Analysis of Tropical Storm Formation Based on Ensemble Data Assimilation and High-Resolution Numerical Simulations of a Nondeveloping Disturbance

ANDREW B. PENNY
Department of Meteorology, Naval Postgraduate School, Monterey, California

JOSHUA P. HACKER
National Center for Atmospheric Research, Boulder, Colorado

PATRICK A. HARR
Department of Meteorology, Naval Postgraduate School, Monterey, California

(Manuscript received 10 March 2016, in final form 26 June 2016)

ABSTRACT

A nondeveloping tropical disturbance, identified as TCS025, was observed during three intensive observing periods during The Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC)/Tropical Cyclone Structure-2008 (TCS-08) field experiment. The low-level circulation of the disturbance was relatively weak, asymmetric, and displaced a considerable distance from the midlevel circulation. An ensemble of high-resolution numerical simulations initialized from global model analyses was used to further examine TCS025. These simulations tended to unrealistically overdevelop the TCS025 disturbance. This study extends that work by examining the impact of assimilating in situ observations of TCS025 and dual-Doppler radial velocities from the airborne Electra Doppler Radar (ELDORA) using the Data Assimilation Research Testbed (DART) ensemble data assimilation system.

The assimilation of observations results in a more accurate vortex structure that is consistent with the observational analysis. In addition, forecasts initialized from the state of the ensemble after data assimilation exhibit less development than both the control simulation and an ensemble of forecasts without prior data assimilation.

A composite analysis of developing and nondeveloping forecasts from the ensemble reveals that convection was more active in developing simulations, especially near the low-level circulation center. This led to larger diabatic heating rates, spinup of the low-level circulation from vorticity stretching, and greater alignment of the low- and midlevel vorticity centers. In contrast, nondeveloping simulations exhibited less convection, and the circulation was more heavily impacted by vertical wind shear.

1. Introduction

Tropical storm formation relies on the interaction of physical processes that span multiple spatial and temporal scales. Although the large-scale environmental conditions necessary for tropical cyclogenesis are generally well understood, convective-scale processes, which are thought to play a vital role during storm formation (e.g., Hendricks et al. 2004; Montgomery et al. 2006), are difficult to observe at adequate spatial and temporal resolution using current observing platforms. As a result, large uncertainty still exists in identifying the key processes necessary for tropical storm formation, and the scientific community has yet to converge upon a single theory to explain the process. For a brief review of tropical storm formation theories, see Sippel and Zhang (2008), Braun et al. (2010), and Davis and Ahijevych (2012). Furthermore, deep moist convective processes are inherently nonlinear and have limited predictability, such that small uncertainties in convective-related parameters of the initial state can lead to rapid error growth and divergent model solutions.

It is perhaps not surprising that forecasting tropical storm formation has proven challenging for numerical weather prediction. Halperin et al. (2013) evaluated the
ability of global forecast models to accurately forecast tropical storm formation in the North Atlantic basin based on model output from 2004 to 2011. Halperin et al. (2013) found that false alarms, misses, and improperly timed genesis forecasts were common, and that the false alarm rate generally increased with increasing forecast lead time.

Experiments using high-resolution mesoscale models also have concluded that inherent predictability can limit skill in forecasting tropical storm formation. Zhang and Sippel (2009) found that initial condition differences smaller in magnitude than typical observation uncertainty can significantly impact the characteristics of deep moist convection. This can limit tropical cyclone predictability by significantly changing the evolution as a modeled storm develops, or it can suppress storm formation entirely. Similarly, the predictability of tropical storm formation in numerical experiments has been shown to be sensitive to the presence of vertical wind shear (Zhang and Tao 2013) and the parameterization of physical processes (Nicholls and Montgomery 2013; Bao et al. 2012; Penny et al. 2016).

Given the inherent uncertainties related to modeling tropical storm formation, Zhang and Sippel (2009), Zhang et al. (2009), and Sippel and Zhang (2010) suggest the need for an ensemble-based probabilistic forecasting approach. A number of recent studies have proven the utility of assimilating in situ and radar observations within an ensemble Kalman filter (EnKF) data assimilation system to create ensemble analyses. These EnKF analyses can improve initial conditions for high-resolution modeling experiments. Torn and Hakim (2009) found that the assimilation of dropwindsondes deployed into Hurricane Katrina (2005) during the Hurricane Rainband and Intensity Change Experiment (RAINEX) helped to reduce the position and intensity error of the ensemble mean for short-range forecasts. Compared to experiments in which only conventional observations were assimilated, Jung et al. (2012) showed that the initial position and subsequent track forecasts of Typhoon Sinlaku (2008) were improved when targeted dropwindsondes were assimilated. Zhang et al. (2009) and Sippel and Zhang (2010) demonstrated that forecasts of Hurricane Humberto (2007) initialized from EnKF analyses after the assimilation of Weather Surveillance Radar-1988 Doppler (WSR-88D) radial velocity observations more closely matched the best-track position and intensity values compared to simulations without data assimilation or when three-dimensional variational (3DVAR) data assimilation was used, and that these forecasts were able to capture the rapid development of the storm. Weng and Zhang (2012) showed that assimilating airborne Doppler radar observations resulted in a more accurate dynamic and thermodynamic structure of Hurricane Katrina (2005), and that initializing forecasts from the resulting EnKF analysis helped to reduce forecast track errors. Zhang et al. (2011) found that for cases from 2008 to 2010, forecasts initialized from EnKF analyses that utilized near-real-time airborne Doppler radar data generally outperformed the intensity forecasts of the Hurricane Weather Research and Forecasting (HWRF) model, Geophysical Fluid Dynamics Laboratory (GFDL) model, and the official forecasts from the National Hurricane Center (NHC).

In this study, we build upon previous work in examining tropical storm formation through the assimilation of Doppler radar data, but from the perspective of a nondeveloping tropical disturbance. This paper serves as a follow-on study to Penny et al. (2015) and Penny et al. (2016) in the investigation of the nondeveloping tropical disturbance “TCS025” observed during the Observing System Research and Predictability Experiment (THORPEX) Pacific Asian Regional Campaign (T-PARC)/Tropical Cyclone Structure-2008 (TCS-08) field experiment (Elsberry and Harr 2008). In the observational component of the study, Penny et al. (2015) examined flight-level, dropwindsonde, and Electra Doppler Radar (ELDORA) observations of TCS025. These data indicated that TCS025 had a relatively weak, asymmetric, and vertically misaligned circulation. Although periods of deep convection were observed, Penny et al. (2015) concluded that the unfavorable structure of TCS025 coupled with a surrounding environment characterized by moderate vertical wind shear and large horizontal flow deformation allowed low equivalent potential temperature (θe) air to negatively impact the inner-core thermodynamic environment. As a result, convection remained weak, even during a period when vertical wind shear weakened. The lack of deep convection limited further organization and kept the disturbance from developing further.

Penny et al. (2016) used a multiphysics ensemble based on the Advanced Weather Research and Forecasting (WRF-ARW) model to further examine aspects of the nondevelopment of TCS025. Similar to other high-resolution modeling studies (e.g., Fritz and Wang 2013), it proved difficult to simulate the nondevelopment of TCS025, as the majority of ensemble members forecast development. Penny et al. (2016) found that the outcome of the TCS025 simulations was extremely sensitive to the representation of convection and latent heating; simulations that experienced overdevelopment exhibited stronger convection and larger diabatic heating rates, while convection was relatively weak in the single nondeveloping simulation.

Given the challenges in correctly simulating the nondevelopment of TCS025 documented by Penny et al. (2016), this study examines the impact of assimilating
in situ and radial velocity observations collected during TCS025 using the Data Assimilation Research Testbed (DART) ensemble data assimilation system (Anderson et al. 2009). It is found that data assimilation results in a more realistic circulation structure compared to global model analyses (Penny et al. 2015), and the sea level pressure (SLP) of forecasts initialized after data assimilation showed less development compared to both the control simulation and ensemble forecasts for which no data assimilation was performed. A composite analysis indicates that developing members are less susceptible to the impacts of vertical wind shear, and exhibit both stronger convection and larger diabatic heating rates compared to nondeveloping members.

The remainder of this paper is organized as follows: section 2 outlines the methodology, section 3 examines the impact of data assimilation, sections 4 and 5 provide a physical interpretation and discussion of the results, and section 6 presents the conclusions.

2. Methodology

Although five aircraft missions over three intensive observing periods (IOPs) were conducted to observe TCS025, the time period encompassing the data collected during the second IOP was chosen for the ensemble data assimilation experiment. Data collected during the second IOP early on 28 August include ELDORA, producing superior coverage near the circulation center (Fig. 1). The TCS025 circulation was moving relatively slowly at this time, and global analyses indicated the system was experiencing some organization (not shown).

a. Observations

Observing platforms available during the second IOP of TCS025 and within the largest model domain are summarized in Table 1, along with the number of horizontal wind observations available for assimilation. ELDORA radial velocities and cloud-track winds compose most of the observations available for assimilation. The ELDORA is a dual-Doppler radar (Hildebrand et al. 1996; Wakimoto et al. 1996) that was mounted on the tail of the Naval Research Laboratory (NRL) P-3 during T-PARC/TCS-08 (Lee et al. 2009).

The enormous data volume collected during the second and third IOPs demands several data thinning steps prior to assimilation. Thinning reduces the computational expense and allowed the spatial resolution of the data to more closely match the background analysis grid. After the radial velocity data is corrected for aircraft motion (Testud et al. 1995) and quality controlled following the procedure outlined by Bell et al. (2013), every fifth “sweep” of the fore and aft antennas is retained (each sweep contained about 4 s of data collected during one rotation of the antenna). This increases the observation spacing along the flight track from about 600 m to 3 km.

Individual radar sweeps are thinned further by only retaining every fifth beam in elevation. Radial velocity and spectrum width, which is a measure of the standard deviation of radial velocity within a single radar volume, were averaged along each beam using a stride (bin size) of 20. An example of a sweep, before and after thinning, is given in Fig. 2. Assuming there are enough independent samples in each radar volume, the spectrum width can be averaged by taking the square root of the average of the variances (M. Bell 2013, personal communication). For a horizontal beam, averaging along the beam increases the observation spacing from 150 m to about 3 km. Finally, only radial velocity observations with a corresponding spectrum width \( \leq 6 \text{ ms}^{-1} \) are retained, removing radial velocities with large uncertainty and thus minimal impact on the analysis.

The NRL P-3 flight-level data (1 Hz) were thinned by retaining every twenty-fifth set of observations (horizontal wind, temperature, and humidity). With a typical ground speed of about 135 m s\(^{-1}\), the observation spacing

![Fig. 1. Summary of airborne observations near 0000 UTC 28 Aug 2008. The blue line corresponds to the flight path of the USAF WC-130J and the red line is the flight path of the NRL P-3, which housed the ELDORA. Open circles denote dropwindsonde locations and wind barbs correspond to the wind at 1500 m AGL. Numbers in parentheses in the inset correspond to the number of dropwindsondes deployed by each aircraft during the mission. The black dashed curve defines the circulation center at 850 hPa from the ECMWF YOTC analyses from 0000 UTC 27 Aug–0000 UTC 29 Aug.](image-url)
of the thinned NRL P-3 flight-level data is about 3.4 km. Since the U.S. Air Force (USAF) 53rd Weather Reconnaissance Squadron (Hurricane Hunters) WC-130J flight-level data has 30-s temporal resolution, additional thinning is not required as the WC-130J ground speed is equivalent to a spacing of 4.7 km. Dropwindsonde data collected from the NRL P-3 and USAF WC-130J flights are thinned by only retaining every fifth set of observations in the vertical, which corresponds to a vertical spacing of about 40 m.

b. Model configuration

The deterministic forecasts of TCS025 by Penny et al. (2016) were conducted within the framework of a multiphysics ensemble without data assimilation. These forecasts tended to overdevelop the disturbance in the majority of cases. For this study, a configuration was chosen similar to that used in the Penny et al. (2016) simulation that employed the Purdue–Lin microphysics scheme (Lin et al. 1983; Chen and Sun 2002), which accounts for the mixing ratios of water vapor, liquid cloud water, rain, ice, snow, and graupel. Based on minimum SLP, the degree of development for the Purdue–Lin microphysics case was near the average of the developing cases, and serves as the control simulation.

Ensemble forecasts of TCS025 were conducted using the WRF-ARW nonhydrostatic mesoscale model version 3.4.1 (Skamarock et al. 2008). Domains were triply nested using 27-, 9-, and 3-km horizontal grid spacing with 45 vertical levels. The 3-km grid was large enough to encompass most of the development region for TCS025.

For the 27- and 9-km domains, cumulus convection was parameterized using the Kain–Fritsch scheme (Kain 2004). Cumulus convection was explicitly represented in the 3-km domain. The Rapid Radiative Transfer Model (RRTM; Mlawar et al. 1997) and Dudhia (1989) were used to handle longwave and shortwave radiation, respectively. In addition, the Yonsei University (YSU) boundary layer (Hong et al. 2006) and the fifth-generation Pennsylvania State University–NCAR Mesoscale Model similarity theory surface layer schemes (Skamarock et al. 2008) were employed.

Initial and lateral boundary conditions were derived from the 0.25° European Centre for Medium-Range
Weather Forecasts (ECMWF) analyses that were also used in the observational study of TCS025 (Penny et al. 2015). A comparison of dropwindsonde data and ECMWF analyses (not shown) revealed good agreement during this period, and the representation of pertinent large-scale features in the WRF-ARW deterministic forecasts appeared more realistic when initialized from the ECMWF analysis rather than other global analyses. A 96-member ensemble was created by randomly drawing from the WRF-VAR default background error covariance following Torn et al. (2006), to perturb the initial and lateral boundary conditions derived from the ECMWF analyses.

c. Data assimilation

DART facilitates ensemble data assimilation with a serial-observation processing implementation of the ensemble adjustment Kalman filter (Anderson 2003). During the period 1200 UTC 27 August–0600 UTC 28 August 2008, hourly assimilation proceeds with all available observations valid within ±25 min of the hour. The state vector is formed from all three domains. Forward operators are computed on the highest-resolution domain, and analysis increments are computed and applied on all three domains. The adaptive covariance inflation scheme from Anderson (2007) is used to ensure the forecast (prior) distribution sufficiently overlaps the observation likelihood. Sample statistics are from a 96-member ensemble.

The radar radial velocity forward operator in DART required that aircraft coordinates, radar beam azimuth, and elevation angles were paired with each airborne ELDORA observation. The forward operator uses hydrometeor species accounted for by the bulk microphysics scheme in the WRF-ARW model to determine an additional component of radial velocity due to precipitation fall speed. The fall speed was then added to the model-derived radial velocity, which allowed the forward operator to map the model wind speed to observation space in a manner representative of actual radar observations during the assimilation cycle.

Localization to mitigate sampling error is applied with the fifth-order piecewise polynomial from Gaspari and Cohn (1999), their Eq. (4.10), and is a function of space alone. Localization improves computational efficiency, and ensures that far-removed grid points are not impacted by observations that sample flows dominated by small scales that might be uncorrelated with larger scales. Each observation platform uses a different localization distance. Localization values are summarized in Table 1. The ELDORA radial velocities, which are most numerous and dense, receive the most stringent localization, followed by the research aircraft flight-level observations, satellite cloud-track winds, and dropwindsondes. Longer localization distances were used for the ACARS and radiosonde data.

Initial observation error (standard deviation) values for T-PARC/TCS-08 flight-level and dropwindsonde data come from a function of pressure borrowed from the data assimilation system for the North American Mesoscale Forecast System (NAM). Observation errors for the ELDORA radar were initially based on Bell et al. (2012), and were specified to account for both instrumentation and representativeness error. Errors for each radar radial velocity observation were determined from the spectrum width and elevation angle; radial velocity error was increased for observations at large elevation angles because a significant component of the measured radial velocity might result from hydrometeor fall speed. Probability density functions of the observation error standard deviation are shown in percent of total number of observations in Fig. 3 for observations before all stages of data thinning were completed. Accounting for elevation angle increases the average error standard deviation and broadens the distribution (cf. the cyan and purple lines in Fig. 3b). Observation error values for radiosonde observations and satellite winds were taken from National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) lookup tables.

After an initial data assimilation experiment that resulted in an overdispersive ensemble, observation errors were tuned once following Desroziers et al. (2005), such that the “total spread” and mean squared error of the ensemble mean are equal within sampling error. The total spread is defined here as the ensemble variance plus observation error variance. To improve agreement between total spread and error, observation errors were adjusted downward, consistent with the higher resolution here compared to the NCEP Global Data Assimilation System (GDAS); all observation error variance values were reduced by 50% except for aircraft and dropwindsonde dewpoint observations, which seemed well matched to the spread during the initial experiment. For example, the total spread was initially too large for both dropwindsonde winds and ELDORA radial velocities (Figs. 4a,c), because the observation errors were too large. After reducing the observation error (Figs. 4b,d), the ratio between total spread and error is closer to 1.

Smaller observation errors also increase the analysis increment magnitudes, and can slightly increase the prior ensemble-mean errors. The larger analysis increments are expected because the high-density observations are better fit at analysis time. Slightly larger errors expose the model inadequacy; accuracy of even the thinned observations, near grid scale, are difficult for the model to simulate. As expected, the analysis increments at the
dropwindsonde wind locations are increased at the beginning of the time window for ELDORA observations (approximately 0112 UTC 28 August 2008), with greater impact after reducing the ELDORA radial velocity observation errors (Fig. 4b).

After the final set of observations is assimilated at 0600 UTC 28 August 2008, the analysis ensemble from that time is advanced as an ensemble simulation with perturbed lateral boundary conditions from the ECMWF analyses. The 54-h ensemble simulation ends at 1200 UTC 30 August 2008.

3. Effect of data assimilation on storm development

Here, we document the aggregate response of the data assimilation impacts to the simulation of TCS025, which is that storm development is suppressed. Physical interpretation of the impacts is provided in section 4. Examination of the relative impacts of the various observation sets is beyond the scope of this paper; that analysis is reserved for future work. Although a quantitative comparison with the control simulation is not appropriate, the qualitative differences with respect to the circulation structure are used to examine the utility of the data assimilation. Comparisons are made between the deterministic control simulation (Fig. 5a), and storm-centered composites of the 96-member ensemble after data assimilation (Fig. 5c), but configured identically to the ensemble data assimilation experiment. Because the high-resolution simulations of Penny et al. (2016) were extremely sensitive to the representation of convection with respect to the development of TCS025, the additional ensemble experiment was conducted without data assimilation to assess whether the random perturbations used to create the initial and lateral boundary conditions of the ensemble might have impacted convection, and altered the development outcome.

At 0600 UTC 28 August, a closed low-level circulation was present in the control simulation near 19°N, 152.5°E, and it was vertically aligned with the midlevel circulation (not shown). This is also evident in the storm-relative frame of the control simulation (Fig. 5a) and the storm-centered composite from the ensemble without data assimilation (Fig. 5c), as the vorticity in the middle levels is centered over the SLP minimum. In contrast, the midlevel vorticity and circulation in the storm-centered composite from the ensemble with data assimilation (Fig. 5b) is displaced a considerable distance to the southeast of the SLP minimum, indicating that the mid- and low-level circulations were poorly aligned in the vertical. These structural characteristics from the ensemble that employed data assimilation are in agreement with the findings from the observational analysis of TCS025 by Penny et al. (2015), and indicate that the data assimilation helped to constrain the model state to be more in line with observations by disrupting the vertical coherence of the circulation.
Based on minimum SLP (Fig. 6), the data assimilation reduces the forecast probability that a storm will develop (Fig. 6a). The cyclone does not deepen in the ECMWF analysis (cyan), but it does deepen in the WRF-ARW control simulation (orange) as the central pressure lowers from 1006 to 990 hPa just before the end of the...
The entire ensemble with data assimilation (Fig. 6a) begins with a SLP lower than the ECMWF analysis, which could be expected from resolution arguments alone. But in the first 6 h after assimilation, minimum pressure in all of the ensemble members rebounds to a mean of 1004 hPa, closer to the control simulation and ECMWF analysis. During the last few hours of the simulation, the ensemble is highly skewed, with a larger number of nondeveloping storms; the mean is approximately 12 hPa above the control simulation near the end of the forecast period. In contrast, the experiment without data assimilation (Fig. 6b) exhibits far greater variability in minimum SLP values, as well as a larger number of members that develop the disturbance. The final mean minimum SLP of the experiment without data assimilation is similar to that of the control simulation. This demonstrates that the forecast improvements with respect to the nondevelopment of TCS025 result from data assimilation rather than the random perturbations used to create the ensemble.

To more clearly identify differences between developing and nondeveloping members from the ensemble with data assimilation, quartiles, denoted by colors in Fig. 6a, were defined according to the minimum SLP during the period between 0000 UTC 29 August and the end of the simulations. The result is eight members in
the first (developing) quartile, hereafter Q1, and six in the fourth (nondeveloping) quartile, hereafter Q4. Only members that remain in the same quartile during the period between 0000 UTC 29 August and the end of the simulations are included in the quartile means. The mean of the developing quartile is close to the control simulation, owing to the long tail of low pressures in the ensemble distribution.

The impact of data assimilation is also evident when comparing storm tracks (Fig. 7). The track in the control simulation is generally to the east of the ECMWF analysis track, while the 96-member ensemble mean most closely follows the ECMWF track, although its position at 1200 UTC 28 August ($T + 6$ h) is a considerable distance north of the ECMWF position. In addition, large track differences between the developing (Q1) and nondeveloping (Q4) quartile means are clear. The Q1 track is displaced east and is closest to the control track early in the forecast, while the Q4 mean is west of the ECMWF analysis and the 96-member ensemble-mean positions. By 1200 UTC 29 August, the tracks of the Q1 and Q4 means converge with the ECMWF and 96-member ensemble mean tracks, and follow a north-northwest trajectory.

4. Physical interpretation

Physical differences between the Q1 and Q4 composite means help reveal reasons for the differences between them. Differences in the model-derived cloud-top temperature and SLP during the first 6 h of forecast integration are shown in Fig. 8 at 3-h intervals. At 0600 UTC 28 August (Fig. 8a), the Q1 mean possesses an area of colder cloud-top temperature (i.e., deep convection) near 20°N, 158°E, but the SLP is only slightly lower (~0.5 hPa) compared to the Q4 mean just to the northwest of this area near 21°N, 157°E (Fig. 8a). A larger area of lower SLP exists in the Q4 mean at this time (e.g., near 17°N, 155°E), but model-derived cloud-top temperature differences indicate a lack of deep convection in the Q4 mean. By 0900 UTC (Fig. 8b), the area of colder cloud-top temperature has expanded southward in the Q1 mean, and the magnitude of the differences has increased. The SLP at 0900 UTC is lower in the Q1 mean over a broad area, with a local minima relative to the Q4 mean of ~1.5 hPa near the area of largest cloud-top temperature difference. By 1200 UTC (Fig. 8c), the area of lower SLP in the Q1 mean has consolidated near 22°N, 155°E, and colder cloud-top temperatures are present to the east and south of this area.

An examination of storm-centered cloud-top temperature for the Q1 and Q4 means at 1200 UTC 28 August (Figs. 9a,b) provides additional evidence of this difference in convection near the SLP center; cloud-top temperatures present to the east and south of this area in the Q1 mean (Fig. 9a) are representative of enhanced curved banding of the convection. In contrast, convection in the Q4 mean (Fig. 9b) appears much weaker, especially close to the center.

At 1200 UTC 28 August, the minimum SLP in the Q1 composite (Figs. 9a,c,e) is lower than in the Q4 composite (Figs. 9b,d,f) (1005 vs 1006 hPa), and the SLP contours in the Q1 composite are more tightly spaced. The 500-hPa relative vorticity in the Q1 composite (Fig. 9c) is vertically aligned above the SLP center, and wind vectors at 500 hPa (Fig. 9c) indicate that the midlevel circulation is slightly displaced southwest of the SLP center. In contrast, most of the 500-hPa relative vorticity is displaced south of the SLP center in the Q4 composite (Fig. 9d), and wind vectors at 500 hPa (Fig. 9d) reveal that the midlevel circulation in the Q4 composite...
is displaced a considerable distance to the south-southwest of the SLP center.

Storm-centered 500-hPa diabatic heating rates at 1200 UTC 28 August (Figs. 9e,f) correspond to where deep convection and latent heating are expected based on the cloud-top temperature comparisons in Figs. 9a,b. The largest diabatic heating rates in the Q1 mean (Fig. 9e) occur east of the SLP center where the winds...
are strongest. In contrast, the largest diabatic heating in the Q4 mean (Fig. 9f) was south of the SLP center, although areas with large diabatic heating rates are also present to the east.

Given the large difference in the initial track locations between the Q1 and Q4 composites, differences in the environmental flow were analyzed by examining the 200–850-hPa vertical wind shear in the storm-relative...
At 0600 UTC 28 August (forecast initialization time) (Figs. 10a,b), westerly vertical wind shear is present north of the SLP center in both cases, and is associated with a westward-moving tropical upper-tropospheric trough (TUTT) cell. The vertical wind shear associated with this TUTT cell was stronger in the Q4 mean (Fig. 10b), and extends farther south and much closer to the SLP center than in the Q1 mean (Fig. 10a).

At 1200 UTC 28 August (Figs. 10c,d), the vertical wind shear north of the SLP center is from the southwest, as the TUTT cell had continued to move west after 0600 UTC. Vertical wind shear is weaker north of the SLP center in the Q1 mean (Fig. 10c). Similar to conditions at 0600 UTC, the SLP center is in a more favorable region of reduced vertical wind shear compared to the Q4 mean (Fig. 10d).

The vertical wind shear pattern in the Q1 mean resembles that of an upper-level anticyclone centered over the SLP center, which McBride (1981) and McBride and Zehr (1981) found to be a characteristic of developing disturbances based on an analysis of radiosonde composites. This structure is also evident in the Q4 mean, but appears to be weaker and is not centered over the SLP center. The enhanced vertical wind shear south of the SLP center in the Q4 mean (Fig. 10c) is consistent with the development of a TUTT cell, which can be associated with the formation of a low-pressure system. The vertical wind shear pattern in the Q1 mean (Fig. 10c) is indicative of a more stable atmospheric environment, where the SLP center is located in a region with reduced vertical wind shear.
of the SLP center in the Q1 mean corresponds to strong upper-level outflow from the deep convection in that region.

Given the notable differences in the representation of convection between the Q1 and Q4 means (e.g., Figs. 9a,b), individual members that exhibit a similar degree of development to that of their respective quartile means were selected from Q1 and Q4 for further examination of convective characteristics and circulation budget tendencies. Contoured frequency by altitude diagrams (CFADs; see Yuter and Houze 1995) of vertical velocities (Figs. 11a,b) and diabatic heating rates (Figs. 11c,d) were constructed from model output within the region 16°–23°N, 151°–157°E (green-boxed region in Fig. 8) for the 6-h period of 0600–1200 UTC 28 August.

The vertical velocity CFADs (Figs. 11a,b) indicate that strong updrafts (>6 m s⁻¹) occur more frequently near 500 hPa in the Q1 member (Fig. 11a) compared to the Q4 member (Fig. 11b), and the frequency of less-intense updrafts is greater at this level. At the second maximum in vertical velocity near 250 hPa, the frequency of updrafts >5.0 m s⁻¹ is slightly larger in the nondeveloping Q4 member (Fig. 11b). Strong downdrafts are also more prevalent in the Q1 member near 700 hPa (Fig. 11a).

Similar diagrams of diabatic heating rates (Figs. 11c,d) indicate that the largest heating rates (>120 K h⁻¹) occur more frequently near 500 hPa in the Q1 member, and that grid cells with positive, but less-intense heating rates are also far more abundant. This matches the pattern observed for vertical velocity in that stronger updrafts correspond with larger diabatic heating rates. Large low-level diabatic cooling rates also occur more frequently in the Q1 member where downdrafts are stronger. These contrasts in diabatic heating match the patterns observed in the mean composites of 500-hPa diabatic heating rates (Figs. 9e,f).

Circulation budgets were also computed over the same area and time period as the CFADs in Fig. 11 for the individual members from the Q1 and Q4 quartiles (Fig. 12). The circulation budget was computed in terms of area-averaged vorticity following Davis and Galarneau (2009); contributions from the vertical stretching of vorticity, horizontal eddy fluxes of vorticity anomalies, and tilting of horizontal vorticity into the vertical plane were explicitly computed, while tendencies from friction were implicitly accounted for in a residual term. Tendency profiles in Fig. 12 indicate considerable differences in vorticity evolution between Q1 and Q4 during this period. A net positive tendency below 500 hPa is evident in the developing Q1 member (Fig. 12a), which is largely a result of the vertical vorticity stretching. The stretching tendency is maximum near the surface and negative above 500 hPa, especially near 200 hPa where upper-level divergence dominates (not shown). Contributions from horizontal eddy fluxes of vorticity anomalies are largely negative in the Q1 member throughout the depth of the troposphere, but not enough to significantly offset the positive tendency from stretching. The tilting tendency is mostly negligible except for positive tendencies near 300 hPa.

The largest net positive tendency in Q4 occurs near 550 hPa and is almost entirely due to stretching. This suggests that stratiform precipitation processes are more active in the Q4 member, leading to enhanced midlevel convergence and positive vorticity tendencies in the midtroposphere. While the stretching tendency for the Q4 member is large in the midlevels, it is smaller in magnitude in the lower levels. This results in a smaller net positive vorticity tendency in the lower troposphere for the Q4 member. The net tendency in the Q4 member is strongly negative above 500 hPa due to the negative tendencies from stretching and horizontal eddy fluxes.

In addition to the ensemble members that remain within a particular quartile following 0000 UTC 29 August, the average low- and midlevel thermodynamic conditions of ensemble members that transitioned from the middle two quartiles to Q1 (Q23→Q1) or to Q4 (Q23→Q4) are also analyzed. This allows for a comparison between ensemble members in the middle two quartiles that exhibit a deepening or weakening trend, but are not obviously predisposed to either tendency based on the initial conditions at 0600 UTC 28 August.

A transition was defined as the time period encompassing a monotonic change of at least two quartiles over 18 h ending in either Q1 (developing transition) or Q4 (nondeveloping transition). Membership of the transitioning groups is not constant over time, and members that experience a developing transition could later be included with the nondeveloping transitions (or vice versa) if conditions warrant. Given that the number of transitions decreases over time as the spread between the minimum SLP of the quartile means increases, insufficient transitions exist to create a meaningful composite past 1800 UTC 29 August.

The thermodynamic conditions of the transitioning members and those remaining in Q1 and Q4 were analyzed by averaging low- and midlevel θ_e and diabatic heating rates within a 2° box from the SLP center (Fig. 13). Except for low-level diabatic heating rates (Fig. 13c), the θ_e and diabatic heating rates are largest for the Q1 mean. Thermodynamic conditions of the transitioning members remained close to the 96-member ensemble mean values. The average θ_e and diabatic heating rates are largest for developing transitions in the midlevels (Figs. 13a,b), but
slightly larger for nondeveloping transitions in the low levels (Figs. 13c,d).

A closer examination of midlevel diabatic heating rates (Fig. 13a) reveals a spike in the heating rates following 0600 UTC 28 August that is associated with an initial convective burst. This is especially evident in the Q1 mean; the Q4 mean shows a slight decline in diabatic heating rates over this period. After 1200 UTC 28 August, the midlevel diabatic heating rates increase for all groups. Apart from a temporary decline in the Q4 mean diabatic heating near 1200 UTC 29 August, the rates remain fairly steady until 0000 UTC 30 August when they begin to decline for all groups near the end of the forecast period.

The rank ordering of average 700–400-hPa $\theta_e$ (Fig. 13b) is similar to the midlevel diabatic heating (Fig. 13a). The initial convective burst, most evident in the Q1 mean, corresponds with a sharp rise in midlevel $\theta_e$ followed by a more gradual increase over the remaining forecast period. Despite a temporary decline in midlevel $\theta_e$ in the Q4
mean, an initial increase in midlevel $\theta_e$ also occurs in the other means. Following 0600 UTC 29 August, midlevel $\theta_e$ remains fairly steady in the ensemble mean, but decreases slightly in the Q4 mean near the end of the forecast period.

The evolution of diabatic heating and $\theta_e$ in the lower troposphere appears more complex (Figs. 13c,d). During the initial spike in the midlevel diabatic heating rates (Fig. 13a) and $\theta_e$ (Fig. 13b) in the Q1 mean, low-level diabatic cooling (Fig. 13c) also increases. Except for a brief period near 0000 UTC 29 August, low-level cooling is largest in the Q1 mean. Likewise, the midlevel diabatic heating rates (Fig. 13a) and low-level diabatic cooling rates (Fig. 13c) are larger for developing transitions than for nondeveloping transitions. This suggests that the increased convective activity in the Q1 mean and the developing transitions causes evaporative downdrafts leading to larger diabatic cooling near the surface. This process also appears to be at work in the Q4 mean, as midlevel diabatic heating rates (Fig. 13a) temporarily increase at 0000 and 1200 UTC 29 August while low-level diabatic cooling rates (Fig. 13c) also increase. Although low-level diabatic cooling becomes less prominent following 0000 UTC 29 August for all groups (Fig. 13c), the low-level cooling remains largest for the developing quartile (Q1) followed by the developing transitions (Q23→Q1).

The average low-level $\theta_e$ (Fig. 13d) is very similar for all cases at the initialization time of 0600 UTC 28 August. After 1200 UTC 28 August, the low-level $\theta_e$ begins to increase for all groups and remains largest in the Q1 mean. The magnitude of low-level $\theta_e$ in the developing transitions (Q23→Q1) is also less than in the non-developing transitions (Q23→Q4) throughout almost the entire compositing period. This suggests that the low-level thermodynamics of members experiencing a developing transition are vulnerable to the diabatic cooling from evaporative downdrafts during episodes of deep convection. It is likely that the lack of an organized circulation in these members (relative to the Q1 members) allows low-$\theta_e$ air to negatively impact the low-level thermodynamic environment near the storm center.

5. Discussion

Based on the mean SLP of ensemble forecasts initialized at 0600 UTC 28 August after data assimilation, less development of the TCS025 disturbance occurs relative to the deterministic control forecast. However, there is a considerable amount of spread in the degree of development among the ensemble members. The spread allows for comparison of the physical characteristics of the developing and nondeveloping members.

At forecast initialization time, there are only small differences in the SLP (Fig. 8) within the development region between members that remain in the developing quartile (Q1) and those in the nondeveloping quartile (Q4). There are large differences in the nature of convection based on the model-derived cloud-top temperature; deep convection is more prevalent in the
Q1 mean, and this quickly leads to surface pressure falls and system-scale organization. The vertical alignment of the circulation is better in the Q1 mean, and development of a robust anticyclone and outflow occurs above the SLP center. In contrast, convection is weaker in the Q4 mean as indicated from the analysis of vertical velocity and diabatic heating (Fig. 11), and the largest positive stretching tendencies occur in the midlevels (Fig. 12).

It is not entirely clear why convection is more active in the Q1 mean, and clarification is not possible with the 3-h time interval of the ensemble output. It is possible that the development outcome for this case is sensitive to the location of the disturbance at the forecast initialization time, since developing members tend to follow a track that was east of the nondeveloping members (see Fig. 7). The vertical wind shear analysis (Fig. 10) provides some support for this reasoning: the vertical wind shear is stronger near the beginning of the forecast period over the SLP center in the Q4 mean, which is detrimental for continued development. On the other hand, given that convection in the Q1 mean is considerably stronger, upper-level divergence and outflow resulting from the convection likely reduces the vertical wind shear magnitude in the vicinity of the convection to allow convective heating to be concentrated over the SLP center.

In the multiphysics ensemble experiments of TCS025 by Penny et al. (2016), the developing simulation follows a track similar to the Q1 mean, east of the non-developing simulation. The fact that the developing and

---

**Fig. 13.** Average thermodynamic quantities ($\pm 2^\circ$ from the minimum SLP position) for members in quartile 1 (solid red), quartile 4 (solid purple), quartiles 2 and 3 transitioning to quartile 1 (red dashed), and quartiles 2 and 3 transitioning to quartile 4 (purple dashed). (a) Average 700–400-hPa diabatic heating rate (K h$^{-1}$), (b) average 700–400-hPa $\theta_e$ (K), (c) average 1000–950-hPa diabatic heating rate (K h$^{-1}$), and (d) average 1000–850-hPa $\theta_e$ (K). Black lines correspond to the 96-member ensemble mean values. The number of members in each of the transitioning groups is indicated along the bottom left in (d).
nondeveloping simulations of Penny et al. (2016) have identical initial conditions and only differ with respect to the choice of microphysics parameterization scheme, suggests that the storm track was instead more sensitive to the nature of convection and organization of the disturbance.

The impact of deep convection on the low-level thermodynamic structure is clear. Although midlevel $\theta_e$ and diabatic heating rates are larger for the developing members and developing transitions that experienced deep convection, the diabatic cooling near the surface is also more pronounced. This indicates that evaporative downdrafts have a larger impact on the low-level thermodynamic structure for developing members, but these impacts are either not important or are overwhelmed by the positive contributions from convection. Davis (2015) suggests that the interaction of vertical wind shear and low-level cold pools resulting from evaporative downdrafts can help to organize updrafts, which is beneficial for storm formation. For members in the Q1 group that are already well organized and deepening, the diabatic cooling does not appear to significantly impact low-level $\theta_e$.

However, low-level $\theta_e$ is more heavily impacted for developing transition members that experience low-level evaporative cooling caused by deep convection; the magnitude of low-level $\theta_e$ in the developing transitions was less than in the nondeveloping transitions. This suggests that the typical incubation period of an incipient tropical disturbance may be partly governed by the time scale over which the low-level thermodynamic structure is no longer appreciably impacted by evaporative downdrafts; a gradual process of sustained convection and circulation development may be necessary to ensure that evaporative downdrafts do not overwhelm the boundary layer before surface fluxes (latent and sensible) can increase.

Based on the analysis of the high-resolution multi-physics ensemble of TCS025, Penny et al. (2016) concluded that the development outcome was extremely sensitive to the nature of convection. Small differences in the updraft strength and diabatic heating rates led to significant differences in the organizational structure of the disturbance, as well as the development outcome and final intensity. Additional experiments were conducted by Penny et al. (2016) using a configuration similar to the control simulation in this study, but with diabatic heating artificially adjusted by a small percentage. A diabatic heating rate of 95% of the control simulation still led to overdevelopment, while 90% inhibited development altogether. These results suggest that tropical cyclogenesis in numerical models can be extremely sensitive to the representation of diabatic heating. The ensemble data assimilation experiments in this study confirm that finding.

6. Conclusions

In the ensemble data assimilation experiments of the nondeveloping TCS025 disturbance, the assimilation of in situ observations and ELDORA radial velocities results in a circulation structure that better matches observations, and leads to an improvement in the minimum SLP compared to the deterministic control forecast; the ensemble mean minimum SLP is about 12 hPa higher than the control at the end of the forecast period, and is much closer to the nondeveloping solution of the ECWMF analysis. The ensemble forecast probability that a storm will develop drops considerably compared to an ensemble experiment using random initial and lateral boundary perturbations.

To assess key differences between developing and nondeveloping members, quartiles were defined based on the minimum SLP. A comparison between composite means reveals that Q1 (developing) members exhibit stronger and more widespread convection early in the forecast compared to Q4 (nondeveloping) members. As a result of the deep convection, the development and strengthening of an upper-level anticyclone near the SLP center is more prominent in the Q1 mean. In addition, the vertical alignment of the circulation is better in the Q1 mean, and less vertical wind shear occurs near the SLP center early in the forecast period, which is partly due to its location relative to the TUTT cell.

A closer examination of a developing member revealed that the midlevel diabatic heating rates and positive low-level circulation tendencies are large. Low-level spinup is driven primarily by vertical vorticity stretching in the lower troposphere, which is indicative of a convective divergence profile. In contrast, stratiform precipitation is more prevalent than deep convection in the nondeveloping member examined, such that the largest positive circulation tendencies from stretching occur in the midlevels, while the net low-level tendency is slightly negative.

These ensemble data assimilation experiments support the findings from Penny et al. (2016) in that the outcome of tropical cyclogenesis forecasts can be extremely sensitive to the modeled representation of convection. This sensitivity advocates the use of stochastic physics within an ensemble framework for tropical storm forecasting, such that the uncertainties related to the impacts from convection (i.e., diabatic heating) are represented over appropriate spatial and temporal scales. This would enable the ensemble to better encompass realistic modes of tropical storm formation or nonformation.
For TCS025, properly representing the vertical misalignment of the circulation is important for improving the probabilistic forecasts of tropical cyclogenesis. Despite the sensitivities to convection, and the apparent difficulty of correctly forecasting nondevelopment using deterministic high-resolution models (e.g., Fritz and Wang 2013; Penny et al. 2016), the assimilation of airborne ELDORA radar data and other observations clearly improves the forecast. Unfortunately, high-density observations are often not available in operations, but are likely needed to improve forecasts of tropical storm formation by enabling high-resolution numerical models to better resolve the mesoscale circulation structure of an incipient disturbance.

Acknowledgments. This research was funded by National Science Foundation Grant ATM-0736003 and by the Office of Naval Research Marine Meteorology Grant N0001412AF0002. Support for JPH was provided by the Office of Naval Research Basic Research Initiative Award N0001410WX20059. We thank Michael M. Bell from the University of Hawai‘i at Mānoa for his help with the quality control of the ELDORA data. The comments and suggestions from the anonymous reviewers are appreciated and led to an improved manuscript.

REFERENCES


Lee, W. C., M. Bell, C. Wolff, E. Loew, and M. Donovan, 2009: THORPEX Pacific Asian Regional Campaign (TPARC) and Tropical Cyclone Structure (TCS08) ELDORA data quality report. NCAr/Earth Observing Laboratory, 9 pp.


——, and R. Zehr, 1981: Observational analysis of tropical cyclone formation. Part II: Comparison of non-developing


