.5ms}$ depends only on two physical parameters: V_{max} and $(1/2)fR_{17.5ms}$. The second parameter is a velocity scale that combines information about outer size ($R_{17.5ms}$) and storm center latitude (f).[2] Figure 2a displays the C15 model prediction for $M_{max}/M_{17.5ms}$ as a function of V_{max} and $(1/2)fR_{17.5ms}$. For this calculation, following C15, in the inner region we set the ratio of surface exchange coefficients $C_k/C_d = 1$ constant, and in the outer region we set the drag coefficient $C_d = 0.0015$ constant and the radiative-subsidence rate $w_{cool} = 2$ mm s^{-1} constant; C_d and w_{cool} principally modulate the structure of the far outer circulation and so have a minimal effect, while C_k/C_d is uncertain in nature and is commonly set to a value near one (e.g., Tang and Emanuel 2012). The $M_{max}/M_{17.5ms}$ monotonically decreases, indicating a greater loss of angular momentum, when moving from bottom left to top right in the figure, i.e., for higher intensities (V_{max}) and for higher latitude and/or outer size $[(1/2)fR_{17.5ms}]$. The lone exception is at very high V_{max} and

---

[1] The planetary angular momentum term can be written as $(\Omega\sin\phi r) \times r$, where the term in parentheses is the azimuthal velocity of Earth's rotation projected onto the local vertical. Thus, even at the outer edge of the storm where the circulation vanishes ($V = 0$), an air parcel still has non-zero absolute angular momentum (except if the storm center is on the equator).

[2] Theoretically, the second parameter should also be multiplied by the quantity C_d/w_{cool}, which is the ratio of the drag coefficient to the clear-sky free tropospheric subsidence rate due to radiative cooling outside of the convective inner core. This term can be neglected though as it may be taken as approximately constant from storm to storm; this assumption was also made for predicting the minimum central pressure in Chavas et al. (2017).
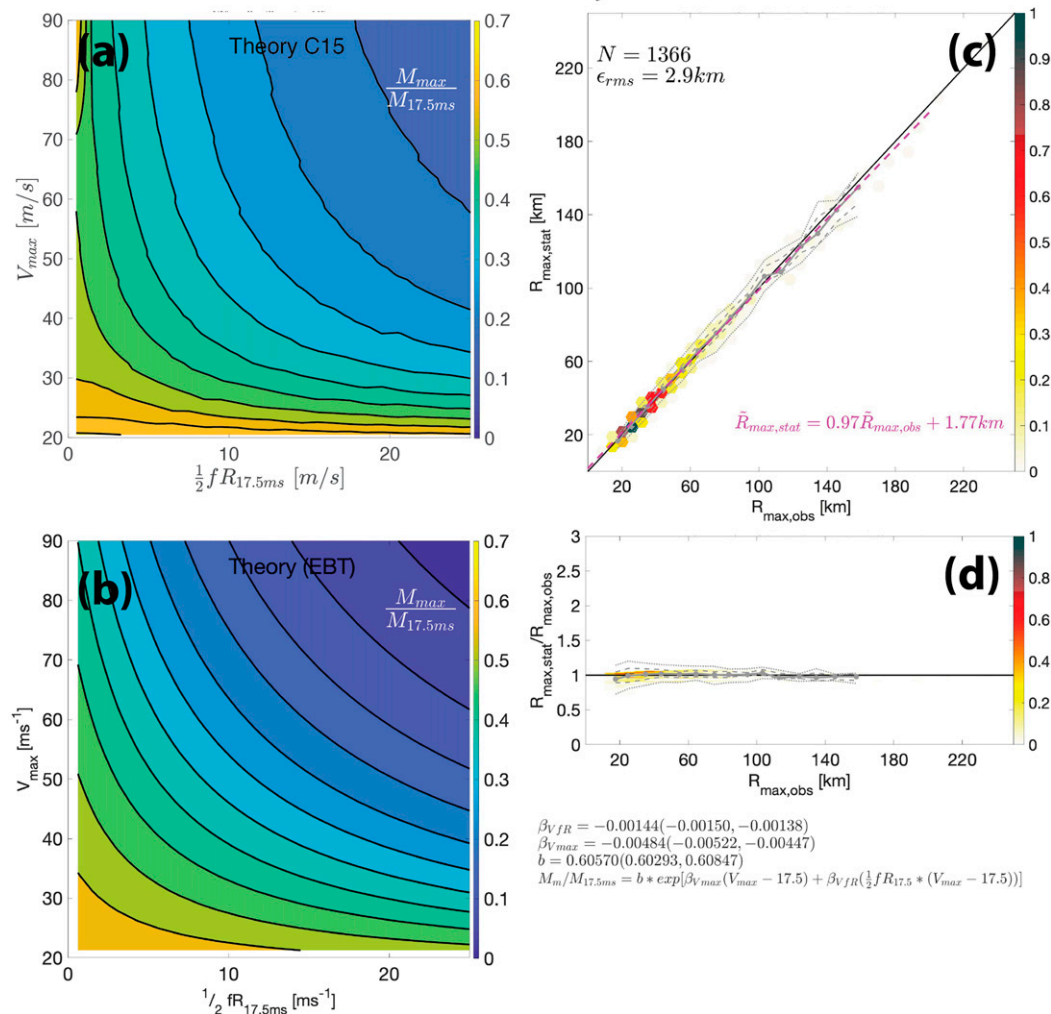
Theory

FIG. 2. Theory and empirical model prediction fit to theory. (a) Theoretical distribution of angular momentum loss fraction, $M_{max}/M_{17.5\,ms}$, as a function of two velocity scales, V_{max} and $(1/2)fR_{17.5\,ms}$, given by the TC wind structure model of Chavas and Lin (2016). (b) Empirical model prediction of $M_{max}/M_{17.5\,ms}$ [Eq. (6)] fit to the theory shown in (a); fitting is performed by first interpolating the theory of (a) to the values of $[V_{max}, (1/2)fR_{17.5\,ms}]$ in the EBT dataset for closest analog to observations. (c) Empirical model prediction of R_{max} (y axis) vs "observed" R_{max} (x axis) from Eqs. (2)–(4) using the model [Eq. (6)] shown in (b); color = relative frequency (i.e., sample size in hexagon divided by largest sample size value on plot); black line = 1-to-1 line; gray solid/dashed/dotted lines denote conditional median, interquartile range, and 5%–95% range, respectively, within 10-km bins of the observed value starting from 0; and pink dashed line + equation = linear regression of the conditional median prediction ($\tilde{R}_{max,stat}$) vs observed ($\tilde{R}_{max,obs}$; gray dots). (d) As in (c), but for fractional error relative to the known value. Deviation of the slope of the pink line from the black 1-to-1 line in (c) indicates systematic bias.

very small $(1/2)fR_{17.5ms}$ (top-left corner of the plot) where the angular momentum loss fraction can actually increase with increasing intensity; we return to this briefly below.

The theory (Fig. 2a) can in principle be used directly to model $M_{max}/M_{17.5ms}$. However, doing so will inherently incorporate any biases in the theoretical model structure relative to that found in real storms in nature. Indeed, C15 showed that the model captures the first-order radial structure, but it also has nonnegligible biases, particularly an over-

estimation of wind speeds at intermediate radii beyond R_{max}. Moreover, C15 theory lacks a simple analytic solution and must be solved numerically, which makes it less practical for everyday use.

Instead, we use the underlying theoretical basis of 15 theory in the design of a simple empirical model for $M_{max}/M_{17.5ms}$. We choose a log-link linear regression model for $M_{max}/M_{17.5ms}$ that depends on V_{max} and $(1/2)fR_{17.5ms}$ as follows:

$$\ln\left(\frac{M_{\max}}{M_{17.5\,\text{ms}}}\right) = \beta_0 + \beta_{V\max}(V_{\max} - 17.5\,\text{m s}^{-1})$$

$$+ \beta_{VfR}(V_{\max} - 17.5\,\text{m s}^{-1})\left(\frac{1}{2}fR_{17.5\,\text{ms}}\right) + \epsilon, \quad (5)$$

where $\beta_{V\max}$ and β_{VfR} are the regression coefficients to $(V_{\max} - 17.5\,\text{m s}^{-1})$ and $(V_{\max} - 17.5\,\text{m s}^{-1})[(1/2)fR_{17.5\text{ms}}]$, respectively; β_0 is a constant; and ϵ is the model residual error.

We can solve Eq. (5) for $M_{\max}/M_{17.5\text{ms}}$ to yield our final model equation:

$$\frac{M_{\max}}{M_{17.5\,\text{ms}}} = b\exp\left[\beta_{V\max}(V_{\max} - 17.5\,\text{m s}^{-1})\right.$$

$$\left. + \beta_{VfR}(V_{\max} - 17.5\,\text{m s}^{-1})\left(\frac{1}{2}fR_{17.5\,\text{ms}}\right)\right], \quad (6)$$

where $b = e^{\beta_0}$ and we drop the error term. Model coefficients b, $\beta_{V\max}$, and β_{VfR} will be estimated from observational data. Coefficient estimation is performed using the MATLAB function "fitglm" using the link option "log."

We choose a log-link model (natural logarithm on the left hand side) because $M_{\max}/M_{17.5\text{ms}}$ is a positive definite quantity, which an exponential model always reproduces but is not guaranteed when using standard linear regression. We choose the intensity predictor to be $(V_{\max} - 17.5\,\text{m s}^{-1})$ and the $(1/2)fR_{17.5\text{ms}}$ predictor to be multiplied by $(V_{\max} - 17.5\,\text{m s}^{-1})$ to better capture the nonlinear dependence on the two parameters that can be seen visually in Fig. 2a: the sensitivity of $M_{\max}/M_{17.5\text{ms}}$ to $(1/2)fR_{17.5\text{ms}}$ increases moving from lower to higher intensity (larger $V_{\max}$); at $V_{\max} - 17.5\,\text{m s}^{-1}$, this sensitivity is zero because $(M_{\max}/M_{17.5\text{ms}}) = 1$ by definition regardless of the value of $(1/2)fR_{17.5\text{ms}}$. The latter is a fundamental constraint that is not specific to the C15 theory. The form of our model explicitly builds in this nonlinear dependence. These assumptions allow us to incorporate the underlying physics constraining $M_{\max}/M_{17.5\text{ms}}$ into our empirical model.

We first demonstrate that Eq. (6) is appropriate for modeling $M_{\max}/M_{17.5\text{ms}}$. This is accomplished by applying Eq. (6) to C15 theory itself shown in Fig. 2a, and then predicting $R_{\max}$. This allows us to compare the empirical predictions for $M_{\max}/M_{17.5\text{ms}}$ and $R_{\max}$ with the "true" values given by the theory. We emphasize that the theory is *not* the truth, but rather provides a reasonable first-order representation of the relationship among our structural parameters. This offers a first test of our empirical model in which all parameters are known. To fit Eq. (6), we calculate $M_{\max}/M_{17.5\text{ms}}$ by interpolating the theory to the values of $[V_{\max}, (1/2)fR_{17.5\text{ms}}]$ in the Extended Best Track dataset discussed below; this provides the closest analog to observations. Figure 2b displays the empirical prediction of $M_{\max}/M_{17.5\text{ms}}$ [Eq. (6); regression coefficients are listed in Table 2]. Equation (6) can closely reproduce the relatively complex structure found in the theory. Figure 2c compares the final empirical model prediction of $R_{\max}$ to the theoretical value predicted by the C15 model, and Fig. 2d displays fractional errors as a function of the known $R_{\max}$. Systematic bias is defined as the slope of the linear

regression of the conditional median predicted value ($\tilde{R}_{\max,\text{stat}}$) versus observed value ($\tilde{R}_{\max,\text{obs}}$); a slope of 1 indicates no systematic bias. These results demonstrate that our empirical model can reliably reproduce the "true" $R_{\max}$ with very small error across all values of $R_{\max}$, spanning a range from 10 to over 200 km, and with nearly zero systematic bias (pink dashed line, Fig. 2c; linear regression slope of 0.97). Note that the empirical model does not reproduce the increase in $M_{\max}/M_{17.5\text{ms}}$ at very high $V_{\max}$ and very small $(1/2)fR_{17.5\text{ms}}$ in the theory. Overall, this outcome indicates that the form of our empirical model for $M_{\max}/M_{17.5\text{ms}}$ is well suited for the task and may be applied to real data below.

We also tested two other forms of log-link regression model: 1) linear model with two predictors, $(V_{\max} - 17.5\,\text{m s}^{-1})$ and $(1/2)fR_{17.5\text{ms}}$; and 2) nonlinear model with three predictors $(V_{\max} - 17.5\,\text{m s}^{-1})$, $(1/2)fR_{17.5\text{ms}}$, and $(V_{\max} - 17.5\,\text{m s}^{-1})$ $(1/2)fR_{17.5\text{ms}}$. The linear model is viable, as it can capture the first-order pattern of monotonic decrease from bottom left to top right (online supplemental Fig. S01b), but it exhibits a larger systematic bias (supplemental Fig. S01c). This larger bias arises because by definition it cannot capture the nonlinear dependence on the predictors and so it misses the detailed structure in $M_{\max}/M_{17.5\text{ms}}$ at low intensities. Thus, the added complexity of Eq. (6) compared to a linear model is both valuable for explaining variations in $M_{\max}/M_{17.5\text{ms}}$ and useful for making predictions (Shmueli et al. 2010). Results from the linear model will be included in the discussion below for context and comparison. In contrast, the three-predictor nonlinear model is rejected because it is more complex than Eq. (6) but does not perform better in terms of error or bias. Moreover, while it can also reproduce the theoretical distribution of $M_{\max}/M_{17.5\text{ms}}$, when it is applied to observations it produces a qualitatively different structure that is concave down rather than concave up because the coefficient of the nonlinear term is found to be positive rather than negative. Thus, in the context of the results for both the linear and two-variable nonlinear model, the added complexity of the three-variable model is at best unhelpful and may in fact be detrimental.

To summarize, our model predicts $R_{\max}$ from $R_{17.5\text{ms}}$ via Eqs. (2)–(4) and (6), whose coefficients are estimated from data in the Results section below [Eq. (7)]. An optional final bias adjustment is given by Eq. (8) below as well. The physical basis of our model offers three key advantages: 1) Choosing angular momentum loss fraction as the predictand constrains the model (in theory) to be a value the range (0, 1]; 2) Theory indicates that this predictand should decrease smoothly and monotonically toward zero; 3) This monotonic dependence can be reduced to two physical predictors. We use these advantages to define the form of our empirical model. Ultimately, the empirical basis of our model acknowledges that we do not fully understand the details of how angular momentum is lost in the inflow approaching $R_{\max}$; it is undoubtedly more complex than idealized theory, and the representation and implications of these complexities remains highly uncertain. Hence, using data to directly estimate the true dependence allows us to capture the final outcome as found in nature despite the current gaps in our understanding.

c. Observational data

We test our model against data for the North Atlantic basin from the Extended Best Track v2021–03–01 (Demuth et al. 2006, hereafter EBT) which is by far the longest and most widely used wind structure database available. The EBT dataset provides 6-hourly estimates of storm location, storm intensity, storm type, wind radii (including $R_{17.5ms}$) in four quadrants (northeast, northwest, southeast, southwest), and R_{max} (single value) for the life cycle of each storm during the period 1988–2020 in the North Atlantic. We calculate $R_{17.5ms}$ as the mean of all available nonzero values in each quadrant then multiplied by a factor of 0.85. This reduction factor accounts for the fact that $R_{17.5ms}$ is defined operationally as the outermost value within each quadrant, and hence the mean value has been estimated to be approximately 15% smaller (DeMaria et al. 2009). For 2004 onward, all data except R_{max} are final best track data that are reanalyzed by the National Hurricane Center post-storm (Landsea and Franklin 2013); prior to 2004, wind radii data are not reanalyzed post-storm and are simply taken from the Automated Tropical Cyclone Forecast (ATCF; Sampson and Schrader 2000). All R_{max} data are subjective forecaster estimates based on available aircraft and remotely sensed data in near–real time and are not reanalyzed post-storm. The 2021–03–01 version of EBT replaces some advisory-based data contained in the a-deck CARQ (Combined Automated Request Query) entries with superior best track data contained in the b-deck (G. Chirokova 2021, personal communication). Our results presented below are nonetheless quantitatively similar when using the preceding version with data only through 2018.

To provide the optimal data subset for our work, we use the years 2004–20, corresponding to the period in which the wind radii are best tracked (Knaff et al. 2021). Moreover, we focus our analysis in the western half of the North Atlantic basin to limit ourselves to the subset of storms that were of immediate interest to forecast agencies and hence were most likely to have garnered dedicated attention from forecasters and reconnaissance (Demuth et al. 2006). Specifically, we filter the dataset as follows: 1) storm center longitude < 50°W (to focus on the dominant aircraft reconnaissance region); 2) ϕ < 30°N (to minimize effects from extratropical transition); 3) storm center distance from coast ≥ $R_{17.5ms}$ (to minimize land effects on inner-core structure); 4) at least 3 valid quadrant values of $R_{17.5ms}$ (to ensure a reasonable estimate of the azimuthal-mean value); 5) $0 < M_{max}/M_{17.5ms} \leq 1.1$ (to remove unphysical values while allowing for values slightly above 1 that may be due to noise); 5) 6 km ≤ R_{max} ≤ 250 km (to remove extreme outliers); and 6) $V_{max} \geq 20$ m s^{-1} (to retain storms with intensities above 17.5 m s^{-1}). All data are converted to MKS units. The distance from the coast is taken directly from the EBT database, which is defined using a coastline that includes all large islands (the Leeward and Windward Islands and the Azores are excluded). Figure 3a displays a map of the resulting data subset, and Figs. 4a–c displays histograms of V_{max}, R_{max}, and $R_{17.5ms}$. We do not remove any translation speed effect on V_{max}, though we tested the method of Lin and Chavas (2012) which had
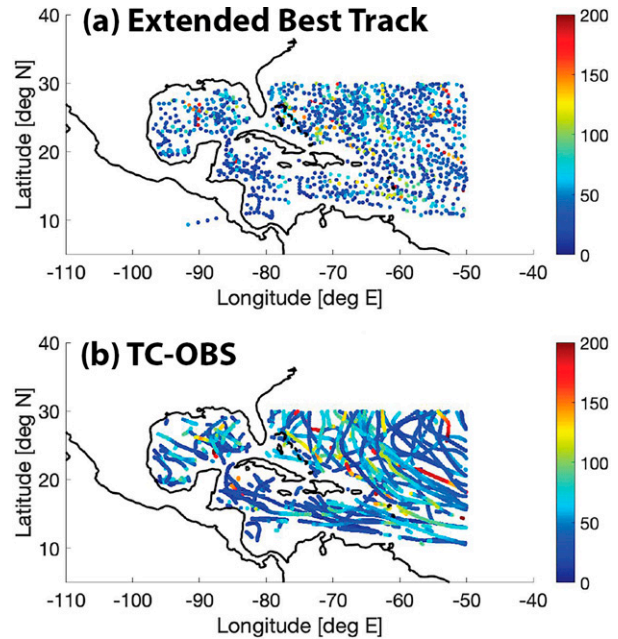

FIG. 3. Maps of each dataset used in this study; color denotes R_{max} (km). Raw datasets are (a) Extended Best Track (NOAA CIRA) and (b) TC-OBS (v0.40; Vigh et al. 2016). Data filters are listed in the text.

minimal effect on the final model structure. How best to represent this effect is uncertain and it can yield odd outcomes for weak, fast-moving storms, and hence we elect to avoid it altogether. The final sample size is $N = 1366$. This sample size is much larger than the number of degrees of freedom in our model (three: the two predictor velocities and the initial angular momentum $M_{17.5ms}$), which greatly limits the potential for overfitting.

To add robustness to the choice of dataset, we perform an identical analysis with the same filters using the Tropical Cyclone Observations-Based Structure Database v0.40 (TC-OBS; Vigh et al. 2016) for the North Atlantic. TC-OBS merges aircraft and satellite data with Extended Best Track data to provide a more objective and observationally constrained estimate of TC location, intensity, and sustained near-surface wind radii and R_{max} at hourly temporal resolution. The years 2004–14 are used in our analysis to align with the start of best tracking of wind radii in EBT; 2014 is the final year of the latest version of the database. For storm intensity and storm central latitude, we simply use the linearly interpolated values from the Best Track database (variables "BT_Vmax_interp" and "BT_lat_interp"). For storm structure, $R_{17.5ms}$ is calculated as the mean of all available nonzero values in each quadrant (variable "TCOBS_wind_radii"), while near-surface R_{max} has a single value (variable "TCOBS_maximum_sustained_surface_wind_radius"). When aircraft data are present, all wind radii and R_{max} are an optimally weighted blend of 1) the radius of the slant-adjusted flight level wind observations, 2) the radius estimated from the Stepped Frequency Microwave Radiometer
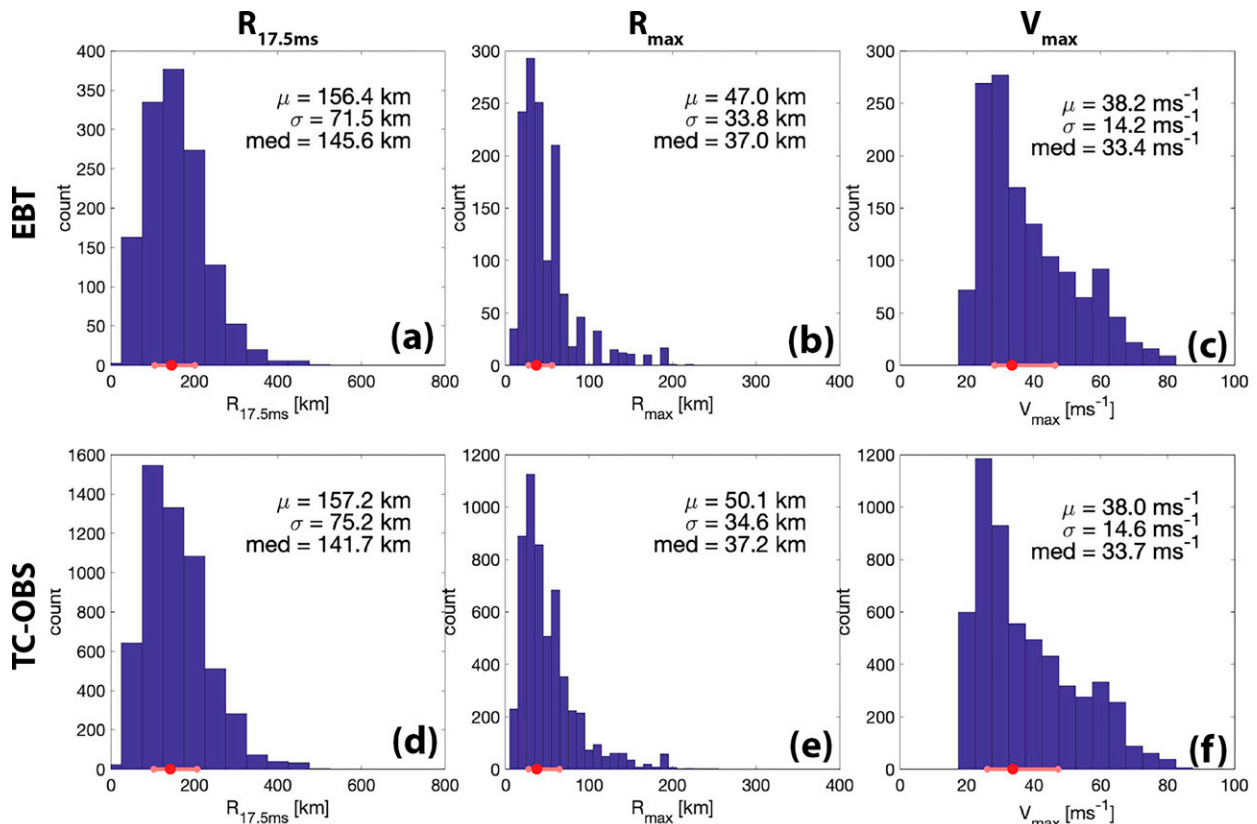
FIG. 4. Histograms of $R_{17.5ms}$, R_{max}, and V_{max} from (a)–(c) EBT and (d)–(f) TC-OBS. Dots and bars along the x axis denote the median and interquartile (25th–75th percentile) ranges.

(SFMR; Uhlhorn et al. 2007), and 3) the radius of maximum wind value from b-deck file or wind radius from the EBT database. When no nearby aircraft data are available, this parameter is relaxed toward the interpolated b-deck value (for R_{max}) or EBT wind radius value with an e-folding time scale of 4 h. Figure 3b displays a map of the resulting data subset ($N = 5574$), and Figs. 4d–f display histograms of V_{max}, R_{max}, and $R_{17.5ms}$. The TC-OBS subset is statistically similar to that of EBT.

Note that $R_{17.5ms}$ varies by more than a factor of 10 between smallest and largest values, whereas f is restricted to vary by a factor of less than 3 between 10° and 30°N. Hence, the parameter $(1/2)fR_{17.5ms}$ varies primarily through variations in $R_{17.5ms}$. Nonetheless, we apply our model to EBT data poleward of 30°N below to evaluate its performance for high-latitude cases as well.

d. Comparison and validation

To demonstrate the utility of our model we compare the performance of our model prediction of R_{max} against three existing predictive models used operationally: 1) the polynomial empirical model of Knaff et al. (2015) [their Eq. (1)], which depends only on V_{max} and latitude; 2) the multi-satellite-platform tropical cyclone surface wind analysis (MTCSWA; Knaff et al. 2011); and 3) a two-regime statistical

method based on storm-centered infrared (IR) data (IR-2R). We test all methods against observed EBT values for the 2018–20 Atlantic hurricane seasons (the 2020 values are preliminary); for our model, we first refit the model coefficients to the EBT dataset excluding the 2018 and 2019 seasons to ensure an out-of-sample test.

The MTCSWA is a real-time surface wind analysis that combines satellite atmospheric motion vectors below 600 hPa, oceanic vector winds from scatterometry, 2D balanced winds derived from microwave sounding instruments (Bessho et al. 2006), and an IR-based flight-level wind proxy (Knaff et al. 2015) as described in Knaff et al. (2011). The surface winds are estimated every three hours using all of the available data, and its R_{max} estimates are routinely available to operational centers. The routine availability and long history of MTCSWA provide a nice comparison with a currently available technique.

The IR-2R method has not been formally documented elsewhere, but has been running in real time since 2017. Here we briefly document how that baseline model was developed. The developmental data were divided into two regimes: 1) storms with intensities less than 33 m s^{-1}, and 2) storms with TC intensities ($\geq$33 m s^{-1}). The R_{max} estimates used for development (1995–2014) are based on a flight-level analysis of aircraft reconnaissance flights described in Knaff et al. (2015). The initial set of potential predictors are based on

TABLE 1. Regression coefficients and statistics related to the IR-2R two regime models. Regime 1 is for intensities less than 33 m s^{-1} and regime 2 for intensities above that threshold. Predictors are current intensity (V_{max}), sine of storm center latitude [sin(lat)], IR principal component 2 (PC2), IR-based estimate of TC size (radius of 5-kt wind, R_{5kt}), percent of pixels colder than $-50°$C (PC50), and PC50^2. The intercept (a) and regression coefficients are provided along with the percent variance explained R^2 and the mean absolute error (MAE) associated with the dependent fit (km).

Model	a	V_{max}	sin(lat)	PC2	R_{5kt}	PC50	PC50^2	$R^2 \times 100$	MAE
Regime 1	4.23	-9.1×10^{-3}	8.6×10^{-1}	1.8×10^{-1}				10	36.7
Regime 2	-6.67	9.9×10^{-3}			7.8×10^{-2}	7.7×10^{-2}	1.4×10^{-4}	49	24.8

current storm characteristics (V_{max} and latitude) and those derived from infrared satellite imagery that measure size, convective vigor, and radial structure. IR-based TC size is defined as an estimate of the radius of 5-kt wind (R_{5kt}) described in Knaff et al. (2014); convective vigor is measured by the percent of pixels colder than $-10°$, $-20°$, $-30°$, $-40°$, $-50°$, and $-60°$C calculated within 200 km of the TC center. The radial structure was estimated by azimuthally averaged principal components of the IR brightness temperature field, also described in Knaff et al. (2014), which provide radial wavenumbers 0–4. The regression models for both regimes are each multiple linear regression models whose predictors were determined using a leaps and bounds approach (Furnival and Wilson 2000) that systematically tests all possible regressions of 1, 2, 3, … variables to identify the equation with the best performance (variance explained). Variable selection is stopped when adding an additional predictor results in an increase in explained variance less than 2%. No independent verification or retraining was conducted, as this model is meant to be a baseline for the performance of other models. The resulting models are as follows: the regime 1 model retains a predictor related to current intensity, latitude, and principal component 2 (radial wavenumber 1). The regime 2 model is a function of current intensity, TC size (R_{5kt}), the percent pixels with brightness temperatures less than $-50°$C (PC50) and PC50^2. Both models are fit to the ln of the R_{max} and units are kilometers. The two regimes are blended together using linear weighting between intensities of 23 and 33 m s^{-1}. Table 1 provides the regression coefficients and statistics related to the linear fit.

3. Results

a. Model results

We now apply our model to the EBT dataset. Figure 5a displays the observed distribution of $M_{max}/M_{17.5ms}$ [Eq. (2)]. The $M_{max}/M_{17.5ms}$ tends to decrease toward higher values of V_{max}

and/or $(1/2)fR_{17.5ms}$, consistent with theory. Figure 5b shows the regression model [Eq. (6)] fit to the observed data, given by

$$\frac{M_{max}}{M_{17.5ms}} = 0.699 \exp\left[-0.00618(V_{max} - 17.5 \text{ m s}^{-1}) \right.$$
$$\left. - 0.00210(V_{max} - 17.5 \text{ m s}^{-1})\left(\frac{1}{2}fR_{17.5ms}\right)\right]. \quad (7)$$

The regression model coefficients are also provided in Table 2. The coefficients for the two predictors are negative, which provides quantitative confirmation that $M_{max}/M_{17.5ms}$ tends to decrease toward higher values of each predictor. The dependence on $(1/2)fR_{17.5ms}$ has wider uncertainties, which arises in part because $(1/2)fR_{17.5ms}$ has a smaller range of values than V_{max}, but also indicates more variability in the data. The structure of the dependence of the model fit to EBT (Fig. 5b) looks remarkably similar to that for the theory (Fig. 2b). Indeed, the EBT coefficients are qualitatively similar to theory, but they are quantitatively a bit different: this is precisely the bias in the theory that we avoid by using an empirical model with coefficients estimated directly from data.

Figure 5c compares the final prediction of R_{max} against the theoretical value, and Fig. 5d shows fractional errors as a function of the observed R_{max}. Model error and bias are presented in Table 3. The model can consistently capture the first-order variability in R_{max} over a wide range of values from 15 to 200 km. There is a slight systematic bias in the prediction of R_{max} (linear regression slope of 0.76) in which the model overestimates R_{max} at small values and underestimates it at large values, with the transition point at approximately 40 km (pink dashed line). Above this threshold, the systematic underestimation is roughly constant at approximately 20% of the observed R_{max}. Below this threshold, overestimation can increase to very large values as R_{max} becomes very small; this is unsurprising given that estimates in the observed R_{max} itself carry substantial uncertainty whose relative errors

TABLE 2. Regression coefficients for our log-link nonlinear regression model [Eq. (6)] fit to theory, Extended Best Track ($N = 1366$), and TC-OBS ($N = 5574$). The 95% confidence intervals are shown in parentheses. See text for details.

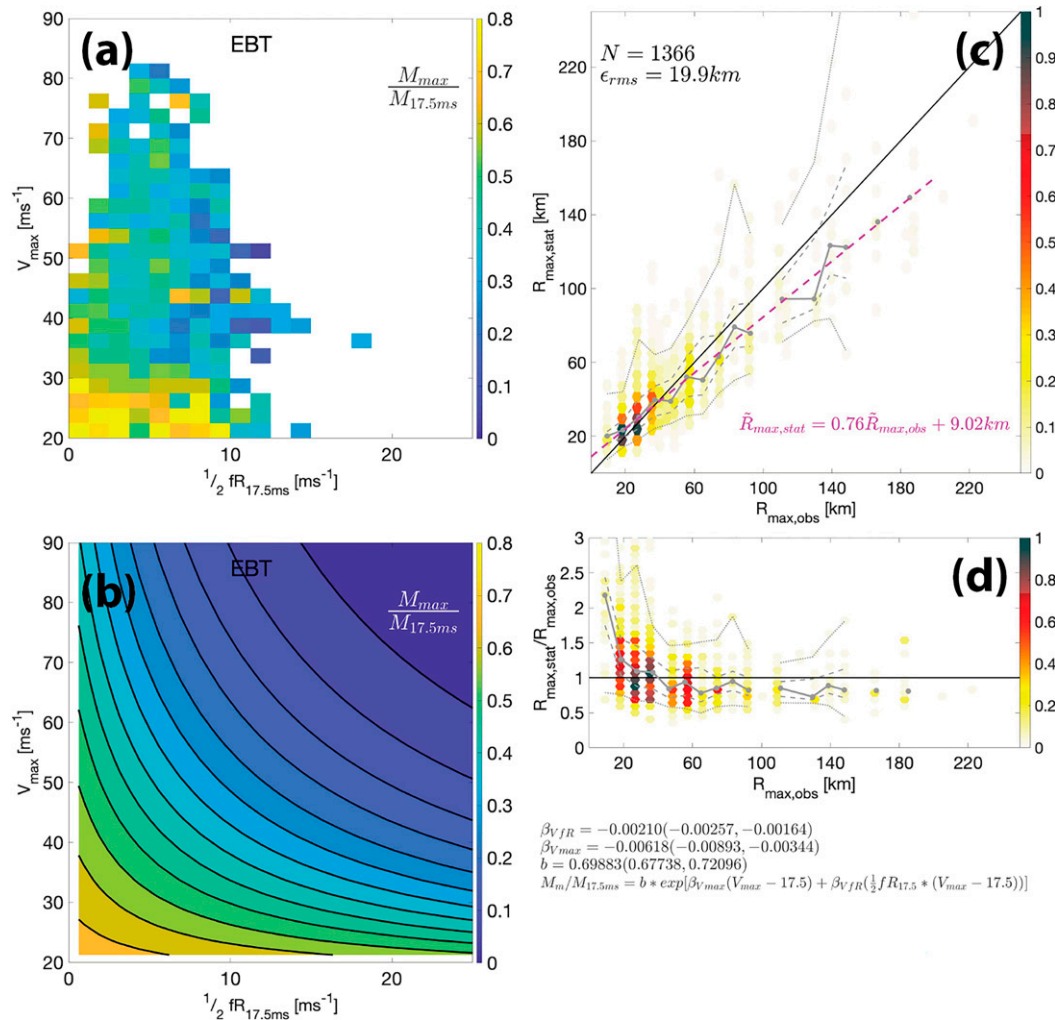| Coefficient	Theory	Extended Best Track	TC-OBS
β_{Vmax}	-4.84×10^{-3} (-5.22×10^{-3}, -4.47×10^{-3})	-6.18×10^{-3} (-8.93×10^{-3}, -3.44×10^{-3})	-6.57×10^{-3} (-7.89×10^{-3}, -5.25×10^{-3})
β_{VfR}	-1.44×10^{-3} (-1.50×10^{-3}, -1.38×10^{-3})	-2.10×10^{-3} (-2.57×10^{-3}, -1.64×10^{-3})	-1.84×10^{-3} (-2.06×10^{-3}, -1.62×10^{-3})
b	0.606 (0.603, 0.608)	0.699 (0.677, 0.721)	0.735 (0.724, 0.746)

Extended Best Track

FIG. 5. Empirical model prediction fit to Extended Best Track dataset for the North Atlantic 1988–2018; aesthetics as in Fig. 2. (a) Observed distribution of angular momentum loss fraction, $M_{max}/M_{17.5ms}$. (b) Empirical model prediction of $M_{max}/M_{17.5ms}$ fit to the data shown in (a). (c) Empirical model prediction of R_{max} (y axis) vs observed R_{max} (x axis) using the empirical model shown in (b). (d) As in (c), but for fractional error relative to the observed value.

will be magnified as R_{max} becomes very small. This uncertainty is at least in part due to the tendency to discretize operational estimates into 5 n mi bins (1 n mi = 1.852 km), though it may also indicate real physical variability within compact inner cores of TCs. Moreover, there is evidence of slight overestimation of R_{max} in the EBT (Combot et al. 2020).

TABLE 3. Comparison of model performance statistics. For systematic bias, a slope of one is unbiased.

Model	rmse (km)	Systematic bias (slope)
Our model [Eq. (7)]	19.9	0.76
Two-predictor linear log-link	18.3	0.66
Knaff et al. (2015) [Eq. (1)]	33.0	0.12

Table 3 also compares our model performance statistics against both the linear version of our model, with predictors $(V_{max} - 17.5\ \text{m s}^{-1})$ and $(1/2)fR_{17.5ms}$, as well as the Knaff et al. (2015) model. The nonlinear model yields a lower systematic bias than the linear model (supplemental Fig. S02c). As noted earlier, the linear model can capture the qualitative pattern of $M_{max}/M_{17.5ms}$ but the linear dependence on the predictors prevents it from capturing the nonlinear structure at lower intensities (supplemental Fig. S02b). Our model also performs substantially better than the model of Knaff et al. (2015), whose conditional best-fit slope is only marginally higher than zero, indicating strong systematic bias, due in part to the fact that it predicts a relatively limited range of variability in R_{max} between 25 and 70 km.
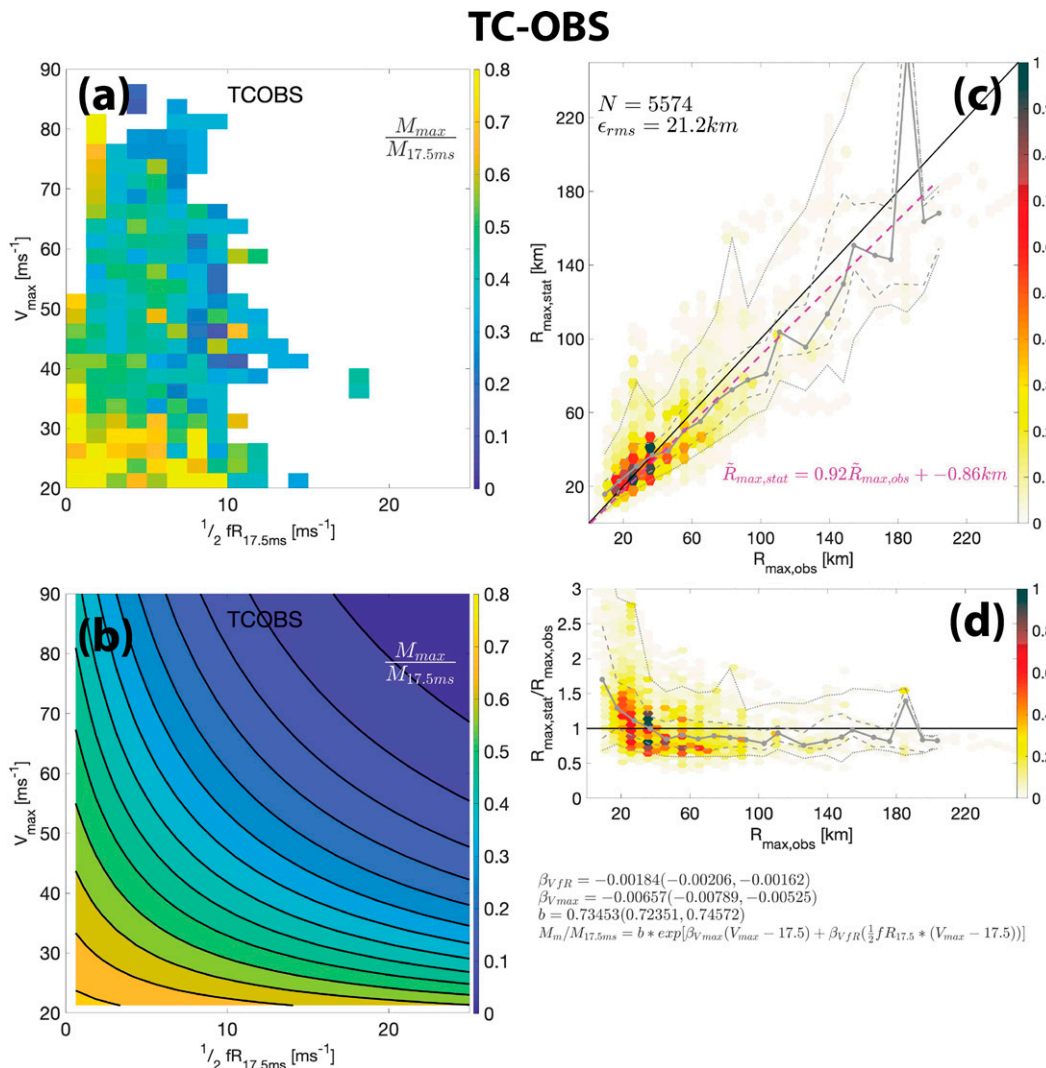
TC-OBS

FIG. 6. Empirical model prediction fit to TC-OBS dataset for the North Atlantic 2004–14; aesthetics as in Fig. 5.
Results are quantitatively similar to those for EBT shown in Fig. 5.

Quantitatively similar results to EBT are found in TC-OBS (Fig. 6). The empirical model structure and coefficient estimates are quantitatively similar to those of EBT (Table 2), and the final prediction of R_{max} yields similar errors and systematic biases. This result lends greater confidence in the robustness of our empirical modeling results.

One notable deviation from the empirical model is the apparent slight *increase* in $M_{max}/M_{17.5ms}$ moving toward very high intensities at small values of $(1/2)fR_{17.5ms}$ evident in both the EBT (Fig. 5a) and TC-OBS (Fig. 6a) databases. This behavior is not predicted by the empirical model presented here, but it does show up in a similar portion of the phase space in the C15 theory (Fig. 2a) as noted above. Physically, this behavior indicates that at very high intensity and for small and/or high-latitude storms, less angular momentum is lost to the surface between $R_{17.5ms}$ and R_{max} at higher intensity than

at lower intensity. More generally, this result may suggest an important change in the qualitative behavior of the boundary layer and its interaction with the surface, which may be due to e.g., a reduction in the drag coefficient at high wind speeds, which is a topic of ongoing debate (Richter et al. 2021; Richter and Stern 2014; Donelan et al. 2004). Here our analysis suggests such effects occur specifically at very small $(1/2)fR_{17.5ms}$. More in-depth interpretation of this result and its manifestation in C15 theory lies well beyond the scope of this work.

b. Application to a historical case

We next provide an example application of the model to Hurricane Michael (2018; Fig. 7), which formed at 1800 UTC 6 October 2018 and made landfall in Florida as a category-5 storm at 1800 UTC 10 October. Michael gradually and
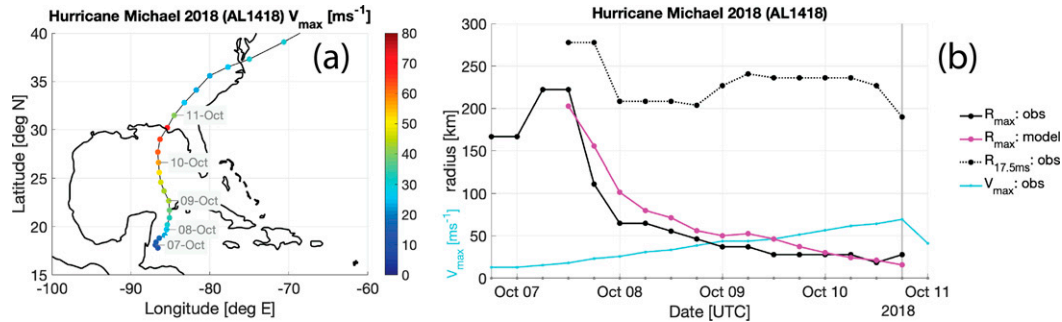
FIG. 7. Example application of our model to Hurricane Michael (2018). (a) Map of storm track with best track V_{max} (color dots); dates shown at 0000 UTC. (b) Predicted R_{max} (pink), EBT R_{max} (black solid), EBT $R_{17.5ms}$ (black dashed), and V_{max} (light blue) prior to landfall (light gray line).

monotonically intensified during its life cycle leading up to landfall (Fig. 7a). Figure 7b compares the observed R_{max} against the prediction from our empirical model. Michael's R_{max} was initially greater than 200 km and then decreased rapidly to 64 km over the 12-h period from 1200 UTC 7 October to 0000 UTC 8 October before decreasing gradually with time until landfall; this evolution is predicted well by our model. During the early 12-h period of rapid decrease in R_{max}, storm outer size ($R_{17.5ms}$) also decreased significantly from 278 to 208 km while storm intensity (V_{max}) increased from 18 to 26 m s^{-1}. Thus, R_{max} is predicted to decrease due to both of these effects occurring simultaneously: contraction of the inner core with intensification and shrinking of the storm as a whole. Thereafter, R_{max} decreases gradually with time as the storm gradually intensifies, whereas outer size ($R_{17.5ms}$) remains relatively constant with values between 208 and 241 km and f increases only modestly as Michael moves northward. Thus, R_{max} is predicted to decrease principally due to the increase in intensity. Note that R_{max} changes much more rapidly with changes in intensity for TCs at lower intensity, a behavior captured by our model.

c. Application to higher-latitude storms

It is useful to further evaluate how our model performs when applying it to data poleward of 30°N. Though we trained the model equatorward of 30°N to avoid the messier details associated with extratropical interactions, the theory is in principle valid at any latitude. Figure 8 shows the result of applying Eq. (7) to our EBT dataset filtered in the same way as above except now for the latitude band 30°–50°N. The model performs reasonably well overall given the greater uncertainties both observationally and theoretically at these higher latitudes, with slightly greater rmse (27.4 km) and systematic bias (conditional slope of 0.66). The transition from overestimate to underestimate occurs at an observed value of approximately 60 km, with median fractional errors toward higher observed values again remaining relatively constant at an underestimate of approximately 25%. These results suggest that the model is suitable for application at higher latitudes as well. The wider error at higher latitudes likely reflects the increased complexity of storm interaction with

baroclinic and land features, as well as subtropical storm formation, which can enhance storm asymmetries (Hart 2003; Guishard et al. 2009). Such asymmetries may induce biases in estimating the azimuthal-mean R_{17ms} particularly when approaching the mid-Atlantic and northeastern U.S. coastline where data may only exist on one side of the storm.

d. R_{max} versus latitude

Finally, we examine statistical changes in R_{max} and other storm parameters with latitude (Fig. 9). Observed median R_{max} increases with latitude, as has been noted in past studies, from 28 km south of 15°N to 74 km north of 35°N (solid black line). This behavior is quantitatively well captured by our model (solid pink line), including poleward of 30°N.

For the input parameters, median $R_{17.5ms}$ increases substantially with latitude (118 km south of 15 N; 222 km north of 40°N), whereas median V_{max} is nearly constant with latitude with a slight decrease from 33.4 m s^{-1} within 15°–30°N to 30.8 m s^{-1} moving north of 30°N. This suggests that R_{max} tends to increase statistically with latitude principally because $R_{17.5ms}$ increases significantly with latitude. We test this hypothesis explicitly using our model. First, our model can reproduce the increase in latitude in a simpler manner by applying it directly to the median values of each input parameter within each bin (pink X marks). This yields a similar increase with latitude; the values are systematically biased high relative to the true median within each bin, an indication of the nonlinearity of the problem. We then perform the same model prediction from median values but holding $V_{max} = 30$ m s^{-1} constant (pink triangles). The result is a very similar increase in R_{max} with latitude as before, indicating that the small decrease in V_{max} with latitude is not important for this behavior. Finally, we further hold f constant at its value at 20°N (pink circles) to fully isolate the effect from the increase in outer size alone. The result is only a modest effect on the trend with latitude, indicating that variations in f are also not important. Taken together, then, R_{max} increases with latitude predominantly because outer size ($R_{17.5ms}$) increases with latitude.
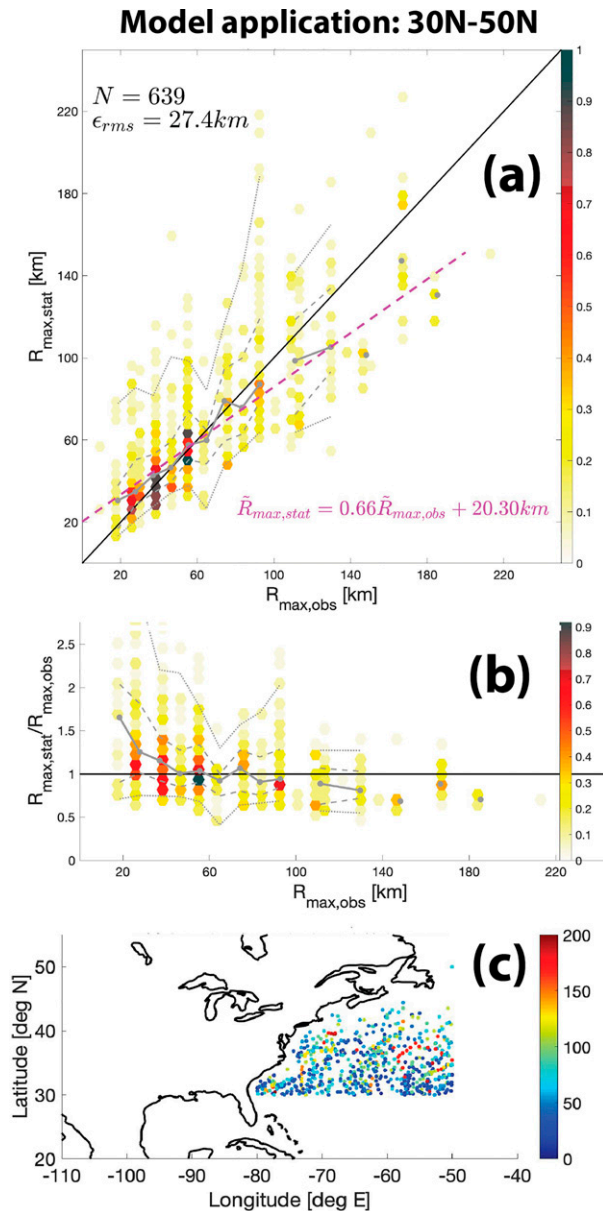
Model application: 30N-50N

FIG. 8. (a),(b) As in Figs. 5c,d, but for our model prediction of R_{max} [from Eq. (7)] applied to North Atlantic EBT data between 30° and 50°N. (c) Map of data points; color denotes R_{max} (km).

e. Comparison with existing predictive models

We next compare out-of-sample predictions from our model against three existing operational models: the Knaff et al. (2015) model, MTCSWA, and IR-2R. We use the EBT datasets for test period of 2018–20 using the same filters as were applied above. To provide a true out-of-sample test of our model, we refit our model coefficients to the EBT dataset excluding the years 2018–20. The resulting coefficients ($\beta_{V\max} = -0.006\,07$, $\beta_{VfR} = 0.001\,94$, $b = 0.688$) are very similar to the

values for the full dataset (Table 2). Because this sample size is much smaller, we define systematic bias using a simple linear regression fit to the predicted versus observed R_{max}.

Overall, our model performs substantially better than all three existing predictive models (Fig. 10). Our model has an out-of-sample rmse of 18.2 km and systematic bias slope of 0.70, comparable to the full-dataset results of Fig. 5. As found above for the full dataset (Table 3), the Knaff et al. (2015) model has relatively little predictive power. Meanwhile, MTCSWA and IR-2R both exhibit substantially larger rmse (37.4 and 29.4 km, respectively) and systematic biases (0.40 and 0.49, respectively) than our model.

Between the two remote sensing-based models, IR-2R does not predict R_{max} values much larger than about 120 km, though it does have a smaller error than MTCSWA. The MTCSWA analysis, on the other hand, has much larger predicted values though often associated with moderate rather than large values of R_{max} and hence yield large errors.

f. Optional final bias adjustment

For the purposes of prediction, we may take one final step and adjust our model prediction to remove the systematic bias. This is done by solving the equation for the linear regression on the conditional median (pink line, Fig. 5c) for $\tilde{R}_{max,obs}$, which we will denote here as our bias-adjusted final prediction $R_{max,statadj}$:

$$R_{max,statadj} = \frac{1}{0.76}(R_{max,stat} - 9.02\,\text{km}). \qquad (8)$$

Doing so eliminates the systematic bias (Fig. 11), though it slightly increases the overall rmse to 25.8 km. This outcome is expected: by definition, this multiplicative adjustment will increase the range of predicted values within each observational bin, and in particular will increase absolute errors in overestimated predictions more strongly than it will decrease the absolute errors in underestimated predictions. Such behavior is an example of the "bias-variance trade-off," in which adding this additional complexity to our model reduces bias but amplifies errors associated with noise in a dataset (Shmueli et al. 2010). Since the goal of our model is to provide a reasonable first-order prediction of R_{max} across a wide range of values, this bias adjustment may be desirable even at the expense of a modest increase in rmse.

Ultimately, this bias adjustment has no obvious physical explanation and is solely a means of improving the final prediction. It may be most useful for risk prediction models where the objective is to produce R_{max} values that are statistically similar to observations rather than make predictions of individual values directly from noisy observations.

4. Conclusions

The tropical cyclone radius of maximum wind (R_{max}) is critical for estimating the magnitude and footprint of wind, surge, and coastal rainfall hazards. However, R_{max} is noisy in observations and poorly resolved in reanalyses and climate models. In contrast, wind radii from the outer circulation are much
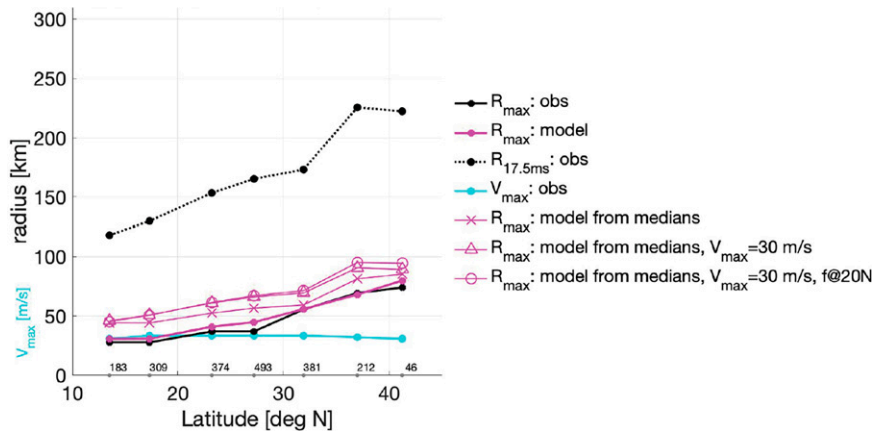
FIG. 9. The model predicts the statistical increase in R_{max} with latitude found in observations. Median values of predicted R_{max} (pink), EBT R_{max} (black solid), EBT $R_{17.5ms}$ (black dashed), and V_{max} (light blue) vs median latitude in 5° latitude bins; lowest latitude bin includes all data south of 15°N and highest latitude bin includes all data north of 40°N. Also shown, model prediction direct from bin-median values of latitude, V_{max}, and $R_{17.5ms}$ (pink X); this same prediction, but holding $V_{max} = 30$ m s^{-1} constant (pink triangle); and further holding f constant at its value at 20°N (pink circle). The latter tests demonstrate how the statistical increase in R_{max} with latitude is due to the increase in $R_{17.5ms}$ with latitude.

less sensitive to such issues. The radius of 34-kt wind ($R_{17.5ms}$) is the outermost wind radius that is routinely estimated operationally and has been reanalyzed postseason for the Best Track database since 2004.

Here we have presented a simple model for predicting R_{max} from $R_{17.5ms}$ that combines the underlying physics of TC radial structure with empirical data. The model is given by Eqs. (2)–(4) and (7), with an optional final bias adjustment given by Eq. (8). Our approach uses the physics of angular momentum loss in the TC inflow to reduce the relationship between the two radii to a dependence on two simple physical parameters, and then uses observational data to estimate the model coefficients. The form of the empirical model is chosen purposefully to incorporate the mathematical constraints that are imposed by the physics of our problem.

Our findings are summarized as follows:

1) Our model offers a promising first-order prediction of R_{max} from outer size (here $R_{17.5ms}$) in real tropical cyclones. The model only requires as input quantities that are routinely estimated operationally (V_{max}, $R_{17.5ms}$, and storm center latitude).
2) Model results are consistent between Extended Best Track and TC-OBS historical tropical cyclone databases for the North Atlantic.
3) Model performance exceeds existing operational models for predicting R_{max}.
4) The model predicts the observed statistical increase in R_{max} with latitude and demonstrates that this is principally driven by the statistical increase in $R_{17.5ms}$ with latitude.

Our model is fast and straightforward to implement, and it has value for a range of applications, and the operational

setting. The model can provide simple, observationally constrained first-order estimates of R_{max} for TC forecasting. This estimate could then be further refined in the presence of additional data as needed. The simplicity of the model may further allow one to quantify how uncertainties in observational estimates of input parameters, such as R_{34kt}, translate to uncertainties in the prediction of R_{max}. The model can also be used to estimate R_{max} for TC risk applications using TC track models that are either empirical (Yonekura and Hall 2011) or downscaled from climate models [Emanuel et al. 2006; Jing and Lin 2020; Lee et al. 2018). Given V_{max} and R_{max}, a simple parametric model (e.g., the C15 model presented here or that of Willoughby and Rahn (2004)] may be used to model the entire TC wind field in a manner that optimally captures the strongest winds in the vicinity of R_{max}, which is ideal for hazard modeling. Finally, the model may also be useful for downscaling estimates of R_{max} from coarser resolution weather models with grid spacings on the order of R_{max} itself (Davis 2018).

Here we have applied our model to $R_{17.5ms}$ given its numerous advantages noted above. However, our model is not specific to this choice and can be easily applied to any other desired wind radius if given sufficient data. Doing so simply requires re-estimating the model coefficients from the new dataset. Alternatively, one could also choose to first model the relationship between the new wind radius and $R_{17.5ms}$ and then apply our model as presented here. To link $R_{17.5ms}$ to a larger wind radius in the outer circulation, a model similar to that presented here could be used but with only the single predictor $(1/2)fR_{new}$ since the outer circulation structure is largely independent of V_{max} (Frank 1977; Merrill 1984; Chavas and Lin 2016). Another option is to directly apply the outer wind structure component of the C15 model since it has
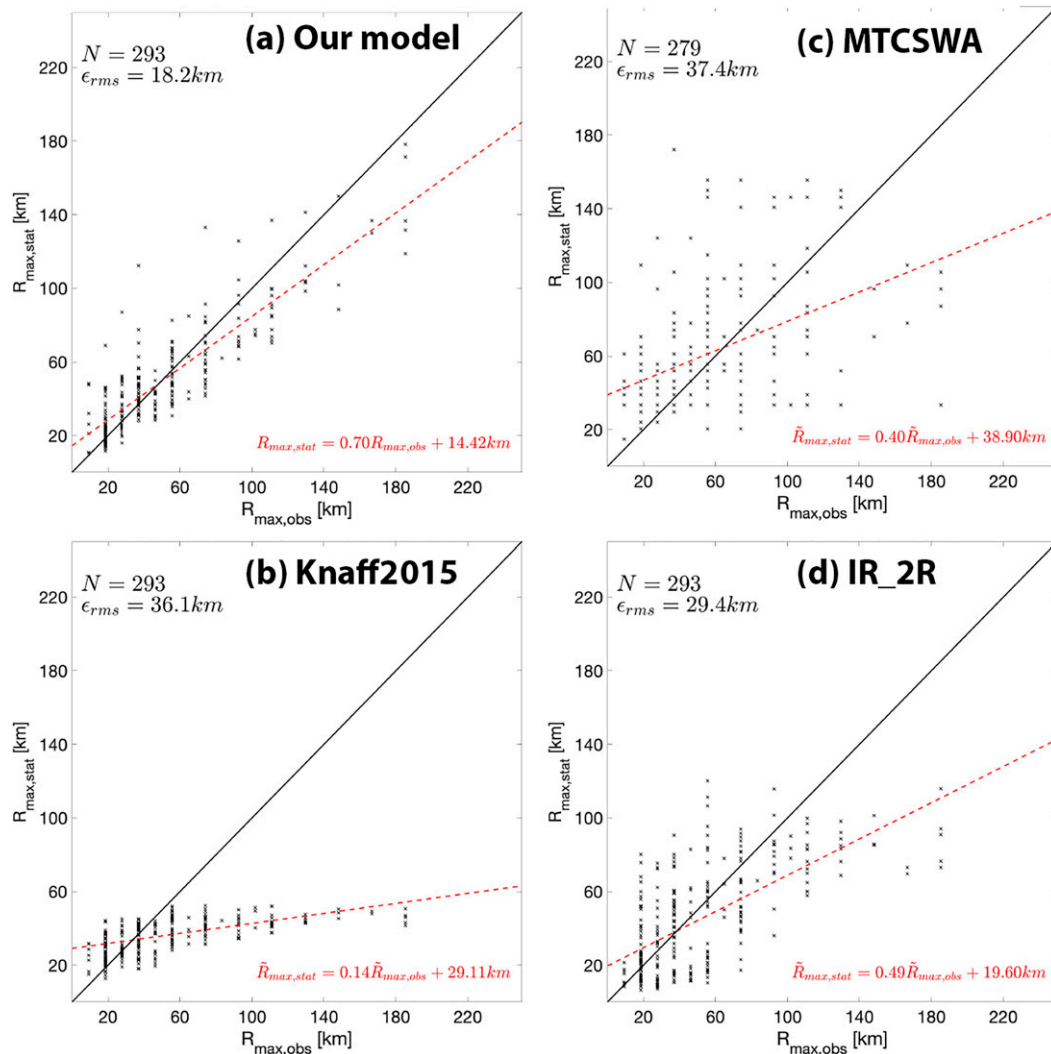
Fig. 10. Scatterplots of out-of-sample predicted vs observed R_{max} for EBT data 2018–20. (a) Our model (coefficients refit to EBT excluding 2018–20), (b) Knaff et al. (2015) model, (c) MTCSWA, and (d) IR-2R. Black line = 1-to-1 line, and red dashed line = linear fit to data.

been shown to provide an excellent representation of the TC outer wind structure when compared to QuikSCAT data (Chavas et al. 2015).

While we have focused on the outermost available wind radius (R_{34kt}), one could also apply the model framework using wind radii associated with stronger wind speeds such as R_{50kt} and R_{64kt}, which are also routinely estimated operationally and are available for more intense storms whose maximum wind speeds exceed this threshold. Note though that observational uncertainty of these wind radii may be greater as some observational instruments lose accuracy moving toward higher wind speeds above 34 kt. We do not take this step here, but this could be a useful avenue of future work to better constrain the estimate of R_{max} using wind radius data closer to R_{max} itself.

We highlight that the quantity $M_{max}/M_{17.5ms}$ analyzed here represents the radially integrated effect of surface drag on an

air parcel, which may hold useful information for helping to constrain the variation of C_d in the inner core of a TC. To our knowledge this quantity has yet to be applied in such a context.

The model as presented here was applied to two historical datasets owing to the highly variable methods for estimating wind radii, especially at the surface. One promising new dataset on the horizon is surface wind field data estimated from synthetic aperture radar (SAR) (Mouche et al. 2017; Zhang et al. 2021), which has the potential to provide a direct estimate of the near-surface wind field that comes from a single consistent source with minimal bias at all wind speeds in a TC. The SAR dataset is currently limited in sample size but has shown substantial promise for TC near-surface wind field estimation (Combot et al. 2020). As this dataset grows, applying our model to SAR data is likely to be a fruitful avenue for refining parameter estimates in future work.
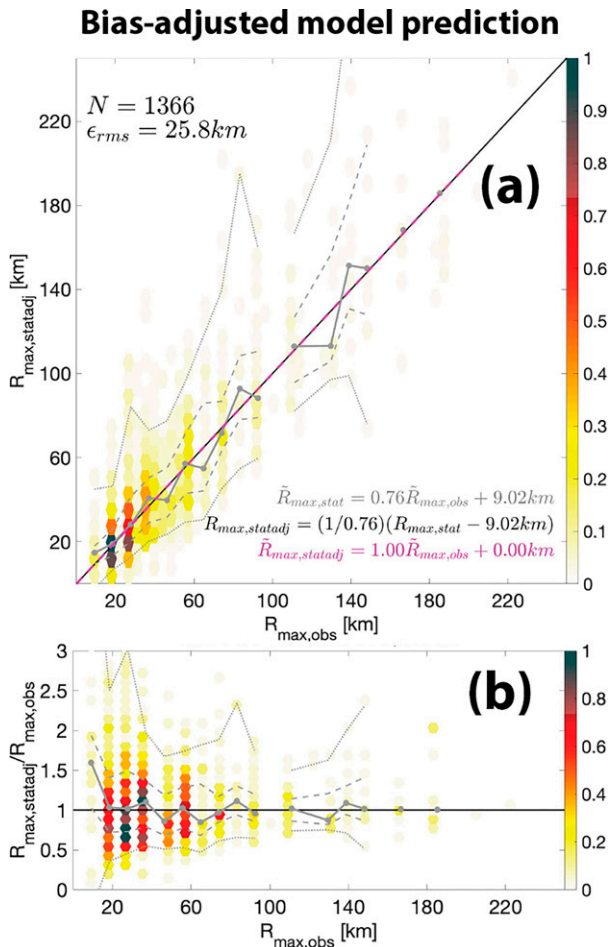
Bias-adjusted model prediction

$$\tilde{R}_{max,stat} = 0.76\tilde{R}_{max,obs} + 9.02km$$
$$R_{max,statadj} = (1/0.76)(R_{max,stat} - 9.02km)$$
$$\tilde{R}_{max,statadj} = 1.00\tilde{R}_{max,obs} + 0.00km$$

FIG. 11. As in Figs. 5c,d, but adjusted to remove systematic bias in prediction of R_{max}. Bias adjustment is given by Eq. (8).

Data availability statement. Code for the final predictive model for Rmax presented here, as well as data and code for the figures in this manuscript are available at https://doi.org/doi:10.4231/3FQ4-1E11. All raw data analyzed in this study are freely available from the Extended Best Track website at https://rammb.cira.colostate.edu/research/tropical_cyclones/ tc_extended_best_track_dataset/ and the TC-OBS website at https://verif.rap.ucar.edu/tcdata/historical/.

REFERENCES

Bessho, K., M. DeMaria, and J. A. Knaff, 2006: Tropical cyclone wind retrievals from the advanced microwave sounding unit: Application to surface wind analysis. *J. Appl. Meteor. Climatol.*, **45**, 399–415, https://doi.org/10.1175/JAM2352.1.

Chavas, D. R., and N. Lin, 2016: A model for the complete radial structure of the tropical cyclone wind field. Part II: Wind field variability. *J. Atmos. Sci.*, **73**, 3093–3113, https://doi.org/10.1175/JAS-D-15-0185.1.

——, ——, and K. Emanuel, 2015: A model for the complete radial structure of the tropical cyclone wind field. Part I: Comparison with observed structure. *J. Atmos. Sci.*, **72**, 3647–3662, https://doi.org/10.1175/JAS-D-15-0014.1.

——, K. A. Reed, and J. A. Knaff, 2017: Physical understanding of the tropical cyclone wind-pressure relationship. *Nat. Commun.*, **8**, 1360, https://doi.org/10.1038/s41467-017-01546-9.

Chen, J., and D. R. Chavas, 2020: The transient responses of an axisymmetric tropical cyclone to instantaneous surface roughening and drying. *J. Atmos. Sci.*, **77**, 2807–2834, https://doi.org/10.1175/JAS-D-19-0320.1.

Cocks, S. B., and W. M. Gray, 2002: Variability of the outer wind profiles of western North Pacific typhoons: Classifications and techniques for analysis and forecasting. *Mon. Wea. Rev.*, **130**, 1989–2005, https://doi.org/10.1175/1520-0493(2002)130<1989: VOTOWP>2.0.CO;2.

Combot, 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.

Courtney, J., and J. A. Knaff, 2009: Adapting the Knaff and Zehr wind-pressure relationship for operational use in tropical cyclone warning centres. *Aust. Meteor. Oceanogr. J.*, **58**, 167–179, https://doi.org/10.22499/2.5803.002.

Davis, C., 2018: Resolving tropical cyclone intensity in models. *Geophys. Res. Lett.*, **45**, 2082–2087, https://doi.org/10.1002/2017GL076966.

DeMaria, M., J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting*, **24**, 1573–1591, https://doi.org/10.1175/2009WAF2222286.1.

Demuth, J. L., M. DeMaria, and J. A. Knaff, 2006: Improvement of Advanced Microwave Sounding Unit tropical cyclone intensity and size estimation algorithms. *J. Appl. Meteor. Climatol.*, **45**, 1573–1581, https://doi.org/10.1175/JAM2429.1.

Donelan, M., B. Haus, N. Reul, W. Plant, M. Stiassnie, H. Graber, O. Brown, and E. Saltzman, 2004: On the limiting aerodynamic roughness of the ocean in very strong winds. *Geophys. Res. Lett.*, **31**, L18306, https://doi.org/10.1029/2004GL019460.

Emanuel, K., 2004: Tropical cyclone energetics and structure. *Atmospheric Turbulence and Mesoscale Meteorology*, E. Fedorovich, R. Rotunno, and B. Stevens, Eds., Cambridge University Press, 165–192, https://doi.org/10.1017/CBO9780511735035.010.

——, and R. Rotunno, 2011: Self-stratification of tropical cyclone outflow. Part I: Implications for storm structure. *J. Atmos. Sci.*, **68**, 2236–2249, https://doi.org/10.1175/JAS-D-10-05024.1.

——, S. Ravela, E. Vivant, and C. Risi, 2006: A statistical deterministic approach to hurricane risk assessment. *Bull. Amer. Meteor. Soc.*, **87**, 299–314, https://doi.org/10.1175/BAMS-87-3-299.

Frank, W. M., 1977: The structure and energetics of the tropical cyclone. Part I: Storm structure. *Mon. Wea. Rev.*, **105**, 1119–1135, https://doi.org/10.1175/1520-0493(1977)105<1119:TSAEOT>2.0.CO;2.

Furnival, G. M., and R. W. Wilson, 2000: Regressions by leaps and bounds. *Technometrics*, **42**, 69–79, https://doi.org/10.1080/00401706.2000.10485982.

Gentry, M. S., and G. M. Lackmann, 2010: Sensitivity of simulated tropical cyclone structure and intensity to horizontal resolution. *Mon. Wea. Rev.*, **138**, 688–704, https://doi.org/10.1175/2009MWR2976.1.

Guishard, M. P., J. L. Evans, and R. E. Hart, 2009: Atlantic subtropical storms. Part II: Climatology. *J. Climate*, **22**, 3574–3594, https://doi.org/10.1175/2008JCLI2346.1.

Hart, R. E., 2003: A cyclone phase space derived from thermal wind and thermal asymmetry. *Mon. Wea. Rev.*, **131**, 585–616, https://doi.org/10.1175/1520-0493(2003)131<0585:ACPSDF>2.0.CO;2.

Irish, J. L., and D. T. Resio, 2010: A hydrodynamics-based surge scale for hurricanes. *Ocean Eng.*, **37**, 69–81, https://doi.org/10.1016/j.oceaneng.2009.07.012.

——, ——, and J. J. Ratcliff, 2008: The influence of storm size on Hurricane Surge. *J. Phys. Oceanogr.*, **38**, 2003–2013, https://doi.org/10.1175/2008JPO3727.1.

Jing, R., and N. Lin, 2020: An environment-dependent probabilistic tropical cyclone model. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001975, https://doi.org/10.1029/2019MS001975.

Kepert, J. D., 2017: Time and space scales in the tropical cyclone boundary layer, and the location of the eyewall updraft. *J. Atmos. Sci.*, **74**, 3305–3323, https://doi.org/10.1175/JAS-D-17-0077.1.

Knaff, J. A., and R. M. Zehr, 2007: Reexamination of tropical cyclone wind–pressure relationships. *Wea. Forecasting*, **22**, 71–88, https://doi.org/10.1175/WAF965.1.

——, and C. R. Sampson, 2015: After a decade are Atlantic tropical cyclone gale force wind radii forecasts now skillful? *Wea. Forecasting*, **30**, 702–709, https://doi.org/10.1175/WAF-D-14-00149.1.

——, M. DeMaria, D. A. Molenar, C. R. Sampson, and M. G. Seybold, 2011: An automated, objective, multiple-satellite-platform tropical cyclone surface wind analysis. *J. Appl. Meteor. Climatol.*, **50**, 2149–2166, https://doi.org/10.1175/2011JAMC2673.1.

——, S. P. Longmore, and D. A. Molenar, 2014: An objective satellite-based tropical cyclone size climatology. *J. Climate*, **27**, 455–476, https://doi.org/10.1175/JCLI-D-13-00096.1. Corrigendum. *J. Climate*, **28**, 8648–8651, https://doi.org/10.1175/JCLI-D-15-0610.1.

——, ——, R. T. DeMaria, and D. A. Molenar, 2015: Improved tropical-cyclone flight-level wind estimates using routine infrared satellite reconnaissance. *J. Appl. Meteor. Climatol.*, **54**, 463–478, https://doi.org/10.1175/JAMC-D-14-0112.1.

——, 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.tcrr.2021.09.002.

Kossin, J. P., J. A. Knaff, H. I. Berger, D. C. Herndon, T. A. Cram, C. S. Velden, R. J. Murnane, and J. D. Hawkins, 2007: Estimating hurricane wind structure in the absence of aircraft

reconnaissance. *Wea. Forecasting*, **22**, 89–101, https://doi.org/10.1175/WAF985.1.

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

Lee, C.-Y., S. J. Camargo, F. Vitart, A. H. Sobel, and M. K. Tippett, 2018: Subseasonal tropical cyclone genesis prediction and MJO in the S2S dataset. *Wea. Forecasting*, **33**, 967–988, https://doi.org/10.1175/WAF-D-17-0165.1.

Lin, N., and D. Chavas, 2012: On hurricane parametric wind and applications in storm surge modeling. *J. Geophys. Res. Atmos.*, **117**, D09120, https://doi.org/10.1029/2011JD017126.

Lonfat, M., R. Rogers, T. Marchok, and F. D. Marks, 2007: A parametric model for predicting hurricane rainfall. *Mon. Wea. Rev.*, **135**, 3086–3097, https://doi.org/10.1175/MWR3433.1.

Lu, P., N. Lin, K. Emanuel, D. Chavas, and J. Smith, 2018: Assessing hurricane rainfall mechanisms using a physics-based model: Hurricanes Isabel (2003) and Irene (2011). *J. Atmos. Sci.*, **75**, 2337–2358, https://doi.org/10.1175/JAS-D-17-0264.1.

Merrill, R. T., 1984: A comparison of large and small tropical cyclones. *Mon. Wea. Rev.*, **112**, 1408–1418, https://doi.org/10.1175/1520-0493(1984)112<1408:ACOLAS>2.0.CO;2.

Mouche, A. A., B. Chapron, B. Zhang, and R. Husson, 2017: Combined co- and cross-polarized SAR measurements under extreme wind conditions. *IEEE Trans. Geosci. Remote Sens.*, **55**, 6746–6755, https://doi.org/10.1109/TGRS.2017.2732508.

Mueller, K. J., M. DeMaria, J. Knaff, J. P. Kossin, and T. H. Vonder Haar, 2006: Objective estimation of tropical cyclone wind structure from infrared satellite data. *Wea. Forecasting*, **21**, 990–1005, https://doi.org/10.1175/WAF955.1.

OFCM, 2020: National Hurricane Operations Plan. Rep. FCM-P12-2020, 171 pp.

Palmén, E., and H. Riehl, 1957: Budget of angular momentum and energy in tropical cyclones. *J. Meteor.*, **14**, 150–159, https://doi.org/10.1175/1520-0469(1957)014<0150:BOAMAE>2.0.CO;2.

Penny, A. B., L. P. Alaka, C. L. Fritz, J. Rhome, and A. A. Taylor, 2021: Improvements to the probabilistic storm surge model (P-Surge). *34th Conf. on Hurricanes and Tropical Meteorology Storm Surge*, virtual, Amer. Meteor. Soc., 6C.2, https://ams.confex.com/ams/34HURR/meetingapp.cgi/Paper/373360.

Reed, K. A., and C. Jablonowski, 2011: Assessing the uncertainty in tropical cyclone simulations in NCAR's Community Atmosphere Model. *J. Adv. Model. Earth Syst.*, **3**, M08002, https://doi.org/10.1029/2011MS000076.

Richter, D. H., and D. P. Stern, 2014: Evidence of spray-mediated air–sea enthalpy flux within tropical cyclones. *Geophys. Res. Lett.*, **41**, 2997–3003, https://doi.org/10.1002/2014GL059746.

——, C. Wainwright, D. P. Stern, G. H. Bryan, and D. Chavas, 2021: Potential low bias in high-wind drag coefficient inferred from dropsonde data in hurricanes. *J. Atmos. Sci.*, **78**, 2339–2352, https://doi.org/10.1175/JAS-D-20-0390.1.

Riehl, H., 1950: A model of hurricane formation. *J. Appl. Phys.*, **21**, 917–925, https://doi.org/10.1063/1.1699784.

Rotunno, R., and G. H. Bryan, 2012: Effects of parameterized diffusion on simulated hurricanes. *J. Atmos. Sci.*, **69**, 2284–2299, https://doi.org/10.1175/JAS-D-11-0204.1.

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.

Schenkel, B. A., N. Lin, D. Chavas, M. Oppenheimer, and A. Brammer, 2017: Evaluating outer tropical cyclone size in reanalysis datasets using QuikSCAT data. *J. Climate*, **30**, 8745–8762, https://doi.org/10.1175/JCLI-D-17-0122.1.

Shea, D. J., and W. M. Gray, 1973: The hurricane's inner core region. I. symmetric and asymmetric structure. *J. Atmos. Sci.*, **30**, 1544–1564, https://doi.org/10.1175/1520-0469(1973)030<1544:THICRI>2.0.CO;2.

Shmueli, G., and Coauthors, 2010: To explain or to predict? *Stat. Sci.*, **25**, 289–310, https://doi.org/10.1214/10-STS330.

Sitkowski, M., J. P. Kossin, and C. M. Rozoff, 2011: Intensity and structure changes during hurricane eyewall replacement cycles. *Mon. Wea. Rev.*, **139**, 3829–3847, https://doi.org/10.1175/MWR-D-11-00034.1.

Stern, D. P., J. D. Kepert, G. H. Bryan, and J. D. Doyle, 2020: Understanding atypical midlevel wind speed maxima in hurricane eyewalls. *J. Atmos. Sci.*, **77**, 1531–1557, https://doi.org/10.1175/JAS-D-19-0191.1.

Tang, B., and K. Emanuel, 2012: A ventilation index for tropical cyclones. *Bull. Amer. Meteor. Soc.*, **93**, 1901–1912, https://doi.org/10.1175/BAMS-D-11-00165.1.

Uhlhorn, E. W., P. G. Black, J. L. Franklin, M. Goodberlet, J. Carswell, and A. S. Goldstein, 2007: Hurricane surface wind measurements from an operational Stepped Frequency Microwave Radiometer. *Mon. Wea. Rev.*, **135**, 3070–3085, https://doi.org/10.1175/MWR3454.1.

——, B. W. Klotz, T. Vukicevic, P. D. Reasor, and R. F. Rogers, 2014: Observed hurricane wind speed asymmetries and relationships to motion and environmental shear. *Mon. Wea. Rev.*, **142**, 1290–1311, https://doi.org/10.1175/MWR-D-13-00249.1.

Vickery, P. J., and D. Wadhera, 2008: Statistical models of Holland pressure profile parameter and radius to maximum winds of hurricanes from flight-level pressure and H*Wind data. *J. Appl. Meteor. Climatol.*, **47**, 2497–2517, https://doi.org/10.1175/2008JAMC1837.1.

Vigh, J. L., E. Gilleland, C. L. Williams, D. R. Chavas, N. M. Dorst, J. M. Done, G. J. Holland, and B. G. Brown, 2016: A new historical database of tropical cyclone position, intensity, and size parameters optimized for wind risk modeling. *32nd Conf. on Hurricanes and Tropical Meteorology*, San Juan, Puerto Rico, Amer. Meteor. Soc., 12C.2, https://ams.confex.com/ams/32Hurr/webprogram/Paper293910.html.

Willoughby, H., and M. Rahn, 2004: Parametric representation of the primary hurricane vortex. Part I: Observations and evaluation of the Holland (1980) model. *Mon. Wea. Rev.*, **132**, 3033–3048, https://doi.org/10.1175/MWR2831.1.

——, R. Darling, and M. Rahn, 2006: Parametric representation of the primary hurricane vortex. Part II: A new family of sectionally continuous profiles. *Mon. Wea. Rev.*, **134**, 1102–1120, https://doi.org/10.1175/MWR3106.1.

Wing, A. A., S. J. Camargo, and A. H. Sobel, 2016: Role of radiative–convective feedbacks in spontaneous tropical cyclogenesis in idealized numerical simulations. *J. Atmos. Sci.*, **73**, 2633–2642, https://doi.org/10.1175/JAS-D-15-0380.1.

Xi, D., N. Lin, and J. Smith, 2020: Evaluation of a physics-based tropical cyclone rainfall model for risk assessment. *J. Hydrometeor.*, **21**, 2197–2218, https://doi.org/10.1175/JHM-D-20-0035.1.

Yonekura, E., and T. M. Hall, 2011: A statistical model of tropical cyclone tracks in the western North Pacific with ENSO-dependent cyclogenesis. *J. Appl. Meteor. Climatol.*, **50**, 1725–1739, https://doi.org/10.1175/2011JAMC2617.1.

Zhang, B., Y. Lu, W. Perrie, G. Zhang, and A. Mouche, 2021: Compact polarimetry synthetic aperture radar ocean wind retrieval: Model development and validation. *J. Atmos. Oceanic Technol.*, **38**, 747–757, https://doi.org/10.1175/JTECH-D-20-0035.1.