

Autoregressive Modeling for Tropical Cyclone Intensity Climatology

CHIA-YING LEE

International Research Institute for Climate and Society, Columbia University, Palisades, New York

MICHAEL K. TIPPETT

*Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York,
and Center of Excellence for Climate Change Research, Department of Meteorology,
King Abdulaziz University, Jeddah, Saudi Arabia*

ADAM H. SOBEL

*Department of Applied Physics and Applied Mathematics, Columbia University,
New York, New York, and Lamont-Doherty Earth Observatory,
Columbia University, Palisades, New York*

SUZANA J. CAMARGO

Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

(Manuscript received 19 December 2015, in final form 17 May 2016)

ABSTRACT

An autoregressive model is developed to simulate the climatological distribution of global tropical cyclone (TC) intensity. The model consists of two components: a regression-based deterministic component that advances the TC intensity in time and depends on the storm state and surrounding large-scale environment and a stochastic forcing. Potential intensity, deep-layer mean vertical shear, and midlevel relative humidity are the environmental variables included in the deterministic component. Given a storm track and its environment, the model is initialized and then iterated along the track. Model performance is evaluated by its ability to represent the observed global and basin distributions of TC intensity as well as lifetime maximum intensity (LMI). The deterministic model alone captures the spatial features of the climatological TC intensity distribution but with intensities that remain below 100 kt ($1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$). Addition of white (uncorrelated in time) stochastic forcing reduces this bias by improving the simulated intensification rates and the frequency of major storms. The model simulates a realistic range of intensities, but the frequency of major storms remains too low in some basins.

1. Introduction

For assessing tropical cyclone (TC) risk, it is necessary to estimate the risk of the most severe storm with a nonnegligible probability of making landfall at a given location. Even a very small probability may be nonnegligible if the impact of the event would be sufficiently high. In most, if not all, locations, the historical record alone is inadequate to characterize these

low-probability, high-impact events. This is even more true in a changing climate, when the historical record may not be representative of future conditions.

Global climate models (GCMs) can simulate many features of the TC climatology. Recent improvements in both computational power and understanding of atmospheric dynamics and physics have led to the development of climate models with impressive ability to predict the interannual variability of TC activity (e.g., Vitart et al. 2007; Zhao et al. 2009, 2010; Chen and Lin 2013; Camargo and Wing 2016). Nevertheless, few climate models are capable of simulating the most intense storms, and the simulated distribution of lifetime maximum intensity (LMI) in these models is often deficient

Corresponding author address: Chia-Ying Lee, International Research Institute for Climate and Society, Columbia University, 202 Monell, 61 Route 9W, Palisades, NY 10964.
E-mail: clee@iri.columbia.edu

(Zhao et al. 2009; Manganello et al. 2012, 2014; Wu et al. 2014; Oouchi et al. 2006). The LMI distribution is a basic measure for understanding and describing the TC climatology in both current and future climates (e.g., Emanuel 2000; Kossin et al. 2014; Zhao et al. 2009; Manganello et al. 2014). Unlike other natural hazards whose rarity increases with intensity—for example, earthquakes (Hristopulos and Mouslopoulou 2013) and tornadoes (Feuerstein et al. 2005)—the LMI distribution is bimodal with peaks located at about 50 kt ($1 \text{ kt} \approx 0.51 \text{ m s}^{-1}$) and at 110–130 kt (Kossin et al. 2013; Lee et al. 2016). In observations, the storms that undergo rapid intensification (RI) at some point during their lifetime are responsible for the second peak of the LMI distribution (Lee et al. 2016). Therefore, one explanation for the failure of most climate models to simulate the bimodal LMI distribution is their failure to simulate RI, the dramatic strengthening of a TC in a short time. High spatial resolution appears to improve the simulation of intensification rate and the LMI distribution (Manganello et al. 2012; Murakami et al. 2012). Using a coupled GCM with 25-km horizontal grid spacing, Murakami et al. (2015) successfully modeled hurricane intensities in categories 4 and 5 on the Saffir–Simpson scale with frequencies comparable to observations but with a unimodal LMI distribution.

However, apart from any limitations in their fidelity, high-resolution GCM simulations are computationally expensive, and that expense may limit their use in sampling the full range of low-probability, high-impact events. Alternative approaches of estimating TC risk, using statistical or simplified dynamical models, are desirable as a complement to GCM simulations.

Statistical models based on empirical equations are commonly used for prediction of TC activity on seasonal and longer time scales. Instead of simulating storms directly, these models usually predict their frequency or intensity-associated metrics, such as accumulated cyclone energy (ACE) or power dissipation index (PDI) from climate indices and/or large-scale forcing (e.g., Gray 1984; Klotzbach and Gray 2009; Klotzbach and Oliver 2015; Wang et al. 2009; Davis et al. 2015). These approaches do not directly calculate the probabilities of the most rare, extreme landfalling events, however.

One set of approaches for estimating the risk of low-probability, high-impact events involves generating large numbers of synthetic storms using methods whose computational cost is much lower than that of high-resolution GCMs. Such approaches simulate TC genesis, track, intensity, and size using varying combinations of statistical and dynamical methods. Some models simulate the storm track and intensity using stochastic models trained on the historical hurricane data as a

function of location, storm characteristics, and/or environmental parameters. These include the models developed by Hall and Jewson (2007), Hall and Yonekura (2013), Nakamura et al. (2015), and AIR Worldwide Corporation (AIR 2015). These models' use of the historical record or their choices of environmental parameters may restrict their use to the current climate. The use of fixed sea surface temperature thresholds (e.g., Hall and Jewson 2007; Hall and Yonekura 2013), for example, is likely inappropriate under climate change (e.g., Johnson and Xie 2010; Yoshimura et al. 2006; Knutson et al. 2008). The approach of Emanuel (2006) incorporates more physics, simulating TC intensity with an axisymmetric dynamical model. This dynamical model is coupled to a statistical–dynamical track model and embedded in an explicit three-dimensional representation of the large-scale climate that comes from either a reanalysis dataset or a (possibly low resolution) climate model. The more complete representation of the dependence of storm intensity on the environment makes this approach well suited to the task of quantifying the influence of anthropogenic climate change (e.g., Emanuel 2013).

We are working to develop a new statistical–dynamical downscaling system forced by large-scale environmental fields. Our intent is similar to that of Emanuel (2006) and Emanuel et al. (2006), whose work inspires this research. An essential element is the representation of TC intensity. Our approach here is based on the observed relation between TC intensification and the large-scale environment and thus uses less physics than that of Emanuel (2006) and Emanuel et al. (2006) but more than those of Hall and Yonekura (2013), Nakamura et al. (2015), and AIR (2015). We include an explicit dependence on the local atmospheric environment at each point in the storm's evolution, requiring an explicit set of three-dimensional fields to represent that environment, as in Emanuel's approach.

The development of the empirical TC intensity model in Lee et al. (2015) was a first step in developing such an intensity model. We pointed out that a multiple linear regression model widely used operationally [e.g., Statistical Hurricane Intensity Prediction Scheme (SHIPS); DeMaria et al. 2005] is essentially a form of statistical–dynamical downscaling that predicts intensity change based on storm characteristics and large-scale environment. With a similar configuration but differences in detail, a short-lead TC intensity model was built in that study. In the present study, we add a stochastic component to the model of Lee et al. (2015) and use it to simulate TC intensity through the complete storm life cycle as a function of ambient forcing with given observed storm tracks. The model is autoregressive, in that

the linear stochastic model is initialized at the beginning of the storm and iterated over the TC's lifetime with subsequent intensity values depending on earlier ones as well as the environment. The stochastic component draws from the empirical errors of the deterministic model and makes it possible for the complete model to simulate storms with realistic intensities.

After developing the stochastic model, we model the TC intensity climatology in the current climate using both the tracks and the environment taken from observation-based datasets. We focus on the global and basin LMI distributions, as well as on the complete distribution of intensity (not just the lifetime maxima).

This study is organized as follows. The datasets used, the environmental conditions considered, how LMI is defined, etc. are described in section 2. In section 3, a stochastic system is developed by extending the multiple linear regression intensity model used in Lee et al. (2015). The model performance in simulating the LMI distribution is discussed as well. The overall model performance in simulating the TC intensity distribution is shown in section 4. Results and findings are summarized in section 5.

2. Data

The best-track dataset HURDAT2, produced by the National Hurricane Center (NHC), is used for North Atlantic (ATL) and eastern North Pacific (EPC) hurricanes (Landsea and Franklin 2013; NHC 2013). Data produced by the Joint Typhoon Warning Center (JTWC) are used for storms in the western North Pacific (WPC), Indian Ocean (IO), and Southern Hemisphere Ocean (SH) (Chu et al. 2002; JTWC 2014). The best-track data include 1-min maximum sustained wind, minimum sea level pressure, and storm location every 6 h. Storm LMI is defined as the maximum sustained wind speed during the storm's life cycle. We use all the recorded data, including storms with tropical depression (TD) strength.

Large-scale environmental variables examined here are calculated from the $2.5^\circ \times 2.5^\circ$ monthly mean European Centre for Medium-Range Weather Forecasts interim reanalysis (ERA-Interim; Dee et al. 2011; ECMWF 2013).¹ Monthly data are used instead of higher-frequency (e.g., daily) data because we have previously shown that monthly data are adequate for a statistical TC intensity model (Lee et al. 2015). We consider environmental parameters representing the critical conditions for TC intensification: potential intensity (PI; Bister and Emanuel 2002; Camargo et al.

2007), 800–200-hPa-deep layer mean vertical wind shear (SHRD; Chen et al. 2006), 500–300-hPa midlevel and boundary layer mean relative humidity (midRH and RHbl), total column water vapor (rhCol; Tippett et al. 2011), convective available potential energy (CAPE; Bechtold et al. 2014), and 200-hPa zonal wind (U200) and divergence (div200). Additionally, upper-ocean structures from the approximately $1.8^\circ \times 1.8^\circ$ NOAA/NCEP Environmental Modeling Center (EMC) Climate Modeling Branch (CMB) Global Ocean Data Assimilation System (GODAS; Behringer and Xue 2004; IRI Data Library 2013) are used to calculate the upper-100-m mean ocean temperature (T100; Price 2009).

The monthly reanalysis fields are first linearly interpolated to the forecast day and storm location. The interpolated fields are averaged over a disk extending 500 km from the storm center for PI, over a circle with radius of 1000 km for div200, and over an annulus extending 200–800 km around the storm center for the remaining predictors. In the end, the predictors are averaged over the forecast interval (e.g., 12 h). Section 3 describes the procedure for selecting which environmental variables are used as predictors.

Data from 1981 to 1999 are used for the model development, and data from the period 2000–12 are used for evaluating model performance. All the climatological distributions shown here (i.e., all figures) use TC data from 2000 to 2012. Throughout the study, we use a Saffir–Simpson scale to categorize storm strength. Major TCs are defined as category-3–5 storms (LMI > 96 kt), and intense storms are category-4 or category-5 storms (LMI > 113 kt).

3. The global stochastic intensity model and LMI distribution

The probabilistic multiple linear regression model (MLR) in Lee et al. (2015) was developed to predict ATL TC intensity at 12-h intervals up to five days in advance. Here we apply the same MLR methodology to TC intensity in all basins. The statistics of storm characteristics and the surrounding environments vary substantially among different basins. For example, storms in the IO tend to be weaker than those in the other four basins owing to their shorter lifetimes (e.g., Lee et al. 2016); PI is usually stronger in the WPC than in the ATL because of the warmer sea surface temperature (Emanuel 1986). Despite these differences, we expect the physical laws controlling TC intensification to be the same in all basins. To the extent that those laws can be captured in the statistical relationships between storm behavior and the surrounding environment, then we can expect a single MLR to perform well in all basins.

¹ ERA-Interim was $0.75^\circ \times 0.75^\circ$, but we interpolated to a $2.5^\circ \times 2.5^\circ$ grid.

On the other hand, the MLR is at best a crude approximation of the large-scale physics, and its deficiencies likely vary by basin. Allowing the MLR to vary by basin would address such inadequacies to the degree that they can be parameterized by the large-scale environment. Operationally, statistical TC intensity models use different predictors and coefficients in different basins (DeMaria et al. 2005; Knaff et al. 2005; Knaff and Sampson 2009). However, as we will show below, there is no notable disadvantage in using a single global MLR instead of using a set of basin-dependent MLRs, at least with our approach and purpose. Furthermore, for climate downscaling (the ultimate goal of our project), a single global model is the natural and probably better choice because as climate changes, basin differences may change as well. Moreover, fitting basin-dependent MLRs introduces more sampling variability. Therefore, in this study we focus on a global model, and results from basin-dependent models are only mentioned for comparison purposes.

The formal structure of the model is as follows:

$$V_{t+12h} - V_t = L(V_t, V_{t-12h}, X_t, X_{t+12h}) + \varepsilon_{t+12h}, \quad (1)$$

where V is the TC intensity, X represents the environmental variables related to TC intensification, L is a linear regression function (deterministic), and ε is the stochastic forcing. The subscripts indicate time. For example, V_t means TC intensity at time t while V_{t-12h} means TC intensity 12 h before time t . This model is essentially a second-order vector autoregressive model with time step equal to 12 h and environmental variables as exogenous inputs. The choice of order 2 follows from allowing the TC intensity at the next time step to depend on both the current intensity and rate of change of the intensity. We refer to the first term on the right-hand side of Eq. (1) as the deterministic part of the model and the second as the stochastic. The deterministic part is constructed using the MLR methodology. For each storm, Eq. (1) is initialized with the first observed intensity and then iterated with stochastic forcing and the observed environment varying along the track.

In what follows, we describe the development and behavior of the components on the right-hand side of Eq. (1). Definitions of various combinations of models and their acronyms are listed in Table 1.

a. Short-lead intensity model

In this section, we describe the environmental variables selected to be included in the model and their performance for forecasts with lead times up to 5 days. We begin with the initial pool of predictors representing large-scale forcing listed in section 2, along with storm intensity at the current time and its change over the

TABLE 1. Descriptions of models.

Model	Description
	Short-lead intensity models
Persistence	The persistence model in which intensity change is set to 0.
MLR	A multiple linear regression model with 6 (7) selected predictors for ocean (near land) points.
basin-MLR	Similar to MLR but for a set of multiple linear regression models that are trained with individual basinwide data, respectively.
	Deterministic models
MLR	Iterating 12-h multiple linear regression model. We continue using MLR because it is essentially a form of the regression model.
MLR3	Iterating 12-h multiple linear regression model that is similar to the MLR but with additional nonlinear predictors. The highest power of predictors is cubic.
	Stochastic models
MLR&wn	MLR with white noise stochastic forcing.
MLR3&wn	MLR3 with white noise stochastic forcing.
basin-MLR&wn	MLR3 trained with individual basinwide data and with white noise stochastic forcing.
persistence&wn	Persistence model with white noise stochastic forcing.

previous 12 h. Predictors are chosen using forward selection for each basin and lead time (not shown). While the precise predictors chosen can vary by basin and lead time, five predictors are important in all basins and at all leads: initial storm intensity V_0 [one can see it as V_t in Eq. (1)], previous intensification rate dV/dt , translation speed $trSpeed$, difference of PI and initial storm intensity dPI_V_0 , and SHRD, in agreement with Lee et al. (2015), who found robust dependence of ATL TC intensification on these predictors, regardless of reanalysis dataset. Therefore, these five predictors are initially selected.

Humidity parameters are not selected initially since they are not useful predictors in all basins at all lead times. Nevertheless, as moisture has been found to be an important factor in controlling the level of TC activity in response to climate change (Emanuel 2008; Camargo et al. 2014), we do include midRH, one of the two moisture variables considered. Large-scale subsurface oceanic conditions impact TC intensification (Schade and Emanuel 1999; Wu et al. 2007) and structure (Lee and Chen 2014) through storm self-induced sea surface temperature cooling. Projected changes in subsurface stratification can affect TC activity in a changing climate (Huang et al. 2015). However, the oceanic parameter T100 used here did not significantly improve the performance at any lead time and therefore is not included in the deterministic model.

TABLE 2. Coefficients of 12-h MLR; ocean and near land refer to the model for ocean and near-land points, respectively, and a dash indicates that the item is not applicable.

Models	V_0	dVdt	trSpeed	dPI_ V_0	SHRD	midRH	dLandMask	dVdt ²	(dPI_ V_0) ²	(dPI_ V_0) ³
MLR ocean	-0.11	0.52	0.03	0.15	-0.09	0.06	—	—	—	—
MLR near land	-0.52	0.34	0.002	0.003	-0.04	0.0012	-0.17	—	—	—
MLR3 ocean	-0.15	0.49	0.05	0.31	-0.10	0.08	—	0.03	0.18	-0.41
MLR3 near land	-0.54	0.43	-0.007	-0.05	-0.04	0.002	0.12	-0.18	-0.09	0.12

Intensity changes at landfall are considered separately. Operationally, an empirical exponential decay model is used to predict intensity change during landfall (Kaplan and DeMaria 1995). Here we use separate multiple linear regression models for ocean and near-land points. We define a predictor dLandMask representing the difference in the surface type between initial and forecast time. The surface type LandMask is defined by averaging a $0.5^\circ \times 0.5^\circ$ resolution land–sea mask (the value of land points is 0 while ocean points are -1) in a 300-km radius of the storm location. The 300-km radius is a proxy for the size of the strong storm circulation. Data points with LandMask values greater than -0.5 have more than 50% of the storm circulation over land and are not used. LandMask values between -1 and -0.5 are labeled as near land while those with value of -1 are labeled as ocean.² Two models are then developed: one with six predictors (V_0 , dVdt, trSpeed, dPI_ V_0 , SHRD, and midRH) for ocean points and a second one with dLandMask as an additional predictor for cases with either initial or forecast location in the near-land group. This approach allows different coefficients for the near-land and ocean points (e.g., coefficients for the 12-h forecast are shown in Table 2).

We first develop a global model for TC intensity without stochastic forcing (called MLR). The MLR is tested using independent data from the period 2000–12, and its overall performance is measured using mean absolute error (MAE) and root-mean-square error (RMSE), calculated against best-track data in each TC basin. MAE and RMSE results are similar to each other and therefore we show only MAE (Fig. 1). The skill of the MLR is compared to that of the persistence model, which predicts no change in the storm intensity. The MAE of the persistence model, in addition to providing a baseline error level, describes the average magnitude of TC intensity change in each basin as a function of time window. Changes in TC intensity are the largest in the WPC, followed by the SH. The

frequent and dramatic strengthening of TCs in the WPC and the SH may be related to the inner-core convection, which often is poorly handled by linear statistical intensity prediction models (Knaff et al. 2005). In the IO, the error in the persistence model drops after 84-h lead time, indicating smaller intensity change in long leads. A possible explanation is the shorter lifetime for IO storms. The MLR MAE in each basin is in general less than that of the persistence model, and in that sense the MLR is skillful in all basins. While the MAE of the MLR is the largest in the WPC and SH at almost all lead times (except for 12 h), the MLR provides the greatest advantage over the persistence model and has the lowest normalized error in these basins as well (blue and yellow lines in the Fig. 1c). On the other hand, the MLR for the ATL is more accurate but provides less advantage over the persistence model than in other basins. The 12-h MAE in the ATL is even slightly bigger than that of the persistence model. The advantage of the MLR increases with increasing lead time in all basins, except the IO. The MLR, trained with global data, does not take into account the shorter lifetime of IO storms. The dependent MAE (not shown) is close to the MAE in Fig. 1b and is only slightly smaller, and gives no indication of overfitting.

We also developed a set of basin-dependent models (basin-MLR), which have the potential to perform better than the (global) MLR because they allow basin-dependent relations with the large-scale environment. We use the same set of predictors in all basins to maintain model consistency but allow their coefficients to vary by basin. Similar to the global case, the basin-MLR models have smaller MAEs than those of the persistence model (not shown). For WPC and SH storms, the MLR and basin-MLR have similar MAEs. The MLR performs slightly better than the basin-MLR for IO storms for 12–36-h lead forecasts. For long leads, the basin model in the IO has an advantage, presumably because it recognizes the short storm lifetime. Other uncertainties of the basin-MLR for IO are the poor quality of best-track data (especially before 1990; Chu et al. 2002) and the small sample size. The use of global training data in the MLR is expected to help improve both. For ATL storms, the difference between global and basin MAEs is small,

²The -0.5 threshold and 300-km radius are chosen arbitrarily, but with these criteria, the model is capable of modeling the land impact on TC intensity change to some degree.

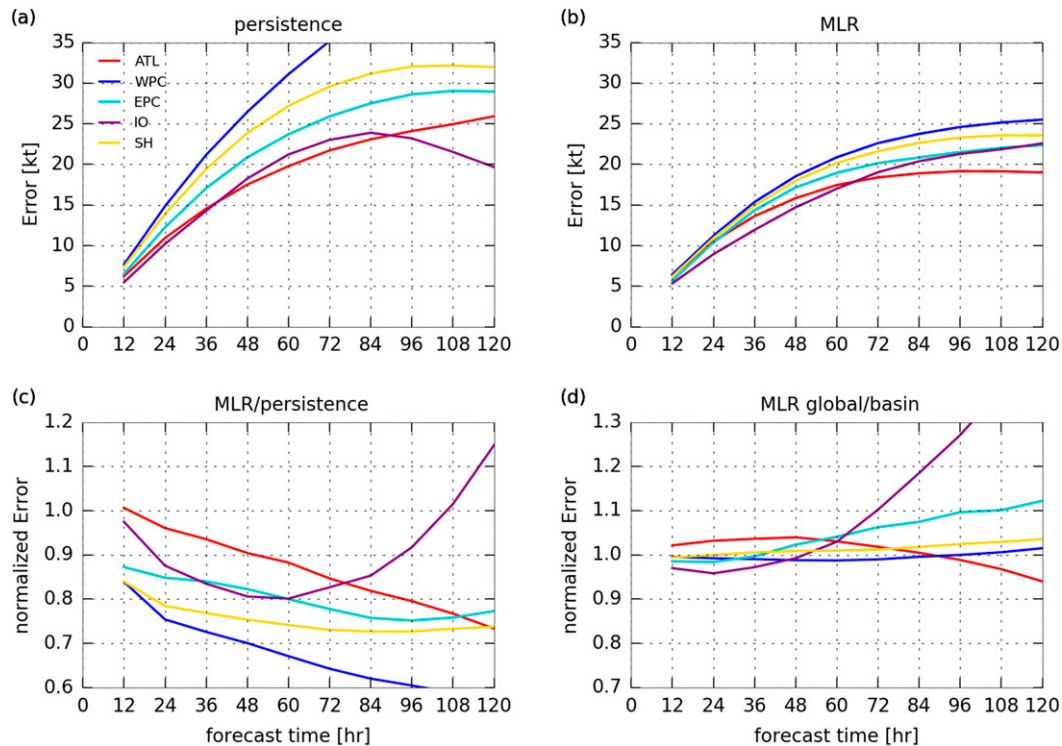


FIG. 1. Mean absolute error (MAE) of the intensity predictions from (a) the persistence model and (b) the MLR. (c) MLR MAE normalized by that from the persistence model. (d) MLR MAE normalized by that of the basin-dependent MLR. Colors indicate basins: ATL (red), WPC (blue), EPC (cyan), IO (purple), and SH (yellow).

approximately 5% at all leads. For EPC storms, the global MLR MAE is about 10% larger than that from basin-MLR for 120-h lead time, which is an error of only 2 kt. We can conclude that the global MLR performs similarly to the basin-MLR in most basins.

We compare the MAE of the MLR and basin-MLR with that of the persistence model only. Direct comparison with operational tools for predicting TC intensity change (e.g., SHIPS) is not entirely appropriate because the MLR uses monthly averaged reanalysis environment information and best-track storm information, while operational tools must use forecast fields and track. Moreover, the purpose of the MLR here is to provide the deterministic part of Eq. (1) rather than to develop a new forecasting tool. In any case, the error level of the MLR model in the ATL is roughly comparable to that of SHIPS (Lee et al. 2015).

b. Deterministic model—12-h MLR

The deterministic model in Eq. (1) uses a 12-h time step. The predictor coefficients are given in Table 2. As an example of the model's behavior, the blue line in Fig. 2a shows the predicted lifetime intensity of Typhoon Fanapi. The model in Eq. (1) is initialized on 0600 UTC 13 September 2010 with an initial intensity of

15 kt and $dVdt = 0$ and is subsequently iterated with the stochastic term ε_{t+12h} set to zero. The modeled intensity is close to observations up to 16 September but is lower than the observed intensity after that. Repeating this calculation for all observed TCs during the period 2000–12 shows that the intensification rates of the deterministic model (blue bars in Fig. 3) are never greater than $10 \text{ kt} (12 \text{ h})^{-1}$ or less than $-15 \text{ kt} (12 \text{ h})^{-1}$. The probability density function (PDF) of MLR intensification rate has a peak at $5\text{--}10 \text{ kt} (12 \text{ h})^{-1}$ while the observed one is at $0\text{--}5 \text{ kt} (12 \text{ h})^{-1}$, though with a tail at high intensities, which the deterministic model lacks entirely. The deterministic model does not produce storms with LMI larger than 100 kt, and its LMI PDF has a strong peak at about 50 kt (blue dashed line in Fig. 4).

c. Stochastic model—uncorrelated empirical error

The deterministic model is designed to minimize the squared error of 12-h forecasts by explaining as much variability as can be linearly related to the initial storm information and surrounding environment. The variability unexplained by the MLR may be due to other processes (e.g., nonlinear, mesoscale or small scales) and are accounted for in principle by the stochastic term in Eq. (1). Rather than characterizing the stochastic

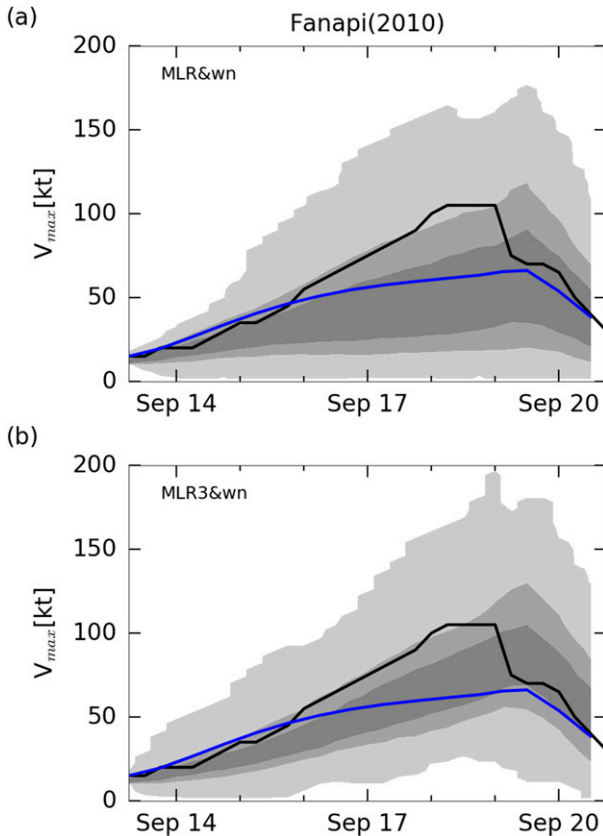


FIG. 2. The observed (black line) and the simulated (blue line) intensity for Typhoon Fanapi (2010) from (a) MLR&wn and (b) MLR3&wn. Acronyms indicate the choice of the deterministic (MLR or MLR3) and the stochastic (wn) components. Detail descriptions are listed in Table 1. Blue lines show results from the deterministic component only, while gray shadings are probabilities from the 400 realizations from the stochastic system. Gray-shaded boundaries represent the 0th, 10th, 25th, 75th, 90th, and 100th percentiles.

forcing term by a parametric distribution fitted to the residuals, we draw from the training period errors conditional on the initial intensity V_0 , based on our finding that the error depends on the initial storm intensity (Lee et al. 2015). This stochastic forcing is not correlated in time, and we refer to it as MLR&wn, where wn stands for white noise.

The gray area in Fig. 2a shows the PDF of Typhoon Fanapi's lifetime intensity based on 400 realizations of the MLR&wn model. The observed intensity is at the edge of the 90th percentile of the distribution of realizations. Applied to all storms from 2000 to 2012, the MLR&wn simulated global LMI distribution includes major storms but underestimates their frequency. Only a few storms become intense TCs, and the LMI distribution is unimodal (gray lines in Fig. 4). Compared to the observed LMI distribution (black line), the MLR&wn

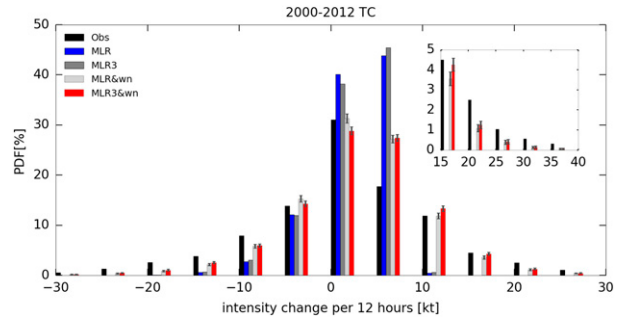


FIG. 3. (a) PDF of the simulated and observed 12-h global TC intensity change. Data are binned in 5-kt interval with the observed data plotted on the starting value. For stochastic systems, the ensemble averaged data are used and the gray line shows the range from all members. Colors indicate different datasets: observations (black), MLR (blue), MLR3 (gray), MLR&wn (light gray), and MLR3&wn (red). Acronyms indicate the choice of the deterministic (MLR or MLR3) and the stochastic (wn) components. The inset shows a zoom-in of the right tail.

distribution is biased toward zero in part because the lowest LMI in the best-track datasets is 25 kt and in the MLR&wn is 15 kt. This LMI value corresponds to simulated storms that never intensify since 15 kt is the lowest initial intensity in the best track.

d. Further model improvement—nonlinear terms

Adding stochastic forcing to the MLR extends the upper range of simulated storm intensity from category 2 to category 5. However, the number of major storms is still underpredicted and there is a leftward bias in the simulated LMI distribution. To improve the model further, we reexamine the model assumptions.

We first examine the assumption of a linear relation between the predictors and TC intensity changes. We follow the same method used in Tippett et al. (2011), in which the coefficient for a predictor is allowed to vary with the value of the predictor. To the extent that the variation is small, the linear assumption is valid. Results (not shown) suggest that the linear assumption is adequate for all variables except dV/dt and dPI_{V_0} . The coefficient of dV/dt is smaller when dV/dt is negative (Fig. 5a), perhaps reflecting the fact that strong weakening is limited. When dV/dt is positive, its coefficient varies little. A positive relationship between dV/dt and its regression coefficient indicates that a quadratic term of dV/dt should be included in the MLR. Similarly, both quadratic and cubic terms in dPI_{V_0} should be included (Fig. 5b). With these nonlinear terms, the coefficients of the deterministic model are still estimated in a linear framework at each time step, and the resulting deterministic model is called MLR3 hereafter, with 3 indicating the highest power of predictor used. The deterministic forecast of Typhoon Fanapi by the MLR3 shows some

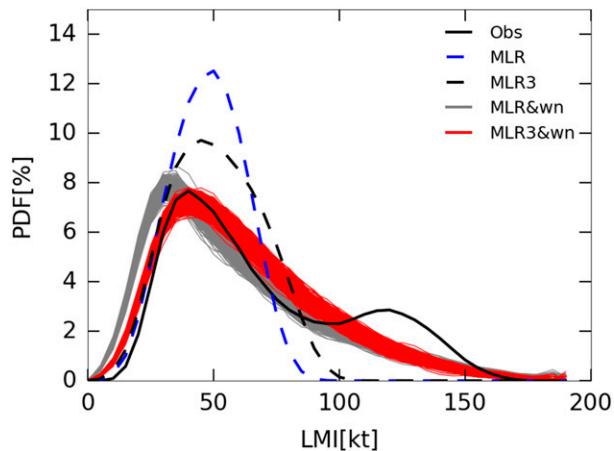


FIG. 4. PDFs of the global LMI from observations (solid black line), the deterministic models (dashed lines), and the stochastic systems (thin lines). Experiments are labeled in the legend.

improvement (blue line in Fig. 2b) but still does not capture the rapid intensification on 16–18 September. The MLR3 does not produce storms with LMI larger than 100 kt (black dashed line in Fig. 4), although it performs better than the MLR. In short-lead forecasting mode, the MAE from the MLR3 is close to that of the MLR in all basins with some modest (less than 10%) improvement in the ATL, EPC, IO, and SH basins and a little worsening in the WPC (not shown).

Adding the nonlinear terms in the deterministic part of equation leads to a new stochastic model called MLR3&wn, for which the range of the 25th–75th percentiles of the PDF of intensity for Typhoon Fanapi becomes narrower and shifts toward the observed intensity (Fig. 2b). The MLR3&wn successfully represents the first peak of the LMI PDF (gray lines in Fig. 4e), with the nonlinear terms preventing storms from being weaker than 15 kt and correcting the leftward bias in the simulated LMI distribution.

Next, we examine the white noise assumption in the MLR&wn model, in which the stochastic term is uncorrelated in time. In fact, forecast errors could be correlated in time. For example, the 12-h forecast errors of the Geophysical Fluid Dynamics Laboratory hurricane model and Hurricane Weather Research and Forecasting Model both have 12-h lag correlations of about 0.45. On the other hand, the 12-h lag correlation is close to 0.1 in the SHIPS model. Here, the 12-h lag correlation of the MLR model error is small and insignificant (black dotted line in Fig. 6). The lag correlation of ε^2 (black dashed line in Fig. 6) indicates that while the errors are uncorrelated in time, they are not independent. This dependence is related to the fact that the magnitude of errors depends on the initial intensity, which is serially

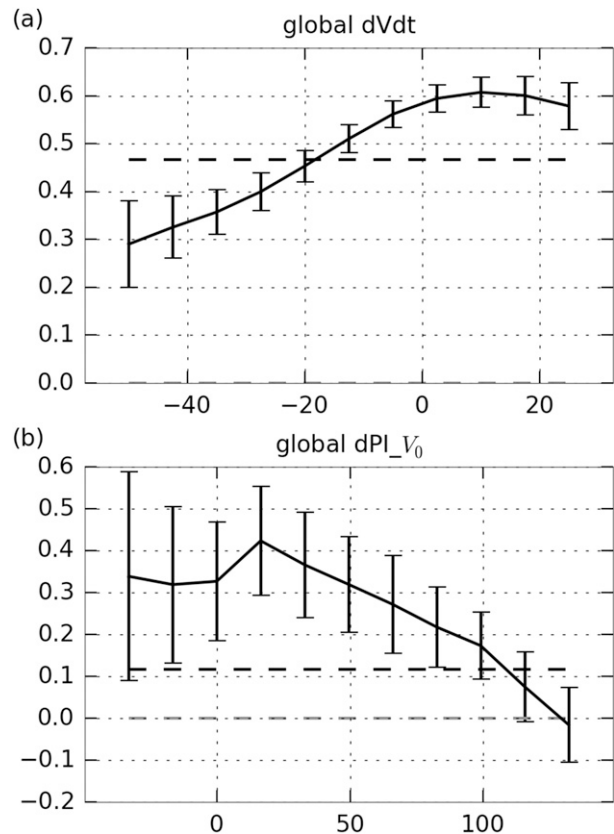


FIG. 5. Correlation coefficients of (a) $dVdt$ and (b) dPI_{V_0} from a locally weighted polynomial regression model. Black dashed line shows the value used in the MLR, and zero is marked by the gray dashed line.

correlated. We plan to include the effect of this correlation in our model in a future study.

e. Quantitative verification—RPSS

We additionally assess the probabilistic quality of the simulated distribution of LMI values. For each storm, we count the number of simulated LMI values in each 5-kt bin from 30 to 150 kt and compare this simulated LMI frequency distribution with the observed LMI value using the rank probability skill score (RPSS). The rank probability score (RPS) measures the sum-squared differences of forecast and observed cumulative occurrence. Smaller values of RPS indicate better forecasts. RPSS compares the average RPS to that of a baseline model, and positive RPSS values indicate more skill than the baseline model. The baseline model used here consists of a deterministic persistence model and the white noise stochastic model, denoted as persistence&wn. RPS and RPSS values are then stratified according to observed LMI value. RPS values from the persistence&wn model (black line in Fig. 7a) increase with increasing observed LMI, indicating the increasing

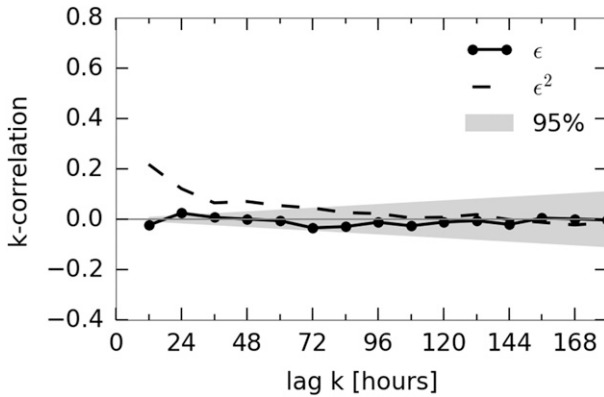


FIG. 6. Lag correlation coefficients of the forecast errors ϵ and ϵ^2 . The gray shading indicates the 95% confidence band.

difficulty in predicting LMI. A similar tendency is found in the RPS of the MLR&wn and MLR3&wn models. These two models are both skillful relative to the persistence&wn model (Fig. 7b), although the RPSS values drop with increasing LMI. The advantage of using the MLR as well as MLR3 (either together or individually) is modest for storms with LMI greater than 60 kt. Inclusion of nonlinear terms gives some small improvement (comparing the RPSS values of the MLR3&wn and MLR&wn). The basin-MLR3&wn (the system trained in individual basins; thin light-blue lines) has the same RPSS as the MLR3&wn (red lines), showing that, by this measure, there is no advantage in using basin-dependent models rather than a single global model.

One clear deficiency in the MLR3&wn-simulated LMI distribution (as well as in other stochastic simulations) is the missing second peak at high intensities. The observed bimodality in LMI is related to RI and reflects two types of storms: those that undergo RI sometime in their lifetime (RI storms) and those that do not (non-RI storms; Lee et al. 2016). The definition of RI that best achieves this separation is an increase of 35 kt in the maximum wind speed in 24 h. The MLR alone allows storms to intensify by at most $10 \text{ kt} (12 \text{ h})^{-1}$ (Fig. 3). The MLR3 shows no significant improvement on the intensification rate, but the stochastic error term allows intensification as large as $30 \text{ kt} (12 \text{ h})^{-1}$. The simulated frequency of intensification rates between 15 and $20 \text{ kt} (12 \text{ h})^{-1}$ is quite close to observations, but the frequency of higher intensification rates is still underpredicted. The addition of nonlinear terms primarily improves the distribution of non-RI storms with little change in the distribution of RI storms (not shown). The MLR3&wn reproduces the LMI distribution best in the ATL and EPC (Fig. 8). The deficiency of the model in generating RI storms stands out the most in WPC and SH because the frequencies of RI and intense storm occurrence are greatest there.

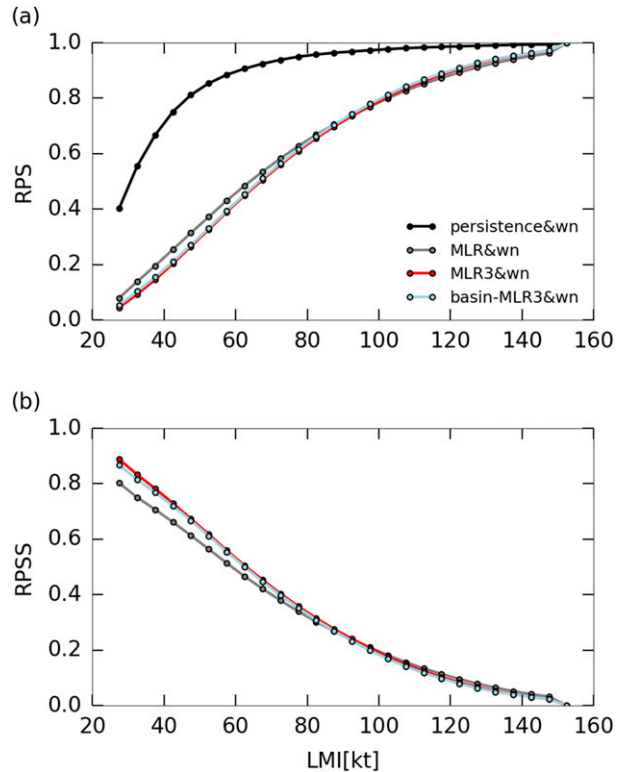


FIG. 7. (a) RPS for measuring the accuracy of probabilistic prediction of LMI from the persistence&wn (black), the MLR&wn (gray), the MLR3&wn (red), and the basin-MLR3&wn (light blue) as a function of LMI. (b) RPSS measuring the skill of probabilistic prediction of LMI from the MLR&wn, the MLR3&wn, and the basin-MLR3&wn with respect to the persistence&wn.

Models with dependent data (not shown) show similar improvement at each development stage, and the model limitation in representing the second peak of the LMI distribution remains. Similar to the short-lead models, the overall performance of the stochastic models using dependent data is close to that with independent data.

4. TC intensity distribution

The complete distribution of TC intensity V (over all 12-h periods, as opposed to LMI) is another meaningful characteristic of the TC intensity climatology. We begin with the PDF and cumulative density function (CDF) of V . The observed V PDF has a unimodal distribution with a peak at around tropical storm (TS) strength and is right skewed (black line in Fig. 9a). More than 50% of the recorded storm intensities are weaker than or at TS strength (black line in Fig. 9b); only about 20% of storm intensities exceed category 1 in the Saffir–Simpson scale and less than 10% of the recorded intensities are categories 3–5. Considering only the deterministic components, MLR (blue dashed line) and MLR3 (black dashed

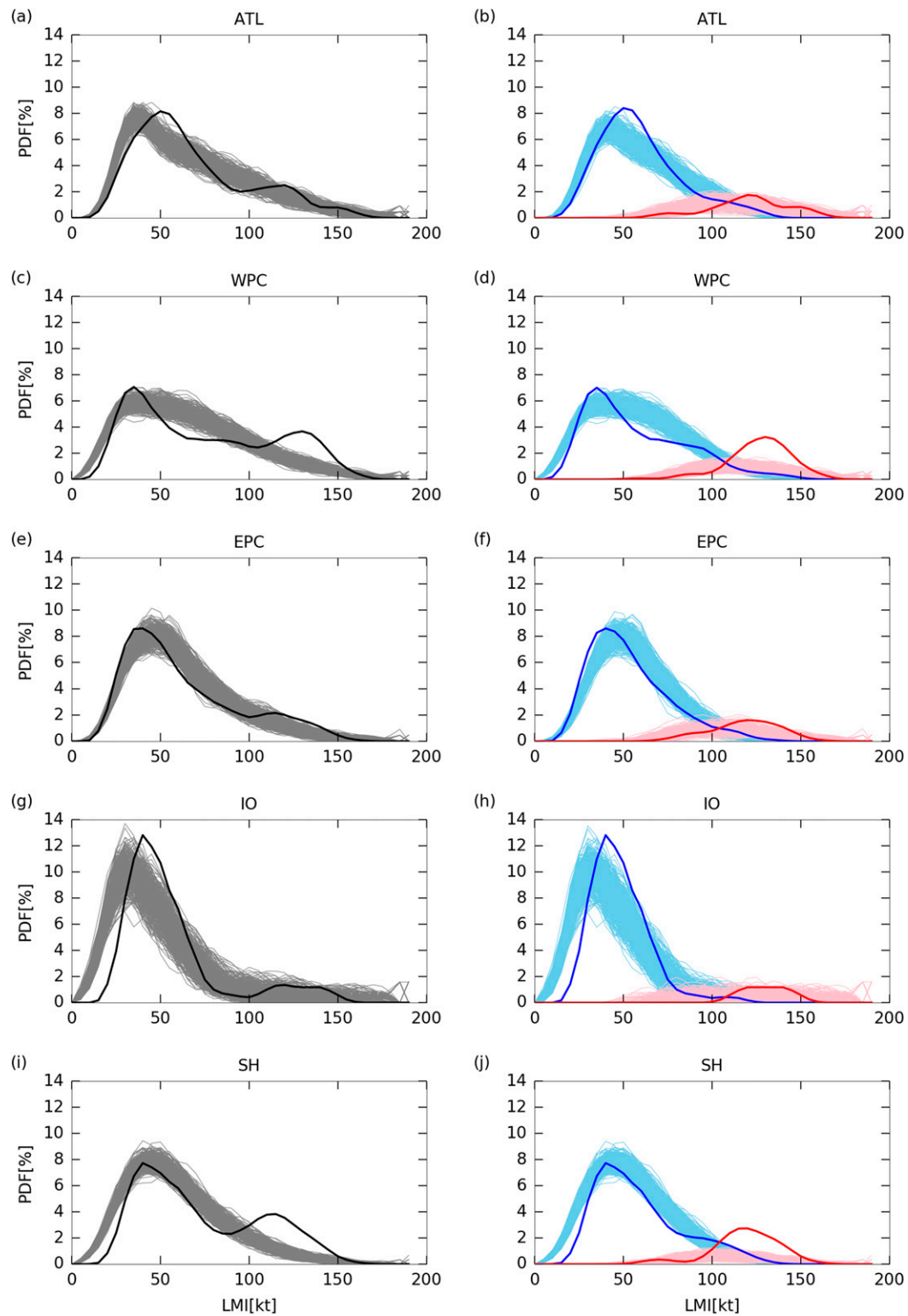


FIG. 8. PDFs of observed (black) and the MLR3&wn simulated (gray) LMI for (a) ATL, (c) WPC, (e) EPC, (g) IO, and (i) SH storms. (b),(d),(f),(h),(j) PDFs for RI (red and pink lines) and non-RI (blue and light blue lines) storms in each basin, respectively. Red and blue lines are from observations while pink and light blue lines are from simulations.

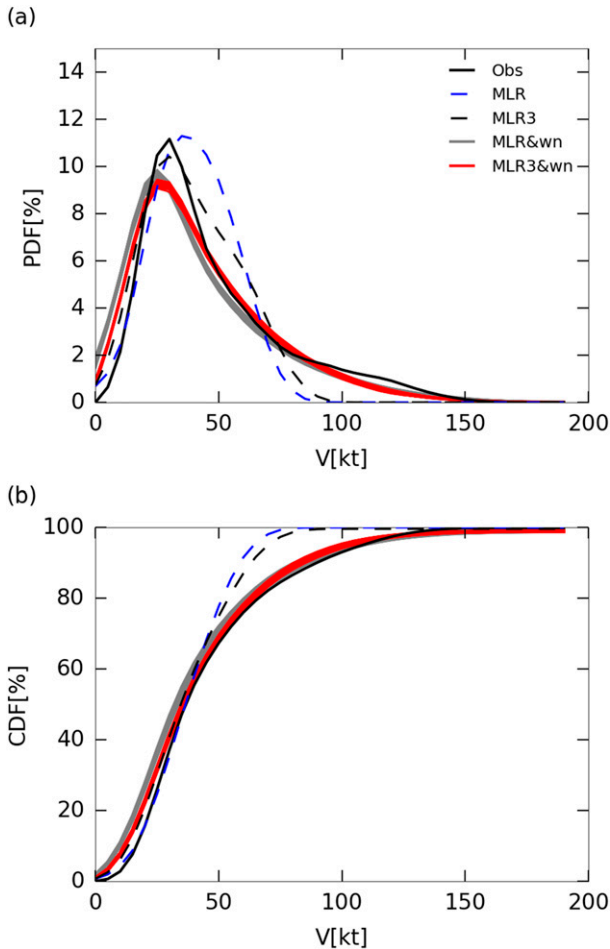


FIG. 9. (a) PDFs and (b) CDFs of TC intensity for all times, not only LMI as shown in Fig. 4.

line) produce more symmetric V PDFs than observed. Both these models only match the observed distribution for $V < 34$ kt. Addition of the stochastic components (MLR&wn and MLR3&wn) produces intensity PDFs that match the observed peak location and much of the tail behavior, at least for intensities up to 100 kt. The frequency of major storm intensity is still underpredicted, in agreement with the results from LMI analyses. The simulated intensity PDFs and CDFs have a leftward bias because some of the simulated storm intensities reach 0 kt while 10 kt is the lowest value in the best-track datasets. Differences in V distributions are fairly small among the stochastic simulations. The advantages of the models with stochastic forcing and nonlinear terms, while not dramatic, are apparent, especially over the range of 50–100 kt. RPSS analyses reflect their improvement on probabilistic storm intensity forecasts (not shown) similarly to Fig. 7.

Observed and modeled TC intensity along historical tracks are shown in Fig. 10 with the intensity indicated

by color. Only the maximum intensity at any given location is visible. Since historical tracks are used in all figures, the only differences in occurrence locations are due to a few instances where modeled storms dissipate early. Consistent with other observational studies (e.g., Knapp et al. 2010), the best-track data (Fig. 10a) show the highest intensities (red areas) where many major storms occur—east of Taiwan, northeast of the Philippines, and over the Caribbean Sea and the Gulf of Mexico. These areas are favorable for intense storms because the sea surface temperatures are usually warm, the upper-oceanic structure is favorable (e.g., Lin et al. 2008), and the large-scale vertical shear is usually weak (Chen et al. 2006). Intensities from the MLR3 (Fig. 10c) show a spatial pattern similar to that found in the observations but with weaker maximum intensities. The similarity between Figs. 10a and 10c indicates that the spatial distribution of strong storms is controlled to a large extent by the large-scale environment. One realization from the MLR3&wn is shown in Fig. 10b and indicates that the stochastic forcing improves the strength of the simulated intensity but slightly worsens some aspects of the spatial distribution. This realization has a couple of unrealistically strong storms in high latitudes in the ATL, a category-3 TC in the central Pacific, and almost no major storms in the Gulf of Mexico. The details of these deficiencies are different in other realizations, as the location of problematic storms is different because the stochastic forcing does not depend on the environment. The persistence&wn model also shows no coherency in the distribution of strong storms (Fig. 10d). In the persistence&wn model, the simulated storm intensities in the EPC and ATL are greater than in other basins because large training errors are more frequent in these two basins compared to the other three basins. The track map from ensemble mean of the MLR3&wn (not shown) is similar to that of the MLR3.

We plot the frequencies of storms with specified intensity ranges in Fig. 11 to compare the observed and simulated distribution of storm intensities at a given location. Figures 11a,c,e,g show maps of the observed annual frequency of TC intensities rated TD, TS, categories 1 and 2, and categories 3–5, respectively. For each range of intensities, this quantity is the product of two factors: average number of storms per unit area (TC frequency) and the fraction of storms in the given range. The TC frequency is the same in both observations and simulations because only historical tracks are used. The TC frequency term has a large role in setting the spatial pattern, especially for weak storms (Fig. 11a). With a higher intensity threshold, the area with relatively high probabilities is more restricted. Category-1 and

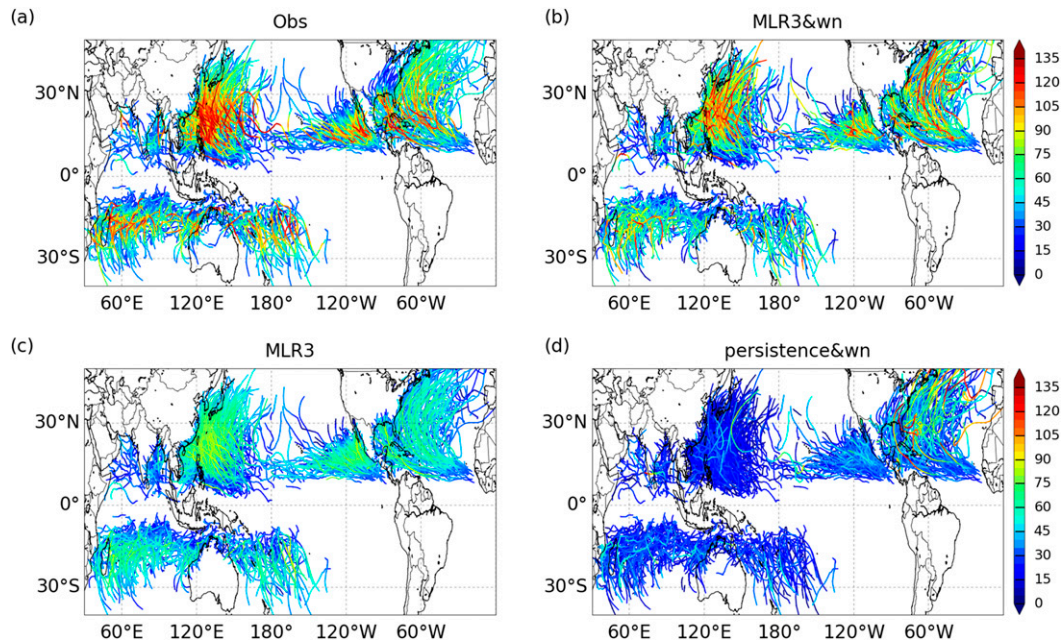


FIG. 10. Global TC tracks color coded by intensity; the intensities plotted are from (a) observations, (b) the MLR3&wn, (c) the MLR3, and (d) the persistence&wn. A member randomly selected from 400 realizations is shown in (b) and (d).

category-2 storms (Fig. 11e) are more likely to occur east of Taiwan, northeast of the Philippines, and west of Mexico. Over the eastern ATL and 15–20°S band, the probability of a category-1 or category-2 storm is 2%–5% $\text{yr}^{-1} (\text{° lat/lon})^{-2}$. The global frequency of major storms is low, but the frequency in the WPC is twice as large as in other basins (Fig. 11g). The MLR3&wn captures most of the features seen in observations (Figs. 11b,d,f,h). Both the MLR3 and the persistence&wn have the right pattern as well, especially for weak storms (Fig. 12). However, the quantitative accuracy in MLR3&wn-simulated probabilities is due to the well-simulated TC intensities. MLR3, with only a deterministic component, gives almost zero probabilities for major storms while it overestimates the frequency of TD, TS, and category-1 and category-2 storms. Persistence&wn underestimates all categories except TD.

In summary, the deterministic (MLR3) model alone is able to represent much of the TC intensity spatial structure. The stochastic terms significantly improve the simulated intensity distribution, at the cost of slightly worsening the simulated spatial structure. MLR3&wn successfully models the observed distribution of TC intensity, both the overall V distribution (PDF and CDF), as well as the spatial patterns. The intensity PDF, compared to the LMI distribution, is less sensitive to model configuration and therefore is a less stringent metric for evaluating our simulation of the TC intensity climatology.

5. Summary

We have developed autoregressive (AR) models for simulating the climatological distribution of TC intensity. The AR models consist of a multiple linear regression (MLR) deterministic component and a stochastic component. The deterministic component is an extension of that from our previous study (Lee et al. 2015), which described a model for TC intensity as a function of environmental variables with given observational tracks. Low-frequency (i.e., monthly averaged) reanalysis from ERA-Interim and best-track data from the NHC and JTWC from 1981 to 1999 are used for model development, and data from 2000 to 2012 are used for model verification. We simulate the complete life cycle of TC intensity given observed genesis location, first recorded storm intensity, and track. The key climatological features discussed here are the distributions of lifetime maximum intensity (LMI), as well as the full (not just maxima) distribution of TC intensities V .

The model is developed globally. We show that, for our application, there is little disadvantage in using a single global model instead of the basin-dependent ones typically developed for operational statistical forecasting. A practical benefit of the global model is that sample size for estimating parameters is increased, which has a benefit in the Indian Ocean where there are relatively few storms. Moreover, the fact that a single

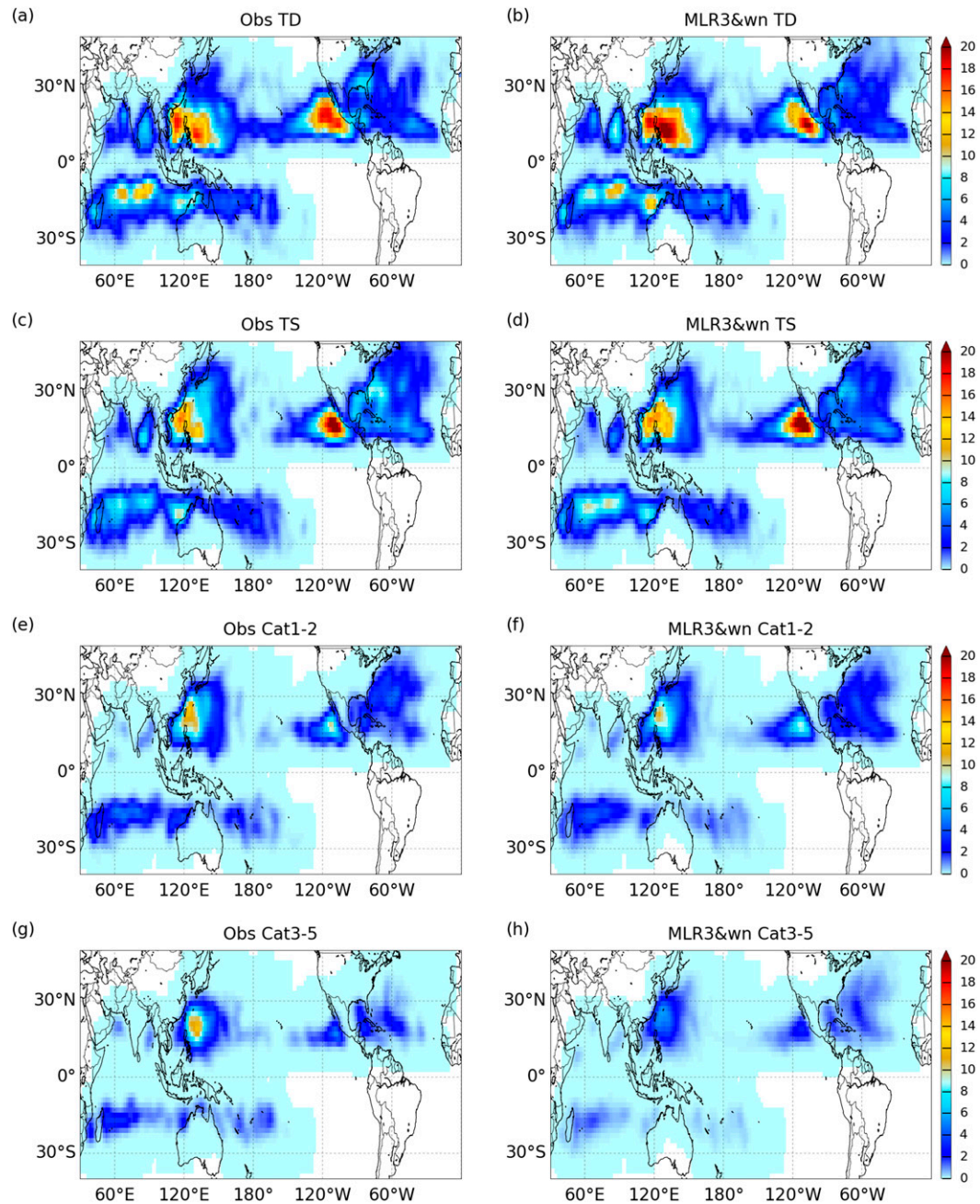


FIG. 11. Normalized probability (probability per year and $^{\circ}$ lat/lon) for storms with various strengths: (a), (b) TD; (c),(d) TS; (e),(f) categories 1 and 2; and (g),(h) categories 3–5 of the Saffir–Simpson scale (left) from observations and (right) from all 400 realizations in the MLR3&wn.

global model is effective provides some evidence that robust physical relationships between TC intensity and environment are being represented. Furthermore, as our goal is climate downscaling and the differences between basins in future climates might not be the same as they are in the current climate, the extra basin-dependent tuning could potentially introduce additional errors. A minimal set of predictors is used, including three

environmental conditions [potential intensity (PI), deep-layer mean shear (SHRD), and midlevel relative humidity (midRH)] and three storm state quantities [initial storm intensity (V_0), change in storm intensity in previous 12 h ($dVdt$), and storm translation speed (trSpeed)]. PI enters the model in the form of the difference between PI and V_0 (dPI_{V_0}). The influence of land is included in a simple way.

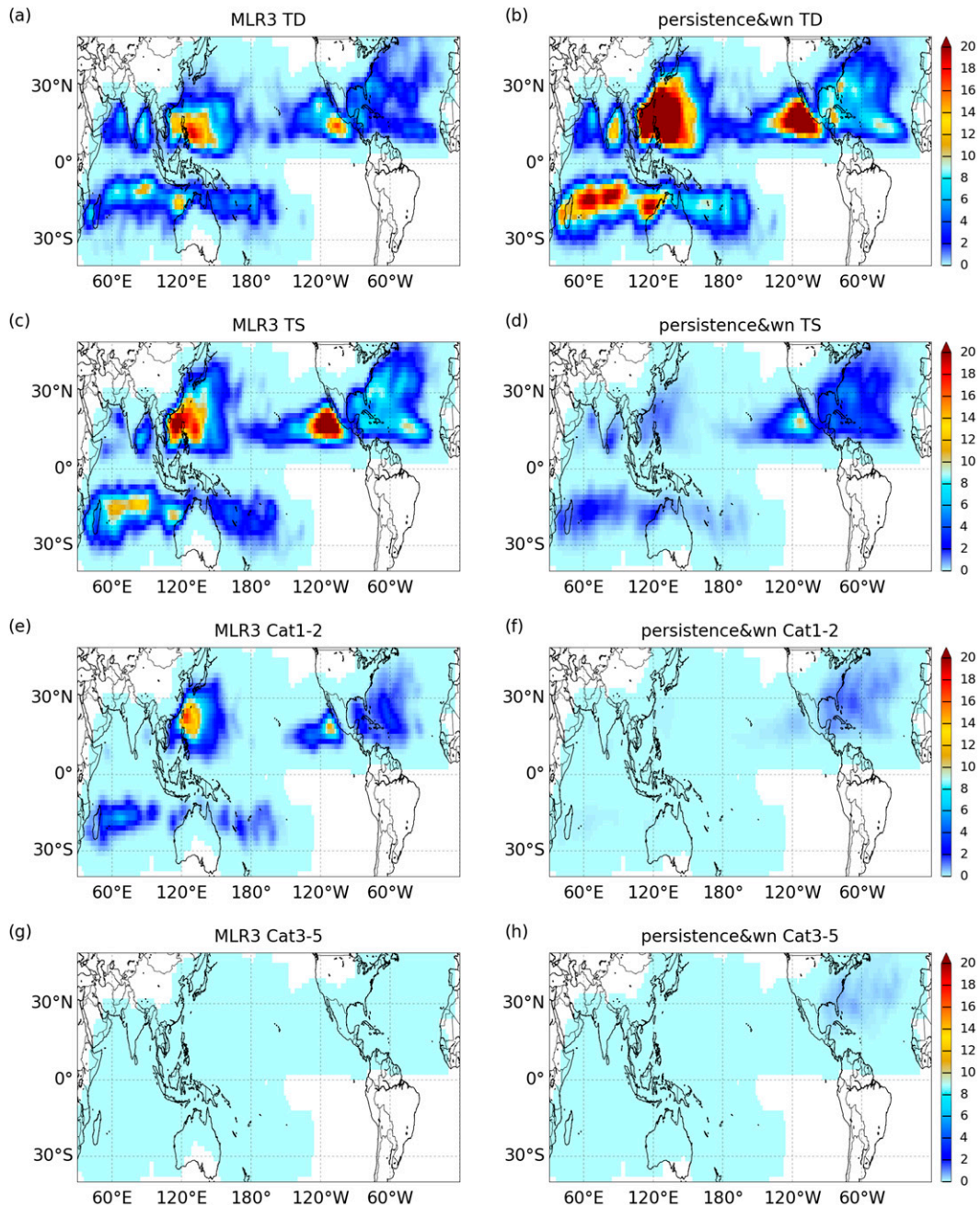


FIG. 12. As in Fig. 11, but for results from the MLR3 and the persistence&wn.

Without stochastic forcing, the simulated distribution of LMI has a negative bias, containing almost no storms with intensities greater than 100 kt and no simulated intensification rates greater than $10 \text{ kt} (12 \text{ h})^{-1}$. The systematic negative bias is significantly reduced by the addition of stochastic forcing. The stochastic forcing is constructed from the residuals of the MLR model. The addition of nonlinear terms for some predictors [$dVdt^2$, $(dPI_{V_0})^2$, and $(dPI_{V_0})^3$] in the deterministic component

(MLR3) allows some model storms to intensify more than $15 \text{ kt} (12 \text{ h})^{-1}$. Inclusion of a stochastic component (MLR3&wn) further improves the distribution of simulated intensification rates and the modeled number of category-3 storms as well. The number of simulated category-4 and category-5 TCs remains too low, however. As a consequence, the simulated global distribution of LMI is unimodal unlike the observed one, which is bimodal. This model deficiency is related to the

model's failure to simulate rapid intensification. With this caveat, MLR3&wn simulates the LMI distribution reasonably well both globally and regionally. The MLR3&wn model reproduces the full (as opposed to just lifetime maxima) observed distribution of storm intensities well, including the PDF, CDF, and spatial structure. Given suitable models for track and genesis, this autoregressive model (MLR3&wn) could be used to downscale tropical cyclone intensity using environmental variables from a low-resolution climate model, keeping in mind deficiencies in the number of rapidly intensifying and high-intensity storms.

Acknowledgments. The research was supported by the Office of Naval Research under the research grant of MURI (N00014-16-1-2073). We thank Dr. Dmitri Kondrashov from UCLA for his suggestion on the importance of including nonlinear terms in the linear regression model. Comments and suggestions from Dr. Chris Landsea and an anonymous reviewer are appreciated.

REFERENCES

- AIR, 2015: The AIR hurricane model: AIR Atlantic tropical cyclone model V15.0.1 as implemented in touchstone V3.0.0. AIR Worldwide Corporation Tech. Rep., 510 pp.
- Bechtold, P., N. Semane, P. Lopez, J.-P. Chaboureaud, A. Beljaars, and N. Bormann, 2014: Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, **71**, 734–753, doi:10.1175/JAS-D-13-0163.1.
- Behringer, D. W., and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. *Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, Seattle, WA, Amer. Meteor. Soc., 2.3. [Available online at <https://ams.confex.com/ams/pdfpapers/70720.pdf>.]
- Bister, M., and K. A. Emanuel, 2002: Low frequency variability of tropical cyclone potential intensity: 1. Interannual to interdecadal variability. *J. Geophys. Res.*, **107**, 4801, doi:10.1029/2001JD000776.
- Camargo, S. J., and A. A. Wing, 2016: Tropical cyclones in climate models. *Wiley Interdiscip. Rev.: Climate Change*, **7**, 211–237, doi:10.1002/wcc.373.
- , A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007: Tropical cyclone genesis potential index in climate models. *Tellus*, **59A**, 428–443, doi:10.1111/j.1600-0870.2007.00238.x.
- , M. K. Tippett, A. H. Sobel, G. A. Vecchi, and M. Zhao, 2014: Testing the performance of tropical cyclone genesis indices in future climates using the HiRAM model. *J. Climate*, **27**, 9171–9196, doi:10.1175/JCLI-D-13-00505.1.
- Chen, J.-H., and S.-J. Lin, 2013: Seasonal predictions of tropical cyclones using a 25-km-resolution general circulation model. *J. Climate*, **26**, 380–398, doi:10.1175/JCLI-D-12-00061.1.
- Chen, S. S., J. A. Knaff, and F. D. Marks, 2006: Effects of vertical wind shear and storm motion on tropical cyclone rainfall asymmetries deduced from TRMM. *Mon. Wea. Rev.*, **134**, 3190–3208, doi:10.1175/MWR3245.1.
- Chu, J.-H., C. R. Sampson, A. S. Levine, and E. Fukada, 2002: The Joint Typhoon Warning Center tropical cyclone best-tracks, 1945–2000. Naval Research Laboratory Tech. Rep. NRL/MR/7540-02-16, 22 pp.
- Davis, K., X. Zeng, and E. A. Ritchie, 2015: A new statistical model for predicting seasonal North Atlantic hurricane activity. *Wea. Forecasting*, **30**, 730–741, doi:10.1175/WAF-D-14-00156.1.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim re-analysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- DeMaria, M., M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the Statistical Hurricane Intensity Prediction Scheme (SHIPS). *Wea. Forecasting*, **20**, 531–543, doi:10.1175/WAF862.1.
- ECMWF, 2013: ERA Interim, monthly means of daily means. ECMWF, accessed 2013. [Available online at <http://apps.ecmwf.int/datasets/data/interim-full-mods/levtype=pl/>.]
- Emanuel, K. A., 1986: An air–sea interaction theory for tropical cyclones. Part I: Steady-state maintenance. *J. Atmos. Sci.*, **43**, 585–604, doi:10.1175/1520-0469(1986)043<0585:AASITF>2.0.CO;2.
- , 2000: A statistical analysis of tropical cyclone intensity. *Mon. Wea. Rev.*, **128**, 1139–1152, doi:10.1175/1520-0493(2000)128<1139:ASAOTC>2.0.CO;2.
- , 2006: Climate and tropical cyclone activity: A new model downscaling approach. *J. Climate*, **19**, 4797–4802, doi:10.1175/JCLI3908.1.
- , 2008: The hurricane–climate connection. *Bull. Amer. Meteor. Soc.*, **89**, ES10–ES20, doi:10.1175/BAMS-89-5-Emanuel.
- , 2013: Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proc. Natl. Acad. Sci. USA*, **110**, 12 219–12 224, doi:10.1073/pnas.1301293110.
- , S. Ravela, E. Vivant, and C. Risi, 2006: A statistical deterministic approach to hurricane risk assessment. *Bull. Amer. Meteor. Soc.*, **87**, 299–314, doi:10.1175/BAMS-87-3-299.
- Feuerstein, B., N. Dotzek, and J. Grieser, 2005: Assessing a tornado climatology from global tornado intensity distributions. *J. Climate*, **18**, 585–596, doi:10.1175/JCLI-3285.1.
- Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part I: El Niño and 30 mb quasi-biennial oscillation influences. *Mon. Wea. Rev.*, **112**, 1649–1668, doi:10.1175/1520-0493(1984)112<1649:ASHFPI>2.0.CO;2.
- Hall, T., and S. Jewson, 2007: Statistical modelling of North Atlantic tropical cyclone tracks. *Tellus*, **59A**, 486–498, doi:10.1111/j.1600-0870.2007.00240.x.
- , and E. Yonekura, 2013: North American tropical cyclone landfall and SST: A statistical model study. *J. Climate*, **26**, 8422–8439, doi:10.1175/JCLI-D-12-00756.1.
- Hristopulos, D. T., and V. Mouslopoulou, 2013: Strength statistics and the distribution of earthquake interevent times. *Physica A*, **392**, 485–496, doi:10.1016/j.physa.2012.09.011.
- Huang, P., I. I. Lin, C. Chou, and R.-H. Huang, 2015: Change in ocean subsurface environment to suppress tropical cyclone intensification under global warming. *Nat. Commun.*, **6**, 7188, doi:10.1038/ncomms8188.
- IRI Data Library, 2013: NOAA NCEP EMC CMB GODAS Monthly Global Ocean Data Assimilation System. International Research Institute for Climate and Society, accessed 2013. [Available online at <http://iridl.ldeo.columbia.edu/SOURCES/NOAA/NCEP/EMC/CMB/GODAS/>.]
- Johnson, N. C., and S.-P. Xie, 2010: Changes in the sea surface temperature threshold for tropical convection. *Nat. Geosci.*, **3**, 842–845, doi:10.1038/ngeo1008.

- JTWC, 2014: Tropical cyclone best track data site. Joint Typhoon Warning Center, accessed 2014. [Available online at http://www.usno.navy.mil/NOOC/nmfc-ph/RSS/jtwc/best_tracks/.]
- Kaplan, J., and M. DeMaria, 1995: A simple empirical model for predicting the decay of tropical cyclone winds after landfall. *J. Appl. Meteor.*, **34**, 2499–2512, doi:10.1175/1520-0450(1995)034<2499:ASEMFP>2.0.CO;2.
- Klotzbach, P. J., and W. M. Gray, 2009: Twenty-five years of Atlantic basin seasonal hurricane forecasts (1984–2008). *Geophys. Res. Lett.*, **36**, L09711, doi:10.1029/2009GL037580.
- , and E. C. J. Oliver, 2015: Modulation of Atlantic basin tropical cyclone activity by the Madden–Julian oscillation (MJO) from 1905 to 2011. *J. Climate*, **28**, 204–217, doi:10.1175/JCLI-D-14-00509.1.
- Knaff, J. A., and C. R. Sampson, 2009: Southern Hemisphere tropical cyclone intensity forecast methods used at the Joint Typhoon Warning Center. Part II: Statistical dynamical forecasts. *Aust. Meteor. Oceanogr. J.*, **58**, 9–18.
- , —, and M. DeMaria, 2005: An operational statistical typhoon intensity prediction scheme for the western North Pacific. *Wea. Forecasting*, **20**, 688–699, doi:10.1175/WAF863.1.
- Knapp, K. R., M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, 2010: The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Amer. Meteor. Soc.*, **91**, 363–376, doi:10.1175/2009BAMS2755.1.
- Knutson, T. R., J. J. Sirutis, S. T. Garner, G. A. Vecchi, and I. M. Held, 2008: Simulated reduction in Atlantic hurricane frequency under twenty-first-century warming condition. *Nat. Geosci.*, **1**, 359–364, doi:10.1038/ngeo202.
- Kossin, J. P., T. L. Olander, and K. R. Knapp, 2013: Trend analysis with a new global record of tropical cyclone intensity. *J. Climate*, **26**, 9960–9976, doi:10.1175/JCLI-D-13-00262.1.
- , K. A. Emanuel, and G. A. Vecchi, 2014: The poleward migration of the location of tropical cyclone maximum intensity. *Nature*, **509**, 349–352, doi:10.1038/nature13278.
- 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, doi:10.1175/MWR-D-12-00254.1.
- Lee, C.-Y., and S. S. Chen, 2014: Stable boundary layer and its impact on tropical cyclone structure in a coupled atmosphere–ocean model. *Mon. Wea. Rev.*, **142**, 1927–1944, doi:10.1175/MWR-D-13-00122.1.
- , M. K. Tippett, S. J. Camargo, and A. H. Sobel, 2015: Probabilistic multiple-linear regression modeling for tropical cyclone intensity. *Mon. Wea. Rev.*, **143**, 933–954, doi:10.1175/MWR-D-14-00171.1.
- , —, A. H. Sobel, and S. J. Camargo, 2016: Rapid intensification and the bimodal distribution of tropical cyclone intensity. *Nat. Commun.*, **7**, 10625, doi:10.1038/ncomms10625.
- Lin, I. I., C.-C. Wu, I.-F. Pun, and D.-S. Ko, 2008: Upper-ocean thermal structure and the western North Pacific category 5 typhoons. Part I: Ocean features and the category 5 typhoons' intensification. *Mon. Wea. Rev.*, **136**, 3288–3306, doi:10.1175/2008MWR2277.1.
- Manganello, J. V., and Coauthors, 2012: Tropical cyclone climatology in a 10-km global atmospheric GCM: Toward weather-resolving climate modeling. *J. Climate*, **25**, 3867–3893, doi:10.1175/JCLI-D-11-00346.1.
- , and Coauthors, 2014: Future changes in the western North Pacific tropical cyclone activity projected by a multidecadal simulation with a 16-km global atmospheric GCM. *J. Climate*, **27**, 7622–7646, doi:10.1175/JCLI-D-13-00678.1.
- Murakami, H., and Coauthors, 2012: Future changes in tropical cyclone activity projected by the new high-resolution MRI-AGCM. *J. Climate*, **25**, 3237–3260, doi:10.1175/JCLI-D-11-00415.1.
- , and Coauthors, 2015: Simulation and prediction of category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. *J. Climate*, **28**, 9058–9079, doi:10.1175/JCLI-D-15-0216.1.
- Nakamura, J., U. Lall, Y. Kushnir, and B. Rajagopalan, 2015: HITS: Hurricane intensity and track simulator with North Atlantic Ocean applications for risk assessment. *J. Appl. Meteor. Climatol.*, **54**, 1620–1636, doi:10.1175/JAMC-D-14-0141.1.
- NHC, 2013: Best track data (HURDAT2). National Hurricane Center, accessed 2013. [Available online at <http://www.nhc.noaa.gov/data/#hurdat>.]
- Oouchi, K., J. Yosimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006: Tropical cyclone climatology in a global-warming climate as simulated in a 20 km mesh global atmospheric model: Frequency and wind intensity analyses. *J. Meteor. Soc. Japan*, **84**, 259–276, doi:10.2151/jmsj.84.259.
- Price, J. F., 2009: Metrics of hurricane–ocean interaction: Vertically-integrated or vertically-averaged ocean temperature? *Ocean Sci.*, **5**, 351–368, doi:10.5194/os-5-351-2009.
- Schade, L. R., and K. A. Emanuel, 1999: The ocean's effect on the intensity of tropical cyclones: Results from a simple coupled atmosphere–ocean model. *J. Atmos. Sci.*, **56**, 642–651, doi:10.1175/1520-0469(1999)056<0642:TOSEOT>2.0.CO;2.
- Tippett, M. K., S. J. Camargo, and A. H. Sobel, 2011: A Poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *J. Climate*, **24**, 2335–2357, doi:10.1175/2010JCLI3811.1.
- Vitart, F., S. Woolnough, M. A. Balmaseda, and A. M. Tompkins, 2007: Monthly forecast of the Madden–Julian oscillation using a coupled GCM. *Mon. Wea. Rev.*, **135**, 2700–2715, doi:10.1175/MWR3415.1.
- Wang, H., J.-K. E. Schemm, A. Kumar, W. Wang, L. Long, M. Chelliah, G. D. Bell, and P. Peng, 2009: A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical seasonal forecast. *J. Climate*, **22**, 4481–4500, doi:10.1175/2009JCLI2753.1.
- Wu, C.-C., C.-Y. Lee, and I. I. Lin, 2007: The effect of the ocean eddy on tropical cyclone intensity. *J. Atmos. Sci.*, **64**, 3562–3578, doi:10.1175/JAS4051.1.
- Wu, L., and Coauthors, 2014: Simulations of the present and late-twenty-first-century western North Pacific tropical cyclone activity using a regional model. *J. Climate*, **27**, 3405–3424, doi:10.1175/JCLI-D-12-00830.1.
- Yoshimura, J., M. Sugi, and A. Noda, 2006: Influence of greenhouse warming on tropical cyclone frequency. *J. Meteor. Soc. Japan*, **84**, 405–428, doi:10.2151/jmsj.84.405.
- Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *J. Climate*, **22**, 6653–6678, doi:10.1175/2009JCLI3049.1.
- , —, and G. A. Vecchi, 2010: Retrospective forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Mon. Wea. Rev.*, **138**, 3858–3868, doi:10.1175/2010MWR3366.1.