Impact of UAS Global Hawk Dropsonde Data on Tropical and Extratropical Cyclone Forecasts in 2016

A. C. KREN

Cooperative Institute for Marine and Atmospheric Studies, University of Miami, and NOAA/Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division, Miami, Florida

L. CUCURULL

NOAA/Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division, Miami, Florida

H. WANG

National Center for Atmospheric Research, Boulder, Colorado

ABSTRACT

A preliminary investigation into the impact of dropsonde observations from the Global Hawk (GH) on tropical and extratropical forecasts is performed using the National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS). Experiments are performed during high-impact weather events that were sampled as part of the NOAA Unmanned Aerial Systems (UAS) Sensing Hazards with Operational Unmanned Technology (SHOUT) field campaigns in 2016: 1) three extratropical systems in February 2016 and 2) Hurricanes Matthew and Nicole in the western Atlantic. For these events, the benefits of GH observations under a satellite data gap scenario are also investigated. It is found that the assimilation of GH dropsondes reduces the track error for both Matthew and Nicole; the improvements are as high as 20% beyond 60 h. Additionally, the localized dropsondes reduce global forecast track error for four tropical cyclones by up to 9%. Results are mixed under a satellite gap scenario, where only Hurricane Matthew is improved from assimilated dropsondes. The improved storm track is attributed to a better representation of the steering flow and atmospheric midlevel pattern. For all cases, dropsondes reduce the root-mean-square error in temperature, relative humidity, wind, and sea level pressure by 3%–8% out to 96 h. Additional benefits from GH dropsondes are obtained for precipitation, with higher skill scores over the southeastern United States versus control forecasts of up to 8%, as well as for low-level parameters important for severe weather prediction. The findings from this study are preliminary and, therefore, more cases are needed for statistical significance.

1. Introduction

One of the goals of the NOAA Unmanned Aerial Systems (UAS) Sensing Hazards with Operational Unmanned Technology (SHOUT; Black et al. 2014; Wick et al. 2018a) project was to conduct observing system experiments (OSEs) to test the impact of supplementary data from a long-endurance UAS platform on tropical and extratropical storm forecasts, with and without a possible data gap in global satellite coverage. The use of supplementary weather observations, both to increase our understanding of weather phenomena such as the tropical cyclone inner core and mature cyclones, and to improve forecast prediction of high-impact storms, has been employed in the meteorological community since the 1980s (Majumdar 2016). The concept of using supplementary or "adaptive" sampling relates to the assimilation of additional observations into numerical weather prediction (NWP) models aimed at improving weather forecast skill. The existence of data-sparse regions, such as over the oceans, may lead to lower NWP forecast skill for high-impact weather events due to an inadequate representation of initial conditions. Forecast skill may, at times, be improved by deploying observations in data-sparse regions and in areas that may be sensitive to large model error growth.

In contrast to the earlier studies, benefits in forecast skill from using supplementary observations have been

---

Corresponding author: A. C. Kren, andrew.kren@noaa.gov

relatively small over the past decade, with the largest improvements found for the prediction of tropical cyclone track (~10% on average with modest improvements on a case-by-case basis; Majumdar 2016). Improved NWP models with higher spatial resolution and more realistic physical parameterizations and more sophisticated data assimilation systems, as well as an increase in the amount of conventional and satellite observations, might have significantly limited the benefits of targeted observations in recent years (Langland 2005; Majumdar et al. 2011; Hamill et al. 2013; Majumdar 2016). The reader is referred to Toth et al. (2000, 2002), Langland (2005), Rabier et al. (2008), Majumdar et al. (2011), and Majumdar (2016) for a thorough review of field campaigns and sampling strategies aimed at improving the forecast skill of high-impact weather events.

Although extensive literature exists on the value of supplementary data to improve NWP forecast skill, little research has been conducted to quantify the potential value of supplementary observations under a potential loss of satellite data coverage. An OSE study by Buizza et al. (2007) found that, in the absence of satellite data in the target regions, using supplemental observations in sensitive areas of the Pacific and Atlantic Oceans can lead to improvements of ~1%–4% for the 500-hPa geopotential heights on day 2 over Europe and North America.

Ultimately, the benefits of assimilating additional observations depend on several factors, including the technique used to select the sensitive regions, the methodology and variables used for verification, the meteorological situation, the NWP model, and the data assimilation system. One of the recommendations from the Data Assimilation and Observing Systems (DAOS; Rabier et al. 2008; Majumdar 2016) working group over the past several years is to test new observing platforms, such as long-endurance UAS, as previous field missions were only able to deploy ~8–30 dropsondes per ~6-h flight (Hamill et al. 2013; Majumdar 2016). The SHOUT project addressed this recommendation by using the National Aeronautics and Space Administration (NASA) Global Hawk (GH; Black et al. 2014; Braun et al. 2016; Christophersen et al. 2017, 2018a,b; Wick et al. 2018a) UAS, a platform that can deploy as many as 88 dropsondes during a single 24-h mission. The long endurance of the GH makes it a potentially highly valuable UAS platform in providing supplementary observations for weather forecasting. A data impact study using the GH has recently been performed by Christophersen et al. (2017) for Hurricane Edouard in 2014.

This study is a preliminary investigation into the benefits of adding dropsonde data from the GH to improve the forecast skill of high-impact weather events that took place in 2016 under SHOUT, namely Hurricanes Matthew and Nicole in October 2016 and all SHOUT El Niño Rapid Response (ENRR; Dole et al. 2018) cases in February 2016.

For these case studies, the impact of GH dropsondes under a potential temporal gap in satellite data is also evaluated, comparable with what has been recently investigated using an observing system simulation experiment (OSSE; English et al. 2018). This secondary component of the study addresses a potential temporal gap between the decommissioned *Suomi National Polar-Orbiting Partnership* (*Suomi-NPP*) and the launch of the *Joint Polar Satellite System-1* (*JPSS-1*), and/or between the end-of-life of *JPSS-1* and the launch of JPSS-2. We note that GH dropsondes were not operationally assimilated in the National Centers for Environmental Prediction (NCEP) global data assimilation system at the time of the study (K. Friedman 2018, personal communication). However, these observations were assimilated in the operational Hurricane Weather Research and Forecasting (HWRF) Model and the European Centre for Medium-Range Weather Forecasts (ECMWF) model. Finally, it should be emphasized that the results of this study provide an initial evaluation into the potential value of the GH UAS to improve NWP skill, and additional case studies would be needed in order to quantify our results.

The experiment setup, including a brief summary of the targeted observation technique used as field mission support during SHOUT, is presented in section 2. Sections 3 and 4 describe the meteorological conditions and the impact of the GH dropsondes for the high-impact weather events investigated. Finally, a summary of our main findings and suggestions for future work are given in section 5.

2. Experiment design

A series of OSE experiments were conducted to provide a preliminary investigation into the impact of GH dropsondes: three Hurricane Matthew flights on 5, 7, and 9 October 2016 during the Hurricane Rapid Response (HRR) campaign, and the 12, 15, and 21–22 February 2016 El Niño Rapid Response (Dole et al. 2018) flights (hereafter denoted as ENRR). These cases are described in detail in sections 3 and 4, respectively. In all cases, the cycling in the data assimilation system started 10 days prior to the availability of dropsondes and ended at the last analysis time where GH dropsonde data were available. First, a control experiment (CTL) includes all of the observations used in NCEP's operational Global Data Assimilation System (GDAS) at the time of the SHOUT 2016 field campaigns. To evaluate the impact of the observations in a satellite data gap scenario, a second experiment (noNPP) excludes the Advanced Technology Microwave Sounder (ATMS) and the Cross-track Infrared Sounder (CrIS) observations available on the *Suomi-NPP* satellite from CTL.

The dropsonde observations (Earth Observing Laboratory 1993; Hock and Franklin 1999; Wick et al. 2018b) of temperature, specific humidity, and vector winds are assimilated into GDAS following the identical methodology at NCEP operations. Namely, dropsonde wind observations within 3 times the radius of maximum wind of tropical cyclones are excluded from assimilation, and there is no account for dropsonde drift. During the February 2016 cases, a dry bias was noted in the upper-level dropsonde observations above 500 hPa. As a result, in coordination with the NCEP Environmental Modeling Center (EMC), we have masked out the specific humidity observations at 500 hPa and above (K. Friedman 2018, personal communication). The dry bias was no longer an issue during the October 2016 flights so no changes were made to the data during this period. GH dropsondes are assimilated in addition to the observations available in the CTL in the GH experiment, and in addition to the observations available in noNPP in experiment GH_noNPP. The comparison of GH with CTL quantifies the added value of GH dropsondes, while differences between GH_noNPP and noNPP quantifies the impact of GH dropsondes if a potential satellite observing gap would occur. Experiments noNPP and GH_noNPP used the same cycling period as GH and CTL.

All of the experiments performed in this study use the 19 July 2017 (Q3FY17) operational version of NCEP's GDAS, version 14.0.0, with horizontal resolution of T670 (~28 km; this is the highest resolution supported on the NOAA Theia Supercomputer) for 0–10-day forecasts, and T574 (~34 km) for 10–16-day forecasts. In operations, the NCEP GDAS is run at T1534 (~13 km) for 0–10-day forecasts. The model has 64 hybrid sigma-pressure levels (Sela 2009), with the model top at approximately 0.27 hPa (55 km). The GDAS analysis system utilizes a hybrid four-dimensional ensemble–variational data assimilation (4DEnVar; Wang and Lei 2014; Kleist and Ide 2015) system with resolution T254L64. More details on the Q3FY17 version can be found online (https://vlab.ncep.noaa.gov/web/gfs/past-implementations). Dropsondes were assimilated in the 6-h time window, the same as in operations. All of the experiments are verified against the ECMWF analyses (ECMWF 2011), which included the assimilation of GH dropsondes. In addition to the verification of mass fields, we compute storm track, sea level pressure (SLP), and wind errors for all hurricanes that occurred during the 5–10 October 2016 period. Storm tracks are generated from Global Forecast System (GFS) output using the Geophysical Fluid Dynamics Laboratory (GFDL) Vortex Tracker, and errors are verified with the National Hurricane Center (NHC) and the Joint Typhoon Warning Center (JTWC) best tracks for Atlantic and western Pacific tropical cyclones. Statistical significance, where appropriate, was computed using a paired t test. It should be pointed out that since forecasts were generated every 6 h, the results likely are correlated. This study does not account for the correlated forecast samples.

As part of our contributions to SHOUT, we provided maps with regions sensitive to model error growth in support of flight mission designs. These maps were primarily based on sensitive regions that were estimated using the ensemble transform sensitivity (ETS) method (Zhang et al. 2016), which is a first-order approximation of the ensemble transform (ET) technique (Bishop and Toth 1999). Briefly, a set of 80 NCEP Global Ensemble Forecast System (GEFS) ensemble forecasts were used to determine which (sensitive) regions had to be observed during the GH flights at a targeting time in order to improve weather forecast skill at a future verification time and region. A dry total energy norm computed from temperature and zonal and meridional winds at pressure levels of 200, 500, and 700 hPa was used to spatially represent the sensitive regions. During the SHOUT 2016 field campaigns, the ETS technique was run in real time four times daily. It is important to emphasize that this information was used as guidance by decision-makers, and in no way determined the final flight patterns, which were determined based on other factors beyond targeting information such as coordination with other aircraft, dynamically active regions of the storm's center, FAA constraints, and results from other targeting approaches provided by other centers. Therefore, this paper does not directly address the ETS targeting method.

3. SHOUT HRR storms

a. Case study description and ETS sensitivity region

The SHOUT team participated in the HRR mission in the fall of 2016 to target tropical cyclones over the western Atlantic Ocean. Three of the research flights focused on evaluating the impact of dropsondes to improve the track and intensity forecast of Hurricane Matthew on 5, 7, and 9 October 2016. Hurricane Matthew was a very high-impact weather event in terms of precipitation, high winds, major flooding, property damage, and loss of life. A total of 34 deaths across five states were directly attributed to Matthew—2 in Florida, 2 in Georgia, 4 in South Carolina, 25 in North Carolina, and 1 in Virginia, with wind and water damage of approximately $10 billion (U.S. dollars; NHC Hurricane Matthew report, April 2017; see https://www.nhc.noaa.gov/data/tcr/AL142016_Matthew.pdf). In this section, we examine the impact of GH dropsondes released during the three Hurricane Matthew science flights on 5, 7, and 9 October 2016 (HRR case). These science flights and accompanying dropsonde
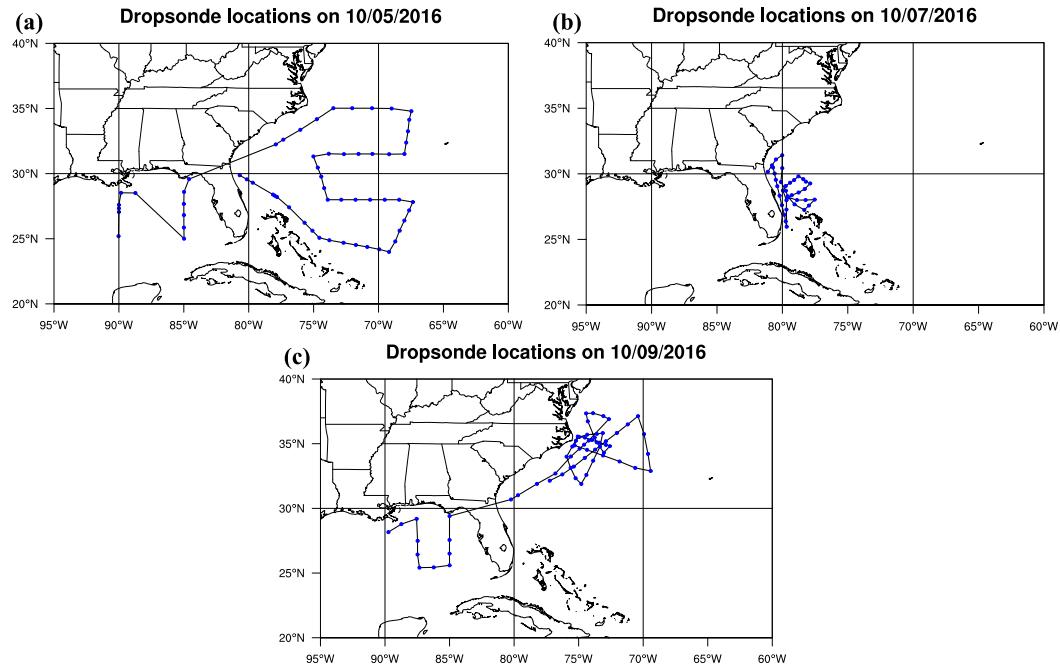
FIG. 1. Dropsonde release locations for the three Hurricane Matthew science flights on (a) 5, (b) 7, and (c) 9 Oct 2016. Dropsonde locations are denoted as blue dots, with flight track denoted by the black lines.

locations are shown in Fig. 1. After a 10-day spinup, cycling for the HRR case started at 0000 UTC 5 October and ended at 1200 UTC 10 October 2016. The verification period spans from 1200 UTC 5 October 2016 through 0000 UTC 10 October 2016 to cover the six analysis cycles where the GH dropsondes were assimilated (1200 UTC 5 October–0000 UTC 6 October 2016, 1200 UTC 7 October 2016, and 1200–1800 UTC 9 October 2016). This case, therefore, examines the cumulative effect of dropsondes on the analysis and forecast.

For the first Matthew flight on 5 October 2016, a verification region was placed over the southeastern United States (25°–35°N, 83°–73°W) to cover an envelope of the potential tracks of Hurricane Matthew. Roughly 5 days prior to Matthew reaching the verification region, there was considerable disagreement in the GFS on the exact track of the hurricane. Some forecasts showed landfall over Miami, Florida, while others showed the hurricane remaining offshore. Two to three days prior to the verification time, the exact westward movement of Matthew and its potential interaction with another tropical cyclone (Hurricane Nicole) were still uncertain.

Figure 2a shows the normalized ETS generated ~5 days in advance of the verification time, with a targeting time (approximate time of sampling) at 0000 UTC 5 October 2016 and a verification time (time at which storm is expected to impact the verification region) at 0000 UTC

7 October 2016, along with the actual GH flight pattern and dropsonde locations for the first Hurricane Matthew science flight. The ETS algorithm indicated large sensitivity in connection with Hurricane Matthew, as well as in the Caribbean, Yucatan Peninsula, the Gulf of Mexico, Cuba, the Ohio valley, and to the east of Florida and the southeastern United States. These sensitivity regions during the first science flight were likely tied to 1) the uncertainty of the track of Hurricane Matthew; 2) the steering flow for Matthew, which was controlled by a ridge of high pressure over the western Atlantic Ocean; and 3) a trough over the eastern United States. Sensitive regions also existed north and east of Matthew, along the western periphery of the ridge of high pressure. Mission managers decided to sample the large-scale steering flow north and west of Matthew and, therefore, did not sample the largest ETS region around Hurricane Matthew. The GH flight path covered the northern Gulf of Mexico and the Atlantic Ocean east of the Carolinas and Florida. While over the Atlantic, it sampled a region of normalized ETS greater than 0.5, along the western flank of the high pressure system connected with the steering flow. Hurricane Matthew reached its greatest strength just offshore of Miami, at 25.2°N and 77.9°W between 1200 and 1800 UTC 6 October 2016, with a minimum central pressure of ~937 hPa and maximum 1-min sustained winds of 138 mi h^{-1}. The last two flights, on 7 and 9 October 2016, primarily sampled in and around the center of Hurricane
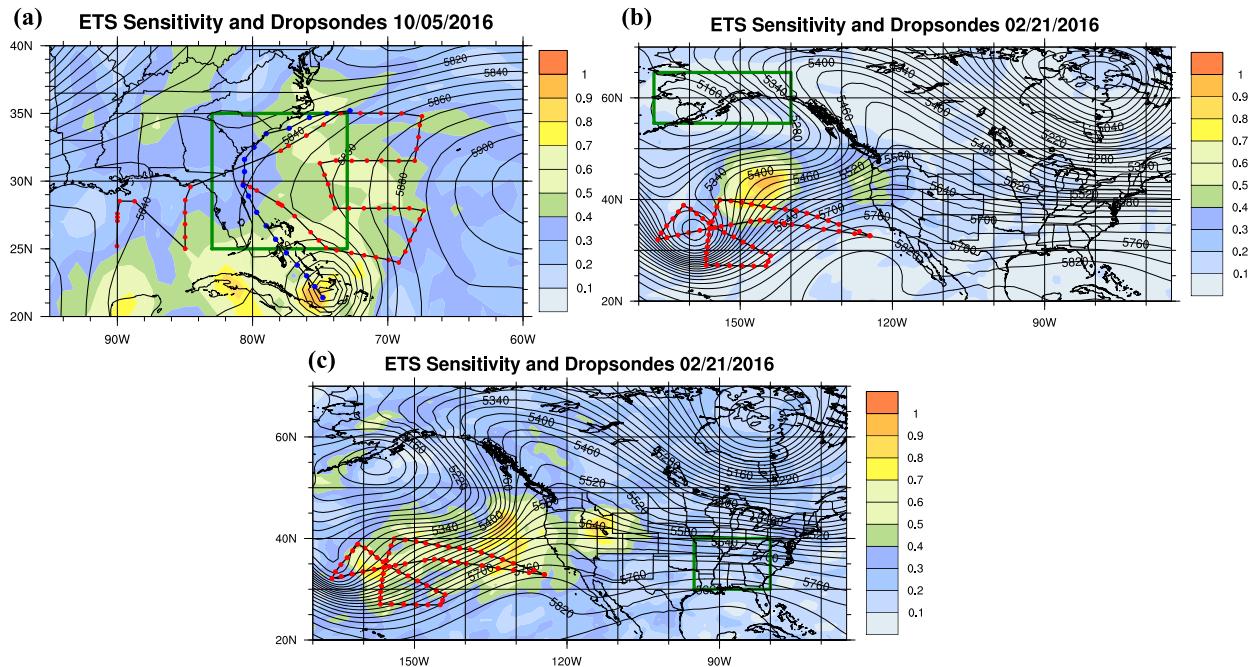
FIG. 2. (a) Normalized ETS sensitivity (shaded contours) during the first Hurricane Matthew science flight on 5 Oct 2016, initialized over four cycles from 1200 UTC 1 Oct to 0600 UTC 2 Oct 2016 for a targeting time of 0000 UTC 5 Oct 2016 and verification time of 0000 UTC 7 Oct 2016. The black line denotes the GH flight pattern and red dots represent the dropsonde release locations on 5 Oct 2016. Also plotted are the mean 500-hPa geopotential height (black contours; m) from 80 GEFS ensemble forecasts. Blue dots denote the NHC best track of Hurricane Matthew, starting at 0600 UTC 5 Oct 2016 and ending during its dissipating stage at 1800 UTC 9 Oct 2016. The green box denotes the verification region (25°–35°N, 83°–73°W). (b) As in (a), but for normalized ETS for the third ENRR flight initialized from 0600 UTC 18 Feb to 0000 UTC 19 Feb 2016 for a targeting time of 0000 UTC 22 Feb 2016 and a verification time of 0000 UTC 24 Feb 2016. The black line denotes the GH flight pattern, and the red dots represent the dropsonde release locations during 21–22 Feb 2016. The green box denotes the verification region over southern Alaska (55°–65°N, 167°–140°W). (c) As in (b), but for normalized ETS sensitivity when placing a verification region over the southeastern United States (30°–40°N, 95°–85°W). Large ETS values denote regions where fast amplifying forecast errors may originate.

Matthew (Figs. 1b,c) as it traveled along the coast of Florida and the Carolinas.

b. Storm track, sea level pressure, and wind errors

The impact of GH dropsondes on the storm track, SLP, and wind of Hurricane Matthew is shown in Fig. 3. Average errors are shown, along with differences between CTL and GH with the 95% confidence interval. Results are shown as a function of forecast lead time over all cycles from 1200 UTC 5 October to 0000 UTC 10 October 2016. Neutral track error is found out to ~60 h (Fig. 3a). After 60 h, a positive impact (reduction in track error) is found in the GH experiment, which persists out to 102 h. The results are statistically significant in the 72–84-h lead times, with an approximately 90-km improvement from the assimilation of dropsondes relative to CTL at 96 h (Fig. 3d). This reduced error is between 7% and 30% after 60 h. As for SLP and wind errors (Figs. 3b,c), the analysis time and first 12 h of the forecast are improved under the assimilation of dropsondes, with an average reduction in error of 0.5 hPa and

1 kt (1 kt ≈ 0.51 m s^{-1}) relative to CTL (Figs. 3e,f). The sign reverses at 18–48-h lead time for SLP and at 18–72-h lead time with respect to wind, with the GH experiment showing increased error compared to CTL, with a significant degradation at 30–36 h (Figs. 3b,c). Differences in SLP and wind errors reverse again after 72 h, with the GH experiment reducing the error by up to 2 hPa and 2 kt, respectively. Overall, most lead times are statistically neutral.

Although the three HRR science flights did not specifically target other tropical cyclones, Hurricane Nicole was present during this same time period out over the western Atlantic. Figure 4 shows the same results as for Matthew but for Hurricane Nicole when averaged over all forecast cycles. As with Matthew, neutral track errors are found at early lead times (Fig. 4a), out to 24 h, with reduced track error in the GH experiment from 36 to 96 h (a difference in error of ~25 km and an improvement of between 4% and 20%). A slight degradation is present from 108 to 120 h, with improvement from the assimilation of dropsondes out to 168 h (~40 km relative
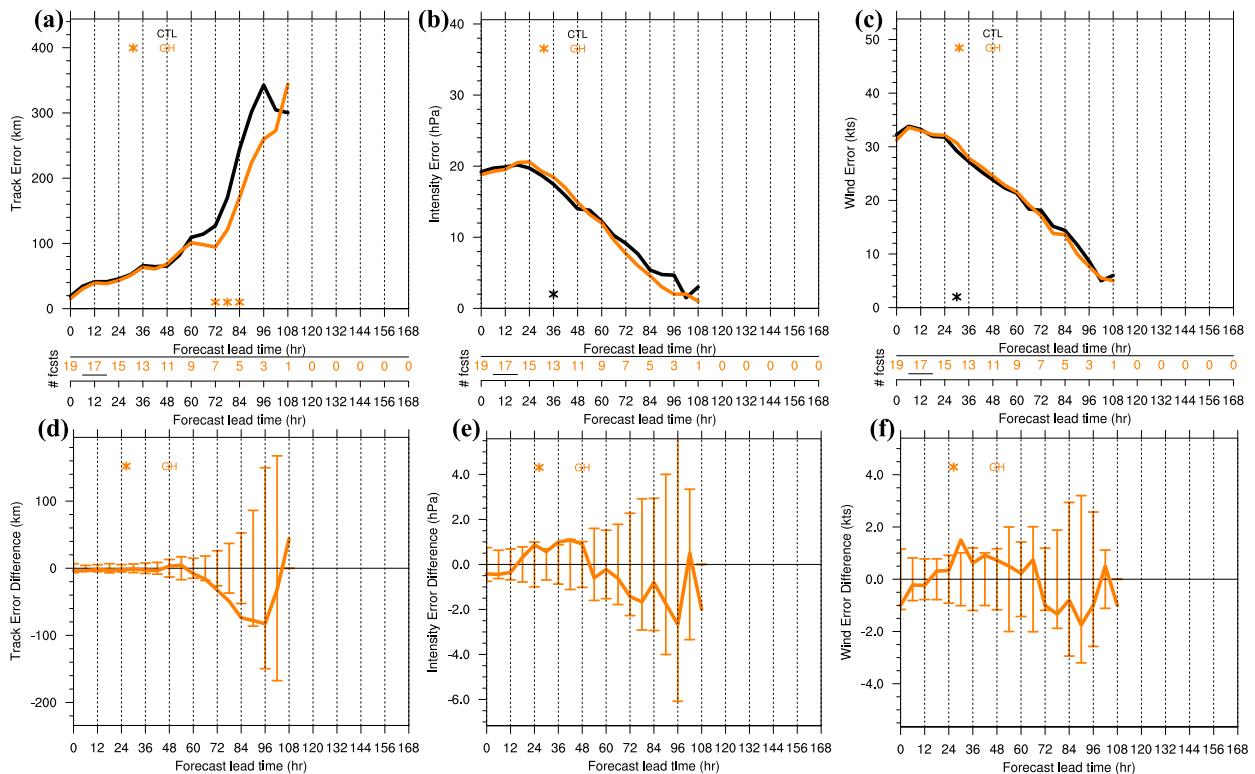
FIG. 3. Hurricane Matthew average absolute errors for CTL (black) and GH (orange) experiments for (a) storm track (km), (b) sea level pressure (hPa), and (c) wind (kt) as a function of forecast lead time out to 168 h. The number of forecasts (denoted as # fcsts) at each lead time is shown. Stars denote that differences between the experiments are statistically significant at the 95% confidence interval using a paired *t* test. Orange stars indicate the GH experiment is better than CTL, while black denotes CTL is better than GH. (d)–(f) Average differences of GH minus CTL for track error (km), sea level pressure (hPa), and wind (kt), along with 95% confidence intervals at each forecast lead time. Lead times where the confidence interval is zero are due to a single forecast; these times are excluded from being considered significant.

to CTL; Fig. 4d). All lead times are not statistically significant. These results indicate that sampling Matthew also had a positive impact on Nicole's track. Similarly, neutral results overall are found for SLP and wind errors as for Matthew (Figs. 4b,c). A statistically insignificant degradation is found in the GH experiment relative to CTL in the first 72 h, with differences of up to 1 hPa and 1 kt (Figs. 4e,f). Neutral or slight reductions in the error with the assimilation of dropsondes are found thereafter out to 168 h, though they are not significant.

Figures 5 and 6 show the same results, but under a potential satellite data gap by comparing noNPP to GH_noNPP. As under the current satellite configuration, the results are similar for Matthew with respect to the track error. That is, neutral errors are found out to 60 h, with an improvement in the GH_noNPP experiment thereafter up to 108 h (Fig. 5a), with statistically significant reductions in error at 72 and 90 h (14%–20% improvement). The differences in track error are ~60 km (Fig. 5d) at 96 h. Similar to Figs. 3b and 3c, SLP and wind errors are reduced in the GH_noNPP

experiment in the first 12–24 h, with insignificant increased error from 30 to 48 h in SLP and from 24 to 66 h in wind from dropsonde assimilation (Figs. 5b,c). As under the current satellite scenario, the greatest improvements from dropsonde assimilation are found at the later lead times under a data gap, with reduced error relative to noNPP out to 96 h for SLP and 108 h in wind. These differences are up to 1.5 hPa and 1.5 kt (Figs. 5e,f), respectively, comparable to those found for GH versus CTL.

Figure 6 shows the results for Nicole. Contrary to the results of GH versus CTL in Fig. 4, neutral to positive improvement in track error is found in GH_noNPP in Fig. 6a out to 96 h. After this, a consistent degradation is present in GH_noNPP out to 168 h by ~50 km, though the result is not statistically significant. These degradations are between 3% and 12%. As evidenced by the 95% confidence interval in Fig. 6d, there is considerable spread for Nicole in comparison to Matthew. In general, most lead times indicate reduced SLP and wind errors in GH_noNPP versus noNPP (Figs. 6b,c), with statistically significant reductions in the error at 12-, 36–48-, and 66-h
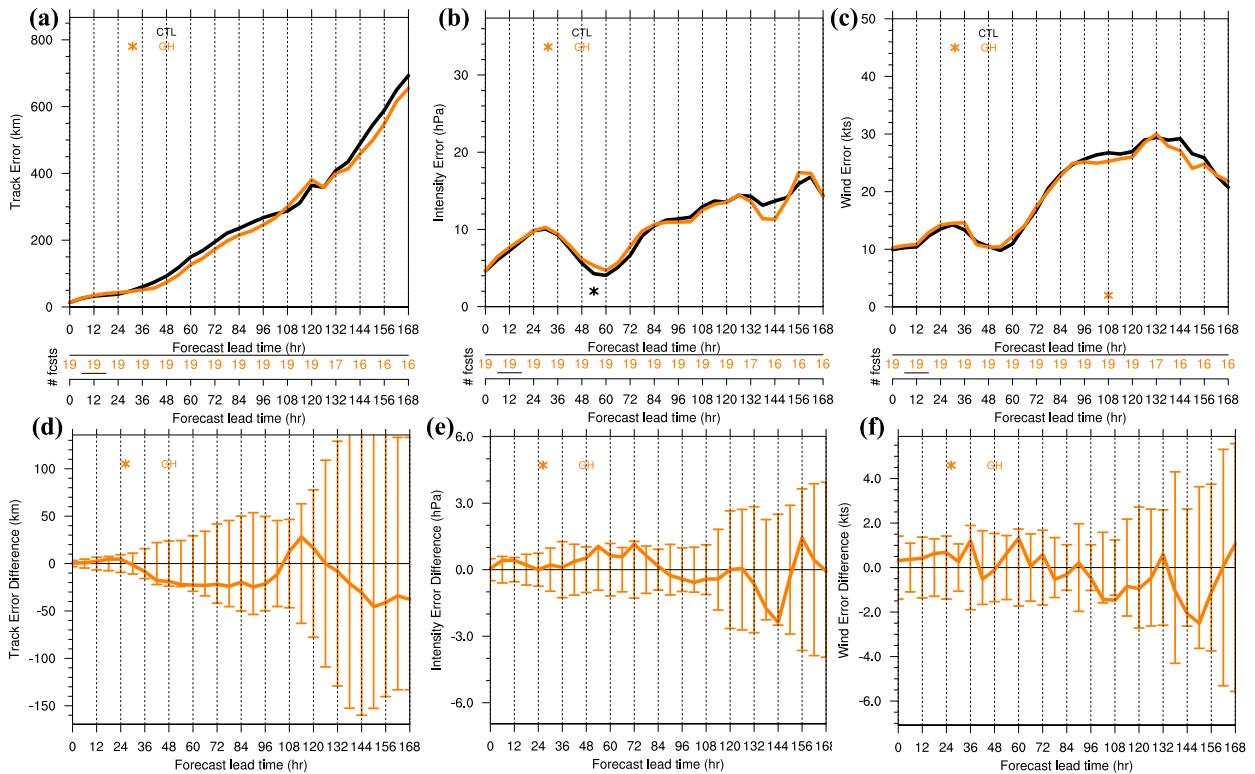
FIG. 4. As in Fig. 3, but for Hurricane Nicole.

lead times. These differences are comparable to those for Matthew in Fig. 5 of 1.5–2 hPa (kt). Thus, with the exception of the Nicole track error, results under a satellite data gap are similar in magnitude to results under the current observational configuration. It is clear from Figs. 3–6 as well that all experiments failed to capture the deepening of Matthew at early lead times; for Nicole, the opposite occurred, in which the experiments showed larger SLP and wind error at the later forecast lead times.

Aberson (2011) examined the impact of localized dropsondes on global tropical cyclones by cycling from 26 August through 2 October 2008 using the GFS model. A total of 18 tropical cyclones occurred during this time period. An approximate 10% improvement in track error was found through 72 h with the assimilation of dropsondes. These results showed that localized dropsondes are likely to positively impact tropical cyclones on a global scale. As a result, we examine this result here by aggregating the statistics of track and wind error for four tropical cyclones that occurred during the HRR cycling period. In addition to Matthew and Nicole in the Atlantic, two tropical cyclones were present in the western Pacific at this time, namely Typhoon Songda and severe Tropical Storm Aere; no cyclones were present in the eastern Pacific.

Figure 7 shows the results for GH versus CTL for track and wind errors. When averaged across all tropical cyclones, a consistent reduction in track error is found at all lead times (with the exception of hours 114–126) from the assimilation of dropsondes, with the exception of neutral results in the first 24 h (Fig. 7a). These differences are statistically significant in the 60–96-h lead times (Fig. 7c), with the largest improvement of ~30 km. These track improvements in the GH experiment are ~6% in the 30–96-h lead times, with maximum improvement of 9% at 96 h, similar to that found in Aberson (2011). For wind errors, Fig. 7b indicates that neutral to degraded results are found in the first 72 h from dropsonde assimilation (statistically significant at 18 and 30–36 h). Thereafter, reduced wind errors overall in the GH experiment are present by about 1.5 kt (Fig. 7d). Results were mixed under a satellite data gap, with increasing track error in the GH experiment and neutral results for wind (results not shown). The only tropical cyclone with improved track error from dropsondes in a satellite data gap was Matthew, indicating the importance of the *Suomi-NPP* data in improving analyses and forecasts on a global perspective.

c. Case study analysis

In this section, we examine how the additional dropsonde observations led to the track forecast improvements evident in Figs. 3 and 4 for both Matthew and Nicole. The cycles that assimilated the dropsondes that showed the
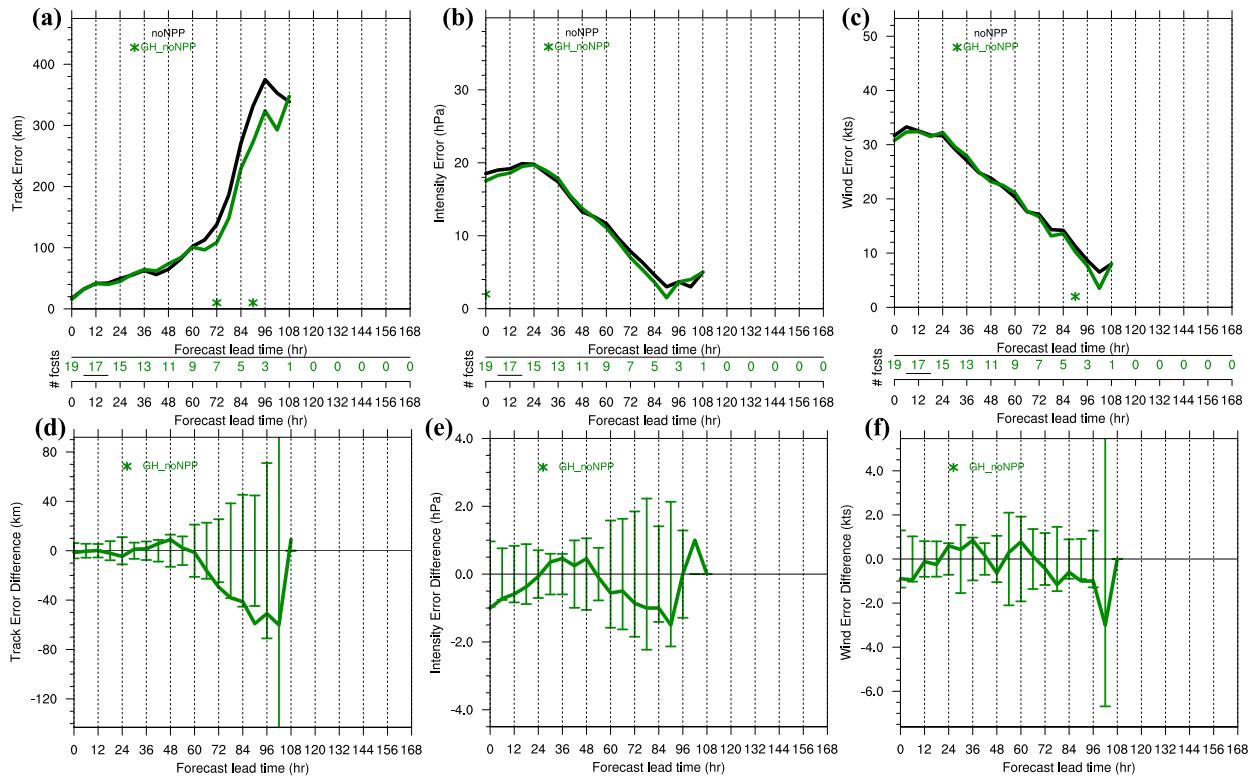
FIG. 5. As in Fig. 3, but for Hurricane Matthew when comparing GH_noNPP (green) to noNPP (black) under a potential satellite data gap.

greatest positive impact on the track were 1800 UTC 5 October and 0000 UTC 6 October 2016 for Matthew, as well as 1200 UTC 7 October and 1200 UTC 9 October 2016 for Nicole. Figure 8 shows the track maps for both Matthew and Nicole at these four initialization times. Both cycles for Matthew (Figs. 8a,b) show that there is marginal improvement in Matthew with respect to its westward movement along the Florida coast. However, a clear improvement is evident from the assimilation of dropsondes at later lead times with a more northern position more in line with the best track.

During the 1200 UTC 7 October 2016 forecast cycle, a blocking ridge of high pressure to the north off the coast of the eastern United States pushed Nicole southward (Fig. 8c) prior to its east-northeast track into the central Atlantic (NHC 2016 Annual Summary, May 2017; see https://www.nhc.noaa.gov/data/tcr/summary_atlc_2016.pdf). Both the CTL and GH forecasts at this cycle failed to recurve Nicole in line with the best track. However, the GH experiment shows a farther north track than CTL at later lead times, leading to an improved storm track. At 1200 UTC 9 October 2016 (Fig. 8d), a consistent farther north track in the GH experiment (closer to the best track) is clear relative to CTL.

Part of the reason for these track improvements appears to be connected with an overall better representation of

the mass fields over the western Atlantic Ocean. Shown in Fig. 9 is the root-mean-square error (RMSE) of 700-hPa wind, height, relative humidity, and SLP for the CTL and GH experiments over all cycles out to 120 h. Results were averaged over the western Atlantic (25°–40°N, 83°–55°W), denoted by the red box in Fig. 10. The 700-hPa level was chosen due to its importance in the steering flow for tropical cyclones. Across all lead times, there is reduced error after the assimilation of dropsondes for all variables, pointing to an improvement in the synoptic pattern. The differences in RMSE between GH and CTL are statistically significant in the 24–72-h lead times for several of the variables. These improvements during this time frame correspond to reduction in error of between 3% and 6% in 700-hPa wind, between 6% and 11% in 700-hPa heights, between 3% and 8% in SLP, and between 3% and 5% in 700-hPa relative humidity.

A spatial examination of the 700-hPa height and wind field for the 0000 UTC 6 October 2016 forecast cycle for Matthew is provided in Fig. 10. At the analysis time (Fig. 10a), Matthew is located over the Caribbean west of 75°W, with Nicole out in the Atlantic at 25°N and 65°W. Analysis differences in the 700-hPa meridional wind field (GH minus CTL) indicate a larger south-to-north-oriented flow in the GH experiment over much of the western Atlantic (Fig. 10b). These differences are on the
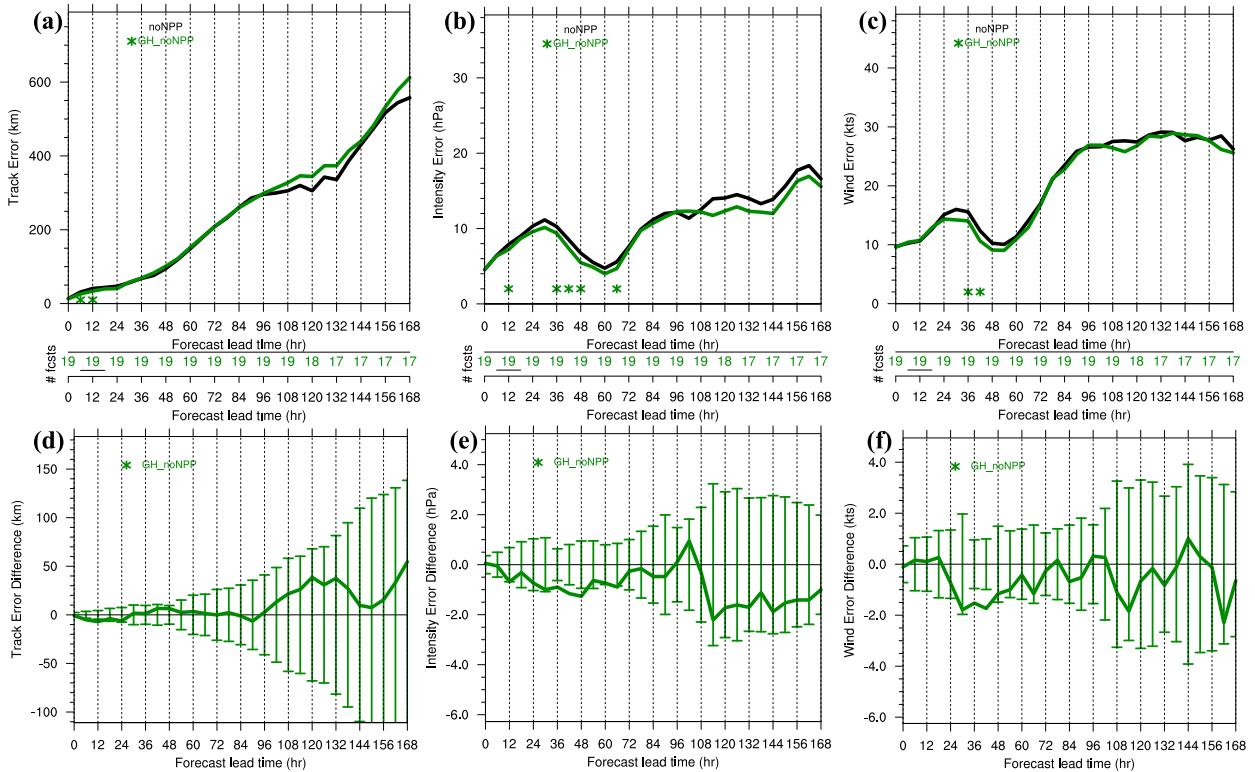
FIG. 6. As in Fig. 5, but for Hurricane Nicole.

order of 1–3 m s^{-1}. A more positive meridional flow in the experiment with dropsondes may have led to the path of Matthew becoming closer to the best track shown in Fig. 8.

The same fields of 700-hPa wind and height are shown in Fig. 10c, but for the 48-h forecast time from the 0000 UTC 6 October 2016 cycle. At this time, Matthew is just offshore from Florida, with Nicole slightly farther to the west and north at 27°N and 67°W. Forecast differences of GH minus CTL for 700-hPa height are shown in Fig. 10d. Although the differences are subtle (only 2–5 m), positive height changes are evident between 65° and 75°W and between 25° and 35°N. This indicates a slightly stronger ridge to the east of Matthew in the GH experiment, which likely helped prevent Matthew from dipping southward to interact with Nicole as compared to the CTL forecast. Thus, a combination of both the improved wind and height field likely played a role in the track improvement for Matthew.

An identical analysis was performed for the 1200 UTC 9 October 2016 forecast cycle for Hurricane Nicole. A stronger ridge was present in the central Atlantic in the GH experiment at both the analysis and forecast times relative to CTL. This likely led to the consistently farther north position of Nicole closer to the best track seen in Fig. 8d (results not shown).

d. Precipitation verification

Precipitation forecasts for the HRR case are quantitatively verified by using observed Stage IV precipitation data (Lin 2011). This dataset is available online (http://data.eol.ucar.edu/dataset/21.093). Precipitation forecasts were quantitatively evaluated by using the equitable threat score (EQTS; Wilks 2006). Because the primary impacts from Matthew were concentrated over the southeastern United States, we compute the EQTS over the domain 22°–42°N and 85°–75°W to cover the total accumulated precipitation from 5 to 9 October (Fig. 11a). All experiments are interpolated onto the observational grid in order to compute differences in accumulated precipitation and generate qualitative and quantitative results.

Figure 11a shows the observed accumulated precipitation from 0000 UTC 5 October through 1800 UTC 9 October 2016 over the southeastern United States. Across North and South Carolina, the observed precipitation approaches 260 mm, a likely cause of the flooding in this region. Figures 11b–g show the EQTS for CTL, GH, noNPP, and GH_noNPP for various thresholds of 24-h accumulated precipitation (from 0.2 to 75 mm) for forecast hours 24–48 (Fig. 11b,c), 48–72 (Figs. 11d,e), and 72–96 h (Figs. 11f,g). With the exception of 0.2–10-mm thresholds in the 24–48-h
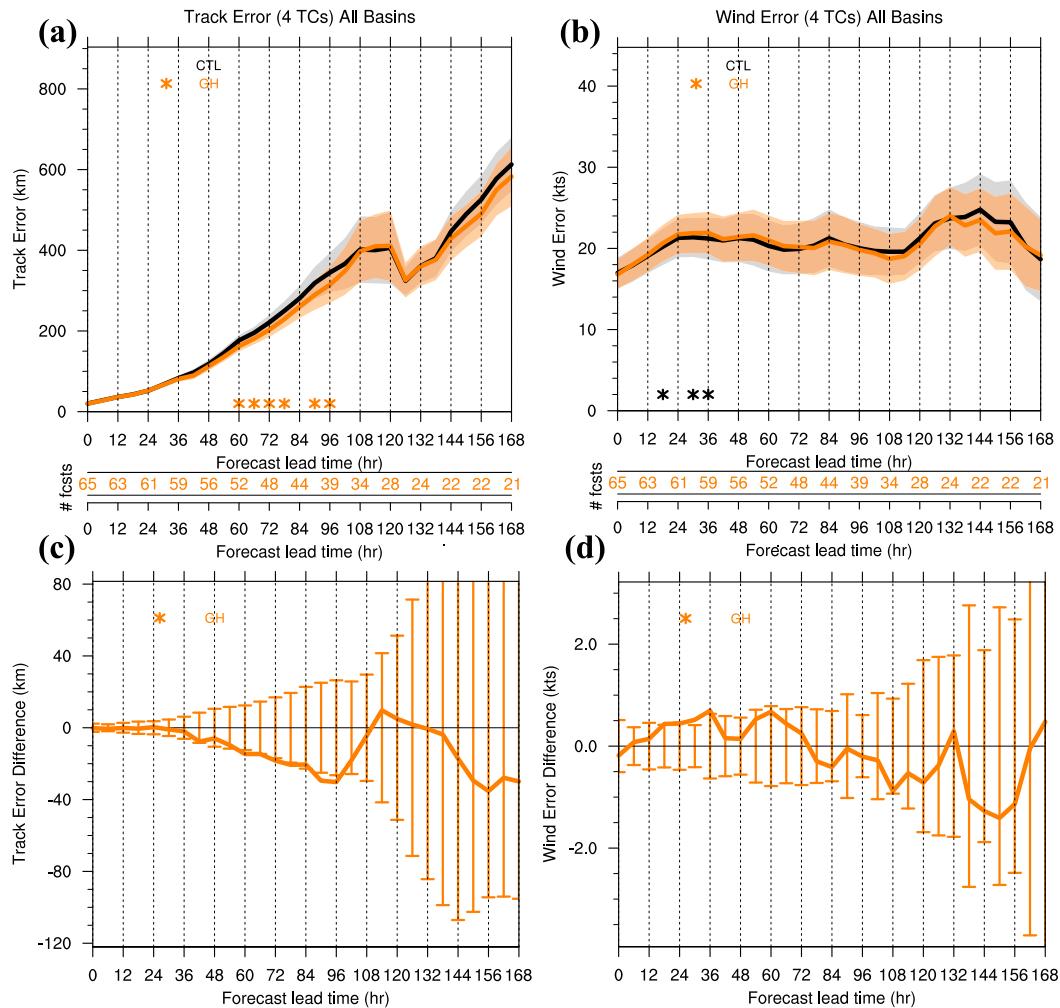
FIG. 7. Same statistics as in Fig. 3 of GH vs CTL, but for all tropical cyclones (four total) during the 5–10 Oct 2016 period of the Atlantic and western Pacific. Shown is the average absolute error for (a) storm track (km) and (b) wind (kt), as well as (c),(d) the differences between the two experiments, as in Fig. 3. Included in the average errors is the 1σ sample standard deviation of each experiment (orange and gray shaded region).

forecast, assimilating GH dropsondes improves the precipitation forecast under both satellite scenarios out to 96 h, as evidenced by the higher EQTS relative to the control forecasts. This improvement is approximately 1%–3% for the 24–48-h time frame, 1%–6% for the 48–72-h lead time, and 3%–9% for 72–96 h. We also examined qualitative differences in forecasted 6-h and accumulated precipitation over the southeastern United States (results not shown). Differences were subtle and likely a result of the reduced resolution of the model in our study.

4. SHOUT El Niño Rapid Response storms

a. Case study description and ETS sensitivity region

During SHOUT-ENRR (Kren et al. 2016; Dole et al. 2018), the focus was to improve the 1–4-day weather forecasts of atmospheric rivers (ARs; Zhu and Newell 1998; Ralph et al. 2004) and high-impact storms impacting the United States during the strong El Niño event. ARs can be defined as long (at times extending thousands of kilometers), narrow (400–500 km in width), and intense regions of water vapor transport concentrated mainly in the lowest 2 km of the atmosphere and located along the warm sector of extratropical cyclones (Zhu and Newell 1998; Ralph et al. 2004; Neiman et al. 2008; Ralph and Dettinger 2011).

During the 3-week mission that took place in February, the GH sampled three winter storms: 1) an AR system on 12 February 2016 that penetrated into the Pacific Northwest and the state of Washington, 2) an event involving a Pacific trough interaction with a cutoff low pressure system and subtropical moisture plume on
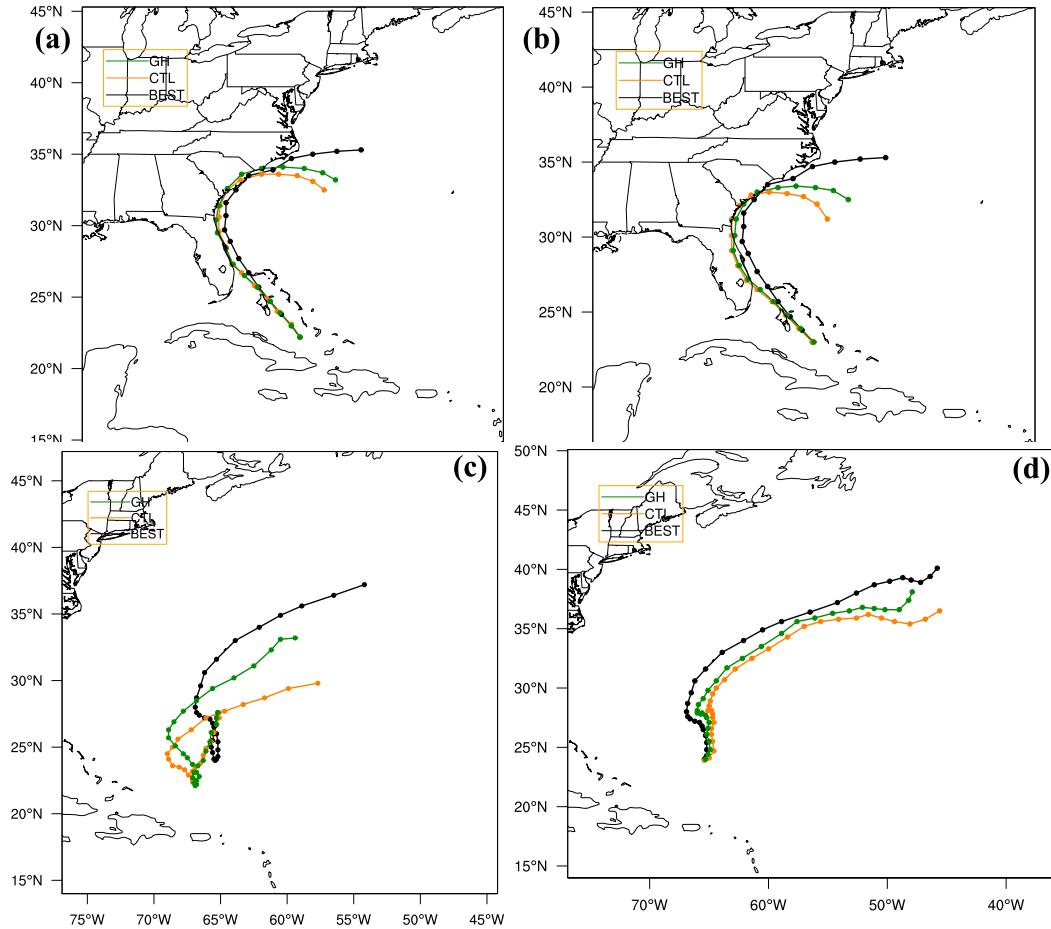
FIG. 8. Track maps for both Hurricanes (top) Matthew and (bottom) Nicole for the two cycles that showed the most positive impact from assimilating GH dropsondes. Black colors denote the best tracks of Matthew and Nicole, while green denotes the GH experiment, and orange is for CTL. The forecast initializations are at (a) 1800 UTC 5 Oct 2016, (b) 0000 UTC 6 Oct 2016, (c) 1200 UTC 7 Oct 2016, and (d) 1200 UTC 9 Oct 2016.

15 February 2016, and 3) a rapidly deepening system and associated AR over the central North Pacific on 21 February 2016 that reached the Gulf of Alaska a few days later. Because of technical difficulties with the dropsonde system, the first two flights only included 2 and 21 dropsondes on 12 and 15 February 2016, respectively. The third mission was more successful; a total of 68 dropsondes were deployed (see Fig. 12). Dropsonde observations for the ENRR cases occurred over seven analysis cycles (1800 UTC 12 February, 1800 UTC 15 February–0000 UTC 16 February 2016, and 1800 UTC 21 February–1200 UTC 22 February 2016). Although verification in this section incorporates all cycles with GH dropsondes, our case study analysis will focus more on the third ENRR mission on 21–22 February 2016 due to the greater sampling density. Contrary to continuous cycling in the HRR case, the verification period for ENRR only includes cycles with

added dropsondes (see above) due to the larger temporal separation between flights.

Sensitivity regions for the third ENRR flight were generated for both a verification region over southern Alaska (55°–65°N, 167°–140°W) and the southeastern United States (30°–40°N, 95°–80°W), shown in Figs. 2b and 2c using a targeting time of 0000 UTC 22 February 2016 and a verification time of 0000 UTC 24 February 2016. These regions were chosen based on the precipitation, snow, and high winds forecasted to impact Alaska at 0000 UTC 24 February 2016, as well as a potential severe weather outbreak in the eastern United States. The largest ETS sensitivity region for targeting Alaska was located northeast of the upper-level trough and low pressure system, in a broad region of the upper-level jet stream and divergent region. It is evident that the GH did not sample a large portion of the sensitivity (Fig. 2b). However, Fig. 2c indicates that by aiming to
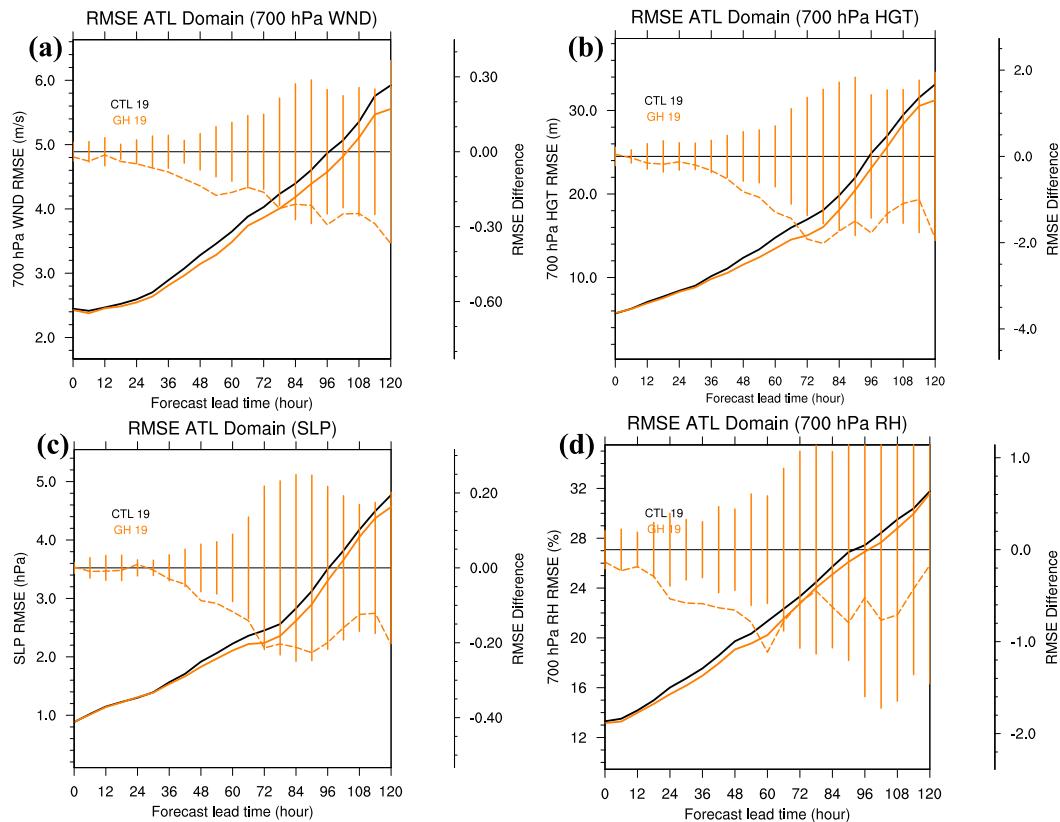
FIG. 9. RMSEs for both the CTL (black) and GH (orange) experiments for all Hurricane Matthew assimilation cycles for (a) 700-hPa wind (m s^{-1}), (b) 700-hPa height (m), (c) sea level pressure (hPa), and (d) 700-hPa relative humidity (%). There are a total of 19 forecasts used in the statistics. Results are shown when averaged over the western Atlantic region (25°–40°N, 83°–55°W). Vertical orange bars denote the 95% confidence intervals using a paired t test to assess the significance of the differences (dashed orange line) of GH minus CTL. Differences outside of the confidence intervals are statistically significant.

improve the southeastern United States, the GH sampling actually covered much of the ETS region present from 160° to 120°W and from 30° to 40°N. For this reason, section 4c will focus on a case study analysis for the third ENRR flight.

b. Verification of mass fields

Since the three ENRR flights affected several regions of the continental United States (CONUS), impacts on the analysis and forecast cycles are first presented over all cycles across the CONUS for the anomaly correlation of SLP and 500-hPa geopotential height, as well as RMSE of 700-hPa wind for both CTL versus GH and noNPP versus GH_noNPP.

The results for the anomaly correlation and RMSE are shown in Fig. 13. Across the majority of the lead times, the results are statistically neutral under both satellite scenarios. In general, higher anomaly correlation skill scores and reduced RMSEs in 700-hPa wind for the first 72 h are present in the dropsonde experiments,

although these improvements are less than 2% and not statistically significant. A statistically insignificant degradation is present after 72 h for 500-hPa height and 700-hPa wind. The one exception is SLP, which shows a consistently higher skill score out to 120 h (Fig. 13a). We also examined other variables and pressure levels among all three ENRR cases; overall, the results were found to be neutral, similar to precipitation forecasts (results not shown).

c. Case study analysis

Although the results were found to be neutral for the three ENRR cases when verified over the CONUS, the impacts from dropsondes for the third ENRR flight provided a marked improvement over the southeastern United States. As mentioned in section 4a, there was severe weather potential a few days after the GH flight on 21–22 February 2016. Figure 14a shows the observed 6-h accumulated precipitation ending on 0600 UTC 24 February 2016 obtained from Lin (2011). An area of
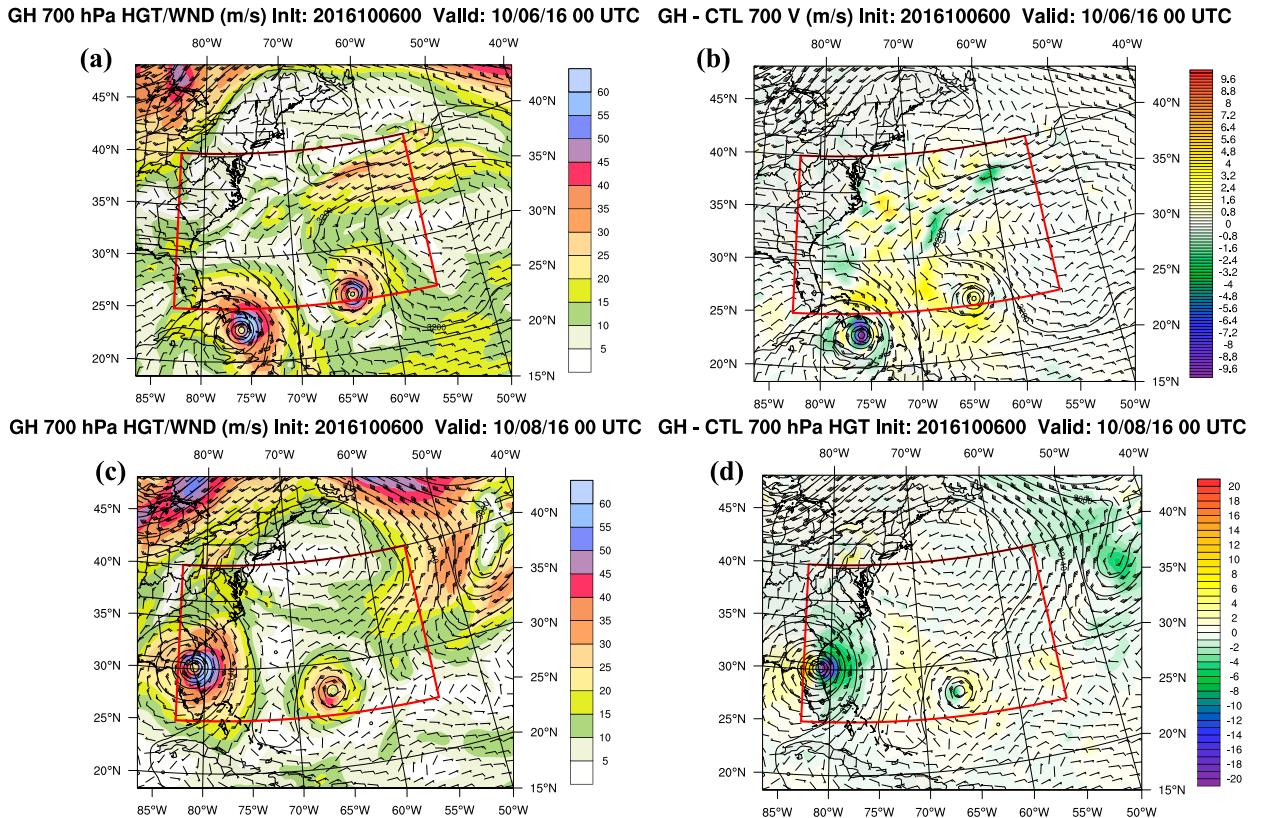
GH 700 hPa HGT/WND (m/s) Init: 2016100600 Valid: 10/06/16 00 UTC GH - CTL 700 V (m/s) Init: 2016100600 Valid: 10/06/16 00 UTC

GH 700 hPa HGT/WND (m/s) Init: 2016100600 Valid: 10/08/16 00 UTC GH - CTL 700 hPa HGT Init: 2016100600 Valid: 10/08/16 00 UTC

FIG. 10. Latitude–longitude plots of the 700-hPa height (m; contoured) and wind (m s^{-1}; shaded every 5 m s^{-1}) field from the GH experiment initialized at 0000 UTC 6 Oct 2016, showing results for the (a) analysis time and the (c) 48-h forecast. The red box denotes the region used to compute the RMSE in Fig. 9 over the western Atlantic (25°–40°N, 83°–55°W). Latitude–longitude plots of GH minus CTL for the same time periods as in (a) and (c), but for differences of (b) 700-hPa meridional wind (m s^{-1}; shaded) and (d) height (m; shaded) with the height field overlaid. The red box is the same as in (a) and (c).

low pressure tracked across Arkansas, northern Mississippi, and Tennessee, accompanied by a cold frontal boundary pushing through southern Louisiana, Mississippi, the panhandle of Florida, and Georgia. Figure 14a clearly shows that several squall lines developed out ahead of the cold front during the evening of 23–24 February 2016, oriented southwest to northeast over the southeastern United States.

Several tornadoes and high winds were reported during the evening hours of 23 February 2016, as indicated in Fig. 14b (results courtesy of the Storm Prediction Center). All told, there were 52 tornadoes and 91 high wind reports across Louisiana, Mississippi, Alabama, Florida, and Georgia. Many of these reports were oriented along a southwest-to-northeast line, likely connected to the squall lines that developed.

Figure 15 presents verification of the RMSEs for the low-level storm environment for the third ENRR flight when averaged over the region in Fig. 2c (30°–40°N, 95°–80°W) where the severe weather occurred. The parameters examined include 850-hPa wind, temperature, and relative

humidity during CTL, GH, noNPP, and GH_noNPP. Under both satellite scenarios, reduced RMSEs are found in the experiments with added dropsonde data, primarily in the 36–90-h time frame. Some of these lead times are statistically significant at the 95% level, such as for 850-hPa relative humidity. The percent reduction in RMSE in the dropsonde experiments relative to CTL and noNPP over the 36–72-h lead time reaches a maximum of ~2%–3% in 850-hPa wind, 5%–7% in 850-hPa temperature, and 4%–5% in 850-hPa relative humidity. Other variables were also examined, such as SLP, which showed reduced error in GH and GH_noNPP versus the control forecasts by up to 6%–8% at 24–72-h lead time (results not shown).

Finally, we examine forecast fields important for the convective environment over the southeastern United States. The surface convective available potential energy (CAPE) and lifted index for the GH experiment, as well as differences of these fields for GH minus CTL, are shown in Fig. 16 for the 42-h forecast initialized at 1200 UTC 22 February 2016. The lifted
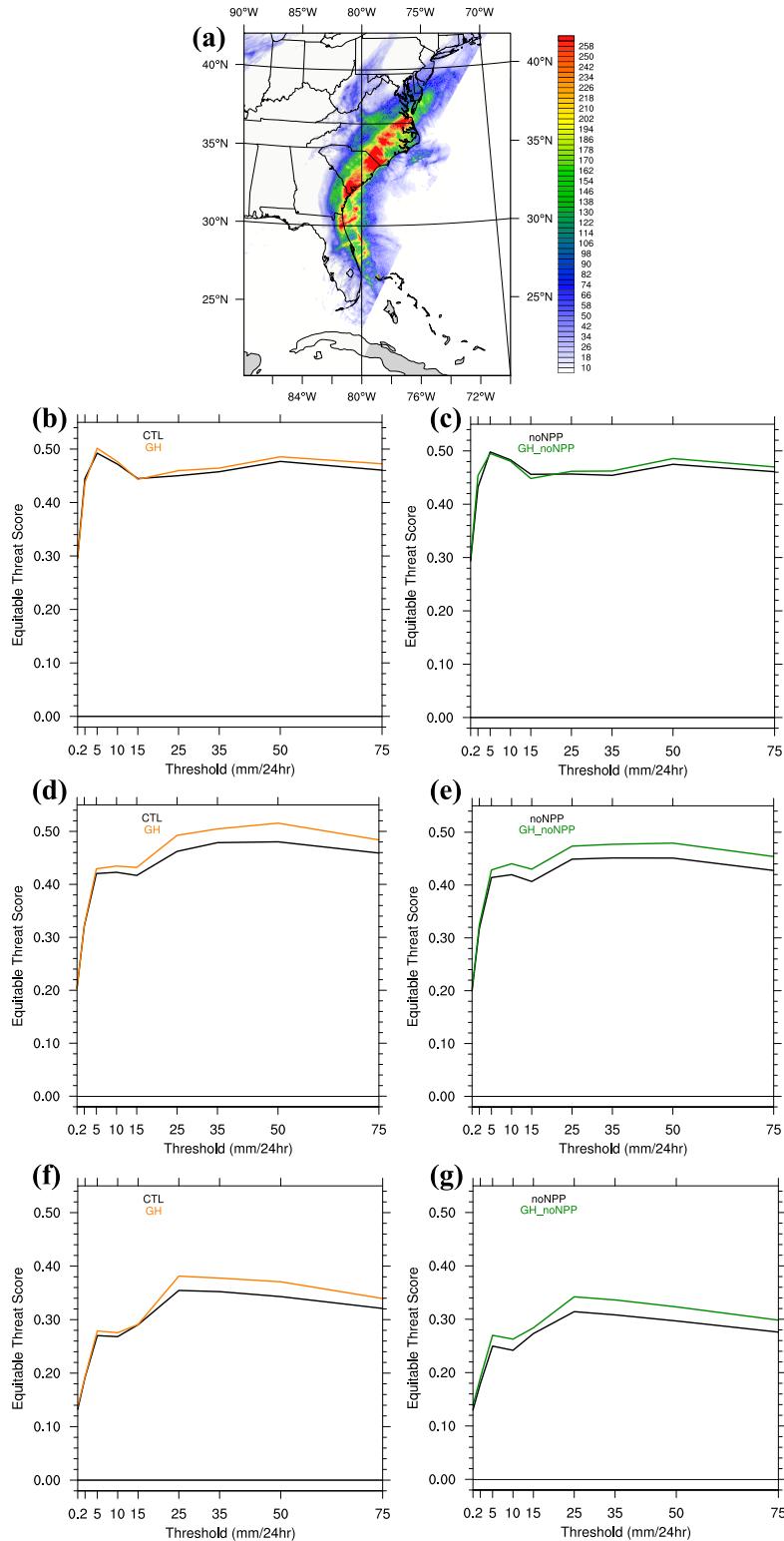
FIG. 11. (a) Accumulated precipitation (mm) from 0000 UTC 5 Oct 2016 through 1800 UTC 9 Oct 2016 from observations (Lin 2011) over the southeastern United States. (b),(d),(f) EQTSs for the 24–48-, 48–72-, and 72–96-h forecasts for various precipitation thresholds of 0.2–75 mm over a 24-h period for the CTL (black) and GH (orange) experiments. (c),(e),(g) As in (b), (d), and (f), but for the GH_noNPP (green) and noNPP (black) experiments under a potential satellite data gap. Results are shown for all cycles and averaged over the domain 22°–42°N and 85°–75°W.
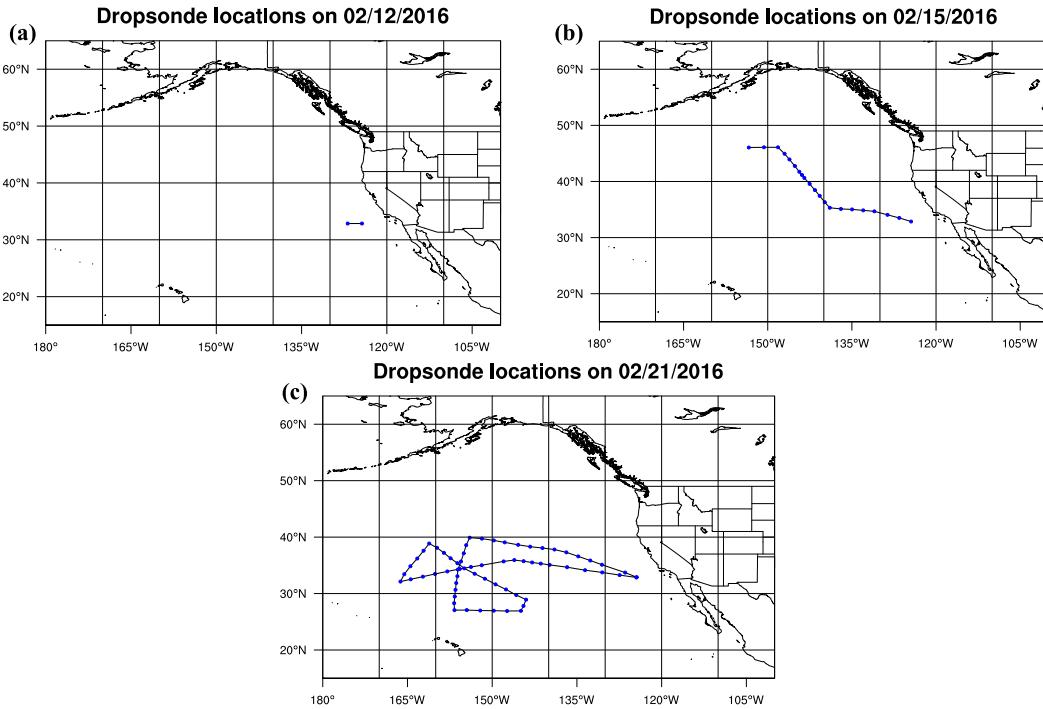
FIG. 12. As in Fig. 1, but for the dropsonde release locations for the three ENRR science flights on (a) 12, (b) 15, and (c) 21–22 Feb 2016.

index is defined as the difference between the environmental temperature and parcel temperature at 500 hPa. Values less than zero indicate an environment conducive to convection. This time period was chosen to match when the severe weather moved through the regions of Louisiana, Mississippi, Alabama, Florida, and Georgia. At the 0600 UTC 24 February 2016 time period in Figs. 16a,b, the GH
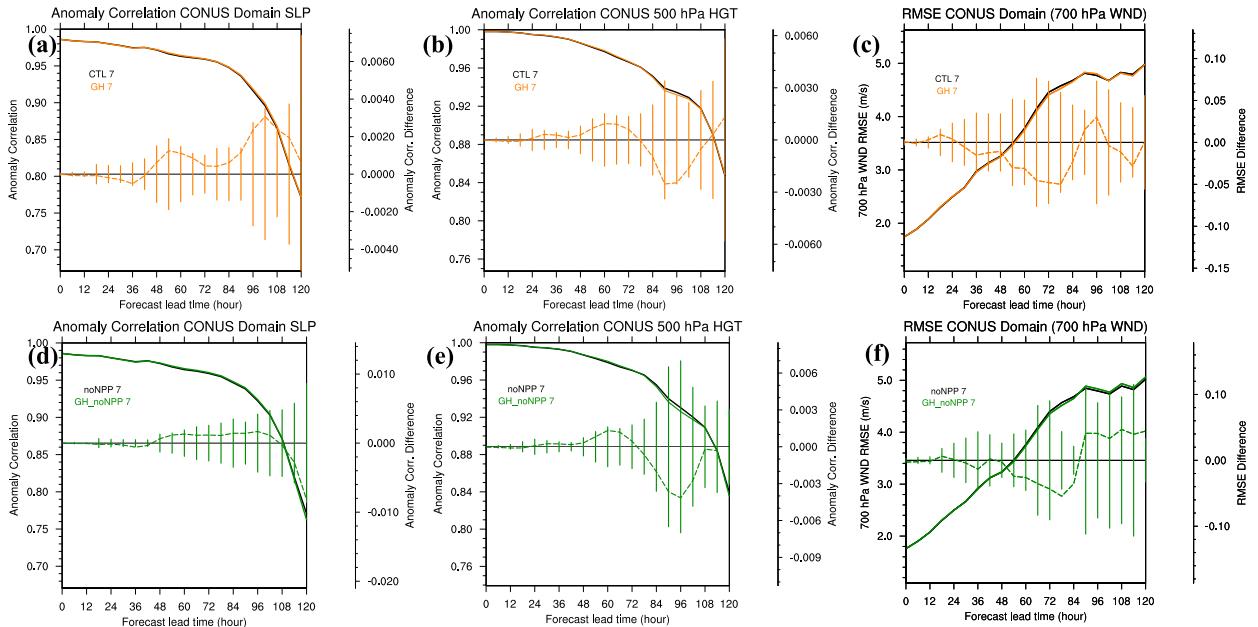
FIG. 13. (a),(b) Anomaly correlation for sea level pressure (hPa) and 500-hPa height, averaged over the CONUS (25°–50°N, 125°–65°W) as a function of forecast lead time for the CTL (black) and GH (orange) experiments. Statistical significance and differences (dashed orange lines) are computed as in Fig. 9. (c) As in (a) and (b), but for the RMSE of 700-hPa wind (m s^{-1}). (d)–(f) As in (a)–(c), but comparing GH_noNPP (green) to noNPP (black) during a possible satellite data gap.
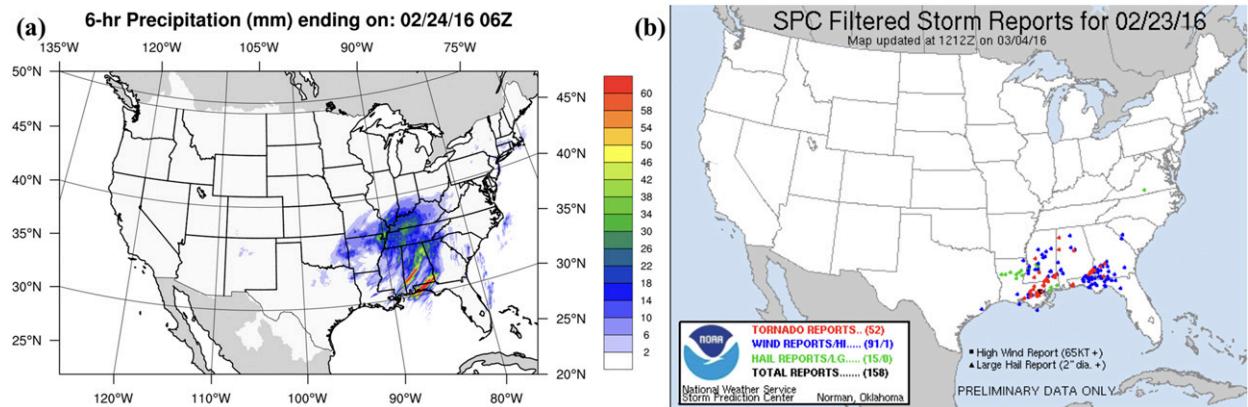
FIG. 14. (a) Instantaneous 6-h precipitation ending at 0600 UTC 24 Feb 2016 (mm) from observations (Lin 2011) over the United States. (b) Storm Prediction Center severe weather reports (filtered) of tornados (red), high winds (blue), and hail (green) over the United States for 23 Feb 2016. (Figure courtesy of the NOAA/NWS/Storm Prediction Center.)

experiment indicates unstable air over the warm sector east of the cold front. The CAPE values are largest over the eastern Gulf of Mexico over $1500 \, \text{J} \, \text{kg}^{-1}$. Farther to the north, in the vicinity of the low pressure system, CAPE values are generally between 400 and $1000 \, \text{J} \, \text{kg}^{-1}$. Correspondingly, lifted indices are less than $-2°\text{C}$ over much of the southeastern United States south of the warm frontal boundary, with values exceeding $-6°\text{C}$ in the Gulf of Mexico.

Differences of GH minus CTL for CAPE and lifted indices are shown in Figs. 16c and 16d. More unstable air is forecasted in the GH experiment relative to CTL, as noted by the positive CAPE differences and negative lifted indices (more unstable). These differences are between 40 and $200 \, \text{J} \, \text{kg}^{-1}$, with lifted indices up to $3°\text{C}$ lower along the cold frontal boundary and on the northwest side of the low pressure system. A region along the south and southwest side of the low
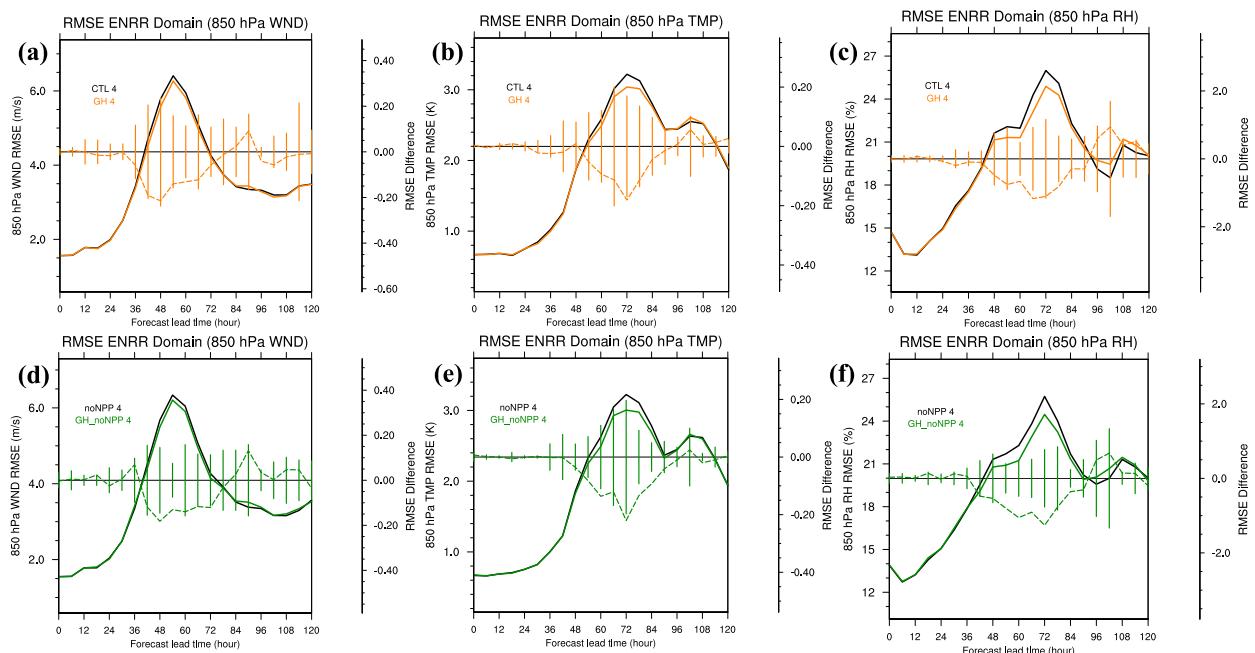

FIG. 15. (a)–(c) RMSE averaged over the southeastern United States (30°–40°N, 95°–80°W) for 850-hPa wind (m s^{-1}), temperature (K), and relative humidity (%) for the CTL (black) and GH (orange) experiments as a function of forecast lead time. Statistics are the same as in Figs. 9 and 13. Results are averaged over all four cycles during the 21–22 Feb 2016 science flight. (d)–(f) As in (a)–(c), but for the GH_noNPP and noNPP experiments under a potential satellite data gap.
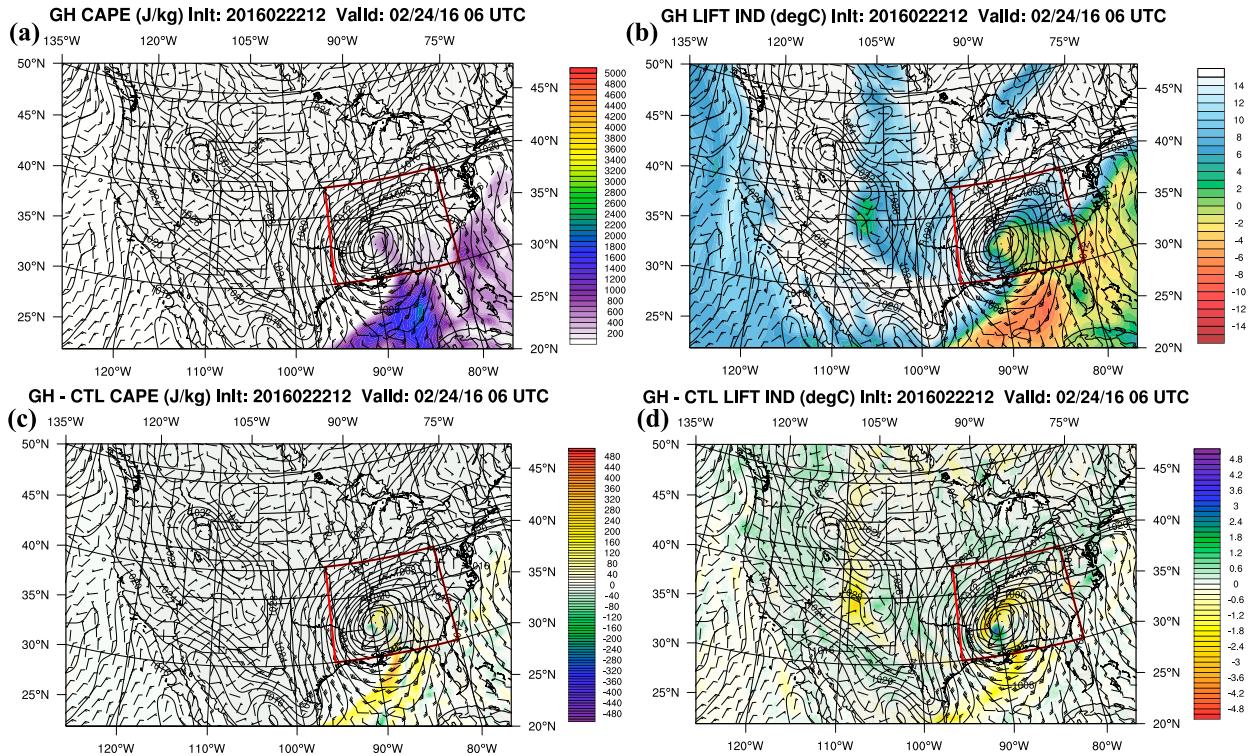
FIG. 16. Latitude–longitude plots of the (a) CAPE (J kg^{-1}; shaded) and (b) lifted index (°C; shaded), wind (m s^{-1}; wind barbs), and sea level pressure (hPa) for the GH experiment initialized at 1200 UTC 22 Feb 2016 for the forecast valid at 0600 UTC 24 Feb 2016 during the last ENRR cycle with dropsondes assimilated. The red box denotes the region used to compute the statistics in Fig. 15 over the southeastern United States (30°–40°N, 95°–80°W). Latitude–longitude plots of GH minus CTL for the same time periods as in (a) and (b), but showing the differences of (c) CAPE (J kg^{-1}; shaded) and (d) lifted index (°C; shaded). The red box is the same as in (a) and (b).

pressure system exhibits more stable air by 80 J kg^{-1} and 1°–2°C. This region may be tied to the dry slot wrapping around the storm. In summary, increased forecast skill over the southeastern United States was attributable to dropsondes released over the eastern Pacific Ocean.

Additionally, we computed the RMSE of CAPE and the 0–6-km above ground level (AGL) bulk wind shear for CTL, GH, noNPP, and GH_noNPP, when verified with ECMWF analysis, under both satellite scenarios. Figure 17 shows the results averaged over the southeastern United States (see Fig. 16) during the third ENRR mission. Although not shown, the ECMWF analysis indicated CAPE values as large as 1050 J kg^{-1} with mean 0–6-km AGL bulk wind shear between 20 and 25 m s^{-1} (39–49 kt) between 1200 UTC 22 February and 1200 UTC 24 February 2016. Although the instability was marginal, the shear that was present during this time was supportive of supercells. Figure 17 shows that while there is an insignificant difference in regard to CAPE, there is a consistent improvement in the wind shear parameter after the assimilation of dropsondes. In general, though, the shear differences

are also not significant. The sampling during this third ENRR flight thus provided an improvement to the low-level environment important for predicting supercells.

5. Summary and discussion

A preliminary data impact study into the value of GH dropsonde data on improving short-range weather forecasts has been investigated under the current satellite platform configuration and a potential gap in satellite data. The focus has been on high-impact storm cases that took place during the 2016 SHOUT field campaigns: 1) Hurricanes Matthew and Nicole over the Atlantic, as well as Typhoon Songda and severe Tropical Storm Aere in the western Pacific (HRR cases) and 2) three winter storms affecting the United States (ENRR). The main conclusions of this study are as follows.

• Under the current satellite configuration, GH dropsondes lead to a statistically significant reduction in the track error for Hurricane Matthew. The track improvement is most evident at later lead times beyond 60 h and is between 7% and 30% compared to CTL. Although the three HRR flights did not target
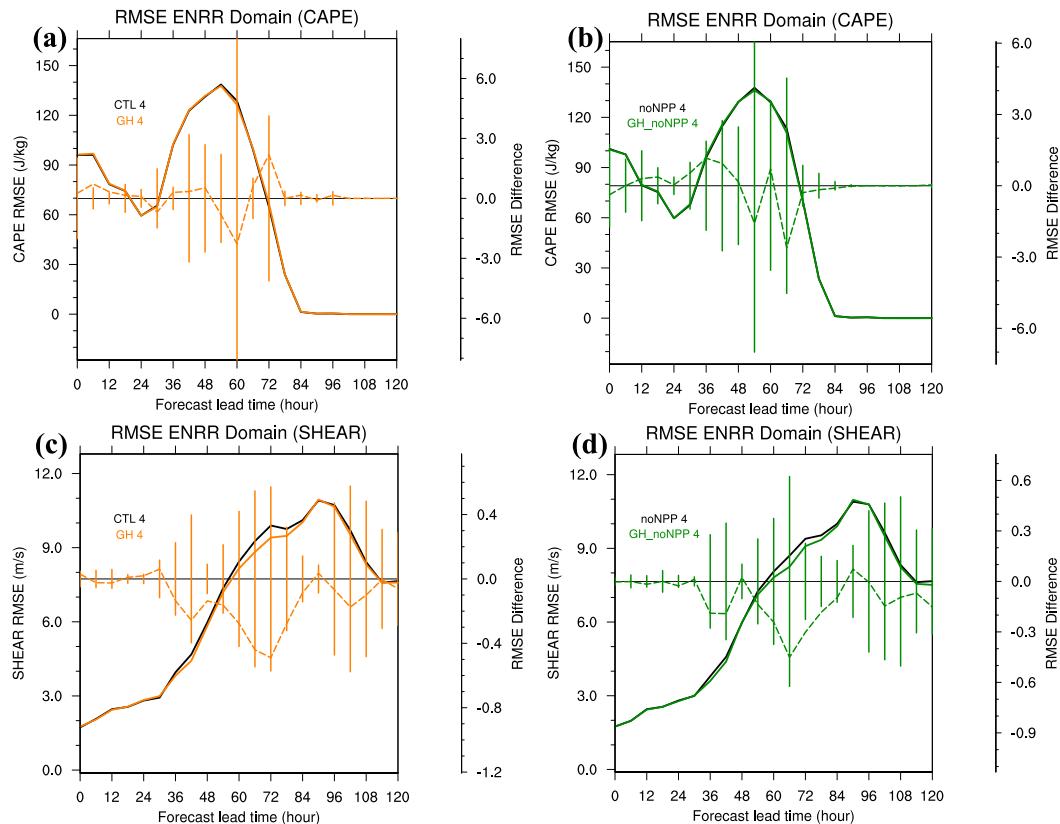
FIG. 17. (a),(b) RMSE averaged over the southeastern United States (30°–40°N, 95°–80°W) of the CAPE (J kg^{-1}) under the current satellite scenario (CTL vs GH) and under a potential satellite data gap (noNPP vs GH_noNPP) as a function of forecast lead time. (c),(d) As in (a) and (b), but for the 0–6 km AGL bulk wind shear (m s^{-1}). Statistics are the same as in Fig. 15. Results are averaged over all four cycles during the 21–22 Feb 2016 science flight.

Nicole, sampling Matthew improved the overall atmospheric pattern and steering flow that was conducive to Nicole's improved storm track. The reduced track error reaches 4%–20% during the 30–96-h time frame and is between 2% and 8% after 120 h.

• Examining four global tropical cyclones that occurred during the 5–10 October time frame over the Atlantic and western Pacific indicates that, on average, track improvements approach 6% during the 30–96-h lead times, with up to 9% at 96 h under the assimilation of dropsondes, similar to Aberson (2011). It has been hypothesized by Sippel et al. (2018) that these high-level dropsonde observations from the GH may be improving the satellite bias correction coefficients in the model, thereby leading to a global impact. On the other hand, Aberson (2011) argues that since the GFS is spectral, local initial condition changes are likely to produce global impacts. Under a satellite data gap, only Hurricane Matthew showed improvement with added dropsonde data, similar to the current satellite configuration. These results point out that while GH

observations were able to partially mitigate a gap in satellite coverage on a regional scale for the HRR mission, it is unlikely to improve global-scale weather. This underscores the importance of Suomi-NPP in global forecast prediction.

• Across most lead times, SLP and wind errors for all tropical cyclones were found to be statistically neutral. Improvements to these fields were found in the dropsonde experiments for both satellite scenarios, primarily after 72 h. The magnitude of these differences were between 1 and 2 hPa (kt).

• Reasons for the improved storm track for both Matthew and Nicole during the three HRR flights appear to be connected to a better representation of the midlevel atmospheric pattern and steering flow at 700 hPa. RMSEs were reduced in the GH experiment relative to CTL by 3%–6% for 700-hPa wind, 6%–11% in 700-hPa geopotential height, and 3%–8% in SLP. Correspondingly, the first Matthew science flight with GH dropsondes initialized a 1–3 m s^{-1} stronger south-to-north flow over the western Atlantic. This coupled

with a stronger ridge during the 48-h forecast is a possible cause for the improved storm track of Matthew.

- Precipitation forecasts verified for all HRR cases over the southeastern United States confirm that added dropsonde observations from the GH increase EQTS values under both satellite scenarios. These improvements are 1%–3% in the 24–48-h time frame, 1%–6% at 48–72-h lead time, and 3%–9% for 72–96-h forecasts.

- For the three ENRR cases investigated during February 2016, results over all cycles were neutral when averaged over the CONUS. Some improvements are found in the 24–72-h lead time, but these results are primarily less than 2% and not significant.

- During the third ENRR flight, which sampled a rapidly intensifying storm system in the central North Pacific, the largest impacts were found to be over the southeastern United States. During this time, a severe weather outbreak occurred over Louisiana, Mississippi, Florida, and Georgia, with a total of 52 tornado reports during the evening of 23–24 February 2016. Sampling the central North Pacific with GH dropsondes improved the downstream low-level storm environment with reduced RMSE in the 0–6 km AGL bulk wind shear. RMSE was also reduced under both satellite scenarios between 4% and 5% in 850-hPa relative humidity, between 2% and 3% in 850-hPa wind, between 5% and 7% in 850-hPa temperature, and between 6% and 8% in SLP at 24–72-h lead times. Forecast increments between GH and CTL for the third ENRR flight indicated a more unstable environment with the assimilation of dropsondes.

Although results are encouraging for the experiments presented, we should stress that they are preliminary and a larger sample size would be necessary to quantify the overall value of the GH UAS. Nevertheless, although not always statistically significant because of the small number of storms and assimilation cycles, the results provide a basis for further research. More recently, Sippel et al. (2018) examined the impact of GH dropsondes for four tropical cyclones in 2016, including Gaston, Hermine, Karl, and Matthew, using the 2017 operational GFS. Over all cases, the track improvements were approximately 12% after 48 h due to the assimilation of dropsondes. Furthermore, global improvements in storm track, similar to the results shown in this paper, were between 6% and 12% over the eastern and western Pacific basins. Additionally, Christophersen et al. (2018a,b) expanded on the number of GH cases, including 10 tropical cyclones between 2012 and 2016, where a total of 10 tropical cyclones were included in the analysis. These experiments were performed using the Hurricane Ensemble Data

Assimilation System (HEDAS; Aksoy et al. 2012) coupled with HWRF. Statistically significant improvements of 10% in track out to 72 h and 10% in SLP up to 108 h were found from the assimilation of dropsondes.

In future studies, we would like to expand on the preliminary results found both in this paper and in Sippel et al. (2018) and Christophersen et al. (2018a,b). In particular, there are instruments on the GH that have yet to be fully utilized in data assimilation, such as the High Altitude MMIC Sounding Radiometer (HAMSR), which measures microwave radiance over 25 spectral channels at high temporal (1.1 s) and spatial (~2 km) resolution (Brown et al. 2011). We plan to develop and implement the capability to assimilate HAMSR radiances into GDAS and run experiments with both HWRF and the preoperational Finite-Volume Cubed-Sphere Dynamical Core GFS (FV3GFS). The combination of HAMSR with GH dropsondes may provide additional forecast benefits. Finally, the impact of targeting was not fully explored in this study, as sensitive regions were only used as guidance for the flight missions and other factors contributed to decisions made concerning the final flight paths. It is possible that sampling a larger portion of the sensitivity region may lead to a greater forecast improvement. A detailed quantitative evaluation of the targeting technique and its impact on predicting high-impact weather events is currently being investigated within an OSSE framework, and the results will be reported upon in the near future.

REFERENCES

Aberson, S. D., 2011: The impact of dropwindsonde data from the THORPEX Pacific Area Regional Campaign and the NOAA hurricane field program on tropical cyclone forecasts in the Global Forecast System. *Mon. Wea. Rev.*, **139**, 2689–2703, https://doi.org/10.1175/2011MWR3634.1.

Aksoy, A., S. Lorsolo, T. Vukicevic, K. J. Sellwood, S. D. Aberson, and F. Zhang, 2012: The HWRF Hurricane Ensemble Data Assimilation System (HEDAS) for high-resolution data: The impact of airborne Doppler radar observations in an OSSE. *Mon. Wea. Rev.*, **140**, 1843–1862, https://doi.org/10.1175/MWR-D-11-00212.1.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748–1765, https://doi.org/10.1175/1520-0469(1999)056<1748:ETAAO>2.0.CO;2.

Black, M., G. A. Wick, and R. E. Hood, 2014: Sensing hazards with operational unmanned technology: NOAA's multi-year plan to deploy the NASA Global Hawk aircraft for high impact weather. *31st Conf. on Hurricanes and Tropical Meteorology*, San Diego, CA, Amer. Meteor. Soc., 5A.7, https://ams.confex.com/ams/31Hurr/webprogram/Paper244777.html.

Braun, S. A., P. A. Newman, and G. M. Heymsfield, 2016: NASA's Hurricane and Severe Storm Sentinel (HS3) investigation. *Bull. Amer. Meteor. Soc.*, **97**, 2085–2102, https://doi.org/10.1175/BAMS-D-15-00186.1.

Brown, S. T., B. Lambrigtsen, R. F. Denning, T. Gaier, P. Kangaslahti, B. H. Lim, J. M. Tanabe, and A. B. Tanner, 2011: The high-altitude MMIC sounding radiometer for the Global Hawk unmanned aerial vehicle: Instrument description and performance. *IEEE Trans. Geosci. Remote Sens.*, **49**, 3291–3301, https://doi.org/10.1109/TGRS.2011.2125973.

Buizza, R., C. Cardinali, G. Kelly, and J.-N. Thépaut, 2007: The value of observations. II: The value of observations in singular-vector-based target areas. *Quart. J. Roy. Meteor. Soc.*, **133**, 1817–1832, https://doi.org/10.1002/qj.149.

Christophersen, H., A. Aksoy, J. Dunion, and K. Sellwood, 2017: The impact of NASA Global Hawk unmanned aircraft dropwindsonde observations on tropical cyclone track, intensity, and structure: Case studies. *Mon. Wea. Rev.*, **145**, 1817–1830, https://doi.org/10.1175/MWR-D-16-0332.1.

——, ——, ——, and S. Aberson, 2018a: Composite impact of Global Hawk unmanned aircraft dropwindsondes on tropical cyclone analyses and forecasts. *Mon. Wea. Rev.*, **146**, 2297–2314, https://doi.org/10.1175/MWR-D-17-0304.1.

——, ——, ——, ——, K. Sellwood, and B. Dahl, 2018b: Impact of Global Hawk dropsondes on tropical cyclone analyses and forecasts. *33rd Conf. on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, Amer. Meteor. Soc., 8D.4, https://ams.confex.com/ams/33HURRICANE/webprogram/Paper338855.html.

Dole, R. M., and Coauthors, 2018: Advancing science and services during the 2015/16 El Niño: The NOAA El Niño Rapid Response field campaign. *Bull. Amer. Meteor. Soc.*, **99**, 975–1001, https://doi.org/10.1175/BAMS-D-16-0219.1.

Earth Observing Laboratory, 1993: NCAR Airborne Vertical Atmospheric Profiling System (AVAPS). Earth Observing Laboratory, UCAR–NCAR, https://doi.org/10.5065/D66W9848.

ECMWF, 2011: ECMWF's operational model analysis, starting in 2011. National Center for Atmospheric Research Computational and Information Systems Laboratory Research Data Archive, accessed 22 August 2017, https://doi.org/10.5065/D6ZG6Q9F.

English, J. M., A. C. Kren, and T. R. Peevey, 2018: Improving winter storm forecasts with Observing System Simulation Experiments (OSSEs). Part 2: Evaluating a satellite gap with idealized and targeted dropsondes. *Earth Space Sci.*, **5**, 176–196, https://doi.org/10.1002/2017EA000350.

Hamill, T. M., F. Yang, C. Cardinali, and S. J. Majumdar, 2013: Impact of targeted Winter Storm Reconnaissance dropwindsonde data on midlatitude numerical weather predictions. *Mon. Wea. Rev.*, **141**, 2058–2065, https://doi.org/10.1175/MWR-D-12-00309.1.

Hock, T. F., and J. L. Franklin, 1999: The NCAR GPS dropwindsonde. *Bull. Amer. Meteor. Soc.*, **80**, 407–420, https://doi.org/10.1175/1520-0477(1999)080<0407:TNGD>2.0.CO;2.

Kleist, D., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational–ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, https://doi.org/10.1175/MWR-D-13-00350.1.

Kren, A. C., H. Wang, J. M. English, T. R. Peevey, and L. Cucurull, 2016: Impact of Global Hawk data from the SHOUT-ENRR 2016 field campaign on an atmospheric river in the central North Pacific. *First Int. Atmospheric Rivers Conf.*, La Jolla, CA, Center for Western Weather and Water Extremes, http://cw3e.ucsd.edu/ARconf2016/Program_Presentations/.

Langland, R. H., 2005: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.*, **131**, 3409–3425, https://doi.org/10.1256/qj.05.130.

Lin, Y., 2011: GCIP/EOP surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV data: Version 1.0. Earth Observing Laboratory, UCAR–NCAR, accessed 29 November 2016, http://data.eol.ucar.edu/dataset/21.093.

Majumdar, S., 2016: A review of targeted observations. *Bull. Amer. Meteor. Soc.*, **97**, 2287–2303, https://doi.org/10.1175/BAMS-D-14-00259.1.

——, and Coauthors, 2011: Targeted observations for improving numerical weather prediction: An overview. WWRP/THORPEX Publ. 15, 37 pp., www.wmo.int/pages/prog/arep/wwrp/new/documents/THORPEX_No_15.pdf.

Neiman, P. J., F. M. Ralph, G. A. Wick, J. D. Lundquist, and M. D. Dettinger, 2008: Meteorological characteristics and overland precipitation impacts of atmospheric rivers affecting the west coast of North America based on eight years of SSM/I satellite observations. *J. Hydrometeor.*, **9**, 22–47, https://doi.org/10.1175/2007JHM855.1.

Rabier, F., and Coauthors, 2008: An update on THORPEX-related research in data assimilation and observing strategies. *Nonlinear Processes Geophys.*, **15**, 81–94, https://doi.org/10.5194/npg-15-81-2008.

Ralph, F. M., and M. D. Dettinger, 2011: Storms, floods, and the science of atmospheric rivers. *Eos, Trans. Amer. Geophys. Union*, **92**, 265–272, https://doi.org/10.1029/2011EO320001.

——, P. J. Neiman, and G. A. Wick, 2004: Satellite and CALJET aircraft observations of atmospheric rivers over the eastern North Pacific Ocean during the winter of 1997/98. *Mon. Wea. Rev.*, **132**, 1721–1745, https://doi.org/10.1175/1520-0493(2004)132<1721:SACAOO>2.0.CO;2.

Sela, J., 2009: The implementation of the sigma pressure hybrid coordinate into the GFS. NCEP Office Note 461, 25 pp., http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on461.pdf.

Sippel, J. A., K. L. Friedman, and V. Tallapragada, 2018: The impact of Global Hawk dropsondes on NCEP hurricane forecasts. *33rd Conf. on Hurricanes and Tropical Meteorology*, Ponte Vedra, FL, Amer. Meteor. Soc., 8D.6, https://ams.confex.com/ams/33HURRICANE/webprogram/Paper338883.html.

Toth, Z., and Coauthors, 2000: Targeted observations at NCEP: Toward an operational implementation. Preprints, *Fourth Symp. on Integrated Observing Systems*, Long Beach, CA, Amer. Meteor. Soc., 186–193.

——, and Coauthors, 2002: Adaptive observations at NCEP: Past, present, and future. *Symp. on Observations, Data Assimilation, and Probabilistic Prediction*, Orlando, FL, Amer. Meteor. Soc., 7.1, https://ams.confex.com/ams/pdfpapers/26036.pdf.

Wang, X., and T. Lei, 2014: GSI-based four-dimensional ensemble–variational (4DEnsVar) data assimilation: Formulation and

single-resolution experiments with real data for the NCEP Global Forecast System. *Mon. Wea. Rev.*, **142**, 3303–3325, https://doi.org/10.1175/MWR-D-13-00303.1.

Wick, G., J. Dunion, and J. Walker, 2018a: Sensing hazards with operational unmanned technology: Impact study of Global Hawk unmanned aircraft system observations for hurricane forecasting, final report. NOAA Tech. Memo. OAR-UAS-002, 92 pp., ftp://ftp1.esrl.noaa.gov/et6/sat/shout/data_impact/TM_SHOUT_Impact_Assessment_FINAL.pdf.

——, T. F. Hock, P. J. Neiman, H. Vömel, M. L. Black, and J. R. Spackman, 2018b: The NCAR–NOAA Global Hawk dropsonde system. *J. Atmos. Oceanic Technol.*, **35**, 1585–1604, https://doi.org/10.1175/JTECH-D-17-0225.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Academic Press, 627 pp.

Zhang, Y., Y. F. Xie, H. L. Wang, D. H. Chen, and Z. Toth, 2016: Ensemble transform sensitivity method for adaptive observations. *Adv. Atmos. Sci.*, **33**, 10–20, https://doi.org/10.1007/s00376-015-5031-9.

Zhu, Y., and R. E. Newell, 1998: A proposed algorithm for moisture fluxes from atmospheric rivers. *Mon. Wea. Rev.*, **126**, 725–735, https://doi.org/10.1175/1520-0493(1998)126<0725:APAFMF>2.0.CO;2.