Lagrangian Analysis of Tropical Cyclone Genesis Simulated by General Circulation Models Compared with Observations

CHANWOO SONG, SUNGSU PARK, AND JIHOON SHIN
School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea

(Manuscript received 31 July 2019, in final form 19 February 2020)

ABSTRACT

As a contribution to understanding the genesis of tropical cyclones (TCs), we compared TC genesis processes simulated by the Seoul National University Atmosphere Model version 0 with a Unified Convection Scheme (SAM0) and the Community Atmosphere Model version 5 (CAM5) with those from the ERA-Interim reanalysis (ERA-Interim, hereafter ERAI) and best track observations. In contrast to previous studies that estimated the TC genesis potential using the Eulerian mean environmental conditions, we calculated the probability of a pre-existing weak cyclonic embryo vortex (EV) developing into a TC by analyzing changes in the environmental conditions along the EV trajectories. Our analysis indicates that the spatial distribution and annual cycles of TCs obtained from the SAM0 and ERAI are similar to those obtained from the best track observation data. With the exception of the mesoscale convective organization and associated variables, most environmental variables along the trajectories of DEVs (EVs developing into TCs) showed monotonic variations. When EVs were born, environmental conditions of DEVs were significantly different from those of nondeveloping EVs, allowing for the prediction of TC genesis. In general, TC genesis probability increased as the environment became more cyclonic, moist, unstable, and with a weaker wind shear. Rapidly strengthening EVs were more likely to develop into TCs. SAM0 and ERAI have the same combination of environmental variables with the best prediction skill for TC genesis—absolute vorticity at 850 hPa, column saturation deficit, sea surface temperature, vertical shear of horizontal winds between 200 and 850 hPa, and latitude—with similar sensitivities to individual environmental variables, indicating that SAM0 well simulates the observed TC genesis processes.

1. Introduction

Tropical cyclones (TCs) are one of the most destructive weather systems and have tremendous societal impacts. During the last two centuries, about 1.9 million lives have been lost due to TCs (Nicholls et al. 1995) and, even with the advances in technology, the overall damage from TCs is still increasing (Rappaport and Fernández-Partagás 1997; Doocy et al. 2013). Timely and appropriate public preparation, such as construction of shelter or evacuation, based on an accurate risk warning can reduce damages caused by TCs (e.g., Shultz et al. 2005). Once formed, precursor tropical disturbances can develop into fully destructive TCs within two to three days under certain environmental conditions (Zehr 1976), and the accurate forecasting of TC genesis, if possible, can provide communities with additional time to prepare for upcoming threats.

An accurate understanding of the genesis and intensification of TCs is a challenging subject (Gray 1975; Camargo et al. 2007; Nolan 2007). Based on the analysis of observation data, numerical simulations, and theoretical studies, various hypotheses have been suggested regarding the physical mechanisms responsible for the genesis and intensification of TCs. According to hydrostatic and mass continuity equations, a depression in sea level pressure (SLP) is induced by column-integrated divergence of air masses, which in turn is driven by differential vertical and horizontal heating. Thus, understanding the sources of diabatic heating is key to understanding TC genesis. Miller (1958) noted that the minimum SLP in TC is determined by environmental conditions such as sea surface temperature (SST) and near-surface relative humidity. Charney and Eliassen (1964) explained the cumulus- and cyclone-scale motions using the concept of conditional instability of the second kind (CISK) that involves positive feedback processes between latent heat released by convection and frictional convergence of moist air by horizontal
winds. Ooyama (1964, 1969) proposed a similar hypothesis with different convective closures. Charney and Eliassen (1964) assumed that diabatic heating by convection is proportional to moisture convergence while Ooyama (1964) assumed that frictional convergence controls convective mass flux. Moreover, Emanuel (1986) suggested a wind-induced surface heat exchange (WISHE) hypothesis that emphasizes the air–sea interaction involving a positive feedback between surface wind and upward moisture flux from the ocean (Emanuel 1989; Bister and Emanuel 1998; Emanuel 2003). Holland (1997) suggested that an upper-tropospheric warm core can maintain an upward motion and is therefore important for the genesis and intensification of TCs. More recently, Montgomery and Smith (2010, 2014) showed that TCs can intensify without a feedback process between the warm core and upward motion. Although there is a general consensus regarding the fact that latent heat release by convection is key for the genesis and intensification of TCs (Kuo 1965), the detailed mechanism remains debatable (e.g., Fritz and Wang (2014)).

Large-scale atmospheric and oceanic environmental conditions can influence the genesis and intensification of TCs (Miller 1958; Gray 1979; Camargo et al. 2009). As a potential predictor for the TC genesis, Emanuel and Nolan (2004) suggested a so-called genesis potential index (GPI) that combines four environmental variables (absolute vorticity at 850 hPa, relative humidity at 600 hPa, maximum potential intensity, and the magnitude of the vertical shear of horizontal winds between 200 and 850 hPa), and Camargo et al. (2007) showed that the GPI successfully reproduces the observed annual cycle of climatological TC genesis. Wang and Moon (2017) suggested another TC genesis index by combining three anomalous environmental conditions with respect to their climatologies (pressure vertical velocity at 500 hPa, the product of planetary vorticity and relative vorticity at 850 hPa, and the directional vertical shear of zonal winds between 200 and 850 hPa) to diagnose intraseasonal variations of TC genesis.

Interannual variations of TCs are influenced by various climate variability modes such as El Niño–Southern Oscillation (ENSO; Chan 1985; Wang and Chan 2002; Camargo and Sobel 2005; Sobel and Camargo 2005), monsoons (Lander 1994; Briegel and Frank 1997; Chen et al. 1998; Ritchie and Holland 1999; Chen et al. 2004), and the Madden–Julian oscillation (MJO; Madden and Julian 1971; Nakazawa 1988; Liebmann et al. 1994; Duvel 2015). Camargo et al. (2007, 2009) showed that the variations of TCs associated with ENSO and the MJO can be explained by changes in GPI.

In addition to the aforementioned observational studies, efforts have been made to understand TCs using various types of numerical models, ranging from analytical models (Emanuel 1986; Holland 1997; Montgomery and Smith 2010) and large-eddy simulations (Rotunno et al. 2009) to cloud-resolving models (Braun 2002). These studies provide detailed insights into the genesis and intensification processes of TCs, which are difficult to obtain from the observational studies alone. However, most modeling works assumed idealized conditions (e.g., flat sea surface with uniform temperature) and focused on a few specific cases, such that they have a limitation in providing a comprehensive view of the relevant TC processes operating over the entire globe. A general circulation model (GCM) that can simulate the observed properties of TCs (e.g., spatial distribution, annual cycle, and strength) reasonably well, if any, could supplement the weakness of the previous observational and modeling studies. In particular, a GCM can be used to study changes of TC activity in a warm climate induced by the increased concentration of carbon dioxide (e.g., Knutson et al. 2019).

In our study, we analyze the processes related to the genesis of TCs as simulated by a state-of-the-art GCM with a unified convection scheme (UNICON; Park 2014a,b) at a 0.5° latitude × 0.5° longitude horizontal resolution. Park et al. (2019) showed that a GCM together with the UNICON parameterization scheme improved simulations of ENSO, MJO, the diurnal cycle of precipitation, and tropical cyclones with a decent mean climate, which were difficult to simulate in previous GCMs. In contrast to previous observational studies that analyzed climatological TC genesis using environmental conditions at a fixed location (i.e., the Eulerian approach), we calculate the potential for a pre-existing, weak cyclonic embryo vortex (EV) to develop into a TC by analyzing various environmental conditions along the trajectories of individual EVs (i.e., the Lagrangian approach). Since TCs are not stationary in space but move around, the Lagrangian approach can provide more accurate information on the environmental conditions that control changes in TC properties than the Eulerian approach at various TC stages.

The structure of this paper is as follows. Section 2 provides a description of the GCM simulations and observation data including various analysis methods. In section 3, we show that the frequency of TC genesis simulated by a GCM with the UNICON parameterization scheme is similar to that of the observation data, and we examine the impacts of various environmental conditions on the TC genesis probability as well as the transition speed of EVs into TCs. Finally, an optimal set of environmental variables for the prediction of TC genesis is suggested. A summary and conclusions are provided in section 4.
2. Data and analysis methods

a. Observation data and model simulations

The observed TC data used in our study were obtained from the International Best Track Archive for Climate Stewardship (IBTrACS; Knapp et al. 2010) for the period January 1979–December 2014 (36 years). The IBTrACS compiles various information along the track of individual TCs [e.g., minimum SLP and maximum sustained 10-m wind speed around the center of individual TC (U10\text{max})] from two different U.S. agencies, the Joint Typhoon Warning Center (JTWC; Chu et al. (2002)) and the National Hurricane Center (NHC; Landsea and Franklin 2013). Following the Saffir–Simpson scale (Saffir 1977; Simpson and Saffir 1974), the observed TCs are distinguished by their wind intensity as tropical depressions (TD with U10\text{max} < 17 m s\(^{-1}\)), tropical storms (TS with 17 \leq U10\text{max} < 34 m s\(^{-1}\)), and one of five categories of tropical cyclones (TCs 1–5 with 34 m s\(^{-1}\) \leq U10\text{max}). Henceforth, the observed TCs used in our study refer to storms with at least tropical storm intensity (i.e., U10\text{max} \geq 17 m s\(^{-1}\)).

The annual mean genesis number of the observed TCs over the globe for the analysis period (January 1979–December 2014) was 84.5.

In our study, we use the Seoul National University Earth System Model version 0 with a Unified Convection Scheme (SEM0-UNICON; Park et al. 2019), as our GCM. The SEM0-UNICON is a fully coupled GCM, consisting of the Seoul National University Atmosphere Model version 0 with the UNICON parameterization scheme (SAM0-UNICON) for the atmosphere model, and other components (e.g., ocean, land, and sea ice models) identical to those of the Community Earth System Model version 1 (CESM1; Hurrell et al. 2013). The SAM0-UNICON (Park et al. 2019) is based on the Community Atmosphere Model version 5 (CAM5; Neale et al. 2010; Park et al. 2014) with the shallow (Park and Bretherton 2009) and deep convection schemes (Zhang and McFarlane 1995) replaced by the UNICON parameterization scheme (Park 2014a,b) and with a revised treatment of convective detrainment processes (Park et al. 2017). Park et al. (2019) showed that the global mean climate and ENSO as simulated by SAM0-UNICON/SEM0-UNICON were similar to those of the CAM5/CESM1; however, the SAM0-UNICON/SEM0-UNICON substantially improved the simulations of the MJO, diurnal cycle of precipitation, and TCs, with a significant reduction of individual short-wave and longwave aerosol indirect effects. Hereafter, the SAM0-UNICON and SEM0-UNICON will be simply referred to as SAM0 and SEM0, respectively.

Following the simulation settings described in Park et al. (2019), we ran the preindustrial coupled simulation with the SEM0 for 135 years at a horizontal resolution of 0.475° latitude \times 0.625° longitude (nominally, 0.5°) starting from the near-equilibrium state of the 0.95° latitude \times 1.25° longitude preindustrial coupled simulation. The last 100 years of SEM0 simulations are used for our analysis. To examine the impacts of atmosphere–ocean coupling and atmospheric convective parameterization on TC genesis, we also analyzed the Atmospheric Model Intercomparison Project (AMIP) simulations with SAM0 and CAM5 for a 36-yr period (January 1979–December 2014) at the same 0.5° horizontal resolution. All simulations were driven by the forcing datasets provided by phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al. 2016). To assess the accuracy of the GCM-simulated relationship between various environmental conditions and TC genesis, we repeated the same analysis using the ERA-Interim reanalysis (ERA-Interim, hereafter ERAI; Simmons et al. 2007) over the same 36-yr period at a 0.7° latitude \times 0.7° longitude horizontal resolution. Various output variables at 6-hourly time intervals were then analyzed. Murakami (2014) showed that the spatial distribution and annual cycle of TCs derived from the ERAI are similar to those from observation data, although the intensity of the derived TC was weaker than that of the observed TC, mainly due to the coarse horizontal resolution of the ERAI reanalysis (Schenkel and Hart 2012; Hodges et al. 2017). As will be explained in the following section, the detection criteria used to define TCs from the ERAI in our study are very similar to those used by Murakami (2014), with the exception of using warm core strength instead of temperature anomaly.

b. Detection of the embryo vortex and tropical cyclone

For all GCM simulations (SAM0/SEM0/CAM5), we defined the embryo vortex (EV) as a cyclonic vortex with a relative vorticity at 850 hPa (\(\xi_{850}\)) of larger than 2 CVU (1 CVU = 10\(^{-5}\) s\(^{-1}\)) for at least one day. Previous studies have used a TD as a precursor cyclone, rather than an EV, as defined using U10\text{max} from the IBTrACS mentioned above (Camargo and Sobel 2004; Peng et al. 2012; Fu et al. 2012; Smith and Montgomery 2012; Duvel 2015; Komaromi 2013; Zawislak and Zipser 2014). The global-mean annual TD genesis number (~97.4) is not much different from that of the observed tropical storms (~84.5) for a 36-yr period (January 1979–December 2014). As our goal was to forecast the formation of TCs to offer enough lead time for public preparation, we use an EV instead of TD as a precursor cyclone. The global-mean annual EV genesis number (~10000) is much larger than that of TDs. Through trial and error, we found that a
minimum EV detection threshold of 2 CVU is necessary to detect a well-defined cyclonic circulation without noise in the tropical region. We repeated the same analysis with different threshold values (e.g., 4 and 6 CVU) and obtained similar qualitative results (not shown).

To obtain insights into the physical processes responsible for the formation of TCs, we divided all EVs into the following two groups: EVs developing into TCs (DEVs) and EVs not developing into TCs (NEVs). By extending the TC detection methods suggested by Walsh (1997), Hodges et al. (2003), Bengtsson et al. (2006), and Murakami and Sugi (2010), we detected TCs in our GCM simulations using the following five criteria: 1) $\xi_{850} \geq 12.7$ CVU, 2) warm-core strength, $\xi_{850} - \xi_{250} \approx 7.8$ CVU, 3) horizontal wind speed at 850 hPa is larger than that at 300 hPa, 4) conditions 1–3 are satisfied for at least 1.5 days, consecutively, and 5) the land fraction in the 5° latitude × 5° longitude offset grid box in which the TC is generated is smaller than 0.5. The earliest time point at which these five conditions are simultaneously satisfied is defined as the TC genesis time $t_{TC}$. Detection thresholds for $\xi_{850}$ and warm-core strength were obtained through trial and error, to ensure that the global-mean annual number of simulated TCs from the SAM0 AMIP simulations is similar to that of the observed TCs (~85 TCs yr⁻¹) as obtained from IBTrACs. The third criterion effectively filters out the midlatitude cyclones (Murakami and Sugi 2010) while the last criterion was included to filter out strong cyclonic vortices simulated by GCMs in the far western Indian ocean near Saudi Arabia where the observed frequency of TC genesis is very low. All EVs and TCs over the oceans between 35°S and 35°N are analyzed and the identical detection criteria were used for all GCM simulations. We used similar detection criteria for ERAI, with the exception that the threshold value of $\xi_{850}$ for EV detection and the threshold values of $\xi_{850}$ and warm-core strength for TC detection were set to 1.0, 7.9, and 4.9 CVU, respectively. In addition, these conditions were required to persist at least for one day, resulting in the annual mean TC and EV numbers of similar to those of IBTrACs and SAM0. From a composite analysis at $t_{TC}$, we checked that the minimum sea level pressure is located at the center of TC, indicating that our vorticity-based detection of TC genesis works well (not shown).

c. Environmental variables for Lagrangian analyses

Following the EV track, we calculated the mean environmental conditions averaged over the 10.5° latitude × 10.5° longitude domain centered at individual EVs every six hours. We did not filter out the contribution of EVs from the mean environmental conditions defined above for all thermodynamic variables. However, in order to check whether our analysis was sensitive to domain size, we repeated the same analysis by defining the mean environmental conditions over the 5.5° latitude × 5.5° longitude domain and found that the results were similar to those from the 10.5° latitude × 10.5° longitude domain. This implies that the chosen domain size was large enough for the contribution of EVs to be negligible. The analyzed environmental variables are summarized in Table 1 and grouped into the following five categories: dynamics, moisture, vertical stability, wind shear, and other. These variables include the variables related to the TC genesis potential index (GPI) proposed by Emanuel and Nolan (2004) ($\eta_{850}$, RH600, PI, and $V_s$) and the intraseasonal GPI suggested by Wang and Moon (2017) ($\eta_{850}$, $\phi_{850}$, and $V_{z, 2}$). Peng et al. (2012) and Fu et al. (2012) reported that the precipitation rate at the surface (PRECT) is a good predictor for the TC genesis. Based on the fact that the main source of energy for TCs is the latent heat released from the phase change of water vapor, we analyzed various moisture variables as listed in Table 1, where $q_v$ is water vapor specific humidity, $q_s$ is saturation specific humidity, $p_s$ is surface pressure, and $\Delta T_{v}$ ($\Delta T_{\phi}$) used for HMC700 and MC is the difference of zonal (meridional) moisture fluxes along the boundaries on the eastern (northern) and western (southern) edges of the 10.5° × 10.5° domain, with $\Delta x$ ($\Delta y$) denoting the zonal (meridional) width of the 10.5° × 10.5° domain. The stability variables $p_{min}$, convective available potential energy (CAPE), and SST comprise the maximum potential intensity (PI; Emanuel 1995; Bister and Emanuel 1998). The formation of TCs is associated with the onset of deep convection and the release of latent heat (Kuo 1965). The convection schemes developed by Kuo (1965, 1974) and Tiedtke (1989) use horizontal moisture convergence (e.g., HMC700 and MC) to diagnose the onset of deep convection. Park (2014a,b) showed that the parameterization of mesoscale convective organized flow driven by subgrid cold pools and its feedback on convective updraft is important for simulations of shallow to deep convection transition and other related phenomena, such as the MJO and diurnal cycle of precipitation. The variable $\Omega$ as computed in UNICON (subgrid mesoscale convective organization) is defined as the fractional area of subgrid cold pools within the planetary boundary layer (PBL) with a virtual potential temperature $\theta_v$ at least 0.2 K lower than the grid-mean $\theta_v$ [Eq. (72) of Park (2014a)], which controls the radius and mass flux of convective updraft plumes [Eqs. (75) and (76) of Park (2014a)]. In addition, $\Delta \theta_{c, 41}$ and $\Delta q_{c, 41}$ are the perturbations of condensate potential temperature $[\theta_c = \theta - (L_v/C_p/\pi)q_{c} - (L_i/C_p/\pi)q_{i}]$ and total
specific humidity \(q_i = q_v + q_l + q_i\) with respect to the grid-mean values in the rising branch of subgrid mesoscale organized flow, and both are used to compute the initial thermodynamic properties of convective updraft plumes at the surface [Eq. (73) of Park (2014a)]. The variable \(WC = \xi_{850} - \xi_{250}\) refers to the warm core strength. All variables listed in Table 1 are the mean environmental variables averaged over the 10.5° × 10.5° domain centered at individually tracked EVs.

d. Research questions

Our study focuses on answering the following questions:

1) Do the GCMs (SAM0, SEM0, and CAM5) used in our study reasonably simulate key properties of observed TCs?

2) What environmental conditions control the probability for EVs to develop into TCs?

3) What environmental conditions control the transition speed of EVs into TCs?

4) What is the best combination of environmental variables for the prediction of TC genesis from EVs?

To answer the first question, we compared GCM-simulated TC tracks and passage frequencies, near-surface wind speeds along individual TC tracks (Fig. 1), and the

<table>
<thead>
<tr>
<th>Group</th>
<th>Variable</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamics</td>
<td>(\xi_{850})</td>
<td>Relative vorticity at 850 hPa</td>
<td>Gray (1975)</td>
</tr>
<tr>
<td></td>
<td>(\eta_{850})</td>
<td>Absolute vorticity at 850 hPa = planetary vorticity ((f) + \xi_{850})</td>
<td>Emanuel and Nolan (2004)</td>
</tr>
<tr>
<td></td>
<td>(\eta_{850})</td>
<td>(f \times \xi_{850})</td>
<td>Emanuel and Nolan (2004)</td>
</tr>
<tr>
<td>Moisture</td>
<td>RH_{600}</td>
<td>Relative humidity at 600 hPa</td>
<td>Gray (1975); Emanuel and Nolan (2004)</td>
</tr>
<tr>
<td></td>
<td>(\delta_{600})</td>
<td>Saturation deficit at 600 hPa (\delta_{600} = q_{v,600} - q_{s,600})</td>
<td>Bretherton et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>(Q_v)</td>
<td>Column water vapor, (Q_v = \int_{p_s}^{p_*} q_v(dp/g))</td>
<td>Kuo (1965); Tiedtke (1989)</td>
</tr>
<tr>
<td></td>
<td>(r_e)</td>
<td>Column relative humidity, (r_e = \left[ \int_{p_s}^{p_<em>} q_v(dp/g) \right] / \left[ \int_{p_s}^{p_</em>} q_v(dp/g) \right] )</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>(\delta_e)</td>
<td>Column saturation deficit, (\delta_e = \left[ \int_{p_s}^{p_<em>} (q_v - q_v)(dp/g) \right] / \left[ \int_{p_s}^{p_</em>} q_v(dp/g) \right] )</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td>Moisture</td>
<td>MSE_{v}</td>
<td>Column moist static energy, (MSE_v = \int_{p_s}^{p_*} (C_v T + g z + L_v q_v)(dp/g))</td>
<td>Charney and Eliassen (1964)</td>
</tr>
<tr>
<td></td>
<td>MC</td>
<td>Moisture convergence, (MC = \int_{p_s}^{p_*} (\Delta \nabla q / \Delta x + \Delta \nabla q / \Delta y)(dp/g) + LHF/L_v)</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>LHC_{700}</td>
<td>Horizontal MC below 700 hPa (MC_{700} = \int_{p_s}^{p_*} (\Delta \nabla q / \Delta x + \Delta \nabla q / \Delta y)(dp/g))</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>(P_{min})</td>
<td>Minimum potential surface pressure</td>
<td>Emanuel (1995); Emanuel and Nolan (2004)</td>
</tr>
<tr>
<td></td>
<td>CAPE</td>
<td>Convective available potential energy, (CAPE = (g/\theta_{ref}) \int_{\Delta z}^{\Delta z} (\theta_e - \bar{\theta}_e) \Delta z)</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>SST</td>
<td>Sea surface temperature</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td>Shear</td>
<td>(V_s)</td>
<td>Wind shear between 200 and 850 hPa (V_s = \sqrt{(u_{200} - u_{850})^2 + (v_{200} - v_{850})^2})</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>(V_{zs})</td>
<td>Directional zonal wind shear between 200 and 850 hPa (V_{zs} = u_{200} - u_{850})</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td>Other</td>
<td>(\omega_{500})</td>
<td>Pressure vertical velocity at 500 hPa</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>PRECT</td>
<td>Precipitation rate at the surface</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>PRECC</td>
<td>Convective precipitation rate at the surface</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>PRECTL</td>
<td>Stratiform precipitation rate at the surface</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>(\Delta\theta_e)</td>
<td>Perturbation of (\theta_e) in the subgrid mesoscale upflow within the PBL</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>(\Delta q_{250})</td>
<td>Perturbation of (q_2) in the subgrid mesoscale upflow within the PBL</td>
<td>Emanuel (1986)</td>
</tr>
<tr>
<td></td>
<td>PBLH</td>
<td>Planetary boundary layer height</td>
<td>Emanuel (1986)</td>
</tr>
</tbody>
</table>

Table 1. Environmental variables used for TC genesis analysis. The 28 variables are collated into five groups: dynamics, moisture, stability, shear, and other. References using corresponding variables for diagnosing TC genesis or other climate research are also indicated.
annual cycle of TC genesis number (Fig. 2) with those from observation data. To answer the second question, we classified all EVs as either DEVs or NEVs, and compared their spatial distributions (Fig. 3), Lagrangian evolution of various environmental conditions along their tracks (Fig. 4), and the probability density functions (PDFs) of various environmental conditions at the moment of initial EV identification, $t_{EV}$ (Fig. 5). Then, we calculated the frequency of TC genesis (GFq) from EVs as a function of various environmental conditions at $t_{EV}$ (Fig. 6). To answer the third question, we calculated the time taken for individual
DEVs to develop into TCs ($\Delta t_{ET} = t_{TC} - t_{EV}$, where $t_{TC}$ is the time when an EV is identified as a TC) and examined the relationship between $\Delta t_{ET}$ and various environmental conditions at $t_{EV}$ (Fig. 7). To answer the fourth question, we examined all possible sets of environmental variables generated by the combination of environmental variables in each five categories shown in Table 1; identified the best combination diagnosing TC genesis probability (GPr) at $t_{EV}$ (Table 2); evaluated its performance in terms of predicting the spatial distribution (Fig. 8) and temporal variations in the frequency of TC genesis (Fig. 9); and examined the impacts of individual environmental variables of the best combination on TC genesis (Fig. 10).
3. Results

a. Comparison of TC genesis frequencies from simulations and observation data

Figure 1 shows the annual tracks and passage frequencies of the observed TDs and TCs at various development stages (TS and five categories of tropical cyclones) obtained from the IBTrACS during January 1979–December 2014 (Figs. 1a,b), and simulated DEVs and TCs from ERAI (Figs. 1c,d), SAM0 (Figs. 1e,f), and CAM5 (Figs. 1i,j) during the same period, and SEM0 (Figs. 1g,h) during the equilibrium state (years 100–135). In contrast to ERAI and GCMs, TCs in the IBTrACS are defined using maximum near-surface wind speed, based on the Saffir-Simpson scale (Saffir 1977; Simpson and Saffir 1974), and the number of TDs in the IBTrACS was much smaller than the number of EVs from the GCMs (~10 000 yr⁻¹) or that the candidate storm is no longer “tropical” after having completed the extratropical transition in the mid-latitude regions. These biases are similar to ones in other climate models with a similar horizontal resolution (Shaevitz et al. 2014; Walsh et al. 2013; Wehner et al. 2015; Reed et al. 2015; Kim et al. 2014; Scoccimarro et al. 2017; Li and Sriver 2018; Wing et al. 2019).

As we used identical criteria for detecting EVs and TCs in all GCM simulations, the differences between results from SAM0 and CAM5 reflect the impacts of convection scheme on TC activity. Similarly, the differences between results from SEM0 and SAM0 reflect the impacts of atmosphere–ocean coupling, greenhouse

**FIG. 3.** Global distributions of (left) the annual-mean number of embryo vortices (nEV), (center) the number of the EVs that develop into TCs (nDEV), and (right) TC genesis frequency (GFq = nDEV/nEV) in the 5° latitude × 5° longitude grid boxes obtained from (a)–(c) ERAI, (d)–(f) SAM0, (g)–(i) SEM0, and (j)–(l) CAM5. Global mean values of nEV and nDEV are shown at the top of individual plots. Pattern correlations and RMSE errors between ERAI and GCMs are shown on the top left and top right of individual plots, respectively. In these plots, nDEV and GFq in each grid box are defined based on the location where EVs were initially identified, rather than where they developed into TCs. Only the grid boxes with a number of EVs larger than 100 are used for computing the pattern correlations and RMSE errors.
gases, and aerosols on TC activity. Over the North Atlantic Ocean, SEM0 simulates weaker TC activity than SAM0, which is due in part to colder SST in SEM0 (Park et al. 2019). The impact of the convection scheme on global TC activity was quite substantial. CAM5 simulated too many TCs near the equator and too few TCs over the North Atlantic Ocean, which was remedied by implementing the UNICON parameterization scheme. UNICON also suppressed unrealistic TC activity in CAM5 over the Southern Hemispheric eastern equatorial Pacific and Atlantic oceans in association with the improved simulation of the double intertropical convergence zone (ITCZ). Overall, SAM0 shows the best performance in reproducing the observed TC tracks, even better than ERAI.
Figure 2 shows the annual cycle of the TC genesis number (nDEV for ERAI and GCMs) in several selected regions indicated in Fig. 1. The annual cycle of TC number generated by CAM5 is somewhat unrealistic with too many TCs during the boreal winter in both hemispheres when minimal observations take place. Even with adjusted TC detection criteria that make the CAM5-simulated global-mean annual TC genesis number similar to that of the observed TCs ($\sim$85 TCs yr$^{-1}$), CAM5 still simulates too many TCs during the boreal winter with a spatial distribution of TC passage frequency similar to Fig. 1j (not shown). SAM0 remedied this problem and simulated more accurate annual cycles than CAM5. With an exception over the North Atlantic...
Ocean where SEM0 had negative SST biases (Park et al. 2019), SEM0 generated more TCs than SAM0. Given the fact that the observed SST during the preindustrial period is generally colder than the present-day SST, the substantial increase of the global mean TC number by 40% from SAM0 to SEM0 implies an important role of atmosphere–ocean coupling and global warming on TC genesis. Overall, SAM0 and ERAI reproduced the observed annual cycle of TC genesis number from IBTrACS very well.

Figure 3 shows the annual-mean number of EVs (nEV), DEVs (nDEV), and TC genesis frequency.
(GFq = nDEV/nEV) in each 5° latitude × 5° longitude grid box, as obtained from ERAI and GCMs. The EVs are frequently generated in the ITCZ/SPCZ and the central-eastern portions of major oceans, with the former frequently developing into TCs while the latter rarely developed into TCs. Over the western Pacific Ocean, EVs were less frequently initiated than in the other regions. However, in the off-equatorial portion of the western Pacific Ocean, EVs frequently developed into TCs, particularly in the Northern Hemisphere between 5° and 15°N. This indicates that an efficient partitioning of environmental energy into a small number of EVs is
more important for the genesis of TCs than the existence of many EVs. Over the western Pacific Ocean, atmosphere–ocean coupling and global cooling substantially increased nDEV and GFq, presumably due to the combined effects of local feedback and nonlocal interactions. Another region with a very high GFq was the far eastern portion of the Pacific ITCZ. As indicated by global pattern correlations and root-mean-square-errors (RMSEs), SAM0 reproduced spatial distributions of nDEV and GFq from ERAI much better than CAM5, which simulated too many DEVs in the eastern Pacific ITCZ but too few DEVs over the western Pacific Ocean. Based on the similarity between the observed and simulated annual cycles as well as the global distributions of

**Table 2.** List of combinations of environmental variables with the highest prediction skill for TC genesis frequency. Prediction skill is defined as the difference of the predicted TC genesis probabilities between DEVs and NEVs [prediction skill = GPr(DEVs) – GPr(NEVs). See section 3d for more details). To construct the multivariate joint bins in each combination, each environmental variable was divided into the five bins so that the same number of EVs were assigned to each bin, resulting in 5^5 multivariate joint bins, for the combinations consisting of five environmental variables. In this table, V5 (V5c) is the best combination, consisting of five environmental variables for ERAI, SAM0, and SEM0 (CAM5). As a reference, the combinations constituting the genesis potential indices of EN04 (η850, RH600, PI, and Vs) and WM17 (η850, vs, and ω500) are also shown.

<table>
<thead>
<tr>
<th>Data</th>
<th>Combination</th>
<th>Prediction skill</th>
<th>Dynamics</th>
<th>Moisture</th>
<th>Stability</th>
<th>Shear</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERAI</td>
<td>V5</td>
<td>0.138</td>
<td>η850</td>
<td>δc</td>
<td>SST</td>
<td>Vs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EN04</td>
<td>0.080</td>
<td>η850</td>
<td>RH600</td>
<td>PI</td>
<td>Vs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WM17</td>
<td>0.081</td>
<td>η850</td>
<td>δc</td>
<td>SST</td>
<td>Vs</td>
<td>ω500</td>
</tr>
<tr>
<td>SAM0</td>
<td>V5</td>
<td>0.151</td>
<td>η850</td>
<td>δc</td>
<td>SST</td>
<td>Vs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EN04</td>
<td>0.102</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WM17</td>
<td>0.086</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEM0</td>
<td>V5</td>
<td>0.173</td>
<td>η850</td>
<td>δc</td>
<td>SST</td>
<td>Vs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EN04</td>
<td>0.107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WM17</td>
<td>0.103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAM5</td>
<td>V5c</td>
<td>0.099</td>
<td>η850</td>
<td>δc</td>
<td>SST</td>
<td>Vs</td>
<td>PRECL</td>
</tr>
<tr>
<td></td>
<td>EN04</td>
<td>0.042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>WM17</td>
<td>0.042</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 8.** (left) Global distributions of annual-mean TC genesis probability (GPr) as predicted by (top) V5, (middle) EN04, and (bottom) WM17 for ERAI, as well as (right) their differences with annual-mean TC genesis frequency (GFq) shown in Fig. 3c for ERAI. The pattern correlations and RMSE errors between GFq and individual GPr are also shown. Grid boxes with a number of EVs less than 100 are not used for the analysis.
FIG. 9. (a)–(i) Annual cycle and (j) interannual variation of annual TC genesis number (= GFq × nEV) obtained from ERAI (bars) and the predicted annual TC genesis number (= GPr × nEV) obtained from V5 (red), EN04 (blue), and WM17 (green). In addition, correlation coefficients and RMSE errors (within parentheses) between bars and individual lines are also shown when the related environmental variable is divided into 5 and 10 bins. Solid (dashed) lines indicate the results using 5 (10) bins.
TCs, we expect that our analysis can provide useful insights into the TC genesis processes observed in nature. In the following section, we will examine statistical relationships between TC genesis and various environmental conditions using the results from ERAI and GCMs.

b. Environmental conditions controlling the TC genesis

Figure 4 shows composite environmental conditions averaged over the $10.5^\circ$ latitude $\times 10.5^\circ$ longitude domain along the trajectories of DEVs and NEVs as obtained from ERAI, SAM0, SEM0, and CAM5. Some variables showed systematic differences in their magnitudes between datasets (e.g., dynamic variables of CAM5; CAPE of ERAI; and $d_c$, $v_{500}$, and PRECT) but their variations along the trajectories were similar regardless of the datasets. Almost all environmental variables are good predictors for forecasting TC genesis during the 3–5 days prior to TC genesis at $t_{TC}$ (see the shaded period in Fig. 4). That is, the differences between the environmental variables for DEVs and NEVs during the 3–5-day period
prior to TC genesis are statistically significant at the 95% confidence level from the Student’s $t$ test. The forecast skill tends to decrease with an increase in lead time from TC genesis. From the shaded period to $t_{TC}$, most environmental variables along DEVs showed monotonic variations. However, variables associated with mesoscale convective organization $(\Omega, \Delta \theta, \Delta \phi, \Delta q, \Delta T)$ reached peak values one or two days prior to TC genesis. This indicates that mesoscale convective organization associated with subgrid cold pools accelerates early stage TC genesis, but slows down final stage TC genesis, presumably, by stabilizing the environment through a compensating subsidence induced by strong convective updraft. This negative feedback and associated weakening of deep convective activity persisted for several days, even after TC genesis, as shown by the rapid decreases in both $\Omega$ and convective precipitation rate (not shown) after $t_{TC}$.

Figure 5 shows the probability density functions of environmental variables at $t_{EV}$ for DEV (solid) and NEV (dotted), respectively. Some variables show systematic offsets in the position of peak PDF values between datasets. For example, in the case of pressure vertical velocity at 500 hPa (Fig. 5m) and total precipitation rate (Fig. 5n), the peak PDFs of DEV in CAM5/ERAI are separated from those of SAM0/SEM0. However, the overall PDF shapes in various datasets are similar to each other. As indicated by the $t$ scores, the differences in the mean values of environmental variables between DEV and NEV simulated by CAM5 are less pronounced than those of ERAI/SAM0/SEM0. With the exception of CAPE and $V_850$, the PDFs of DEVs can be represented by a single mode Gaussian distribution with a certain skewness and are well separated from those of NEVs in most variables in a statistically significant way. This indicates that some of the destinations of the EVs were already determined when they were born, facilitating the forecasting of TC genesis at $t_{EV}$. The strongest predictor for TC genesis in each group is the relative vorticity at 850 hPa $\xi_{850}$ for dynamics, column water vapor $Q_c$ for moisture, SST for stability, $V_{zs}$ for shear, and convective precipitation rate for other group (not shown), which, as will be shown later, differ from the variables constituting the strongest combination for TC genesis $(\eta_{850}, \delta, \sigma, \Omega, \rho, \phi, \theta, \Delta q)$ (i.e., strong PI values near the equator), the negative correlation between PI and GFq (Fig. 6h). However, if PI is anticorrelated with $/j$ (i.e., strong PI values near the equator), the negative correlation between PI and GFq in the regime of $PI > 75$ may simply reflect the well-known fact that TC genesis is suppressed near the equator due to the absence of the planetary vorticity (Fig. 6p). The decrease of GFq with an increase in $\xi_{850}$, $R_{H600}$, $Q_c$, HMC700, SST, PRECT, and $\Omega$, but increases with a decrease in $\delta_{600}$, $\delta$, $V_850$, and $\omega_{500}$. Conceptually, this is consistent with the environmental conditions associated with the onset of deep convection, indicating that TC genesis is accompanied by strong convective activity. The ERAI and SAM0 show a tendency for GFq to approach a certain asymptotic value as $\eta_{850}$ increases, which is consistent with the findings of Tippett et al. (2011). Some variables exert a nonmonotonic (e.g., PI, $V_{zs}$, and $/j$) or seemingly nonmonotonic (e.g., CAPE and SST) influence on GFq. As an example, GFq increases with PI when $PI < 75 \text{ m s}^{-1}$ but decreases when $PI > 75 \text{ m s}^{-1}$. If PI is uncorrelated with other environmental variables, this indicates a nonmonotonic dependency of TC genesis on PI (Fig. 6h). However, if $\eta$ is uncorrelated with $/j$ (i.e., strong PI values near the equator), the negative correlation between PI and GFq (Fig. 6h) may simply reflect the well-known fact that TC genesis is suppressed near the equator. The decrease of GFq with an increase in $\eta_{850}$ and SST values may reflect a similar relationship. This indicates that when combined with other appropriate variables, environmental variables with nonmonotonic impacts on GFq can be good predictors for TC genesis. As will be shown in section 3d, the combination containing SST and $/j$ shows the strongest predictive power for TC genesis. Consistent with $V_850$, GFq decreases as the magnitude of $V_{zs}$ increases (note $V_s = |V_{zs}|$; see Table 1), but interestingly, maximum GFq occurs when $V_{zs}$ is slightly negative rather than zero, indicating that TC genesis is more likely to occur under a weak easterly wind shear with height. The strongest TC genesis predictors for each group are $\xi_{850}$ for dynamics, $Q_c$ for moisture, SST for stability, $V_{zs}$ for shear, and convective precipitation rate (not shown) for other group, and these are similar to the predictors in Fig. 5.

By comparing developing TDs with nondeveloping TDs, Peng et al. (2012) and Fu et al. (2012) noted that developing TDs have a higher water vapor content, more precipitation, stronger relative vorticity in the
lower troposphere, and warmer SST than nondeveloping TDs. Furthermore, Komaromi (2013) and Zawislak and Zipser (2014) showed that developing TDs have a higher midtropospheric relative humidity, more precipitable water, and weaker vertical wind shear than nondeveloping TDs. These results are consistent with our findings that GFq increases with a decrease of \(V_c\) and the increase in \(\xi_{850}\), \(RH_{850}\), \(Q_v\), and PRECT. Komaromi (2013) and Smith and Montgomery (2012) found that developing TDs have lower CAPE values than nondeveloping TDs. In addition, Nolan (2007) reported that an increase in atmospheric instability does not necessarily lead to an increase in TC genesis. These studies support our finding that GFq decreases as CAPE and PI increase in strong CAPE > 500 and PI > 75 conditions.

A couple of notable aspects are as follows: for some variables, CAM5 simulations showed different behaviors from those of ERAI/SAM0/SEM0 (e.g., \(\eta_{850}\), \(RH_{850}\), \(\theta_{850}\), PI, SST, and \([f]\)) indicating that the reason why CAM5 simulates unrealistic TCs (see Figs. 1 and 2) is that it simulates abnormal relationships between environmental conditions and TC genesis, and not just abnormal environmental conditions; a systematic offset exists between ERAI and SAM0/SEM0 in the magnitudes of some environmental variables (e.g., \(\delta_j\)), such that a parametric forecast equation for TC genesis derived from SAM0/SEM0 based on these environmental variables cannot be directly applied to the prediction of TC genesis in ERAI; CAM5 generated more TCs in low-latitude regions than ERAI/SAM0/SEM0 (Fig. 6p); and the sensitivity of GFq to stratiform precipitation rate (not shown) differs greatly between datasets, which may serve as an important clue to improving future physics parameterization schemes of SAM0/SEM0.

c. Environmental conditions controlling the transition speed of EVs into TCs

Figure 7 shows the dependency of the transition time of DEV into TC (\(\Delta t_{ET} = t_{TC} - t_{EV}\)) on environmental conditions at \(t_{EV}\). To compare the strength of various environmental variables as single predictors controlling the transition speed of DEV into TC, we showed the difference of the averages of three maximum and three minimum \(\Delta t_{EV}\), for each environmental variable in each plot. In general, the environmental variables exerting monotonic impacts on GFq have similar monotonic impacts on \(\Delta t_{ET}\) with opposite signs, indicating that rapidly intensifying EVs have a higher probability of developing into TCs. The only exception is \(V_c\), which has a low impact on \(\Delta t_{ET}\) even though it strongly suppresses GFq. On the other hand, the environmental variables that exert nonmonotonic impacts on GFq (e.g., stability factors and \([f]\)) do not have clear impacts on \(\Delta t_{ET}\), although there is a tendency for \(\Delta t_{ET}\) to decrease as SST or \([f]\) increases and \(V_c\) decreases for SEM0. The strongest single predictor for \(\Delta t_{ET}\) in each group is \(\xi_{850}\) for dynamics, HMC\(_{700}\) (or \(Q_v\)) for moisture, and PRECT for the “other” group. In the same environment, EVs in CAM5 are transformed into TCs much faster than in SAM0/SEM0 (ERAI is between CAM5 and SAM0/SEM0), indicating that CAM5 has stronger restoring processes than SAM0/SEM0 into a certain equilibrium state. This could be due to the quasi-equilibrium assumptions explicitly employed into the closures of CAM5’s shallow (i.e., CIN-closure) and deep (i.e., CAPE-closure) convection schemes, which do not exist in the process-based unified convection scheme (UNICON) of SAM0/SEM0.

d. Optimal multivariate environmental conditions for TC genesis

With the goal of finding the best combination of environmental variables for the prediction of TC genesis, we generated all possible combinations of environmental variables by selecting one variable from each group in Table 1 (a total of \(3 \times 9 \times 4 \times 2 \times 10 = 2160\) combinations). Then, for each combination, we divided individual environmental variables into five bins in such a way that (approximately) the same number of EVs are assigned to each bin. We then formed a set of multivariate joint bins (i.e., a total of \(5^5\) joint bins for combinations consisting of five environmental variables), computed the TC genesis frequency \(\Gamma(n) = \frac{nEV(j)/nEV}{nEV(j)}\) in the \(j\)th multivariate joint bins (\(j = 1,..., 5^5\)), and defined the TC genesis probability, \(\Gamma(n)\), of the individual \(n\)th EV (\(n = 1,..., nEV\)) as \(\Gamma(n)\), where \(j\) is the joint index containing the \(n\)th EV. By dividing each environmental variable into five bins, individual multivariate joint bins contained on average \(\sim 100\) EVs for ERAI/SAM0/CAM5, which enabled the calculation of a reliable \(\Gamma(n)\) with small influence of the sampling error. Finally, the prediction skill of individual combinations is measured by dividing all EVs into DEVs and NEVs and calculating the difference of the average \(\Gamma(n)\) between the two groups (e.g., prediction skill = \(\Gamma(n)_{DEV} - \Gamma(n)_{NEV}\) in Table 2). A combination that predicts TC genesis in a perfect way would have \(\Gamma(n)_{DEV} = 1\) and \(\Gamma(n)_{NEV} = 0\), such that its prediction skill is 1. If a certain combination predicts that all DEVs would develop into TCs and some NEVs would also develop into TCs, its associated \(\Gamma(n)_{DEV} = 1\) and \(\Gamma(n)_{NEV} > 0\). If we define the prediction skill by only using \(\Gamma(n)_{DEV}\), the combination has a prediction skill of 1, even though it failed to predict that all NEVs would not develop into TCs. Thus, we defined the prediction skill as the difference between \(\Gamma(n)_{DEV}\) and \(\Gamma(n)_{NEV}\), rather than by only using \(\Gamma(n)_{DEV}\). Therefore, our definition of
prediction skill differs slightly from those used in other studies focusing on the probability for TDs to develop into TCs (i.e., GPr\textsubscript{DEV}). However, we note that our definition of prediction skill is mainly controlled by GPr\textsubscript{DEV} rather than GPr\textsubscript{NEV}.

Table 2 presents the resulting best combinations of the five environmental variables for ERAI, SAM0, and SEM0 (V5, which stands for five variables) as well as CAM5 (V5c, which stands for five variables for CAM5). Furthermore, the prediction skills of the two combinations are also shown, consisting of the environmental variables constituting the genesis potential indices suggested by Emanuel and Nolan (2004) (EN04: η\textsubscript{SSO}, RH\textsubscript{600}, PI, and \(V_s\)) and Wang and Moon (2017) (WM17: η\textsubscript{SSO}, \(V_z\), and \(ω\textsubscript{500}\)). The best combination for SAM0/SEM0 consists of η\textsubscript{SSO}, \(δ_s\), SST, \(V_s\), and \(f]\), which is identical to that of ERAI, indicating that SAM0/SEM0 realistically captures the TC genesis processes that exist in nature. To check the robustness of the best combination, we repeated the same analysis by dividing individual environmental variables into six bins instead of five bins, which, with the exception that \(V_s\) in CAM5 was replaced by \(V_z\), produced the same combinations as shown in Table 2. The second best combination in all ERAI/SAM0/SEM0 was obtained by replacing SST with PI (not shown). The best combination for CAM5 consists of different variables to those of ERAI/SAM0/SEM0 and has a much lower skill score, indicating that TC genesis processes simulated by CAM5 are somewhat unrealistic. Overall, V5 has a substantially higher prediction skill than EN04 which performs as similar as or slightly better than WM17. Although not shown, the best combination with four environmental variables were obtained by excluding SST from V5 with the resulting prediction skills of 0.122/0.130/0.160 for ERAI/SAM0/SEM0, respectively. These are higher than the prediction skills of EN04, consisting of the same number of environmental variables. In the case of ERAI, when η\textsubscript{SSO} (\(δ_s\), SST, \(V_s\), \(f]\)) is excluded from V5, the prediction skill drops by 47% (33, 12, 11, and 22) from that of V5, indicating that the dynamic variable is the most important factor and the moisture variable is more important than the stability or shear variables for predicting TC genesis. Although \(f]\) is a part of η\textsubscript{SSO}, \(f]\) serves as a standalone predictor and is more important than the stability and shear variables.

Figure 8 shows the global distributions of annual-mean TC genesis probability (GPr) predicted by V5, EN04, and WM17 for ERAI and the difference between the annual-mean GPr and TC genesis frequency (GFq) shown in Fig. 3c. All combinations successfully reproduce the overall patterns of annual-mean TC GFq. In terms of spatial correlation and RMSE error, V5 shows the best performance and EN04 has a slightly better performance than WM17. The strengths and weaknesses of individual combinations are more clearly represented by the difference plots. A few notable improvements in V5 over EN04 and WM17 are the suppression of the biases over the western North Pacific and Indian Oceans, and also along about 30°N, which, as will be shown in Fig. 10, is mainly due to \(f]\). All combinations tend to underestimate the observed GFq over the eastern Pacific Ocean and western Pacific Ocean at around 150°E, between 10° and 20°N, where the observed GFq is at its maximum.

Figure 9 shows the annual cycles and interannual variations of TC genesis number as predicted by V5, EN04, and WM17, and compared with the TC genesis number diagnosed from ERAI. All combinations reproduce the annual cycle of the observed TC number reasonably well. Overall, V5 shows a better performance than EN04 and WM17. Compared to the other regions, all combinations show relatively poor performances over the Indian Ocean. Compared with the annual cycle, interannual variations are less well reproduced, particularly those related to WM17. As expected, all combinations tend to underestimate the interannual maximum and minimum TC numbers in ERAI, which is slightly remedied by dividing individual environmental variables into 10, rather than 5, bins (dotted lines). For practical applications, we could derive a simple parametric equation for TC genesis probability by estimating the best-fitting lines of the TC GFq regressed on individual environmental variables in Fig. 6, and then combining them appropriately, as was done by Emanuel and Nolan (2004). This will be reported on in future research.

Figure 10 shows the global distributions of the differences in annual-mean GPr obtained from V5 and its subcombinations generated by excluding one environmental variable from V5. If a new environmental variable incorporated into a specific combination is a random number or perfectly correlated with any of the pre-existing environmental variables, the addition of the new variable will not change the prediction skill. Thus, this figure can be interpreted as presenting the independent contributions of individual environmental variables (i.e., excluding contributions from other environmental variables) to TC genesis probability.

In general, η\textsubscript{SSO} suppresses (enhances) GPr in the equatorial (off-equatorial) area and the spatial pattern of ΔGPr(η\textsubscript{SSO}) is similar to that of V5’s GPr shown in Fig. 8a, indicating that η\textsubscript{SSO} is a major environmental variable controlling TC genesis. The moisture variable \(δ_s\) substantially increases (decreases) GPr over the eastern and western Pacific Ocean (Arabian Sea and northwestern Australia) and also, to a lesser degree, the South Pacific convergence zone and eastern equatorial
Atlantic Ocean (southern Indian Ocean), where climatological $\delta_v$ was large (small) (not shown). This indicates that TC genesis probability increases as $\delta_v$ decreases and is consistent with the results shown in Fig. 6e. Similar to $\eta_{850}$, SST increases GPr over the western Pacific Ocean and to a lesser degree, the Gulf of Mexico and the southwestern Pacific Ocean along $10^\circ$–$15^\circ$S. Over the western Pacific Ocean, a maximum $\Delta$GPr(SST) existed on the northern periphery of the SST warm pool rather than at its center near the equator, presumably as TC genesis is strongly suppressed near the equator due to the impact of other environmental variables, such as $\eta_{850}$ and $|f|$. This is consistent with the nonmonotonic decrease of GPr over the very warm SST shown in Fig. 6j. Similar to $\delta_v$, $V_s$ increases GPr over the eastern and western Pacific Ocean; however, it suppresses GPr over the Bay of Bengal. The main impact of $|f|$ is the suppression of TC genesis over the subtropical western North Pacific and Atlantic along $25^\circ$–$30^\circ$N, which is responsible for the improved performance of V5 over EN04 and WM17 as shown in Fig. 8. Notably, $\eta_{850}$ has a larger impact than $|f|$ on suppressing TC genesis near the equator. The overall sensitivity patterns of SAM0-simulated GPr to various environmental variables constituting V5 (the right panels in Fig. 10) are similar to those from ERAI with the pattern correlations ranging from 0.68 for $V_s$ to 0.79 for $\delta_v$. This implies that SAM0 well mimics the physical processes controlling TC genesis as observed in nature.

4. Summary and conclusions

As a contribution to understanding TC genesis processes, we compared TCs simulated by several GCMs (SAM0/SEM0/CAM5) at a 0.5° horizontal resolution with data obtained from ERAI and IBTrACS observations. In contrast to previous studies that estimated a bulk TC genesis potential using climatological mean environmental conditions at a fixed geographical location (i.e., Eulerian climatology analysis), we calculated the probability of pre-existing weak cyclonic vortices to develop into TCs by analyzing changes in individual EVs and various environmental conditions along their trajectories (i.e., Lagrangian trajectory analysis). The way how TCs are defined in IBTrACS observations differs from that used in our analysis of ERAI and GCM simulations. By adjusting TC detection criteria, we managed global TC numbers diagnosed from ERAI and SAM0 to match with IBTrACS observation data. We expect the statistical relationship between environmental conditions and TC genesis derived from ERAI to be representative of the natural TC development process, as assumed in many other studies (e.g., Walsh et al. 2007). Our study aimed to answer the following four main questions: Do our GCMs reasonably simulate observed TC properties? What environmental conditions control TC genesis probability? What environmental conditions control TC genesis speed? And what is the best combination of environmental variables to predict TC genesis?

Although our models and ERAI underestimated the maximum near-surface wind speed (due, in part, to the coarse horizontal resolution) and some regional deficiencies existed in the diagnosed and simulated TCs (i.e., underestimated TCs over the North Atlantic Ocean), the overall spatial patterns of TC tracks obtained from ERAI and SAM0/SEM0 were similar to those of IBTrACS. With the exception of CAM5, which simulated too many TCs during the boreal winter, the annual cycles of TC genesis number from ERAI/SAM0/SEM0 were also similar to those from the IBTrACS. Overall, SAM0 showed a much better performance than CAM5 in reproducing the spatial distribution and annual cycle of observed TCs. Over the major oceans, EVs were more frequently generated in the central-eastern portions but the EVs in the western portion more easily developed into TCs, implying that the efficient partitioning of environmental energy into a small number of EVs is more important for the genesis of TCs, than the existence of many EVs.

We compared composite environmental conditions along the trajectories of DEVs (EVs developing into TCs) and NEVs (EVs not developing into TCs). At 3–5 days prior to TC genesis, almost all environmental variables examined had significant predictive power regarding TC genesis and this skill increased as the lead time decreased. Along the trajectory of DEVs, most environmental variables showed monotonic variation with lead time. However, the variables associated with mesoscale convective organization reached their peak values one or two days prior to TC genesis and then decreased, implying that mesoscale convective organization accelerates TC genesis during the early stage but decelerates TC genesis at the last stage. At $t_{EV}$ when the EV is initially identified, the PDFs of most environmental variables associated with DEV were well separated from those of NEV, indicating that whether EVs would become TCs, was already predetermined to a certain degree when they are formed, facilitating the forecasting of TC genesis at $t_{EV}$. The genesis frequency of tropical cyclones (GFq) at $t_{EV}$ increased with an increase in relative and absolute vorticities at 850 hPa, relative humidity at 600 hPa, column water vapor, horizontal moisture convergence below 700 hPa, SST, total precipitation rates, and mesoscale convective organization, but increased with a decrease in saturation deficit at 600 hPa, column saturation deficit, vertical wind shear...
between 200 and 850 hPa, and pressure vertical velocity at 500 hPa. These results are consistent with the environmental conditions associated with the onset of deep convection, implying that TC genesis is accompanied by strong convective activity. On the other hand, some variables (e.g., PI, CAPE, $f$, and directional vertical shear of zonal winds between 200 and 850 hPa) had a nonmonotonic impact on GFq. The maximum GFq occurred when the directional vertical shear of zonal winds between 200 and 850 hPa was slightly negative, indicating that TC genesis favors a weak easterly wind shear with height.

Furthermore, we also examined the dependency of the transition time of DEV into TC ($\Delta t_{ET}$) on various environmental conditions. Environmental variables that had a monotonic impact on GFq showed similar monotonic impacts on $\Delta t_{ET}$, implying that rapidly intensifying EVs have a higher probability of developing into TCs. However, $V_8$ did not have a large impact on transition time, even though it strongly suppressed GFq. The EVs simulated by CAM5 developed into TCs much faster than the ones simulated by SAM0/SEM0, implying that CAM5 has a stronger restoring process than SAM0/SEM0. This could be due to the quasi-equilibrium assumptions explicitly employed into the closures of CAM5’s convection schemes.

In both ERAI and SAM0/SEM0, the best combination of environmental variables with the highest predictive power regarding TC genesis (V5) consisted of $\eta_850$, $\delta_8$, SST, $V_8$, and $f$, which indicates that SAM0/SEM0 well simulated the observed TC genesis processes. In terms of contribution to prediction skill, $\eta_850$ was dominant, followed by $\delta_8$. Overall, V5 has a substantially higher prediction skill than EN04, which performed slightly better than WM17. Regarding CAM5, the best combination consisted of variables that differed from those of ERAI/SAM0/SEM0 with a much lower skill score, implying that the TC processes simulated by CAM5 are somewhat inaccurate. All V5, EN04, and WM17 reproduced the spatial patterns and annual cycle of TC GFq as diagnosed from ERAI, with V5 showing the best performance. Compared with the annual cycle, interannual variations are less well reproduced, particularly those related to WM17. We examined the independent contributions of individual environmental variables to TC genesis by comparing annual-mean TC genesis probability (GPr) values obtained from V5 and its subcombinations. Overall, $\eta_850$ suppressed (enhanced) GPr in the equatorial (off-equatorial) area; $\delta_8$ increased (decreased) GPr over the eastern and western Pacific (Arabian Sea and northwestern Australia); SST enhanced GPr on the northern periphery of the SST warm pool in the western Pacific Ocean; $V_8$ increased (decreased) GPr over the eastern and western Pacific (Bay of Bengal); and $f$ suppressed TC genesis over the subtropical western North Pacific and Atlantic. These sensitivity patterns in the ERAI observation data were well simulated by SAM0, indicating that SAM0 successfully simulates the physical processes controlling TC genesis observed in nature.

We identified TC genesis using $\eta_850$ and warm-core strength. Since SAM0 simulates reasonably strong TCs, at least, up to the TC category 1 (see Fig. 1e), it is also possible to detect TC genesis using $U_{10_{\text{max}}}$ as the IBTrACS observation. It turns out that the $U_{10_{\text{max}}}$ criteria ensuring the global-mean annual number of SAM0-simulated TCs similar to the IBTrACS observation was 17.2 m s$^{-1}$, which is very similar to the minimum intensity of the observed tropical storms (17 m s$^{-1}$). About 70% of TCs identified from the $U_{10_{\text{max}}}$ criteria were identical to those identified from $\eta_850$ and warm core strength. The overall properties of SAM0-simulated TCs identified from the $U_{10_{\text{max}}}$ criteria were very similar to those identified from $\eta_850$ and warm core strength. The five environmental variables constituting the best combination for predicting TC genesis (i.e., V5) were also identical. One notable difference was that during boreal winter in the Southern Hemisphere, the $U_{10_{\text{max}}}$ criteria tend to identify too many TCs extending too far eastward.

Our study focused on investigating and understanding TC genesis processes. In the near future, we hope to report on the analysis of postgenesis TC processes, including the sensitivity of TC genesis and postgenesis processes to global warming induced by the increase of global CO2 concentration at a 0.5° latitude × 0.5° longitude horizontal resolution in SEM0. In addition, we are planning to repeat the same analysis using another reanalysis dataset.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the South Korean Government (MEST) (2020R1A2C2007558). We express our thanks to three anonymous reviewers and the editor.

REFERENCES


