Comparison of Reanalysis Data Sets to Comprehend the Evolution of Tropical Cyclones Over North Indian Ocean

P. Malakar\(^1,2\), A.P. Kesarkar\(^1\), J.N. Bhate\(^1\), V. Singh\(^1\), and A. Deshamukhya\(^2\)

\(^1\)National Atmospheric Research Laboratory, Chittoor District, Andhra Pradesh, India, \(^2\)Department of Physics, Assam University, Silchar, Assam, India

Abstract

Several reanalysis data sets are being used for understanding the role of environmental factors controlling tropical cyclones (TCs) evolution. Six reanalysis data sets, namely, European Center for Medium-range Weather Forecast (ECMWF) ERA-Interim (ERAI) reanalysis, Global Forecast System (GFS) analysis, Japan Meteorological Agency’s 55-year reanalysis projects reanalysis (JRA55), Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2) reanalysis, NCEP Climate Forecast System Reanalysis (CFSR), and fifth generation of ECMWF atmospheric reanalysis of global climate (ERA5), have been evaluated for the representation of track, intensity, and structure of 28 TCs which occurred over North Indian Ocean (NIO) during the period 2006–2015. The errors in track, intensity, and minimum sea level pressure (MSLP) of TCs are estimated with respect to the best track data set of India Meteorological Department (IMD). The representation of inner core structure of TCs has been compared. The smallest error in the position of TCs center, MSLP, and maximum wind speed is found in GFS analysis followed by ERA5 and CFSR reanalysis, respectively. GFS and CFSR data sets capture the most intense stages of the TCs followed by the ERA5 data set, while the other three are unable to obtain intensification beyond the severe cyclonic storm stage. The structures of TCs are better represented in GFS analysis followed by ERA5 reanalysis. However, GFS analysis represents early intensification and, in some cases, overprediction of the category of TCs, especially during the most intensified stages (beyond cyclonic storms). Thus, GFS analysis captures the evolution of TCs more realistically, followed by the ERA5 reanalysis data set.

Plain Language Summary

The reanalysis data sets are being used for understanding physical mechanisms associated with genesis and the evolution of complex atmospheric phenomena such as TCs. These are termed as best estimates of the states of the atmosphere. Global circulation models and data assimilation techniques are used to generate such reanalysis data sets. We have compared six reanalysis data sets, namely, ERAI, GFS, MERRA2, JRA55, CFSR, and ERA5, to IMD observation for estimating their applicability for the prediction of the spatiotemporal evolution of TCs. For this purpose, we have considered 28 TCs, which have occurred over NIO during 2006–2015. We have analyzed the error in the location of the center, maximum winds, and minimum sea level pressure in the reanalysis data sets with respect to IMD best track data set. We have also analyzed the vertical profiles of relative vorticity, wind speed, equivalent potential temperature, diabatic heating, and convergence to see how appropriately the reanalysis data sets represent the structure of TCs. The GFS, CFSR, ERA5, and MERRA2 data sets captured the category of TCs beyond the SCS stage. It is also seen that GFS analysis captures the evolution of TCs more realistically followed by ERA5 data set compared to other data sets. However, in some cases, it overpredicts the category of TC and shows early intensification. In general, GFS analysis is found to be most suited for the study of TC evolution over NIO.

1. Introduction

Tropical cyclones (TCs) are one of the devastating disasters causing enormous loss of human lives and property. Therefore, accurate prediction of TC track and intensity is vital for early warning and preparedness purposes. The evolution of TCs, mainly over oceanic regions, is controlled by environmental factors such as sea surface temperature, wind shear, and middle tropospheric moisture as well as their internal dynamics, which are responsible for their structure and intensity changes. The determination of the role of such
environmental parameters on TCs structure and intensity changes is essential to enhance our understanding of TCs, which in turn improves the track and intensity prediction of TCs. High-resolution simulations using mesoscale models and advanced data assimilation systems are recently being used for obtaining a better understanding of track, structure, and intensity variations of TCs (Montgomery et al., 2006; Rajasree, Kesarkar, Bhate, Singh, et al., 2016; Rajasree, Kesarkar, Bhate, Umakanth, et al., 2016). However, it is essential to know whether global reanalysis data sets are useful for understanding the evolution of TCs. Global reanalysis data sets from leading meteorological organizations are generally available with a resolution of 0.5° × 0.5° to 0.125° × 0.125°. They are used for providing initial and boundary conditions to regional weather prediction models. Reanalysis data sets provide complete spatial and temporal data coverage over a longer period (Thorne & Vose, 2010) and have been used to study different climatological aspects of TCs. In recent past, using reanalysis data sets, many studies have been conducted to generate the climatology of TCs, to study the impact of large-scale mechanisms on TC’s evolution (Maloney & Hartmann, 2000; Maloney et al., 2000), and for estimation of TCs power dissipation (Sirver & Huber, 2006), and so forth. Therefore, a better representation of TCs in the reanalysis data set is vital for the application of these data sets for the understanding of the inner core dynamics of TCs and their interaction with climate system (Scoccimarro et al., 2012). Schenkel and Hart (2012) examined the representation of TCs position, intensity, and intensity life cycle in reanalysis data sets over several cyclone formation ocean basins. Their studies showed that TCs representation in different reanalysis data sets varies over different oceanic basins depending on the availability of weather observations to assimilate in a particular data set. They found that, out of the three oceanic basins studied, namely, North Atlantic (NATL), West Pacific (WPAC), and East Pacific (EPAC), EPAC has the highest mean position difference from best track data set and has comparatively weak intensities in reanalysis representation of TCs. They concluded that it could be a result of the relative lack of observations. Their study found that among five reanalysis data sets, namely, European Center for Medium-range Weather Forecast (ECMWF) ERA Reanalysis (ERA)-40, that is, ERA-40(Manning, 2007; Manning & Hart, 2007), ERA-Interim (ERAI, Simmons et al., 2006), Modern-Era Retrospective Analysis for Research and Application (MERRA, Rienecker et al., 2011), Climate Forecast System Reanalysis (CFSR), and Japan Reanalysis-25 (JRA-25), the CFSR and JRA-25 reanalysis appeared to have the strongest intensities and maximum correlations with best track intensities. Murakami (2014) used six reanalysis data sets released after 2004, namely, JRA-25, Japan Reanalysis-55 (JRA55, Ebita, et al., 2011), ERA-40, ERAI, CFSR, and MERRA, to examine the representation of TCs in reanalysis data sets and found that except CFSR reanalysis, all other reanalyses underestimate the annual mean count of TCs. All reanalysis data sets underestimate the TC intensity. However, CFSR and JRA55 data sets could capture the intensification of TCs to Category 1 and above. They concluded that JRA55 represents the best spatial distribution of TCs and that can be a result of the fact that JRA55 reanalysis uses wind retrievals surrounding the TCs in the assimilation system. Also, JRA55 and CFSR show the highest score of the frequency of occurrence, hitting rate, and lower false alarm rate (Murakami, 2014). Hatsuhashika et al. (2006) studied TCs over northern basins over entire globe using the reanalysis storm-relative composite temperature anomalies and compared with the previous observations. They found that in JRA-25 reanalysis, the maximum composite temperature is several degrees lower than that found in the earlier observational studies (Hatsuhashika et al., 2006). Onogi et al. (2007) showed that the representation of temperature anomalies associated with TCs warm core is stronger in WPAC compared to EPAC in both ERA-40 and JRA-25 reanalyses. They concluded that this difference could be due to data sparseness in the EPAC.

Recent available high-resolution reanalysis data sets, namely, MERRA2, JRA55: an improved version of JRA25, ECMWF Reanalysis (ERA)-Interim reanalysis, that is, ERAI, and fifth generation reanalysis, that is, ERA5, CFSR, and Global Forecast System (GFS) analysis, are expected to give better representation of TCs in terms of structure and intensity changes, although these are in too coarse resolution to capture the internal dynamics of TCs. This study aims to find out which reanalysis data set provides a better representation of the evolution of tropical TCs intensity and structure over the North Indian Ocean (NIO). The growth of TCs occurred over NIO during 2006–2015 has been examined using six different reanalysis data sets, namely, ERAI, JRA55, MERRA2, CFSR, and ERA5 reanalysis and GFS analysis data sets with the highest available resolutions and compared with India Meteorological Department (IMD) best track data set. The composite vertical structures of TCs have been analyzed to find out the best-represented structure of TCs among six reanalysis data sets by evaluating equivalent potential temperature (θe), diabatic heating (δθ),
relative vorticity ($\xi$), horizontal wind, and convergence in the inner core region of TCs. The organization of the paper is as follows. Section 2 describes the data and methodology used in the study. Section 3 discusses the results obtained, and a summary of the essential conclusions is presented in section 4.

### 2. Data and Methodology

Table 1 gives a detailed description of reanalysis data sets, their spatial and temporal resolution, the number of available vertical levels, and the parameter used from each data set in this study. In an aim to compare the representation of TCs structure in intensity in different global reanalysis data sets, the finest available resolution of all these data sets has been analyzed, although various resolutions are available in their different release. The improvements in TCs representation in ERA5 over ERAI are also investigated in this study. The use of the data set that provides the best representation of structure and intensity of TCs in NWP models as initial and boundary conditions is believed to simulate TCs most accurately. The representation of TCs in global reanalysis data sets is compared with the TCs representation in GFS analysis data set to investigate the most appropriate initial and boundary conditions in NWP model for TCs simulation over NIO. Regional Specialized Meteorological Centre (RSMC)—Tropical Cyclones, New Delhi, which functions in Cyclone Warning Division, IMD, provides the best track data for all cyclonic disturbances over the NIO. This data set contains track, intensity, mean sea level pressure (MSLP), and central pressure drop for the TCs and estimated utilizing satellite, Radar, and in situ observations over the region. This best track data set is used for comparison of track and intensity representation of TCs in different data sets.

The tracks and intensities of 28 TCs originated over NIO during the period 2006–2015 are derived from six reanalysis data sets listed in Table 1. The supporting information provides the list of the TCs chosen for this study. The results are compared with the IMD best track data set. The tracking algorithm used in the study is based on the CycloTrack v1.0 (Flaounas et al., 2014). This algorithm uses a spatial filter to smooth the vorticity field at 850 hPa, and then a threshold is applied to identify the cyclonic circulations. After that, the vorticity maxima are identified in each cyclonic circulation to detect the centers. The algorithm finds all the TCs track by linking the TCs centers at consecutive time steps, and the best track is identified based on the average differences of relative vorticity between consecutive track points, weighted by their distance. The best track of each cyclone has been visually inspected to verify the final track of the cyclone. The 925-hPa pressure level is the level of maximum surface wind in all reanalysis data sets. The maximum surface wind within a 2° × 2° grid box around the TC center is considered as the intensity ($V_{max}$) of the TC. The differences between reanalysis $V_{max}$ and surface wind estimation by IMD are calculated to verify whether reanalysis

### Table 1

**Description of Different Reanalysis Data Sets and Variables Used, Namely, Zonal Component of Wind (u), Meridional Component of Wind (v), 10-m Wind (u10, v10), Vertical Wind (w), Specific Humidity (q), Relative Humidity (r), Air Temperature (t), and Mean Sea Level Pressure (mslp)**

| Data set         | Resolution      | Variable used            | Reference and source                                                |
|------------------|-----------------|--------------------------|---------------------------------------------------------------------|
| ERAI reanalysis  | 0.125° × 0.125° × 37 | u, v, w, r, q, t, u10, v10, mslp | Dee et al., 2011 http://apps.ERAI.int/datasets/data/interim-full-daily/ |
| GFS analysis     | 0.5° × 0.5° × 26     | u, v, w*, q, t, u10, v10, mslp | Sela (2009, 2010) NCEI at NOAA https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/ |
| JRA55 reanalysis | ~0.56° × 0.56° × 61   | u, v, w*, q, t, u10, v10, mslp | Kobayashi et al. (2015) and Harada et al. (2016)                       |
| MERRA2 reanalysis| 0.5° × 0.5° × 42     | u, v, w*, q, t, u10, v10, mslp | Gelaro et al. (2017) CISL at NCAR https://www.earthdata.nasa.gov/datasets/global-forecast-system-gfs/ |
| CFSR and CFSv2 reanalysis | 0.5° × 0.5° × 37 | u, v, w*, q, t, u10, v10, mslp | Saha et al. (2010) and Saha et al. (2014) CISL at NCAR |
| ERA5 reanalysis  | 0.25° × 0.25° × 37     | u, v, w*, q, t, u10, v10, mslp | Copernicus Climate Change Services (C3S; 2017) and Hersbach et al. (2019) https://www.ERAI.int/en/forecasts/datasets/reanalysis-data-sets/era5/ |

*Vertical velocity calculated from diagnostic omega equation as vertical velocity was not available in the reanalysis data set.*
data sets represent the intensity of TCs appropriately. The identification of different categories of TCs represented in reanalysis data set is classified according to IMD’s TCs intensity scale (included in the supporting information). The difference in the position of the TCs centers (track error: ΔR), differences in intensity (ΔV_{max}), and MSLP (ΔMSLP) between reanalysis data sets and IMD estimates are calculated at each 6-hr interval during the complete life span of 28 TCs selected for this study. The average ΔR, ΔV_{max} and ΔMSLP are computed for different IMD estimated intensity stages, namely, Depression (D), Deep Depression (DD), Cyclonic Storm (CS), Severe Cyclonic Storm (SCS), Very Severe Cyclonic Storm (VSCS), and Extremely Severe Cyclonic Storm (ESCS), of the TCs. Different environmental parameters, namely, wind speed, equivalent potential temperature (\(\theta_e\)), diabatic heating (\(\delta\)), relative vorticity (\(\zeta\)), and convergence, have been analyzed to access how well the TC’s evolutions are represented in reanalysis data set.

3. Results and Discussions

3.1. Track and Intensity Errors With Category of TCs

Figure 1a shows the average ΔR at different intensity stages of TCs along with the standard deviation (SD) and the number of observations (along the secondary Y axis) available at each intensity stage in all TCs considered for this study. The pink, magenta, aqua, violet, beige, and brown bars indicate the average ΔR in ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 reanalysis data sets, respectively. At the D stage of TCs, the averaged ΔR maximum is ~131 km in MERRA2, and the minimum is ~92 km in the ERA5 data set with SD of ~55 and ~57 km, respectively. Beyond the D stage, ERAI shows maximum average ΔR (>101 km till the ESCS stage). However, GFS shows minimum ΔR followed by ERA5 other than the ESCS stage. At the ESCS stage, JRA55 shows minimum average ΔR (~41 km) with least SD (~22 km). Figure 1a shows considerable improvement of TC track representation in ERA5 over ERAI data set. Figure 1a suggests that the overall ΔR in reanalysis data sets decreases with the intensification of TCs. The comparison among all six data sets shows that the overall minimum value of ΔR is in GFS analysis followed by ERA5 and CFSR reanalysis data set, respectively.

Figure 1b shows that the average ΔV_{max} increases with the intensification of the TCs and reaches its peak during the ESCS (~75 kt) stage. ERAI shows the minimum ΔV_{max} at the D and DD stages, followed by JRA55. However, beyond the CS stage, ERAI shows maximum ΔV_{max} followed by JRA55. The ERA5 shows minimum ΔV_{max} beyond the DD stage, followed by the GFS data set. However, during the ESCS stage of TCs, GFS analysis shows minimum ΔV_{max} (~33 kt). Therefore, Figure 1b suggests that overall intensity prediction is better in ERAI during weak stages and in ERA5 during stronger stages beyond the DD stage. However, for the ESCS stage, GFS provides better intensity representation followed by CFSR. The values of ΔV_{max} are quite high in the stronger category of TCs. Therefore, to reduce the values of ΔV_{max}, it is necessary to assimilate realistic satellite wind estimations in the core region of TCs. It will lead to the appropriate representation of the secondary circulation of TCs in reanalysis data sets.

Figure 1c shows the average difference in MSLP (ΔMSLP) between IMD estimates and reanalysis data sets. The figure shows that in all analyses, ΔMSLP increases with the intensity of the TCs. The minimum average ΔMSLP is seen during the D stage (2–4 hPa). It grows with the intensification of TCs and reaches its maximum during the ESCS stage (~31–47 hPa) in all data sets. Overall, ΔMSLP is less in ERA5 followed by GFS and CFSR, respectively. As shown in Figure 1c, all reanalysis data sets simulate the MSLP with a significant error compared to IMD observed value. Thus, the reanalysis data sets have limitations in capturing the pressure drop at the center of the TCs, which may be attributed to lack of weather observations in the neighborhood of the TCs center.

3.2. Track and Intensity Errors With the Evolution of TCs

The representation of the variation of TCs track, intensity, and MSLP errors with the development of TCs in different data sets is investigated in this section. Figures 2a–2f show the average ΔR (dark blue line), ΔV_{max} (light blue), and ΔMSLP (red bar) with corresponding SD estimated in ERAI, GFS, JRA55, MERRA, CFSR, and ERA5 data sets, respectively. These variations are shown at every 6-hr interval starting from the D stage during genesis to the landfall of the TCs life cycle. The bar chart of ΔV_{max} and ΔMSLP is shown along the primary Y axis, whereas the overlapping ΔR line is shown along the secondary Y axis. The black dots represent the number of observations that are used for averaging at each hour and demonstrated along the
primary Y axis. The starting 00 hr of each storm has been defined as the hour when it was first observed as a tropical depression in IMD best track data set and continued till the landfall. The number of observations decreased with the increasing evolution hours, and at 156 hr, only two observations are available. It is seen that the average $\Delta R$ estimated in ERAI (Figure 2a) is in the range $\sim 65$–$126$ km during the entire life span of TCs. The maximum average $\Delta R$ is of $\sim 126$ km at 126 hr with the SD of $\sim 55$ km for the total seven available observations. At the initial hours, values of $\Delta R$ are relatively more and decrease with the increase in observation hours until the 114 hr, and thereafter, it starts increasing. It is consistent with the results mentioned in the previous section indicating the decrease in $\Delta R$ with the intensification of the TCs. Initially, all the TCs are in their weaker stage (D, DD, and CS), which leads to larger values of $\Delta R$ followed by a decrease due to intensification during intermediate hours (SCS, VSCS, and ESCS). During the later hours of their evolution, TCs are in weaker stages causing an increase in the values of average $\Delta R$. Similar to ERAI (Figure 2a), Figures 2b–2f show that the average $\Delta R$s estimated in GFS (Figure 2b), JRA55 (Figure 2c), MERRA2 (Figure 2d), CFSR (Figure 2e), and ERA5 (Figure 2f) data sets also decrease with observation hours and again start increasing after certain period during the late hours. In GFS analysis, average $\Delta R$ estimated is

Figure 1. The average values of (a) $\Delta R$, (b) $\Delta V_{\text{max}}$, and (c) $\Delta MSLP$ composited for different stages of TCs along with respective standard deviations using ERAI (pink), GFS (magenta), JRA55 (aqua), MERRA2 (purple), CFSR (beige), and ERA5 (brown) reanalysis data sets. The red dashed line (along the secondary Y axis) is the number of available records in the IMD best track data set at each stage.

10.1029/2019EA000978
Earth and Space Science
in the range ~44–103 km; during the initial hours (00 to 42 hr), maximum SD in ΔR is ~62 km at 18 hr; and the averaged ΔR for 28 TCs is ~84 km. As a combined effect of intensification and weakening of TCs at intermediate and later hours of their life cycle, there is a decrease followed by an increase in ΔR. The maximum SD is ~89 km at 150 hr when the ΔR has been averaged for two of the observations. Due to a smaller number of records during dissipation stages of TCs, mostly all data sets have large SD in ΔR. The overall average ΔR is smaller in GFS analysis (Figure 2b) as compared to ERAI reanalysis (Figure 2a). Similarly, Figures 2c–2f show the decrease and increase in average ΔR with observation hours in JRA55, MERRA2, CFSR, and ERA5 data sets. The maximum average ΔR in JRA55 is ~101 km at 18 hr with an SD of ~53 km, and the minimum is ~45 km with an SD of 24 km. In MERRA2, the maximum ΔR is ~120 km at 00 hr with an SD of ~53 km, and the minimum is ~66 km with an SD of 39 km at 90 hr. In CFSR, the maximum is ~120 km at 00 hr with an SD of ~63 km, and the minimum is ~42 km with an SD of ~35 km at 108 hr. In ERA5, the maximum ΔR is ~105 km at 00 hr with an SD of ~67 km, and the minimum is ~38 km with an SD of ~33 km at 138 hr. Therefore, Figures 2a–2f show that the overall track error is lesser in ERA5 followed by GFS/CFSR and more in ERAI data set compared to all other data sets. Figures 2a–2f show that for all data sets the average ΔV_{max} (light blue bar) increases with the evolution hours, reaches a peak value at intermediate hours, and again starts decreasing with a further rise in observation hours. At initial hours (till 18 hr), GFS and MERRA2 data sets show more ΔV_{max} followed by ERA5, and ERAI showed less value of ΔV_{max} than other reanalysis data sets. However, during intermediate hours (24 to 144 hr), ERAI and MERRA2 show more ΔV_{max} than another reanalysis. During the whole period of TCs evolution, the minimum average ΔV_{max} varies among different data sets. Till 18 hr of TCs evolution, it is best estimated by ERAI, after 18 till 84 hr by ERA5, after 84 till 114 hr by GFS and CFSR, after 114 till 144 hr, by MERRA2, but after 144 hr again ERAI estimated the minimum followed by GFS data set. Therefore, at the initial phase and at the dissipation phase of TCs lifetime, when TCs are mostly in weaker intensity stage (D to DD), ERAI predicts the intensity better than other reanalysis data sets. However, ERAI and MERRA2 show maximum error in intensity at the matured phases of TCs life span. ERA5 represents TCs intensity with less errors followed by GFS and CFSR data sets during the matured phase. From Figures 2a–2f it is seen that SD in ΔV_{max} is comparatively more at intermediate hours starting from ~48 to 120 hr of observation in all reanalysis estimated values. It is due to the combined effect of the

Figure 2. Average ΔRs (blue line along secondary axis), ΔV_{max} (light blue bar), and ΔMSLP (maroon bar) and their respective standard deviations (along primary axis) at every 6-hourly intervals starting from depression stage till landfall stage for 28 tropical cyclones estimated using (a) ERAI, (b) GFS, (c) JRA55, (d) MERRA2, (e) CFSR, and (f) ERA5 reanalysis data sets. The black dots are the total number of observations present at every 6 hr (along the primary axis). The alternative plot of this figure is included in the supporting information for better clarity of comparison.
intensification and dissipation of the TCs during that period. Figure 2b shows that the maximum average \( \Delta V_{\text{max}} \) in ERAI reanalysis is \( \sim 46 \) kt at 102 hr with an SD of \( \sim 21 \) kt for 11 available observations. Maximum SD is \( \sim 29 \) kt in 17 available observations at 90 hr when average \( \Delta V_{\text{max}} \) is \( \sim 40 \) kt. In GFS analysis (Figure 2b), the maximum \( \Delta V_{\text{max}} \) is \( \sim 21 \) kt with \( \sim 12 \) kt SD in 10 available observations at 120 hr. Maximum SD is \( \sim 20 \) kt at 84 hr of observation with 18 available observations. In initial observation hours starting from 00 to 36 hr, and later hours beginning from 126 to 156 hr, GFS analysis shows more \( \Delta V_{\text{max}} \) compared to other analysis estimated value. However, GFS is less compared to other data sets at intermediate hours. It can be the result of the fact that GFS analysis overestimates the intensity in the weak stage of the TCs, which can also be seen from IMD/GFS contingency table for cyclone intensity (Figure 3b). Also, \( \Delta V_{\text{max}} \) in GFS is less compared to other data sets during intermediate hours. It can be attributed to the appropriate estimation of the intensity of TCs in GFS data set (Figure 3b) during intensified stages. Figure 2c shows that the maximum average \( \Delta V_{\text{max}} \) in the JRA55 analysis is \( \sim 43 \) kt at 90 hr with a maximum SD of \( \sim 27 \) kt for 17 available observations. The JRA55 analysis shows more average \( \Delta V_{\text{max}} \) in intermediate observation hours than MERRA2 and GFS analysis but comparable to ERAI analysis value. Figure 2d shows that maximum average \( \Delta V_{\text{max}} \) in the MERRA2 analysis is \( \sim 27 \) kt at 90 hr with a maximum SD of \( \sim 22 \) kt in 17 available observations. Figure 2e shows that maximum average \( \Delta V_{\text{max}} \) in the CFSR analysis is \( \sim 24 \) kt at 78 hr with a maximum SD of \( \sim 15 \) kt in 19 available observations. Figure 2f shows that maximum average \( \Delta V_{\text{max}} \) in the ERA5 analysis is \( \sim 23 \) kt at 90 hr with a maximum SD of \( \sim 21 \) kt in 17 available observations. Figures 2a–2f show that similar to average \( \Delta V_{\text{max}} \), the average

Figure 3. Contingency table representing TCs intensity in (a) ERAI, (b) GFS, (c) JRA55, (d) MERRA2, (e) CFSR, and (f) ERA5 data sets with respect to intensification stages estimated by IMD. The last column labeled as “N” in each contingency table represents the number of underestimation of different stages below the intensity of depression stage.
Figure 3 helps to understand the MSLP and Vmax errors shown in Figures 2a–2f, as discussed below. Figure 1 showed that the errors in track (intensity and MSLP) decrease (increase) with the intensification of TCs. The overestimation of the weak stage of TCs in GFS, CFSR, ERA5, and MERRA2 data sets explains the higher values of MSLP and Vmax errors during the initial hours of TCs lifetime. The maximum (minimum) MSLP and Vmax errors in ERAI and JRA55 (ERA5, GFS, and CFSR) during intermediate hours can be attributed to the inappropriate (appropriate) estimation of the intensity of TCs in the particular data set (Figure 3) during intensified stages.

3.3. Multicategory Verification of Intensity of TCs

To assess how well a particular data set is representing a specific stage of intensity of the TCs in the reanalysis data sets, we have categorized them according to the IMD scale (supporting information) and compared with IMD best track data set. Figures 3a–3f show the IMD/reanalysis data sets contingency tables of intensity prediction for ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 data sets, respectively. It is seen from these figures that the ERAI data set represents the intensity of TCs more appropriately with a maximum number of hits during the D (followed by JRA55) and DD (followed by ERA5) stages but failed to reproduce the higher intensities of the TCs (SCS, VSCS, and ESCS). As seen from a column labeled as “N” in Figures 3a and 3c, for a reasonable number of times the intensities of TCs have been underestimated as no cyclone in ERAI and JRA55 data sets, respectively.

ERA5 (Figure 3f) represents the CS stage with more numbers of hits followed by MERRA2 (Figure 3d) and CFSR (Figure 3e) data sets, respectively, and during the SCS stage followed by GFS data set. However, ERA5 failed to reproduce the TCs intensity beyond the VSCS stage. The GFS data set represents the VSCS and ESCS stages of TCs with a greater number of hits followed by the CFSR data set. Among all six data sets, GFS and CFSR data sets only can represent the TCs intensity during the ESCS stage, although these data sets overpredict lower intensities (D, DD, and CS) of TCs. Maximum hits in ERAI data set are at the D stage, in GFS data set at the VSCS stage, in JRA55 at the D stage, and in MERRA2, CFSR, and ERA5 data sets at the CS stage. In general, it can be observed that the GFS analysis has a better representation of TC intensities compared to other data sets as it can predict the VSCS and ESCS intensity stages whereas ERA5 gives better prediction in moderate-intensity stages (CS and SCS).

Figure 3 helps to understand the MSLP and Vmax errors shown in Figures 2a–2f, as discussed below. Figure 1 showed that the errors in track (intensity and MSLP) decrease (increase) with the intensification of TCs. The overestimation of the weak stage of TCs in GFS, CFSR, ERA5, and MERRA2 data sets explains the higher values of MSLP and Vmax errors during the initial hours of TCs lifetime. The maximum (minimum) MSLP and Vmax errors in ERAI and JRA55 (ERA5, GFS, and CFSR) during intermediate hours can be attributed to the inappropriate (appropriate) estimation of the intensity of TCs in the particular data set (Figure 3) during intensified stages.

3.4. Temporal Lead and Lag in the Multicategory Intensification of TCs

The delay and early intensification of the different stages estimated in ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 are analyzed and depicted in Table 2. In this table, plus (+) sign and minus (−) sign represent the delayed and early intensity prediction, respectively, in hours. N represents the condition when IMD observed the storm at a particular intensity scale, but reanalysis estimation shows no evidence of this intensity scale, that is, unable to capture the intensity. P represents the different stages of intensity in reanalysis data set even though they are not existing in the IMD best track estimates. X represents the unavailability of certain intensity stages in IMD observation as well as in reanalysis estimation. As seen from this table, ERAI and JRA55 data sets have underestimated the intensities of the TCs as CS, especially during the period when the TCs have intensified beyond the CS stage, whereas GFS, MERRA2, CFSR, and ERA5 data sets can capture the intensities more appropriately. However, it can be seen from Table 1 that ERA5, GFS, CFSR, and MERRA2 data sets predict stages to CS before their actual occurrence in most of TCs. In other words, it can be inferred that these four data sets show an early intensification of TCs. Further, for the D and DD stages, these data sets show early occurrence. ERAI data set shows on time and delayed occurrence at the D and DD stages, respectively, in most of TC cases. JRA55 data set shows correct or delayed occurrence during the D stage and delayed occurrence during the DD stage in most of the TCs. It can also be seen that GFS and MERRA2 data sets overpredicted the SCS, VSCS and ESCS stages instead of the CS stage in case of
| Name of TCS | DD | SCS | VSCS | ESCS | TCs |
|------------|----|-----|------|------|-----|
| Aila       | 0  | 48  | −32  | 8   | 30  |
| Alash      | 34 | 0   | 12   | −24 | 0   |
| Chapla     | 0  | 24  | 6    | 42  | 12  |
| Gani       | 30 | 34  | 0    | 42  | 12  |
| Hudhud     | 18 | 48  | 24   | 48  | 18  |
| Keila      | 0  | 30  | −12  | 18  | 6   |
| Madi       | 0  | 24  | 0    | 24  | 6   |
| Mala       | 0  | 24  | 0    | 24  | 6   |
| Mala       | 0  | 30  | 32   | 42  | 12  |
| Nargis     | 0  | 24  | 0    | 24  | 6   |
| Nilam      | 0  | 24  | 0    | 24  | 6   |
| Phyan      | 24 | 0   | 0    | 24  | 6   |
| Rashmi     | 0  | 24  | 12   | 24  | 12  |
| Thane      | 0  | 24  | 0    | 24  | 6   |

Note: ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 (in hours), + = delay; − = prior. N = not shown in analysis. X = TCs not intensified to that stage; P = overprediction to that stage.
teneight TCs, namely, Helen, Hudhud, Keila, Khaimuk, Nanauak, Nilam, Phyan Rashmi, Viyyaru and Ward. Thus, it can be summarized that even though MERRA2 and GFS data sets are able to predict most of the intensified stages (VSCS and ESCS), there is an early occurrence of these stages in these analyses. However, in CFSR and ERA5 data sets, there is a delayed occurrence of the VSCS intensity stage.

3.5. Structure of TCs

In this section, the composite structures of TCs with respect to different environmental parameters, namely, $\theta_e$, $\delta$, $\xi$, and horizontal wind convergence, in the inner core region of TCs are discussed. The parameters are averaged out in a $2^\circ \times 2^\circ$ grid box around the center of the TCs, and the averaged evolution of the vertical profile of these parameters with respect to IMD observed intensity is analyzed. Figures 4a–4f show the composite structure of evolution of the vertical profile of $\theta_e$ with respect to IMD observed intensity stage.

In ERAI (Figure 4a), MERRA2 (Figure 4d), and ERA5 (Figure 4f), the vertical structure of composite $\theta_e$ with respect to different categories of TCs is seen increasing linearly with height. However, the convective organization requires the middle tropospheric ($600$–$400$ hPa) and boundary layer (below $850$ hPa) warming due to latent heat release causing values of $\theta_e$ to be more at these levels and less in between. This evolution pattern is captured more appropriately in the GFS, MERRA2, and CFSR (Figure 4e) data sets. ERAI, MERRA2, and ERA5 data sets probably overestimate the $\theta_e$ structure of the TCs evolution since $\theta_e$ linearly increases with the height and reaches to its maximum value around $500$–$600$ K at $200$ hPa. In GFS and JRA55 data sets, the maximum values of $\theta_e$ are reaching to $370$ K. The GFS analysis shows increase in values of $\theta_e$ at comparatively lower altitudes from $400$ to $600$ hPa with intensification and fewer values between lower and middle troposphere indicating convective instability with height (Wang, 2012).

Diabatic heating ($\hat{\theta}$) comprising latent heat, sensible heat, and radiative heat transfer in the inner core region of TCs have an essential role in TCs size, structure, and intensity evolution. Figures 5a–5f show the composite $\hat{\theta}$ structure of TCs in ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 data sets, respectively. The $\hat{\theta}$ values captured in all reanalysis data sets are in range from $-1.5$ K(6 hr)$^{-1}$ to $+1.5$ K(6 hr)$^{-1}$ which are very low in comparison with the estimated value of $\hat{\theta}$ (Wang, 2014; Rajasree, Kesarkar, Bhatre, Singh, 2012).
et al., 2016; Rajasree, Kesarkar, Bhate, Umakanth, et al., 2016). The pattern of increase in diabatic heating in the inner core region of TCs with intensification is represented more correctly in GFS analysis (Figure 5b) and CFSR (Figure 5e) followed by ERA5 (Figure 5f) data set compared to other data sets. MERRA2 (Figure 5d) reanalysis shows maximum heating at the CS and SCS stages compared to lower and higher categories of TCs. JRA55 (Figure 5c) data set shows more or less similar heating at all stages of TCs. The boundary layer warming due to latent heat release is appropriately captured in GFS analysis followed by CFSR reanalysis. ERA5 captures the boundary layer warming beyond the SCS stage of the TCs. Figure 5 indicates that the increase in warming due to the increase in surface wind speed followed by the rise in spin up and increase in evaporation feedback in the boundary layer is more appropriately represented in GFS analysis followed by CFSR reanalysis.

Figures 6a–6d represent the evolution of the composite of vertical profiles of horizontal wind speed with the different categories of TCs in ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5 data sets, respectively. All reanalysis data sets capture the vertical pattern of horizontal wind speed with a lower value than the actual classification of the category. It indicates that most of these reanalysis data sets underestimate the horizontal wind speeds even with respect to height. The maximum wind speed within the boundary layer captured in ERAI (Figure 6a) data set is ~12 m s\(^{-1}\) during the CS stage, and it is unable to obtain higher wind speeds in intensified stages beyond the CS stage. The highest values of winds are represented in the GFS analysis, which is about 30 m s\(^{-1}\) during the ESCS stage. JRA55 data set shows that maximum wind speed reaches a value of ~12 m s\(^{-1}\) in the ESCS stage, indicating the underestimation of wind speed within the inner core of the cyclone. However, it shows the systematic increase in wind speed with the intensification of TCs. The maximum wind within the boundary layer captured in MERRA2 reanalysis is ~20 m s\(^{-1}\) during the ESCS stage. In spite of the underestimation of the wind speed, MERRA2 data set captures the evolution of wind flow pattern in TCs inner core but unable to represent the diabatic heating (Figure 5d) occurred due to evaporation feedback. Similarly, CFSR and ERA5 data sets also captured the wind flow pattern, and the maximum wind within the boundary layer captured in CFSR and ERA5 is ~20 m s\(^{-1}\) during the ESCS stage.

Figure 5. Composite vertical structure of diabatic heating (K per 6 hr) in (a) ERAI, (b) GFS, (c) JRA55, (d) MERRA2, (e) CFSR, and (f) ERA5 data sets with respect to intensification stages estimated by IMD.
Figure 6. Composite vertical structure of horizontal wind velocity (m s$^{-1}$) in (a) ERAI, (b) GFS, (c) JRA55, (d) MERRA2, (e) CFSR, and (f) ERA5 data sets with respect to intensification stages estimated by IMD.

Figure 7. Composite vertical structure of $\xi$ ($\times 10^{-5}$ s$^{-1}$) in (a) ERAI, (b) GFS, (c) JRA55, (d) MERRA2, (e) CFSR, and (f) ERA5 data sets with respect to intensification stages estimated by IMD.
Although all reanalysis data sets are found underestimating the distribution of wind speed, GFS analysis followed by ERA5, CFSR, and MERRA2 reanalysis data sets captures higher values of wind speed in the TCs core region over NIO than other data sets. Figures 7a–7d show the evolution of the vertical profile of composite $\xi$ with respect to categories of TCs in ERAI, GFS, JRA55, MERRA2, CFSR, and ERA5, respectively. These figures are consistent with the composite wind structure shown in Figure 6. The high values of $\xi$ are found in ERA5 followed by GFS analysis, CFSR, and MERRA2 reanalyses. All data sets other than ERAI show the increase in values of $\xi$ with respect to the intensification of TCs (Figure 8).

4. Conclusions

In this study, the representations of TCs originated over NIO in reanalysis data sets are analyzed. The differences in track, intensity, and MSLP from IMD best track data sets in all reanalysis representations of TCs have been calculated with the aim to determine the best representation of TCs. The accuracy in intensity prediction and the temporal lead and lag in intensification with respect to IMD observations are compared among all data sets. The hit, false alarm, and early and delayed intensification for different TCs in all reanalysis data sets have also been examined. Also, the composite structure of TCs with respect to several environmental parameters has been analyzed to identify the best representation of TCs inner core structure in reanalysis data sets within the highest available resolution. The study shows that $\Delta R$ decreases with the intensification of TCs. Among all reanalysis data sets, GFS (ERAI) analysis showed a minimum (maximum) $\Delta R$ followed by ERA5 and CFSR reanalyses. The $\Delta R$ in JRA55 and MERRA2 are more or less comparable. During the ESCS stage, the average $\Delta R$ in JRA55 data set is minimum compared to other five reanalyses. The $\Delta R$ in JRA55 and MERRA2 are more or less comparable. During the ESCS stage, the average $\Delta R$ in JRA55 data set is minimum compared to other five reanalyses. The increase in $\Delta$MSLP with intensification suggests that the pressure drop at the center of TCs (with intensification) is not well captured in all reanalysis data sets. From the D stage till the ESCS stage, the ERA5 (ERAI and JRA55) data set showed lower (higher) values of $\Delta$MSLP, whereas JRA55 data set shows minimum $\Delta$MSLP during the D stage. The overall representation of $\Delta$MSLP is appeared to be best in the ERA5, followed by GFS data set. During weak intensity stages (D and DD) of TCs, $\Delta V_{max}$ is least in ERAI, followed by JRA55. Beyond the DD till VSCS stages, the error in intensity is minimum in ERA5 reanalysis.
followed by GFS analysis; however, during the ESCS stage, the minimum $\Delta V_{\text{max}}$ is captured in GFS analysis followed by CFSR reanalysis.

ERAI reanalysis underestimates the intense stages of TCs to the D and DD stages and does not capture the intensification beyond the SCS stage. In few cases, ERAI data set did not capture the occurrence of the D stage during the dissipating phase of the TCs life cycle. GFS, CFSR, ERA5, and MERRA2 data sets overestimate the intensity during the D and DD stages of TCs as CS and SCS. The JRA55 data set underestimates the intensity of TCs and unable to capture the intensification beyond the CS stage. The GFS analysis and CFSR data set can only capture the intensity of TCs beyond the VSCS stage over NIO. ERA5 can capture TCs intensity till the VSCS stage, whereas ERAI failed to capture the intensity beyond the CS stage. This result indicates considerable improvements in intensity prediction of TCs in ERA5 over ERAI data set.

The composite structure analysis of TCs in reanalysis data sets shows that GFS, CFSR, and JRA55 provide a better representation of $\partial_v$ and $\partial_\theta$ within the inner core region of TCs compared to other reanalysis data sets. GFS analysis provides the most appropriate representation of horizontal wind flow patterns, low-level convergence, and upper tropospheric divergence followed by ERA5, CFSR, and MERRA2 data sets, respectively. The relative vorticity structure is best represented in ERA5 data set followed by GFS, CFSR, and MERRA2 data sets, respectively. ERAI data set gives the most weak structure representation of these parameters in the inner core region of TCs. In the inner core region of TCs the values of convergence increase with the intensification of the TCs. The results of the study show that, beyond the CS stage, low level convergence in ERAI data set decreases with the intensification. This result is in contrast to the real TC structure representation because during intense stages the values of convergence should be more than during weaker stages.

In general, we can conclude that the minimum values in the $\Delta R$ and $\Delta V_{\text{max}}$ ($\Delta \text{MSLP}$) are found in GFS analysis (ERA5) followed by ERA5 (GFS). Only GFS and CFSR data sets capture the most intense stages (ESCS) of the TCs. ERA5 and MERRA2 data sets can capture TCs intensity till the VSCS stage, while the other two data sets are unable to obtain intensification beyond the SCS stage. Although GFS analysis represents early intensification and, in some cases, overprediction of the weak category (D, DD, and CS) of TCs, and also during the most intensified stages (beyond the CS stage), the structures of TCs are better represented in GFS analysis as compared to other data sets. Thus, it can be stated that GFS analysis represents the evolution of TCs more realistically and is comparatively better suited for the study of the growth of TCs and for providing initial conditions in the NWP model over NIO. Among other five reanalysis data sets, ERA5 (CFSR) data set can be concluded to provide a better representation of TCs structure (intensity).

It is speculated that the majority of the errors in the spatiotemporal representation of development of TCs are mainly due to the coarse resolution of reanalysis data sets and unknown physical processes that contribute to the evolution of TCs and their parameterization in NWP. However, reanalyses developed based on high-resolution mesoscale models are more in agreement with IMD best track data sets (Rajasree, Kesarkar, Bhat, Singh, et al., 2016; Rajasree, Kesarkar, Bhat, Umakanth, et al., 2016). Therefore, it can be said that there is a need to develop a high-resolution global reanalysis data set that can represent the spatiotemporal evolution of TCs more accurately. Also, there is a need for increasing the number of in situ or satellite weather observations measuring vertical structure and surface parameters, especially over marine areas during genesis and evolution of TCs. It will enable the scientific community to study the development of TCs more realistically.

References

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. https://doi.org/10.1002/qj.828

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Ono, K., & Yoshida, H. (2018). The JRA55: An interim reanalysis data set with a high horizontal resolution. Geoscientific Model Development, 11, 1351–1368. https://doi.org/10.5194/gmd-11-1351-2018

Flaounas, E., Kotroni, V., Lagouvardos, K., & Flouvaonas, I. (2014). CycloTRACK (v1.0)—tracking winter extratropical cyclones based on relative vorticity: Sensitivity to data filtering and other relevant parameters. Geoscientific Model Development, 7, 1181–1192. https://doi.org/10.5194/gmd-7-1181-2014

Gelaro, R., McCarty, W. J., Suárez, M., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA2). Journal of Climate, 30(14), 5419–5428. https://doi.org/10.1175/JCLI-D-16-0758.1

Acknowledgments

Authors are grateful to Director, NARL, and Vice-Chancellor, Assam University, Silchar for providing opportunity and support to carry out this work. Authors are grateful to the Department of Science and Technology (DST), India, for providing Inspire Fellowship to Ms Pampi Malakar which enable her to carry out this research. Authors are thankful to ECMWF, JMA, NASA, NCEP, and NCAR for providing Era-Interim reanalysis, ERA5, JRA55, MERRA2, CFSR, CFSv2, and GFS analysis data sets for this study. ERAI data can be obtained from the link http://apps.era.int/datasets/data/interim-full-daily/ and ERA5 data from the link https://www.era5.com/. The JRA55 can be obtained from the link https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs/. The GFS analysis downloaded from the web link https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/glo-bal-forcast-system-gfs/ is used in this study. The JRA55 is available from the link https://rda.ucar.edu/datasets/ds628.0/ and the MERRA2 can be obtained from the link https://disc.gsfc.nasa.gov/datasets/M2I6NPANA_V5.12.4/.

The CFSR and CFSV2 data sets are available from the link https://rda.ucar.edu/datasets/ds938.0/. The MERRA2 can be obtained from the link https://rda.ucar.edu/datasets/ds628.0/. The MERRA2 can be obtained from the link https://rda.ucar.edu/datasets/ds628.0/. The MERRA2 can be obtained from the link https://rda.ucar.edu/datasets/ds628.0/. The MERRA2 can be obtained from the link https://rda.ucar.edu/datasets/ds628.0/. The MERRA2 can be obtained from the link https://rda.ucar.edu/datasets/ds628.0/.
Harada, Y., Kamahori, H., Kobayashi, C., Endo, H., Kobayashi, S., Ota, Y., et al. (2016). The JRA55 reanalysis: Representation of atmospheric circulation and climate variability. *Journal of the Meteorological Society of Japan. Ser. II*, 94(5), 269–302. https://doi.org/10.2151/jmsj.2016-015

Hatsushika, H., Totsuji, J., & Fiorino, M. (2006). Impact of wind profile retrievals on the analysis of tropical cyclones in the JRA-25 reanalysis. *Journal of the Meteorological Society of Japan*, 84(5), 891–905.

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., Sabater, J. M., Nicolas, J., et al. (2019). Global reanalysis: Goodbye ERA-Interim, hello ERA5. ECMWF Newsletter, 159, 17–24. https://doi.org/10.21957/vf291hehd7

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, 93, 5–48. https://doi.org/10.2151/jmsj.2015-001

Maloney, E. D., & Hartmann, D. L. (2000). Modulation of eastern north Pacific hurricanes by the Madden-Julian oscillation. *Journal of Climate*, 13(9), 1451–1460. https://doi.org/10.1175/1520-0442(2000)013<1451:MOENPF>2.0.CO;2

Maloney, E. D., Hein, J. R., Planet, E., Lett, S., & Hartmann, D. L. (2000). Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian oscillation. *Science*, 2002(2002), 2002–2005. https://doi.org/10.1126/science.287.5460.2002

Manning, D. M. (2007). The utility of the ERA40 Cyclone Phase Space in Trend Diagnosis and North Atlantic Tropical Cyclone Reanalysis. Retrieved from https://purl.fvc.org/eua/id/FSU_migr_etd-2759

Manning, D. M., & Hart, R. E. (2007). Evolution of North Atlantic ERA40 tropical cyclone representation. *Geophysical Research Letters*, 34, L028266. https://doi.org/10.1029/2006GL028266

Montgomery, M. T., Nicholls, M. E., Cram, T. A., & Saunders, A. B. (2006). A vertical hot tower route to tropical cyclogenesis. *Journal of the Atmospheric Sciences*, 63(1), 355–386. https://doi.org/10.1175/JAS3604.1

Murakami, H. (2014). Tropical cyclones in reanalysis data sets. *Geophysical Research Letters*, 41, 2133–2141. https://doi.org/10.1002/2014GL061959.

Onogi, K., Totsuji, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., et al. (2007). The JRA-25 reanalysis. *Journal of the Meteorological Society of Japan*, 85(3), 369–432. https://doi.org/10.2151/jmsj.85.369

Rajasee, V. P. M., Kesarkar, A. P., Bhave, J. N., Singhi, V., Umakanth, U., & Harish Varma, T. (2016). A comparative study on the genesis of North Indian Ocean tropical cyclone Madi (2013) and Atlantic Ocean tropical cyclone Florence (2006). *Journal of Geophysical Research: Atmospheres*, 121, 13,826–13,858. https://doi.org/10.1002/2016JD025412

Rajasee, V. P. M., Kesarkar, A. P., Bhave, J. N., Umakanth, U., Singhi, V., & Harish Varma, T. (2016). Appraisal of recent theories to understand cyclogenesis pathways of tropical cyclone Madi (2013). *Journal of Geophysical Research: Atmospheres*, 121, 8949–8982. https://doi.org/10.1002/2016JD025188

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., et al. (2011). MERRA: NASA’s Modern Era Retrospective Analysis for Research and Applications. *Journal of Climate*, 24(14), 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27, 2185–2208.

Schenkel, B. A., & Hart, R. E. (2012). An examination of tropical cyclone position, intensity, and intensity life cycle within atmospheric reanalysis datasets. *Journal of Climate*, 25(2007), 3435–3475. https://doi.org/10.1175/2011JCLI4208.1

Scoccimarro, E., Gualdi, S., & Navarra, A. (2012). Tropical cyclone effects on Arctic sea ice variability. *Geophysical Research Letters*, 39, GL05298. https://doi.org/10.1002/2012GL05298

Sela, J. (2009). The implementation of the sigma pressure hybrid coordinate into the GFS. NOAA/NCEP Office Note 461. 25 pp.

Sela, J. (2010). The derivation of the sigma pressure hybrid coordinate semi-Lagrangian model equations for the GFS. NOAA/NCEP Office Note 462. 31 pp.

Simmons, A., Uppala, S., Dee, D., & Kobayashi, S. (2006). ERA-Interim: New ERAI reanalysis products from 1989 onwards ERA-Interim: New ERAI reanalysis products from 1989 onwards. *Meteorology Section of ERAI Newsletter*, 07(110), 25–35. https://doi.org/10.21957/pocnec236

Srivastava, R., & Huber, M. (2006). Low frequency variability in globally integrated tropical cyclone power dissipation. *Geophysical Research Letters*, 33, L026167. https://doi.org/10.1029/2006GL026167

Thorne, P. W., & Vose, R. S. (2010). Reanalyses suitable for characterizing long-term trends. *Bulletin of the American Meteorological Society*, 92(1), 65–70. https://doi.org/10.1175/2010BAMS3070.1

Wang, Z. (2012). Thermodynamic aspects of tropical cyclone formation. *Journal of the Atmospheric Sciences*, 69(8), 2433–2451. https://doi.org/10.1175/JAS-D-11-0298.1

Wang, Z. (2014). Role of cumulus congestus in tropical cyclone formation in a high-resolution numerical model simulation. *Journal of the Atmospheric Sciences*, 71(5), 1681–1700. https://doi.org/10.1175/JAS-D-13-0257.1