Comparison of 2018-2021 tropical cyclone track forecasts before and after NOAA G-IV missions

Melissa Piper
University at Albany, State University of New York, pipermel19@gmail.com

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Atmospheric Sciences Commons, and the Meteorology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/2995

This Master's Thesis is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
Comparison of 2018–2021 Tropical Cyclone Track Forecasts Before and After NOAA G-IV Missions

by

Melissa A. Piper

A Thesis
Submitted to the University at Albany, State University of New York
In Partial Fulfillment of
the Requirements for the Degree of
Master of Science

College of Arts and Sciences
Department of Atmospheric and Environmental Sciences
Fall 2022
ABSTRACT

Tropical cyclone (TC) hazards are primarily dictated by the TC position; thus, it is important to produce accurate TC track forecasts. Many different features can influence a TC’s motion, yet these features are not always well sampled by in-situ observations over the open ocean, creating a need for supplemental observation collected via aircraft, including the deployment of dropsondes. In 1997, the National Hurricane Center (NHC) began operational synoptic surveillance missions in the near-storm environments of TCs using the Gulfstream IV-SP jet aircraft (G-IV), with the goal of reducing track forecast errors. In the first 10 years of operational missions, the dropsonde data collected during 176 G-IV missions led to a 10–15% improvement in 0–60-h track forecast errors for forecasts initialized at mission nominal times (Aberson 2010). However, despite the addition dropsondes deployed around the TC core to operations and additional research into optimal targeting strategies, no further research has been published regarding the impacts of G-IV dropsondes on recent TC track forecasts. Therefore, this thesis expands on previous studies by investigating the impacts of G-IV dropsonde data on Atlantic basin TC track forecasts from 2018–2021 when ensemble-based sensitivity was utilized for creating aircraft tracks.

This research investigates the impacts of dropsonde data collected during NOAA G-IV synoptic surveillance missions on position forecasts for Atlantic basin TCs from 2018–2021 by comparing forecasts initialized with dropsonde data available against forecasts without dropsondes and to the forecast initialized 12-h before dropsondes. ECMWF EPS, GEFS, and CLP5 position forecasts are analyzed for 675 forecast initialization times, of which 56 had dropsonde data assimilated at the time of initialization. When compared to the 619 forecast initialization times without dropsondes, the forecast initialization times with dropsondes are found to have lower
average position errors and higher skill relative to CLP5, suggesting that the dropsonde data is likely adding skill to position forecasts. In contrast, when forecasts initialized with dropsondes are compared to forecasts initialized 12-h before dropsondes, the results indicate that differences in skill and average position error are similar to differences seen in randomly selected pairs of forecast initializations from climatology.

Furthermore, four cases studies are conducted to diagnose the potential sources of significant position error reductions in track forecasts initialized with the assimilation of dropsondes. Two potential sources are identified: changes in the environmental steering flow that yield differences in TC motion (hurricanes Zeta (2020) and Jerry (2019)) and a change in the initial position of the TC that placed the TC into a different steering flow regime (hurricanes Marco (2020) and Dorian (2019)).
ACKNOWLEDGMENTS

I am endlessly grateful to everyone who has supported me during this journey. I want to first extend a huge thank you to my advisor, Dr. Ryan Torn, for believing in me and giving me the opportunity to grow as a scientist. I had a lot of fears related to elevating myself to a graduate level of work and with writing this thesis, and I am eternally grateful for his mentorship and the dedication he has to his students. I would also like to thank my second reader, Dr. Kristen Corbosiero, who not only provided additional feedback on this thesis, but also supported me through my first semester as a teaching assistant. There are countless others I wish to thank from this department, including Chaina Porter and Barbara Zampella, who are rockstars with the administrative and travel concerns of the department.

I would not be in the position I am today without the support of my parents, Glen and Debbi Piper. I can’t imagine it was easy for them to send their 18-year-old daughter 1400 miles away to Iowa for college, then even further away to New York for graduate school, but I can never thank them enough for always being there for me and helping me through the highs and lows of the past 6 years. Of course, I must also thank all of the animals currently in my life, including Miss Rory, who I adopted a year into graduate school. Her companionship has been crucial to me during my graduate studies.

Finally, I would like to thank all of the incredible people I have met in the department since I moved to Albany two and a half years ago. I deeply value your friendship and the comradery we have through the table we commandeered on the 3rd floor.
# TABLE OF CONTENTS

ABSTRACT ............................................................................................................................. ii

ACKNOWLEDGEMENTS ......................................................................................................... iv

LIST OF TABLES .................................................................................................................... viii

LIST OF FIGURES ................................................................................................................... ix

1. Introduction .......................................................................................................................... 1

   1.1. Motivation ..................................................................................................................... 1

   1.2. Literature Review ......................................................................................................... 2

   1.3. Research Goals and Thesis Structure ........................................................................... 11

2. Data and Methods ............................................................................................................. 13

   2.1. NOAA G-IV Synoptic Surveillance Missions ................................................................. 13

   2.2. Analysis of Error and Skill Statistics over Multiple Seasons ....................................... 14

      2.2.1. Datasets .................................................................................................................. 14

         2.2.1.1. ECMWF EPS ............................................................................................... 14

         2.2.1.2. GEFS ............................................................................................................. 15

         2.2.1.3. CLP5 ............................................................................................................ 16

         2.2.1.4. HURDAT2 ................................................................................................. 17

      2.2.2. Methodology ......................................................................................................... 18

         2.2.2.1. Selection of Track Forecasts ........................................................................... 18

         2.2.2.2. Subsets of Track Forecast Initialization Times .............................................. 19

       2.2.3. Calculations ........................................................................................................... 20

          2.2.3.1. Track Forecast Errors .................................................................................. 20

          2.2.3.2. Track Forecast Skill .................................................................................... 21
2.2.3.3. Frequency of Superior Performance .......................................................... 21
2.2.4. Bootstrap Resampling .................................................................................. 22
2.3. Case Studies .................................................................................................. 23
   2.3.1. Datasets ................................................................................................... 24
   2.3.2. Methodology .......................................................................................... 25
       2.3.2.1. Track Forecast Errors ...................................................................... 25
       2.3.2.2. Environmental Steering Flow ......................................................... 25
       2.3.2.3. Major Axis Correlation Coefficient ............................................... 27
3. Results ............................................................................................................ 34
   3.1. Analysis of Error and Skill Statistics over Multiple Seasons ....................... 34
       3.1.1. With vs Without Dropsondes ............................................................. 35
       3.1.2. Before vs After Dropsondes ............................................................... 39
   3.2. Case Studies ............................................................................................... 44
       3.2.1. Hurricane Marco (2020) ................................................................. 44
           3.2.1.1. Synoptic Overview ..................................................................... 44
           3.2.1.2. Forecast Analysis ...................................................................... 47
       3.2.2. Hurricane Zeta (2020) ...................................................................... 51
           3.2.2.1. Synoptic Overview ..................................................................... 51
           3.2.2.2. Forecast Analysis ...................................................................... 54
       3.2.3. Hurricane Dorian (2019) .................................................................. 58
           3.2.3.1. Synoptic Overview ..................................................................... 58
           3.2.3.2. Forecast Analysis ...................................................................... 61
       3.2.4. Hurricane Jerry (2019) ..................................................................... 65
3.2.4.1. Synoptic Overview ......................................................................................... 65
3.2.4.2. Forecast Analysis ......................................................................................... 67

4. Conclusions ............................................................................................................. 102

REFERENCES ............................................................................................................... 108
LIST OF TABLES

Table 2.1. The basin, year, storm name, take-off time, and assimilation time for every NOAA G-IV synoptic surveillance reconnaissance mission from 2018–2021.

Table 2.2. Operational ECMWF EPS versions during the period of study, the date each version was implemented, and any notable changes from the previous version.

Table 2.3. Number of named TCs in the Atlantic basin that had at least one valid forecast initialization time and the total number of valid forecast initialization times per year.

Table 2.4. The name, mission take-off time, and mission date of each case study.

Table 2.5. Tropical cyclone name and optimal steering flow layer and radius used for each case.
LIST OF FIGURES

Figure 2.1. An example schematic for the position ellipses and associated major axis used in the MACC calculation. The dots indicate the forecasted position of each ensemble member at 24-h intervals, while the triangles indicate the forecasted ensemble mean position at 24-h intervals. The ellipses show the bivariate normal distribution of the forecast positions at 24-h intervals, as in Hamill et al. (2011). Purple denotes the 24-h locations, cyan denotes 48-h locations, green denotes 72-h locations, red denotes 96-h locations, and magenta denotes 120-h locations. The red arrows indicate the ellipse major axis each 24 h. The black arrows indicate the great circle distance between the ensemble mean position and an example ensemble member position each 24 h. The blue arrows indicate the distance along the major axis, oriented from the ensemble mean position toward an example ensemble member each 24 h.

Figure 3.1. The mean ECMWF EPS (blue), GEFS (magenta), and CLP5 (orange) 72-h track forecast error (km) from the same forecast initialization times in each CLP5 error percentile bin.

Figure 3.2. The mean track forecast error (km) for (a) ECMWF EPS (blue) and GEFS (magenta) and (b) CLP5 (blue) forecasts initialized with dropsondes (circle markers) and without dropsondes (x markers). The number of forecasts within each average error is listed in black text at the top of each panel. The percentage of randomly sampled forecasts from climatology that have lower average track forecast error than the forecasts with dropsondes is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.

Figure 3.3. The average skill relative to CLP5 (%) for ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized with dropsondes (circle markers) and forecasts initialized without dropsondes (x markers). The black text denotes the number of forecasts within the average skill at each forecast hour. The blue (ECMWF EPS) and magenta (GEFS) text denote the percentage of randomly sampled forecasts from climatology that have less skill relative to CLP5 than the forecasts initialized with dropsondes.

Figure 3.4. The mean track forecast error (km) for (a) ECMWF EPS (blue) and GEFS (magenta) and (c) CLP5 (orange) forecasts initialized with dropsondes (circle markers) and 12 h before dropsondes (left-facing triangle markers). The difference in mean track forecast error (km) for (b) ECMWF EPS (blue) and GEFS (magenta) and (d) CLP5 (orange) forecasts initialized with dropsondes and 12 h before dropsondes (circle markers) and the difference between all pairs of forecast initializations in the climatology (diamond markers). The number of forecasts within each average error is listed in black text at the top of each panel. The percentage of randomly sampled pairs of forecasts from climatology that have a lower reduction in track forecast error than the forecasts with dropsondes is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.

Figure 3.5. The frequency of superior performance (FSP; %) for (a) ECMWF EPS, (b) GEFS, and (c) CLP5 forecasts initialized after the assimilation of dropsondes (circle marker) and all “after” forecast initializations in the climatology (diamond marker). The percentage of randomly sampled “after” forecasts from climatology that have a lower FSP than the after dropsondes forecast
initializations is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.

Figure 3.6. The (a) average skill relative to CLP5 (%) for ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized 12 h before dropsondes (left-facing triangle marker) and with dropsondes (circle marker) and (b) the difference in average skill relative to CLP5 (%) between ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized 12 h before and with dropsondes (circle markers) and all pairs of forecast initializations in the climatology (diamond marker). The black text denotes the number of forecasts within the average skill at each forecast hour. The blue (ECMWF EPS) and magenta (GEFS) text denote the percentage of randomly sampled pairs of forecasts from climatology that have a greater reduction in average skill than forecasts with dropsondes.

Figure 3.7. (a), (b) ECMWF EPS and GEFS (c), (d) forecasts of Hurricane Marco initialized at (a), (c) 0000 UTC 22 August 2020 and (b), (d) 1200 UTC 22 August 2020 (gray lines). The dots indicate the location of each ensemble member at 24-h intervals. The colored ellipses show the bivariate normal distribution of the forecast positions at 24-h intervals, as in Hamill et al. (2011). Purple denotes the 24-h locations, cyan denotes 48-h locations, green denotes 72-h locations, red denotes 96-h locations, and magenta denotes 120-h locations.

Figure 3.8. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Marco beginning at 1200 UTC 22 August 2020 and ending at 0000 UTC 25 August 2020. The forecasts initialized at 0000 UTC 22 August 2020 are denoted by left-facing triangle markers, while the forecasts initialized at 1200 UTC 22 August 2020 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 0600 UTC 22 August 2020 and by black squares for the position starting at 1200 UTC 22 August 2020. The dropsonde deployment locations during the 0530 UTC 22 August 2020 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.

Figure 3.9. A portion of the 4:00 PM CDT 22 August 2020 NWS NHC forecast discussion for then Tropical Storm Marco. Adapted from https://www.nhc.noaa.gov/archive/2020/al14/al142020.discus.010.shtml.

Figure 3.10. ECMWF EPS (blue) and GEFS (magenta) mean track error every 6 h for forecasts initialized at 0000 UTC 22 August 2020 (left-facing triangle marker) and at 1200 UTC 22 August 2020 (circle marker) for Hurricane Marco.

Figure 3.11. The ensemble mean environmental steering flow vectors (kt) from the 0000 UTC 22 August 2020 (black barbs) and the 1200 UTC 22 August 2020 (blue barbs) ECMWF EPS forecast initializations for Hurricane Marco valid (a) 1200 UTC 22 August 2020, (b) 0600 UTC 23 August 2020, (c) 0000 UTC 24 August 2020, and (d) 1200 UTC 24 August 2020. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 0000 UTC 22 August 2020 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 1200 UTC 22 August 2020 forecast is denoted by the blue circle marker.
Figure 3.12. The magnitude of the vector difference in the near-TC steering flow (i.e., within 100 km of the TC position) (kts; red) and the distance between the ensemble mean position (km; black) for 6-h forecasts initialized at 0000 UTC 22 August 2020 and 1200 UTC 22 August 2020 for Hurricane Marco. Forecasts are valid from 1200 UTC 22 August 2020 to 0000 UTC 25 August.

Figure 3.13. The major axis correlation coefficient (MACC) for 6-h position forecasts initialized for Hurricane Marco at 0000 UTC 22 August 2020 (left-facing triangle marker) and 1200 UTC 22 August 2020 (circle marker). Forecasts are valid from 1200 UTC 22 August 2020 to 0000 UTC 25 August 2020.

Figure 3.14. As in Figure 3.7, but for Hurricane Zeta initialized at (a), (c) 1200 UTC 25 October 2020 and (b), (d) 0000 UTC 26 October 2020.

Figure 3.15. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Zeta beginning at 0000 UTC 26 October 2020 and ending at 1200 UTC 29 October 2020. The forecasts initialized at 1200 UTC 25 October 2020 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 26 October 2020 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 25 October 2020 and by black squares for the position starting at 0000 UTC 26 October 2020. The dropsonde deployment locations during the 1730 UTC 25 October 2020 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.

Figure 3.16. As in Figure 3.10, but for Hurricane Zeta initialized at 1200 UTC 25 October 2020 (left-facing triangle marker) and at 0000 UTC 26 October 2020 (circle marker).

Figure 3.17. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 25 October 2020 (black barbs) and the 0000 UTC 26 October 2020 (blue barbs) ECMWF EPS forecast initializations for Hurricane Zeta valid (a) 0000 UTC 26 October 2020, (b) 1800 UTC 26 October 2020, (c) 1800 UTC 27 October 2020, and (d) 1200 UTC 28 October 2020. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 25 October 2020 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 26 October 2020 forecast is denoted by the blue circle marker.

Figure 3.18. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 25 October 2020 and 0000 UTC 26 October 2020 for Hurricane Zeta. Forecasts are valid from 0000 UTC 25 October 2020 to 1200 UTC 28 October 2020.

Figure 3.19. The MACC 6-h position forecasts initialized for Hurricane Zeta at 1200 UTC 25 October 2020 (left-facing triangle marker) and 0000 UTC 26 October 2020 (circle marker). Forecasts are valid from 0000 UTC 26 October 2020 to 1200 UTC 28 October 2020.

Figure 3.20. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Dorian beginning at 0000 UTC 28 August 2019 and ending at 0000 UTC 02 September 2019. The
forecasts initialized at 1200 UTC 27 August 2019 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 28 August 2019 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 27 August 2019 and by black squares for the position starting at 0000 UTC 28 August 2019. The dropsonde deployment locations during the 1730 UTC 27 August 2019 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.

Figure 3.21. As in Figure 3.7, but for Hurricane Dorian initialized at (a), (c) 1200 UTC 27 August 2019 and (b), (d) 0000 UTC 28 August 2019.

Figure 3.22. As in Figure 3.10, but for Hurricane Dorian initialized at 1200 UTC 27 August 2019 (left-facing triangle marker) and at 0000 UTC 28 August 2019 (circle marker).

Figure 3.23. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 27 August 2019 (black barbs) and the 0000 UTC 28 August 2019 (blue barbs) ECMWF EPS forecast initializations for Hurricane Dorian valid (a) 0000 UTC 28 August 2019, (b) 1800 UTC 28 August 2019, (c) 1800 UTC 29 August 2019, and (d) 1200 UTC 30 August 2019. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 27 August 2019 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 28 August 2019 forecast is denoted by the blue circle marker.

Figure 3.24. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 27 August 2019 and 0000 UTC 28 August 2019 for Hurricane Dorian. Forecasts are valid from 0000 UTC 28 August 2019 to 1200 UTC 30 August 2019.

Figure 3.25. The MACC 6-h position forecasts initialized for Hurricane Dorian at 1200 UTC 27 August 2019 (left-facing triangle marker) and 0000 UTC 28 August 2019 (circle marker). Forecasts are valid from 0000 UTC 28 August 2019 to 1200 UTC 30 August 2019.

Figure 3.26. As in Figure 3.7, but for Hurricane Jerry initialized at (a), (c) 1200 UTC 20 September 2019 and (b), (d) 0000 UTC 21 September 2019.

Figure 3.27. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Jerry beginning at 0000 UTC 21 September 2019 and ending at 1200 UTC 24 September 2019. The forecasts initialized at 1200 UTC 20 September 2019 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 21 September 2019 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 20 September 2019 and by black squares for the position starting at 0000 UTC 21 September 2019. The dropsonde deployment locations during the 1730 UTC 20 September 2019 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.

Figure 3.28. As in Figure 3.10, but for Hurricane Jerry initialized at 1200 UTC 20 September 2019 (left-facing triangle marker) and at 0000 UTC 21 September 2019 (circle marker).
Figure 3.29. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 20 September 2019 (black barbs) and the 0000 UTC 21 September 2019 (blue barbs) ECMWF EPS forecast initializations for Hurricane Jerry valid (a) 0000 UTC 21 September 2019, (b) 1200 UTC 21 September 2019, (c) 0600 UTC 22 September 2019, and (d) 0000 UTC 23 September 2019. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 20 September 2019 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 21 September 2019 forecast is denoted by the blue circle marker.

Figure 3.30. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 20 September 2019 and 0000 UTC 21 September 2019 for Hurricane Jerry. Forecasts are valid from 0000 UTC 21 September 2019 to 1200 UTC 23 September 2019.

Figure 3.31. The MACC 6-h position forecasts initialized for Hurricane Jerry at 1200 UTC 20 September 2019 (left-facing triangle marker) and 0000 UTC 21 September 2019 (circle marker). Forecasts are valid from 0000 UTC 21 September 2019 to 1200 UTC 23 September 2019.
1. Introduction

1.1. Motivation

Accurately forecasting a tropical cyclone’s (TC’s) track is extremely important, especially when a TC is a threat for making landfall. TC hazards (e.g., storm surge, inland flooding, extreme winds, etc.) are primarily dictated by the TC position; thus, providing accurate track forecasts for landfalling TCs allows for forecasters and local officials to issue warnings and evacuation orders at earlier lead times and for the counties most directly impacted by TC hazards. While TC track forecasts have substantially improved in recent decades, continued improvements in track predictability have begun to plateau (Rappaport et al. 2009; Landsea and Cangialosi 2018). As of the most recent Atlantic hurricane season (2021), TC track forecast errors continue to exhibit less year-over-year improvement than has been seen over recent decades (Cangialosi 2022). Despite these long-term improvements in track predictability, outliers with large track forecast errors still occur, as TC track errors can vary significantly on a year-to-year and storm-to-storm basis (Rappaport et al. 2009; Landsea and Cangialosi 2018). One such example is Hurricane Marco (2020), where a significant westward bias in the official National Hurricane Center (NHC) forecasts and guidance models led to significantly larger forecast track errors than the NHC mean official errors for the previous 5-year period (Beven and Berg 2021). One method to account for the year-to-year and storm-to-storm variability of track errors is to normalize the track errors against a benchmark, such as a climatology and persistence model (CLP5), to infer the skill of the track forecasts (Aberson 1998). In recent years, however, there has not been any further improvements in track forecasting skill in the Atlantic Ocean, suggesting a minimal rate of
improvement or even no further advances in skill, as is seen in position error trends (Landsea and Cangialosi 2018).

Further reductions in track forecast error and improvements in forecast skill might be possible by investigating the sources behind large position errors in outlier cases (e.g., Yamaguchi et al. 2017). Many different features can influence a TC’s path, yet these features are not always well sampled by in-situ observations over the open ocean, creating a need for supplemental observations. One method of collecting supplemental observations of TCs is via aircraft, including the deployment of dropsondes into the environment surrounding a TC. Aircraft reconnaissance missions, however, are expensive and should be justified through observed improvements in TC track forecasts with the data assimilated relative to forecasts without the data, either by reducing track forecast errors, improving track forecast skill, or both. Therefore, further advancements in TC track predictability require an understanding of the impact of aircraft reconnaissance observations on operational models.

1.2. Literature Review

It has been well established that a TC’s motion is primarily driven by the deep-layer wind field (i.e., environmental steering flow), which consists of the remaining flow sources after the wind field associated with the TC vortex is removed (e.g., George and Gray 1976; Chan and Gray 1982; Pike 1985; Aberson and DeMaria 1994). George and Gray (1976) examined TC motion in the western North Pacific basin and found that a TC’s movement corresponds well to the mean 500-hPa wind direction and speed within a $1^\circ$–$7^\circ$ radius from the TC center, regardless of the TC’s position and characteristics. Other studies have suggested that a deeper steering layer better
dictates a TC’s motion. Chan and Gray (1982) explored TC motion in the northwest Pacific, the west Atlantic, and the Australian-South Pacific basins and found that the depth of the steering wind layer often corresponds well to the 500–700-hPa winds averaged within a 5°–7° radius from the TC center, but this layer will vary with the amount of directional wind shear with height (i.e., more directional shear with height results in a deeper steering wind layer). Pike (1985) investigated the use of geopotential thicknesses and heights in TC motion and intensity prediction in the Atlantic basin and noted that the deep-layer mean (1000–100-hPa) height field and the mid-level (700–500-hPa) heights produce the most accurate TC motion forecasts.

While other studies have corroborated the effectiveness of using a deep-layer mean to compute the steering wind, they have also suggested that the exact depth of the steering wind layer corresponds to the intensity of the TC. Dong and Neumann (1986) analyzed the 24-h motion of 920 cases of Atlantic basin TCs from 1965–1977 and found that, in general, track forecast errors are minimized when using a deep-layer average steering flow layer, while the height of the best steering layer depth increases in proportion to the intensity of the TC. Velden and Leslie (1991) employed a simple barotropic model in the Australian basin to expand on the general observed relationship between the steering layer depth and TC intensity found by Dong and Neumann (1986). From their sample of 300 cases, they found that weaker systems (mean sea level pressure > 975 hPa) had the lowest track forecast errors when using the 850–500-hPa layer-average steering flow, while stronger systems (mean sea level pressure < 955 hPa) had the lowest track forecast errors when using the 850–300-hPa layer-average steering flow. Aberson and DeMaria (1994) also explored the relationship between TC intensity and the 850–200-hPa average steering flow in the north Atlantic basin and found that tropical depressions and tropical storms had higher track forecast errors than hurricanes, suggesting that a shallower layer-averaged steering flow better
describes the motion of weaker systems. More recently, Galarneau and Davis (2013) sought to reduce the uncertainty regarding the most appropriate deep-layer steering flow for an individual TC by developing an algorithm which computes the optimal steering flow layer for an individual TC forecast. This algorithm identifies the environmental steering flow by subtracting the nondivergent and irrotational wind vectors associated with the TC vortex from the total wind field and determines the radius and vertical layer where the background flow is most like the observed TC motion (i.e., the optimal steering flow layer). The ability to diagnose the most appropriate deep-layer steering flow for a given TC forecast allows for further reductions in TC track forecast error that are due to misrepresentations of the deep-layer environmental steering flow.

George and Gray (1976) and Chan and Gray (1982) also noted in their findings that Northern Hemisphere TCs deviate ~10°–20° to the left of the surrounding mean flow. This bias was explained in modeling studies by Anthes and Hoke (1975) and Holland (1983) as the result of the cyclonic circulation of a TC in the Northern Hemisphere interacting with the meridional gradient in planetary vorticity. They found that a TC’s circulation can advect planetary vorticity, resulting in a cyclonic circulation anomaly to the left and an anticyclonic circulation anomaly to the right of the TC center. The perturbation winds associated with these anomalies, also referred to as β-gyres, constructively interfere over the center of the TC, resulting in the northwestward motion of the TC. Franklin et al. (1996) verified the results of the aforementioned modeling studies within a sample of 16 observed TCs and suggested that this “β-effect” is a relatively small portion of the factors that influence TC motion compared to synoptic scale features.

Previous studies have also shown that other factors can contribute to TC motion. Wang and Holland (1996) and Wu and Wang (2001) found that diabatic heating associated with TC convection can affect the motion of a TC. Vertical shear in the near-TC environment leads to an
asymmetric distribution of convection, resulting in a low-level cyclonic circulation to the right and a low-level anticyclonic circulation to the left of the TC center. These low-level circulations induce changes in the vertical gradient of asymmetric diabatic heating, moving the TC towards the region of maximum convection. Wong and Chan (2006) have shown that frictional effects from nearby landmasses force asymmetries in the near-surface winds, altering the motion of a TC away from the mean steering flow and toward land. Lastly, Bender et al. (1993) and Wu et al. (2005) identified a small, yet persistent, influence on TC motion due to interactions between a TC and the underlying ocean that result in modifications to the environment of the TC and associated deviations in the expected motion of the TC.

Despite our understanding of the mechanisms driving TC motion, continued improvements in TC predictability have slowed in recent years compared to the rapid improvement seen around the start of the 21st century (Rappaport et al. 2009; Landsea and Cangialosi 2018). To achieve further reductions in track forecast error and improvements in track forecast skill, the focus may require shifting toward identifying difficult track cases, investigating the sources of uncertainty driving those large position errors, and developing methodologies to reduce the uncertainty associated with those sources (e.g., Yamaguchi et al. 2017; Magnusson et al. 2019). Many studies have explored the types of situations which lead to difficult TC track forecasts and determined that the overwhelming majority of these cases involve complex interactions between a TC and its environmental steering flow, such as substantial curvature to the motion (i.e., recurvature). Chan et al. (1980) examined TCs that occurred in the Caribbean Sea from 1961–1977 and found that storms which undergo turning motion experience large vertical wind shear between 900–200 hPa in the direction of storm motion 24–36-h before the onset of the turning. Recently, Bi et al. (2020) analyzed TCs that underwent sudden turning motions from May to November 2000–2019 in the
western North Pacific basin and found that sudden right turns are accompanied by a deepening mid-to-upper-level trough and weakening subtropical ridge, while sudden left turns are accompanied by stronger easterly winds to the north of the TC center that result from an increased meridional pressure gradient between the nearby monsoon trough and subtropical ridge.

Alongside recurvature, previous studies have shown that interactions between a TC and nearby large-scale synoptic features can also result in difficult track forecasts. For example, Carr and Elsberry (2000) analyzed 69 cases of large (300 n mi or 555 km at 72 h) track forecast errors in the tropical western North Pacific basin during 1997 using the Navy Operational Global Atmospheric Prediction System (NOGAPS) and found that large position errors stem from both direct cyclone interactions (e.g., incorrect TC or adjacent cyclone size or separation distance within the model) and indirect cyclone interactions (e.g., ridge modification by a TC and reverse trough formation). Wu et al. (2004) applied a potential vorticity diagnosis to investigate the factors affecting the missed deceleration of the TC, southward track bias, and the resulting large track errors of Typhoon Sinlaku (2002) within the operational National Centers for Environmental Prediction (NCEP) Aviation model. They found that the initial deceleration of the TC was the result of the retreating Pacific subtropical ridge and a deepening midlatitude trough, while the southward track bias was due to the model’s overestimation of the strength of a continental high and underestimation of the strength of the subtropical high. Furthermore, recent case studies of Hurricane Joaquin (2015), which was characterized by unusually large track forecast variability, identified that Joaquin’s track forecasts were highly sensitive to variability in the steering flow within 500–900 km of the TC’s initial position within numerical models (Nystrom et al. 2018; Torn et al. 2018; Alaka et al. 2019).
The ramifications of poor model representation of the steering flow and synoptic features, particularly in the initial conditions, can be severe, especially for TCs that are forecasted to make landfall. For example, official and model guidance track forecasts for Hurricane Rita (2005) at 0000 UTC 20–22 September 2005 showed the storm making landfall in Houston, Texas on 24 September, leading to one of the largest mass evacuation orders in U.S. history. Instead, Hurricane Rita made landfall along the Texas–Louisiana coastal zone, and it was discovered that the operational and model forecast guidance contained a persistent left-of-track bias. Galarneau and Hamill (2015) investigated the factors that contributed to the left-of-track errors and discovered that the track forecast was sensitive to errors in the synoptic-scale flow and the steering-layer depth within the models. While Hurricane Rita may be an extreme example of the socioeconomic impacts of steering-layer and synoptic flow errors in numerical models, this case further demonstrates the need for improvements in difficult TC track forecasts.

One method of improving steering-layer and synoptic flow errors in numerical models is through the assimilation of additional observations in regions that might have a large impact on the track. Efforts have been made to collect additional observations in the synoptic environment surrounding a TC using aircraft reconnaissance. Between 1982–1993, the assimilation of middle- and lower-tropospheric Omega dropwindsondes (ODWs) into numerical models resulted in 16–30% reductions in 12–60-h track forecast errors, which is equivalent to the accumulated improvement in operational track forecast errors seen during the 22-y period from 1970–1991 (Burpee et al. 1996). The results of this study and other similar studies suggested that direct measurements of the near-storm environment had the potential to significantly reduce track forecast errors. Accordingly, these results played a role in the development of a new synoptic surveillance program by the National Oceanic and Atmospheric Administration’s (NOAA) NHC
and Atlantic Oceanographic and Meteorological Laboratory/Hurricane Research Division (AOML/HRD) in 1997, which remains operational as of this thesis.

In 1996, NOAA procured a Gulfstream IV-SP jet aircraft (G-IV) for use in operational synoptic surveillance missions in the near-storm environments of TCs that are potential landfall threat to the continental United States, Puerto Rico, the United States Virgin Islands, and Hawaii. These missions began during the 1997 hurricane season, utilize a new global positioning satellite (GPS) dropwindsonde (Hock and Franklin 1999), and collect profiles of wind, temperature, and humidity from flight-level (13,716 m or 45,000 ft) to the surface. During the first two Atlantic and Eastern Pacific hurricane seasons, there were 24 operational missions (five in 1997, 19 in 1998) that sampled the synoptic environment surrounding a TC. For each mission, Aberson (2002) compared two model runs from three different models, one model run with no assimilation of the dropsonde observations and one model run with all of the dropsonde observations assimilated, to assess the changes in the model track forecasts when dropsonde data is added. They found that the assimilation of the dropsonde data led to small track forecast improvements overall; however, in cases where the storm-motion vector was accurately assessed, model forecasts saw a statistically significant improvement of 14–24%, suggesting that the G-IV missions were utilizing suboptimal sampling and data assimilation procedures. The impacts of the G-IV synoptic surveillance missions were reassessed after the 10th hurricane season that included the G-IV missions as part of NHC operations. Between 1997–2006, 176 G-IV missions were conducted in the Atlantic and Eastern Pacific basins for TCs that were identified as a landfall threat. During this time, several major upgrades to the quality control algorithm and operational models were implemented. Aberson (2010) conducted a similar analysis of the synoptic surveillance missions as Aberson (2002) and found that the G-IV missions led to a 10–15% improvement in 0–60-h track forecast
errors for forecasts initialized at mission nominal times, with the impact of the observations decreasing with increasing lead time.

While the synoptic surveillance data appears to improve TC track forecasts on average, the inclusion of this data does not guarantee that an individual forecast will be improved. For example, during the 2004 and 2005 Atlantic hurricane seasons, there were instances where the assimilation of the G-IV dropsonde data led to large track degradations in the operational models. Aberson (2008) investigated cases from four TCs of large track forecast errors after the assimilation of dropsonde data from using the Global Forecasting System (GFS) model. They found that the track forecast degradations were the result of assimilating erroneous data into the models, issues with the quality control portion of the data assimilation process, and problems with the data assimilation system itself. Given that G-IV missions tend to reduce track forecast errors on average, it is worthwhile to address deficiencies within the G-IV observing system to eliminate the outlier cases and improve the overall statistics of the missions.

One way to address the deficiencies within the G-IV observing system is through the investigation of the optimal targeting strategies utilized when planning the locations of dropsonde deployments. The most optimal targeting strategy would allow for the identification of all sensitive regions in which the assimilation of targeted observations is expected to have the greatest influence on improving track forecasts. Multiple targeting strategies have been devised, such as the ensemble deep-layer mean (DLM) wind variance of the deep-layer steering flow (Aberson 2003), total energy singular vectors (TESVs) that identify where the 48-h track forecast is most sensitive to the initial structure and environment (Peng and Reynolds 2006), adjoint-derived sensitivity steering vectors (ADSSVs) that identify where the environmental steering flow at a particular forecast verifying time is sensitive to the environmental steering flow at the observing time (Wu
et al. 2007), and an ensemble-based sensitivity analysis technique that identifies initial near-storm and environmental features that influence the track forecast at later times (Torn and Hakim 2008). Comparison studies of the different targeting strategies have revealed case-to-case differences between the resultant targeting areas, as well as similarities between certain strategies (e.g., Majumdar et al. 2006; Wu et al. 2009). For example, Wu et al. (2009) compared six different targeting strategies based on 86 cases of two-day forecasts from the northwest Pacific basin in 2006 to identify similarities amongst the sensitivity regions produced by each strategy. They found that TESV-based strategies produce similar targeting areas, the DLM wind variance strategy produces unique targeting areas with respect to the other strategies, and ADSSV and ensemble transform Kalman filter (ETKF) strategies typically produce similar targeting areas, although the sensitivity patterns associated with the ADSSV strategy are more similar to sensitivity patterns produced by the TESV-based strategies than the ETKF strategy. Furthermore, adjoint-based targeting methods that account for data assimilation schemes have been shown to better capture sensitivities associated with synoptic features than ensemble methods (e.g., Wu et al. 2009; Daescu and Todling 2010).

Recently, Ryan et al. (2018) evaluated different targeting configurations within AOML’s high-resolution regional observing system simulation experiment (OSSE) system using multiple sensitivity studies focused on the storm-relative location of dropsondes in the near-TC environment. Multiple G-IV aircraft paths and corresponding dropsonde data were simulated within the OSSE system for a rapidly intensifying TC in the North Atlantic basin over a 13-day lifetime (i.e., a hurricane nature run). The impact of these simulated dropsonde deployments on TC track and intensity forecasts was evaluated relative to a control experiment that represents the current G-IV targeting and assimilation procedures. In all experiments, the assimilation of the
simulated G-IV dropsonde data improved the synoptic environment; therefore, improving the short-term track and intensity forecasts for the hurricane nature run. Most notably, the results of the experiments indicated that the addition of dropsondes in the near-TC environment had the largest impact on the track forecasts. This finding resulted in the implementation of dropsonde releases circumnavigating the TC during G-IV synoptic surveillance missions, which is now the standard in all G-IV missions.

1.3. Research Goals and Thesis Structure

The goal of this thesis is to assess the impact of the dropsonde data collected during G-IV synoptic surveillance missions on the track predictability of recent potential landfalling TCs in the Atlantic basin. While the dropsonde data collected during G-IV missions has been included in studies of all NHC reconnaissance missions, the last study solely focusing on the impacts of G-IV dropsonde data was published in 2010 (Aberson 2010). Unpublished analyses hint that the impact of these dropsondes may not be as high as those found in previous studies (e.g., Aberson 2002, 2003, 2010), potentially due to errors in the environmental and near-TC steering flow, interactions with complex synoptic flow patterns, and improper targeting strategies. At the start of the 2018 hurricane season, dropsonde releases circumnavigating the TC core were added to G-IV operations due to the results of Ryan et al. (2018), which found that dropsondes in the near-TC environment led to the greatest reductions in position errors. Therefore, this research expands on previous studies by quantifying the average track forecast error and skill of TC position forecasts from the European Centre for Medium-Range Weather Forecasts ensemble prediction system (ECMWF EPS) and the NCEP Global Ensemble Forecasting System (GEFS) for all Atlantic basin TCs from
2018–2021, and comparing the results to position forecasts from only the forecasts initialized with dropsonde data from the G-IV synoptic surveillance missions. While the most effective way to do this would be to run simulations where the dropsondes are assimilated into the operational models and compare these forecasts against forecasts where the dropsonde data is not assimilated, this approach is computationally expensive. Thus, the methodology utilized in this thesis provides a means of assessing the impact of the dropsonde data without having to perform computationally expensive experiments.

The track forecasts are divided into two groups: those initialized with and without the assimilation of G-IV dropsondes, and those initialized 12-h before and directly after the assimilation of G-IV dropsondes. It is hypothesized that the assimilation of the dropsonde data reduces the average track errors and adds skill to position forecasts. The potential sources of significant error reductions in individual track forecasts after the assimilation of dropsondes is also investigated using four cases studies of G-IV missions. Large reductions in track error in individual cases are hypothesized to be the result of changes in the environmental steering flow guiding the TC’s motion, a change in the initial position of the TC which leads to a different steering flow regime guiding the TC’s motion, or a combination of both. The remainder of this thesis is organized as follows. The data and methods are described in section 2. The multi-season analysis of TC track errors and skill statistics, and case studies are discussed in section 3. Finally, the results are summarized, and the conclusions of this thesis are presented in section 4.
2. **Data and Methods**

2.1. NOAA G-IV Synoptic Surveillance Missions

This study investigates the impact of dropsonde data collected during NOAA G-IV synoptic surveillance missions on TC track predictability during the 2018–2021 Atlantic basin hurricane season. These missions were chosen for this study because their purpose is to collect additional in-situ observations of the near-TC synoptic environment to reduce track forecast uncertainty of potential landfalling TCs. Therefore, it is worthwhile to quantify the impact of these missions on TC track forecast errors and skill.

In general, these missions are conducted starting two to three days before a TC is projected to make landfall in the United States, Puerto Rico, and the U.S. Virgin Islands, and continue in 12-h intervals as needed. During each mission, 25–30 dropsondes are deployed at 150–200-km intervals as the aircraft circumnavigates the TC and in regions surrounding the TC where the track forecast is sensitive to the environmental conditions. The dropsonde data is analyzed and processed as it is received by the aircraft, then transmitted to global modeling centers to be assimilated into the global and hurricane model runs at 0000 and 1200 UTC (Aberson and Franklin 1999; Aberson 2003). Missions typically begin at either 0530 UTC or 1730 UTC and last roughly 8 h. Dropsonde data from missions that begin at 0530 UTC (1730 UTC) are assimilated into the 1200 UTC (0000 UTC) global model runs. From 2018–2021, 56 NOAA G-IV missions were conducted in the Atlantic Ocean that sampled 14 TCs, while in the Eastern Pacific Ocean there were nine G-IV missions that sampled three TCs. The details of each mission, including take-off and assimilation times, can be found in Table 2.1.
2.2. Analysis of Error and Skill Statistics over Multiple Seasons

2.2.1. Datasets

Four datasets will be used in the analysis of TC track error and skill statistics over multiple seasons: two operational ensemble prediction systems (EPSs) that generate track forecasts, a climatology and persistence model whose track forecasts will be used as a benchmark, and an observational dataset to verify the track forecasts. These datasets are described in the following four sub-sections.

2.2.1.1. ECMWF EPS

The first operational EPS used for the TC track forecasts is the ECMWF EPS. The position forecasts were obtained in BUFR format via FTP (ftp://wmo:essential@diss.ecmwf.int). The ECMWF EPS forecasts consist of 51 members (1 unperturbed control member and 50 perturbed members) at a spatial resolution of approximately 18 km. The model had 91 vertical levels prior to the implementation of version CY47R2 on 11 May 2021, after which the model expanded to 137 vertical levels. Table 2.2 provides a summary of the ECMWF EPS versions in operation during the period of study.

For this study, forecasts initialized at 0000 UTC and 1200 UTC at 12-h forecast intervals through 120-h will be used. The tracker is applied individually to each ensemble member. The ECMWF EPS generates position forecasts at 6-h forecast intervals using a TC tracking algorithm described in Vitart et al. (2012). There are three main steps to the tracker: (1) for each time step at
a resolution of about 100 km, identify any warm-core cyclones; (2) using the full model resolution, identify the surface position and intensity of the warm-core cyclones detected in step 1; and, (3) compile the position sequences for the warm-core cyclones identified in steps 1 and 2. A TC does not have to be present for the entire forecast period for the algorithm to produce a forecast, which allows the track forecasts to capture both the early and later stages of a TC’s life. However, given the low resolution employed in step 1 of the algorithm, small, weak TCs may be missed by the algorithm. Further technical details on the tracking algorithm can be found in Magnusson et al. (2021).

2.2.1.2. GEFS

The second operational EPS used for the TC track forecasts is the NCEP Global Ensemble Forecasting System (GEFS). TC position forecasts for this model were obtained from the publicly accessible NHC FTP data server (https://ftp.nhc.noaa.gov/atcf/archive/). This study uses GEFS forecasts initialized at 0000 UTC and 1200 UTC at 12-h forecast intervals through 120 h. Position forecasts initialized prior to 23 September 2020 used version 11 of the GEFS, which consisted of 21 ensemble members (1 unperturbed control member and 20 perturbed members) at a spatial resolution of approximately 33 km. Forecasts initialized after 23 September 2020 use version 12 of the GEFS. Version 12 increased the spatial resolution to approximately 25 km, increased the ensemble size to 31 members (1 unperturbed control member and 30 perturbed members), and replaced the Global Spectral Model with the global Finite-Volume Cubed-Sphere (FV3) dynamical core (NOAA 2020). Both versions contain 64 vertical levels.
The GEFS uses the NCEP TC tracking algorithm described in Marchok (2002) to obtain TC information from the grids. This tracking algorithm locates a TC center at the time of forecast initialization by first employing a single-pass Barnes analysis of five different primary parameters (minimum sea level pressure, 700- and 850-hPa relative vorticity, 700- and 850-hPa geopotential height) at the original resolution of the model. Then, the Barnes analysis of the primary parameters is repeated four times, with each subsequent analysis done on a grid with half the grid spacing of the previous iteration. A center-guess position is generated from a composite of these analyses, which is then refined by a Barnes analysis of the minimum wind speed at 700 and 850 hPa. For subsequent forecast hours, the center-guess position is calculated through an average of the linear extrapolation of the previous two center fixes and the advection of the current storm according to the composite wind speed of a Barnes analysis of the 850-, 700-, and 500-hPa winds. Throughout this process, various checks are performed to ensure the center position being tracked resembles a TC, and to verify that the center position is likely the storm being targeted and not a close-passing vortex.

2.2.1.3. CLP5

CLIPER5, hereinafter referred to as CLP5, is a climatology and persistence model that provides a five-day TC track forecast based on a least-squares fit to predictors derived from a storm’s past motion, the time of year, and a storm’s motion, location, and intensity at time of initialization (Neumann 1972; Aberson 1998). CLP5 does not contain any information about the current state of the atmosphere; therefore, its track errors to be used as a “no-skill” benchmark for evaluating other forecasts. Because CLP5 uses a storm’s initial location (e.g., the NHC advisory
position) to generate a forecast, forecast hour 0 is close to the best-track position for that time, nearly 0 km. Therefore, in this study, forecast hour 0 will be excluded.

CLP5 track forecasts can also provide insight into the “difficulty” of a track forecast; for example, if CLP5 track errors are unusually low for a particular storm, it can be stated that the storm’s track was “less difficult” to forecast than normal or otherwise did not deviate far from the expected behavior of a TC based on climatology and persistence. The current version of CLP5 and the version used in this study are based on developmental data from 1931–2004 for the Atlantic Ocean (Cangialosi 2022). Twelve-hourly CLP5 forecasts were obtained from the NHC FTP data server (https://ftp.nhc.noaa.gov/atcf/archive/).

2.2.1.4. HURDAT2

The position forecasts from the ECMWF, GEFS, and CLP5 are verified using the second-generation hurricane database (HURDAT2) (Landsea and Franklin 2013). These best tracks represent the official historical record for all tropical and subtropical storms included in the database. HURDAT2 contains the NHC “best tracks” for intensity, central pressure, position, and size of all tropical and subtropical cyclones in the Atlantic Ocean dating back to 1851. Position information is available at every synoptic time (0000, 0600, 1200, and 1800 UTC); however, this study will only be using the 0000 UTC and 1200 UTC information. The best track positions were obtained from the dataset available at https://www.nhc.noaa.gov/data/#hurdat.
2.2.2. Methodology

The following section describes the rules employed in this study to select the track forecast initialization times. These rules generally follow the methods the NHC uses in their post-season forecast verification reports (e.g., Cangialosi 2022).

2.2.2.1. Selection of Track Forecasts

All tropical and subtropical cyclones, hereinafter referred to as ‘TCs’, that reached tropical depression strength or greater in the Atlantic Ocean from 2018–2021 are considered. While there were nine NOAA G-IV reconnaissance missions conducted in the Eastern Pacific Ocean in 2018, this is too small of a sample to draw any robust conclusions. Further, these nine missions cannot be added to the Atlantic Ocean sample, as there are too many synoptic-level steering flow differences, amongst other factors, that lead to climatology and persistence being dissimilar between the Atlantic and Eastern Pacific Oceans. Therefore, this analysis will only include Atlantic Ocean TCs.

Further criteria are applied to the TCs in the sample to identify valid forecast initialization times for use in the multi-season analysis. Only times where a storm was classified as a TC (i.e., subtropical depression (SD), subtropical storm (SS), tropical depression (TD), tropical storm (TS), and hurricane (HU)) in the HURDAT2 will be considered. Forecasts are initialized at 0000 UTC and 1200 UTC starting with the first instance that a storm is classified as a TC. Forecasts continue to be utilized for every 12 h the storm remains a TC, with a maximum of a 120-h forecast. A storm must remain a TC through at least forecast hour 12, or else the forecast initialization time is
excluded as there are no valid times to compare two forecasts beyond the initialization time. Additionally, a position forecast from the ECMWF EPS, GEFS, and CLP5 must all exist for the initialization time to be counted. Occasionally, a forecast initialization time for a weak storm early in its lifetime, or a weak or transitioning storm later in its lifetime, is not picked up by one or both of the ECMWF EPS and GEFS trackers and, thus, no forecast is produced. Further, at least half of the ECMWF EPS and GEFS ensemble members must produce a forecast for the initialization time to count. These restrictions resulted in a sample of 675 initialization times across 73 named TCs from 2018–2021. Of these 675 initialization times, 56 had dropsonde data assimilated at the time of initialization. Table 2.3 shows the distribution of valid initialization times by year.

2.2.2.2. Subsets of Track Forecast Initialization Times

The error and skill analysis will consist of two subsets of the track forecast initialization times, allowing for multiple ways of assessing the impact of NOAA G-IV dropsonde data on position forecasts. First, forecasts initialized with the assimilation of dropsonde data (56 initialization times with dropsondes) and forecasts initialized without the assimilated of dropsonde data (619 initialization times without dropsondes). Dividing the forecast initialization times by those with and without dropsondes allows for the error and skill of dropsonde cases to be analyzed in comparison to all position forecasts. Second, forecasts initialized with the assimilation of dropsonde data (56 initialization times after dropsondes are assimilated) and the forecast initialization 12 h before the assimilation of dropsonde data (56 initialization times before dropsondes are assimilated). By analyzing only the forecast initializations 12 h before and after the dropsonde data is assimilated, the impact on track error and skill for TCs that are sampled by
the G-IV missions can be assessed. While the best method to assess the impacts of the dropsonde data would be to run the operational ensembles with and without the assimilation of the dropsonde data, this method is quite computationally expensive. A less computationally expensive way is to analyze the model output of the forecasts initialized before and after the dropsonde data is assimilated and compare the forecast; therefore, this is the method utilized in this thesis.

2.2.3. Calculations

2.2.3.1. Track Forecast Errors

Given that the purpose of the reconnaissance missions is to reduce TC track forecast uncertainty, track forecast error is a useful metric to assess the impact of dropsonde data on position forecasts. The track forecast error is computed for the ECMWF EPS, GEFS, and CLP5 position forecasts. Track forecast error is defined as the great-circle distance between the HURDAT2 best track position and a forecast position at the forecast verification time (Cangialosi 2022). The great-circle distance is computed using the following equation:

\[ d = a \cos^{-1} [\cos(\phi_1) \cos(\phi_2) \cos(\lambda_1 - \lambda_2) + \sin \phi_1 \sin \phi_2] \]  

where \( d \) is the great-circle distance in kilometers, \( a \) is the radius of the earth (6371 km), \( \phi_1 \) and \( \phi_2 \) are the latitude and longitude of the best track position, and \( \phi_2 \) and \( \lambda_2 \) are the latitude and longitude of the forecast position. For the ECMWF EPS and GEFS forecasts, the mean forecast latitude and longitude is calculated then used within the great-circle distance equation to compute the ensemble mean track forecast error.
2.2.3.2. Track Forecast Skill

One method of assessing the impact of the assimilation of reconnaissance dropsonde data on operational EPS track forecasts is to quantitatively assess the degree of skill of these forecasts and compare the result to the skill of forecasts initialized without the assimilation of dropsonde data. Forecast skill represents a normalization of forecast error against a benchmark and is expressed as a percentage improvement over the benchmark (Cangialosi 2022). For TC track forecast skill, CLP5 forecast errors are commonly employed as the benchmark and will be used in this study to assess the degree of skill of the operational EPS track forecasts. The track forecast skill relative to CLP5 \((RS)\) is given by:

\[
RS = \frac{100(CE - FE)}{CE}
\]

where \(FE\) is operational ensemble forecast error and \(CE\) is the CLP5 forecast error for forecasts issued at the same initialization time for the same valid time. Skillful forecasts are indicated by a positive \(RS\), which occurs when the operational ensemble forecast error is less than the CLP5 forecast error. Skill against the benchmark is considered higher as \(RS\) becomes closer to 100%.

The ensemble mean track forecast skill is computed at each forecast hour using the ensemble mean track forecast error.

2.2.3.3. Frequency of Superior Performance

The frequency of superior performance (FSP) is a useful metric to analyze the subset of track forecast initialization times that compare forecasts 12 h before and after dropsonde data is assimilated into numerical models. The FSP compares two forecasts at different lead times for the
same valid time and indicates the frequency in which one forecast has a lower track error than the other. Therefore, this calculation can be used to help assess the impact of the assimilation of reconnaissance dropsonde data on operational ensemble TC track forecasts because it allows us to compare forecasts initialized at different times that contain forecasts for the same valid time.

The FSP is determined by comparing the ensemble mean track forecast errors before and after forecast initialization times at the same valid time. For each valid time, the ensemble mean track forecast errors of the before and after forecasts are compared, and a point is given to the forecast which has the lower error. Both initialization times must have a forecast for the valid time that is being assessed to count. The result is expressed as the percentage of the forecast initialization times where the after forecast initializations had lower track forecast errors than the before forecast initializations.

2.2.4. Bootstrap Resampling

Of particular importance to the thesis is to assess whether there is a statistically significant difference in track forecast error and skill with the inclusion of reconnaissance dropsonde data in the operational ensembles. To accomplish this, a bootstrap resampling method without replacement is conducted for both subsets of forecast initialization times. In each of the 10,000 repetitions, 56 forecast initialization times are randomly selected from a pool of all forecast initialization times. For the “with vs without dropsondes” subset, this pool of forecast initialization times consists of all 675 valid forecast initialization times, regardless of whether dropsondes were assimilated. For the “before vs after dropsondes” subset, this pool of forecast initialization times consists of 601 forecast initialization times from the entire sample of 675. These 601 forecast
initialization times are those within the entire sample where the forecast initialization time 12 h prior is also within the entire sample, therefore creating 601 pairs of “before vs after all” forecast initializations. This pairing will result in a resampled dataset that has a “before” and an “after” forecast initialization, similar to the “before vs after dropsondes” subset. The mean track forecast error and mean track forecast skill are calculated for each set of 56 forecast initialization times in the resampled datasets, resulting in resampled datasets containing 10,000 mean track forecast error values and 10,000 mean track forecast skill values. For the “before vs after dropsondes” subset, the difference between the pairs of “before” and “after” forecast initializations is found, resulting in resampled datasets containing 10,000 differences in mean track forecast error values and 10,000 differences in mean track forecast skill values. The difference between the “before” and “after” forecast initialization times is considered because it allows us to quantify the impact of the dropsonde data compared to the forecast initialization time 12 h before. Lastly, statistical significance is determined by finding the percentile of the resampled dataset where the mean value of the 56 dropsonde initialization times would lie if the dropsonde value was in the resampled dataset. The percentile represents the randomness of the dropsonde forecast initialization times, such that a percentile close to 0% or 100% indicates statistical significance (i.e., the mean dropsonde value is not due to random chance), depending on the calculation.

2.3. Case Studies

Four case studies of NOAA G-IV reconnaissance missions will be conducted in this thesis to investigate the impacts of the dropsonde data on TC track forecasts and to attempt to source the
cause of track shifts after the dropsonde data is assimilated into the operational ensembles. The four cases studies, described by the mission take-off date and time, are described in Table 2.4.

2.3.1. Datasets

The case studies conducted in this thesis will utilize the previously described ECMWF EPS and CLP5 position forecasts, and HURDAT2 best track dataset. While the ECMWF EPS position forecasts will still be the result of its TC tracker, other forecast fields will be used in the steering flow analysis of the case studies. These additional forecast fields are obtained from The Observing System Research and Predictability Experiment (THORPEX) Interactive Global Grand Ensemble (TIGGE; Bougeault et al. 2010) archive located at ECMWF (http://apps.ecmwf.int/datasets/data/tigge) on a 0.5° grid. These additional forecast fields will be described in further detail in section 2.3.2.2.

For each case study, two forecast initialization times are used: the forecast initialization time coinciding with the assimilation of the dropsonde data into the operational ensembles (i.e., the after dropsonde forecast), and the forecast initialization time 12 h prior to the assimilation of the dropsonde data (i.e., the before dropsonde forecast).
2.3.2. Methodology

2.3.2.1. Track Forecast Errors

The track forecast errors will be computed for the ECMWF EPS and CLP5 position forecasts using the same calculation described in section 2.2.3.1. Additionally, the distance between the ensemble mean position for both forecast initialization times is computed at all forecast hours.

2.3.2.2. Environmental Steering Flow

The environmental steering flow is evaluated for all ECMWF EPS members as a means of identifying the source of track shifts seen after the dropsonde data is assimilated into the operational ensembles following the methodology introduced in Galarneau and Davis (2013). This technique determines the optimal steering flow for a TC, or the vertical layer and radius most like the TC motion within ±12 h of a particular time. First, the TC motion is determined using the position information from the HURDAT2 best track dataset. Next, the TC vortex is separated from the environmental wind field in a multi-step process. In this process, vorticity and divergence are calculated at all vertical levels between 850–200 hPa and at a given radius, then the Poisson equation is applied to determine the associated velocity potential and streamfunction. The solutions for the velocity potential and streamfunction are used to determine the nondivergent and irrotational wind vectors associated with the TC vortex. The last step of the TC removal process is to subtract the nondivergent and irrotational wind vectors from the total wind, resulting in the
background flow in which the TC is embedded. Then, the mean of the remaining background flow is computed over a range of radii combinations within 1°–8° of the TC center and over a set of vertical layers beginning at 850 hPa. Each radius and vertical layer combination is compared to the observed motion, and the radius and vertical layer combination with the lowest motion error is chosen as the optimal steering layer. For the purposes of this study, the optimal steering layer is assumed to be the same for all ensemble members and all forecast initialization times within each individual case. Lastly, the ensemble mean $u$ and $v$ components of the environmental steering flow are computed at every grid point for all times in every forecast by subtracting the TC circulation. The optimal steering layer used for each case study can be found in Table 2.5.

The magnitude of the near-TC environmental steering flow is a useful metric for determining whether there was a change in the environmental steering flow in the forecast initialization after the dropsondes are assimilated. The near-TC steering flow at each forecast hour is calculated from the ensemble mean environmental steering flow at that time. The average $u$ and $v$ components of the environmental steering flow within 100-km of the TC center are computed and result in the near-TC steering flow at that time. The change in near-TC steering flow between the forecast initialization before and with dropsonde data assimilated is found by taking the magnitude of the vector difference of the near-TC steering flow at each valid time in the forecast, via:

$$\Delta SF = \sqrt{(u_2 - u_1)^2 + (v_2 - v_1)^2}$$

(3)

where $\Delta SF$ is the magnitude of the steering flow difference, $u_1$ and $u_2$ are the $u$ components of the near-TC steering flow before and after the dropsonde data is assimilated, respectively, and $v_1$ and $v_2$ are the $v$ components of the near-TC steering flow before and after the dropsonde data is assimilated, respectively.
2.3.2.3. Major Axis Correlation Coefficient

The previous components of the case study analysis work together to diagnose the potential source of position forecast shifts after the assimilation of dropsonde data into the operational ensembles. The final component of the case study, the major axis correlation coefficient (MACC) computation, provides quantitative support to the results of the previous components. The MACC quantifies to what extent the position at the final forecast hour is dependent on the position at earlier times in the forecast. First, a position ellipse is generated from a bivariate normal distribution of the forecast positions of all ensemble members at each forecast hour, as in Hamill et al. (2011). The ellipse is centered on the ensemble mean forecast position and encloses 90% of the probability of the forecast position at a given forecast hour, such that the major axis of the ellipse is in the direction of the largest position forecast variability. Next, the distance between the ensemble mean and each ensemble member forecast position along the major axis of the ellipse is computed at each forecast hour (e.g., Figure 2.1). Lastly, the MACC is computed by taking the Pearson correlation coefficient between the major axis distances at the final forecast hour and the major axis distances at every previous forecast hour. A high MACC value at a given forecast hour indicates that the position at the final forecast hour does depend on the position at that previous time in the forecast. Furthermore, consecutive high MACC values indicate where the position variability at the final forecast hour is solidified (i.e., if the position forecast at the final forecast hour is a cake, consecutive high MACC is where “the cake is already baked”). The statistical significance of the MACC value at each forecast hour is determined by the corresponding p-value. This study utilizes a p-value that determines the statistical significance at the 95% level.
Fig. 2.1. An example schematic for the position ellipses and associated major axis used in the MACC calculation. The dots indicate the forecasted position of each ensemble member at 24-h intervals, while the triangles indicate the forecasted ensemble mean position at 24-h intervals. The ellipses show the bivariate normal distribution of the forecast positions at 24-h intervals, as in Hamill et al. (2011). Purple denotes the 24-h locations, cyan denotes 48-h locations, green denotes 72-h locations, red denotes 96-h locations, and magenta denotes 120-h locations. The red arrows indicate the ellipse major axis each 24 h. The black arrows indicate the great circle distance between the ensemble mean position and an example ensemble member position each 24 h. The blue arrows indicate the distance along the major axis, oriented from the ensemble mean position toward an example ensemble member each 24 h.
Table 2.1. The basin, year, storm name, take-off time, and assimilation time for every NOAA G-IV synoptic surveillance reconnaissance mission from 2018–2021.

<table>
<thead>
<tr>
<th>Basin</th>
<th>Year</th>
<th>Storm Name</th>
<th>Take-off Time</th>
<th>Assimilation Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>2018</td>
<td>Florence</td>
<td>1730 UTC 8 September</td>
<td>0000 UTC 09 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 9 September</td>
<td>0000 UTC 10 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 10 September</td>
<td>1200 UTC 10 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 10 September</td>
<td>0000 UTC 11 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 11 September</td>
<td>1200 UTC 11 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 11 September</td>
<td>0000 UTC 12 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 12 September</td>
<td>1200 UTC 12 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 12 September</td>
<td>0000 UTC 13 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 13 September</td>
<td>0000 UTC 14 September</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2018</td>
<td>Michael</td>
<td>1730 UTC 08 October</td>
<td>0000 UTC 09 October</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 09 October</td>
<td>1200 UTC 09 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2019</td>
<td>Dorian</td>
<td>1730 UTC 25 August</td>
<td>0000 UTC 26 August</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 26 August</td>
<td>0000 UTC 27 August</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 27 August</td>
<td>0000 UTC 28 August</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 28 August</td>
<td>0000 UTC 29 August</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 29 August</td>
<td>0000 UTC 30 August</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 31 August</td>
<td>0000 UTC 01 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 01 September</td>
<td>1200 UTC 01 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 02 September</td>
<td>1200 UTC 02 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 03 September</td>
<td>1200 UTC 03 September</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 04 September</td>
<td>1200 UTC 04 September</td>
</tr>
<tr>
<td>Region</td>
<td>Year</td>
<td>Name</td>
<td>Start Time</td>
<td>End Time</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td>---------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2019</td>
<td>Jerry</td>
<td>1730 UTC 19 September</td>
<td>0000 UTC 20 September</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Isaias</td>
<td>1730 UTC 31 July</td>
<td>0000 UTC 01 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Laura</td>
<td>1730 UTC 01 August</td>
<td>0000 UTC 02 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>1730 UTC 21 August</td>
<td>0000 UTC 22 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>1730 UTC 22 August</td>
<td>0000 UTC 23 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 24 August</td>
<td>1200 UTC 24 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>1730 UTC 24 August</td>
<td>0000 UTC 25 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 25 August</td>
<td>1200 UTC 25 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Marco</td>
<td>0530 UTC 22 August</td>
<td>1200 UTC 22 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 23 August</td>
<td>1200 UTC 23 August</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Delta</td>
<td>1730 UTC 06 October</td>
<td>0000 UTC 07 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 07 October</td>
<td>1200 UTC 07 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>1730 UTC 07 October</td>
<td>0000 UTC 08 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 08 October</td>
<td>1200 UTC 08 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Zeta</td>
<td>1730 UTC 25 October</td>
<td>0000 UTC 26 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 26 October</td>
<td>1200 UTC 26 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>1700 UTC 26 October</td>
<td>0000 UTC 27 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td>Eta</td>
<td>0530 UTC 27 October</td>
<td>1200 UTC 27 October</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 07 November</td>
<td>1200 UTC 07 November</td>
</tr>
<tr>
<td>Atlantic</td>
<td>2020</td>
<td></td>
<td>0530 UTC 08 November</td>
<td>1200 UTC 08 November</td>
</tr>
<tr>
<td>Region</td>
<td>Year</td>
<td>Storm</td>
<td>Dates</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>-------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Atlantic 2021</td>
<td></td>
<td>Elsa</td>
<td>0530 UTC 10 November 1730 UTC 02 July 0000 UTC 03 July</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 03 July 0000 UTC 04 July</td>
<td></td>
</tr>
<tr>
<td>Atlantic 2021</td>
<td></td>
<td>Henri</td>
<td>1730 UTC 19 August 0000 UTC 20 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 20 August 1200 UTC 20 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 20 August 0000 UTC 21 August</td>
<td></td>
</tr>
<tr>
<td>Atlantic 2021</td>
<td></td>
<td>Ida</td>
<td>0530 UTC 27 August 1200 UTC 27 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 27 August 0000 UTC 28 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0530 UTC 28 August 1200 UTC 28 August</td>
<td></td>
</tr>
<tr>
<td>Atlantic 2021</td>
<td></td>
<td>Sam</td>
<td>1730 UTC 26 September 0000 UTC 27 September</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 27 September 0000 UTC 28 September</td>
<td></td>
</tr>
<tr>
<td>Eastern Pacific 2018</td>
<td></td>
<td>Hector</td>
<td>1730 UTC 05 August 0000 UTC 06 August</td>
<td></td>
</tr>
<tr>
<td>Eastern Pacific 2018</td>
<td></td>
<td></td>
<td>1730 UTC 06 August 0000 UTC 07 August</td>
<td></td>
</tr>
<tr>
<td>Eastern Pacific 2018</td>
<td></td>
<td></td>
<td>1730 UTC 07 August 0000 UTC 08 August</td>
<td></td>
</tr>
<tr>
<td>Eastern Pacific 2018</td>
<td></td>
<td>Lane</td>
<td>1730 UTC 19 August 0000 UTC 20 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 20 August 0000 UTC 21 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 21 August 0000 UTC 22 August</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 22 August 0000 UTC 23 August</td>
<td></td>
</tr>
<tr>
<td>Eastern Pacific 2018</td>
<td></td>
<td>Norman</td>
<td>1730 UTC 04 September 0000 UTC 05 September</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1730 UTC 05 September 0000 UTC 06 September</td>
<td></td>
</tr>
</tbody>
</table>

31
Table 2.2. Operational ECMWF EPS versions during the period of study, the date each version was implemented, and any notable changes from the previous version.

<table>
<thead>
<tr>
<th>Date Ranges</th>
<th>ECMWF EPS Version</th>
<th>Implementation Date</th>
<th>Notable Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 July 2017 – 04 June 2018</td>
<td>CY43R3</td>
<td>11 July 2017</td>
<td></td>
</tr>
<tr>
<td>05 June 2018 – 10 June 2019</td>
<td>CY45R1</td>
<td>05 June 2018</td>
<td></td>
</tr>
<tr>
<td>11 June 2019 – 29 June 2020</td>
<td>CY46R1</td>
<td>11 June 2019</td>
<td></td>
</tr>
<tr>
<td>30 June 2020 – 10 May 2021</td>
<td>CY47R1</td>
<td>30 June 2020</td>
<td></td>
</tr>
<tr>
<td>11 May 2021 – 11 October 2021</td>
<td>CY47R2</td>
<td>11 May 2021</td>
<td>Number of vertical levels increased from 91 to 137</td>
</tr>
</tbody>
</table>

Table 2.3. Number of named TCs in the Atlantic basin that had at least one valid forecast initialization time and the total number of valid forecast initialization times per year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Named TCs With at Least One Valid Forecast Initialization Time</th>
<th>Total Number of Valid Forecast Initialization Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>14</td>
<td>157</td>
</tr>
<tr>
<td>2019</td>
<td>15</td>
<td>130</td>
</tr>
<tr>
<td>2020</td>
<td>27</td>
<td>224</td>
</tr>
<tr>
<td>2021</td>
<td>17</td>
<td>164</td>
</tr>
</tbody>
</table>
Table 2.4. The name, mission take-off time, and mission date of each case study.

<table>
<thead>
<tr>
<th>Tropical Cyclone</th>
<th>Mission Take-Off Time</th>
<th>Mission Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hurricane Marco (2020)</td>
<td>0530 UTC</td>
<td>22 August 2020</td>
</tr>
<tr>
<td>Hurricane Zeta (2020)</td>
<td>1730 UTC</td>
<td>25 October 2020</td>
</tr>
<tr>
<td>Hurricane Dorian (2019)</td>
<td>1730 UTC</td>
<td>27 August 2019</td>
</tr>
<tr>
<td>Hurricane Jerry (2019)</td>
<td>1730 UTC</td>
<td>20 September 2019</td>
</tr>
</tbody>
</table>

Table 2.5. Tropical cyclone name and optimal steering flow layer and radius used for each case.

<table>
<thead>
<tr>
<th>Tropical Cyclone</th>
<th>Steering Layer (hPa)</th>
<th>Radius (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hurricane Marco (2020)</td>
<td>500–850</td>
<td>333</td>
</tr>
<tr>
<td>Hurricane Zeta (2020)</td>
<td>300–850</td>
<td>333</td>
</tr>
<tr>
<td>Hurricane Dorian (2019)</td>
<td>300–850</td>
<td>333</td>
</tr>
<tr>
<td>Hurricane Jerry (2019)</td>
<td>300–850</td>
<td>333</td>
</tr>
</tbody>
</table>
3. Results

3.1. Analysis of Error and Skill Statistics over Multiple Seasons

Previous studies have quantified the impact of dropsonde data collected during NOAA G-IV synoptic surveillance missions on TC track forecasts (e.g., Aberson 2002, 2008, 2010); however, there have not been any published analyses on this topic since 2010. Since the last published analysis regarding the impacts of G-IV dropsonde data on TC track forecasts, multiple updates have been made to the G-IV observing system and targeting strategies, including the addition of dropsonde releases circumnavigating the TC core to target the near-TC environment. Therefore, it is worthwhile to perform an updated analysis of the impacts of these dropsondes on TC position forecasts. This thesis will utilize a modified version of the methodology presented in these previous studies, as it is computationally expensive to run the operational models with and without the assimilation of the dropsonde data. In this modified methodology, the model output of the forecasts initialized without and 12 h before the dropsonde data is assimilated are compared to the model output of the forecasts with the dropsonde data.

Instead of analyzing just the TCs that were sampled by the G-IV aircraft, this thesis will utilize forecasts from all Atlantic basin TCs from 2018–2021 to determine the impacts of G-IV missions on position forecasts. It is often assumed that forecast times with high CLP5 position errors (i.e., a difficult forecast) are associated with high position errors in the operational models. Therefore, if a single operational forecast has high position errors, it is commonly assumed that the TC had a difficult track forecast. For example, as it relates to this thesis, if a forecast initialization time where dropsonde data was assimilated had high position errors in the operational
models, then it is commonly assumed that the entire track forecast for that TC was difficult. This assumption is investigated by comparing the CLP5 track forecast errors at forecast hour 72 to the GEFS and ECMWF EPS track forecast errors at forecast hour 72 for all forecast initializations in the climatology. The CLP5 track forecast errors are divided into five bins based on error percentiles. Then, the average of the corresponding GEFS and ECMWF EPS position errors in each bin is found (Figure 3.1). As the CLP5 track forecast errors increase, the mean GEFS and ECMWF EPS error within each bin remains around 200 km; thus, it does not appear that cases with higher CLP5 errors have higher operational model errors. Therefore, the CLP5 errors do not appear to determine the difficulty of an individual track forecast. Instead, conclusions about the difficulty of a track forecast with dropsonde data assimilated need to be drawn based on a seasonal or multi-forecast basis. This result supports the multi-seasonal analysis approach for investigating the impacts of dropsonde data from NOAA G-IV reconnaissance missions on TC track forecasts.

3.1.1. With vs Without Dropsondes

In order to assess the impact of assimilating the in-situ data collected by dropsondes deployed during NOAA G-IV synoptic surveillance missions on TC track forecasts, an analysis of track error and skill statistics was performed over multiple seasons for all forecasts. As discussed in sections 2.2.2.1 and 2.2.2.2, all valid forecast initialization times for Atlantic TCs occurring from 2018–2021 were included in the analysis, separated here by the 56 forecast initialization times where dropsondes are assimilated and the 619 forecast initialization times where dropsondes are not assimilated. To deduce the significance of the results of the with and without dropsonde initialization times, a bootstrap resampling without replacement was performed.
Average track forecast errors for the set with dropsonde initialization times have greater CLP5 position errors and lower GEFS and ECMWF EPS position errors than the without dropsonde initialization times. Figure 3.2 shows the average of the GEFS, ECMWF EPS, and CLP5 mean track forecast errors for the 56 forecast initializations with dropsondes and the 619 forecast initializations without dropsondes at 12-h intervals. The number of forecasts included in the average error at each forecast hour decreases throughout the forecast period for both subsets. GEFS and ECMWF EPS position errors for the with and without dropsonde initialization times are very similar through forecast hour 84, though the average error of the dropsonde initialization times is always lower than the without dropsonde initialization times (Figure 3.2a). Both GEFS and ECMWF EPS forecasts without dropsondes have an average error of 45 km at forecast hour 12 which increases to 384 km and 363 km, respectively, by the end of the forecast period. Similarly, GEFS and ECMWF EPS forecasts with dropsondes have an average error of 31 km at forecast hour 12 which increases more slowly than the forecasts without dropsondes to 283 km and 213 km, respectively. While there is a reduction in average track error at all forecast hours in both the GEFS and ECMWF EPS forecast initializations with dropsondes compared to those without dropsondes, the amount of error reduction in both models increases at the end of the forecast period. Through forecast hour 84, the average error of forecasts with dropsondes is around 25 km lower than the average error of forecasts without dropsondes. After forecast hour 84, the difference in average error between the forecasts with and without dropsondes increases, reaching around 100 km in the GEFS and 150 km in the ECMWF EPS. However, this increase in error reduction may be due in part to decreased numbers of forecasts in later times in the forecast period. Despite this, the error reduction in ECMWF EPS forecasts with dropsondes is statistically significant at the 95th percentile at forecast hours 12, 96, 108, and 120, while the error reduction
in GEFS forecasts with dropsondes is statistically significant at the 95\textsuperscript{th} percentile at forecast hours 12, 24, and 36. This result indicates that the error reduction at these forecast hours is not due to random chance, as fewer than 5\% of randomly sampled forecast initializations from climatology have lower track errors at these forecast hours compared to forecast initializations with dropsondes. Further, for GEFS forecasts at forecast hours 48, 96, and 120 and ECMWF EPS forecasts at forecast hour 24, fewer than 8\% of randomly sampled forecast initializations from climatology have lower track errors compared to forecasts with dropsondes. This result supports the hypothesis that forecast initializations with dropsondes have lower average track errors than those without dropsondes, although it is only statistically significant at forecast hour 12 in both models and forecast hours 96, 108, and 120 in the ECMWF EPS.

While GEFS and ECMWF EPS position errors are lower during dropsonde initialization times, CLP5 position errors increased compared to the without dropsonde initialization times (Figure 3.2b). CLP5 position errors for both sets of forecast initializations rapidly increase during the forecast period; however, forecast initializations with dropsondes increase at a faster rate, reaching an average error of 1136 km (around 200-km greater than without dropsonde forecasts) by the end of the forecast period. This result indicates that track forecasts with dropsondes may be inherently more difficult than those without dropsondes. However, the bootstrap resample reveals that this increase in average error is only statistically significant at the 95\textsuperscript{th} percentile at forecast hour 48, as 96\% of randomly sampled forecast initializations from climatology have lower track errors compared to forecast initializations with dropsondes. Furthermore, after forecast hour 48, 84–94\% of randomly sampled forecast initializations have lower track errors than forecast initializations with dropsondes. Thus, track forecasts with dropsondes are likely more difficult than
those without dropsondes after forecast hour 48, although this result is only statistically significant at forecast hour 48.

The impact of reconnaissance dropsonde data on operational EPS track forecasts can be further assessed by quantifying the degree of skill of track forecasts with dropsondes against a benchmark track forecast (i.e., CLP5) and comparing the result to that of forecasts without dropsondes. Despite the similarity of the average position errors for GEFS and ECMWF EPS forecast initializations with and without dropsondes, there appears to be an increase in both GEFS and ECMWF EPS track forecast skill relative to CLP5 in forecast initializations with dropsondes. Figure 3.3 shows the average GEFS and ECMWF EPS track forecast skill for the 56 forecast initializations with dropsondes and the 619 forecast initializations without dropsondes at 12-h intervals. At forecast hour 12, GEFS forecasts without dropsondes have an average skill relative to CLP5 of 46% (i.e., GEFS mean track forecast errors are 46% lower than CLP5 track forecast errors). Similarly, at forecast hour 12, ECMWF EPS forecasts without dropsondes exhibit a 50% improvement in skill over CLP5. Both models reach their maximum improvement in mean track forecast errors of 69% at forecast hour 60 before decreasing to a 58% improvement (GEFS) and 60% improvement (ECMWF EPS) by forecast hour 120. The same general pattern is seen in the forecast initializations with dropsondes, although the improvement in skill is higher. At forecast hour 12, GEFS forecasts with dropsondes have a 58% improvement in mean track forecast errors against CLP5 while ECMWF EPS forecasts with dropsondes have a 57% improvement. At forecast hour 48, GEFS forecasts with dropsondes reach their maximum of a 79% improvement over CLP5 — 10% greater than the maximum skill of GEFS forecasts with dropsondes. The skill of ECMWF EPS forecasts with dropsondes increases throughout the forecast period, reaching its maximum of 81% at forecast hour 120 — 12% greater than the maximum skill of ECMWF EPS
forecasts without dropsondes. Generally, from 2018–2021, both operational models exhibit a 10% increase in skill in forecasts with dropsondes, suggesting that dropsondes are adding skill to TC track forecasts. The increase in skill seen in GEFS forecasts with dropsondes is significantly significant at the 95th percentile at all forecast hours and is unlikely to be a random occurrence. Meanwhile, the increase in skill of ECMWF EPS forecasts with dropsondes is statistically significant at the 95th percentile at forecast hours 48, 96, 108, and 120. At all other forecast hours, 79–93% of randomly selected forecasts from climatology have less skill than ECMWF EPS forecasts with dropsondes. This result supports the hypothesis that forecast initializations with dropsondes have a higher degree of skill than those without dropsondes, indicating that dropsonde data from NOAA G-IV missions are likely adding skill to TC track forecasts.

3.1.2. Before vs After Dropsondes

As discussed in section 2.2.2.2, track error and skill statistics over multiple seasons are also analyzed for forecast initializations 12 h before and after the dropsonde data is assimilated into the operational models. Comparing the track error and skill of forecasts 12 h before dropsondes are assimilated to forecasts after dropsondes are assimilated allows the impact of the dropsondes to be quantified via the change in track error and skill between the two forecasts. Statistical significance of the differences in track forecast error and skill before and after dropsondes is determined using bootstrap resampling without replacement. Because this analysis focuses on the differences in forecast track error before and after dropsondes, the differences will be computed with respect to the same valid time. Therefore, statistical significance will be determined for forecasts valid at the same time, not for the same forecast hour as done in the previous section.
The GEFS and ECMWF EPS average track forecast error before and after dropsondes is very similar at all forecast hours; however, with respect to forecasts initialized 12 h before dropsondes, there is a reduction in average track error in forecasts after dropsondes at all valid times. Figure 3.4 shows the average GEFS, ECMWF, and CLP5 track forecast errors for the forecast initialization times 12 h before dropsonde data was assimilated and the forecast initialization times after dropsonde data was assimilated (Figure 3.4a, c), as well as the difference in average track error at each valid time in the forecast (Figure 3.4b, d). The GEFS and ECMWF EPS average errors before and after dropsondes are both around 37 km at forecast hour 12 and increase at a similar rate throughout the forecast period, never deviating more than 55 km from each other (Figure 3.4a). This result indicates that there are minimal reductions in average error in both operational models between the forecast initializations 12 h before dropsondes and the forecast initializations after dropsondes at all forecast hours. On the other hand, there is a reduction in average error in GEFS and ECMWF EPS forecasts after dropsondes with respect to forecasts initialized 12 h before dropsondes at all valid times in the forecast (Figure 3.4b). The average track errors for GEFS and ECMWF EPS forecasts valid 12 h after the after dropsonde forecast is initialized is 15–20 km lower in forecasts after dropsondes compared to forecasts before dropsondes. By the 120-h forecast, average track errors are 43 km lower in forecasts after dropsondes compared to forecasts 12 h before dropsondes. Nevertheless, the differences in average error of dropsonde forecast initializations are similar to those found in randomly sampled forecasts, as 13–89% of randomly selected forecasts have a greater reduction in average position error than forecasts with dropsondes. Therefore, the reduction in average error in GEFS and ECMWF EPS forecasts after dropsondes are assimilated is not statistically significant and is likely due to random chance.
Similar results are seen for CLP5 forecasts initialized before and after dropsondes are assimilated (Figure 3.4b, d). CLP5 average track forecast errors at the same lead time are slightly greater in forecast initializations after dropsonde data is assimilated compared to those 12 h before dropsondes (Figure 3.4b). The average track errors of both forecasts are around 70 km at forecast hour 12, and rise at similar rates until forecast hour 84, when after dropsonde forecast errors increase at a slightly greater rate. This result suggests that there are minimal increases in average track forecast error in forecasts initialized after dropsondes are assimilated at all forecast hours. The difference in CLP5 average track forecast error at a given valid time becomes less negative throughout the forecast period, suggesting that the improvements in forecasts initialized after dropsondes are assimilated reduce with increasing lead time (Figure 3.4d); however, these results are not statistically significant, and are likely due to random chance.

The difference in average track forecast error for forecasts initialized at different times that contain forecasts for the same valid time can be more succinctly shown using the frequency of superior performance (FSP) metric. The FSP quantifies the frequency in which one forecast initialization time has lower position errors than another initialization time by comparing position errors for the same valid time. Thus, the FSP of the after dropsonde forecast initializations will provide insight into the frequency the after dropsonde forecasts have lower track errors than the forecast initializations 12 h before dropsondes.

Figure 3.5 shows the FSP for the GEFS, ECMWF EPS, and CLP5 forecasts initialized after dropsondes are assimilated and the 601 “after” forecast initializations from the “before vs after all” subset of forecast initializations. The FSP for all three models is greatest at earlier lead times, then decreases with increasing lead time. ECMWF EPS after dropsonde forecasts had a maximum FSP for forecasts valid 12 h after the dropsondes were assimilated, such that the after dropsonde
forecasts performed better than forecasts initialized 12 h before dropsondes 88% of the time (Figure 3.5a). GEFS after dropsonde forecasts had a similar maximum FSP of 86% for forecasts valid 24 h after the dropsondes were assimilated (Figure 3.5b). Lastly, CLP5 after dropsonde forecasts had a maximum FSP of 88% for forecasts valid 12 h after the dropsondes were assimilated (Figure 3.5c). Despite the after dropsonde forecasts performing better than the before dropsonde forecasts at all lead times, the bootstrap resampling shows minimal statistical significance of these results. None of the after dropsonde FSPs are statistically significant, although the after dropsonde FSP is higher than 90% of randomly sampled “after” forecast FSPs at ECMWF EPS forecasts valid 12, 24, and 60 h after dropsondes were assimilated (Figure 3.5a) and at GEFS forecasts valid 24 h after the assimilation of dropsondes (Figure 3.5b). This result suggests that dropsonde data likely has the greatest impact on ECMWF EPS and GEFS forecasts at earlier lead times. Nevertheless, these results indicate that most reductions in position errors in forecasts after dropsonde data is assimilated are not dissimilar to random chance.

The previous results suggest that the differences in average track forecast error in GEFS and ECMWF forecast initializations are similar to an average set of forecasts. Although the quantitative differences in average track forecast error are likely random, it is still possible that the forecasts initialized after dropsondes are assimilated have more skill than the forecasts 12 h before dropsondes are assimilated. Therefore, it is worth investigating whether there was an increase in the skill of the GEFS and ECMWF EPS track forecasts relative to CLP5 in the forecasts following the assimilation of the dropsondes. Figure 3.6 shows the average GEFS and ECMWF EPS mean track forecast skill for the 56 forecast initialization times 12 h before dropsonde data was assimilated and the 56 forecast initialization times after the dropsonde data was assimilated, as well as the difference in average skill of these forecasts for a given lead time. For both models, the
skill of the after dropsonde forecasts is up to 5% higher than the skill of the before dropsonde forecasts, except for forecast hour 12 where the skill is 12% higher (Figure 3.6a). GEFS and ECMWF EPS mean track forecast errors after dropsondes are up to 80% lower than CLP5 track forecast errors at forecast hour 48 and 120, respectively. This result suggests that there is an increase in skill in both operational models between the forecast initializations 12 h before dropsondes and the forecast initializations after dropsondes at each forecast hour in the forecast period. The average change in GEFS and ECMWF EPS skill relative to CLP5 at the same valid time between the 56 pairs of forecast initializations before and after dropsondes is compared to the average change between all 601 pairs of forecast initializations (Figure 3.6b). GEFS and ECMWF EPS forecasts after dropsondes have a decrease in skill relative to CLP5 compared to forecasts 12 h before dropsondes for forecasts valid 12 and 24 h after the initialization of the after dropsonde forecasts. This result disagrees with the hypothesis that the assimilation of dropsonde data into the operational models would add skill to position forecasts. As the forecast lead time decreases, CLP5 typically has lower position errors because the inputs to the statistical model used to produce CLP5 track forecasts include the current movement, movement during the previous 12- and 24-h periods, and direction of TC motion (Neumann 1972; Aberson 1998). Therefore, CLP5 0-, 12-, and 24-h position forecasts are often on par with the operational model position forecasts, unless there is an ongoing complex steering motion that is unlike steering motions in the Atlantic basin climatology. The decrease in skill relative to CLP5 is statistically significant at the 95% percentile for ECMWF EPS forecasts valid 12 h after the initialization of the after dropsonde forecasts. While not statistically significant, the decrease in skill of GEFS forecasts valid 12 and 24 h after the initialization of dropsonde forecasts and ECMWF EPS forecasts valid 24 h after the initialization of dropsonde forecasts is lower than the decrease in skill of 74–89% of randomly selected skill
differences from climatology. GEFS and ECMWF EPS forecasts after dropsondes exhibit an increase in skill compared to forecasts 12 h before dropsondes for all forecasts valid at lead times greater than 24 h after the initialization of the after dropsonde forecasts; however, this increase in skill is not statistically significant. Therefore, these results indicate that most changes in skill relative to CLP5 in ECMWF EPS and GEFS forecasts initialized after dropsondes are assimilated are likely due to random chance.

3.2. Case Studies

3.2.1. Hurricane Marco (2020)

3.2.1.1. Synoptic Overview

While Hurricane Marco was ultimately an uneventful TC in terms of impacts to the United States Gulf Coast, its track forecast was characterized by an uncommonly large amount of uncertainty. On 10 August 2020, a tropical wave moved westward from the coast of Africa, propagating westward over the central tropical Atlantic Ocean for a few days with poorly organized convection (Beven and Berg 2021). As the wave reached the western tropical Atlantic and passed through the Lesser Antilles on 17 August, convection gradually began to increase but remained disorganized and without a closed circulation. Shortly after reaching the central Caribbean Sea, convection began to better organize and a broad area of low pressure formed on 19 August. The low continued its westward movement toward Central America while its convection gradually became more organized and on 0600 UTC 21 August the low was deemed a
tropical depression over the extreme western Caribbean Sea while moving west-northwestward. Meanwhile, an upper-level trough over the western Gulf of Mexico led to a weakness in the subtropical ridge. As the tropical depression approached this weakness in the subtropical ridge, its motion shifted northwestward keeping the center of the depression offshore of Central America. This track allowed the depression to further intensify, and the storm was declared Tropical Storm Marco at 0000 UTC 22 August. Marco continued to intensify as its center moved through the Yucatan Channel into the southeastern Gulf of Mexico, where Marco briefly became a hurricane on 23 August with a peak intensity of 65 kt. Shortly thereafter, the aforementioned upper-level trough led to an increase in southwesterly upper-level winds which began to weaken Marco due to higher vertical shear. As Marco began to approach the mouth of the Mississippi River from the south-southeast on 24 August, its center became exposed to the southwest of the deep convection. Subsequently, the vertically shallow tropical storm turned westward as it became predominantly steered by a low-level ridge to the north of the cyclone. The center passed just south of the Gulf coastline near the mouth of the Mississippi River around 0000 UTC 25 August before weakening to a tropical depression. Marco was unable to redevelop deep convection and degenerated to a remnant low and then to a trough off the Louisiana coastline by 26 August.

NHC official forecast track errors for Marco were significantly greater than the five-year mean NHC official track errors (Beven and Berg 2021). Several official and ensemble track forecasts exhibited a significant westward bias, forecasting Marco to have a northwestward motion in the Gulf of Mexico rather than its actual north-northwestward motion. ECMWF EPS and GEFS forecasts initialized on 0000 UTC 22 August were characterized by high across-track variability amongst ensemble members with potential landfall positions extending from Mississippi on 25 August to the southern tip of Texas on 27 August (Figure 3.7a, c). Subsequent ensemble forecasts
initialized on 1200 UTC 22 August continued to exhibit high variability amongst ensemble members (Figure 3.7b, d); however, there was a notable eastward shift in the position forecasts with the GEFS and ECMWF EPS now forecasting landfall in Louisiana on 25 August (Figure 3.8). This eastward shift in the ensembles was noted in the 4:00 PM CDT 22 August NHC forecast discussion where it was speculated that the eastward shift in the 1200 UTC guidance was influenced by data collected by the 0530 UTC 22 August NOAA G-IV reconnaissance mission (Figure 3.9). This mention suggests that the in-situ observations collected during this mission impacted the forecasted environmental conditions surrounding Marco, leading to a large eastward shift in track. Therefore, this case study will focus on identifying the processes and features that contributed to the eastward shift in the 1200 UTC 22 August position forecasts.

Given the explicit association of the 0530 UTC 22 August NOAA G-IV reconnaissance mission to the eastward shift in Marco’s track forecasts by the NHC, as well as the overall theme of this thesis, it is of interest to investigate the impacts the data collected during this mission had on the 1200 UTC 22 August ensemble forecasts. By early afternoon on 20 August, tropical depression 14 was identified as a potential landfall threat to the Texas and Louisiana Gulf coast, leading the NHC to schedule the first NOAA G-IV synoptic surveillance mission around the storm with a take-off time of 0530 UTC 22 August. The aircraft departed from Lakeland, Florida and proceeded to deploy 31 dropsondes at roughly 150–200-km intervals as the aircraft circumnavigated the storm in the western Caribbean Sea and in the surrounding environment of the storm in the Gulf of Mexico (Figure 3.8). The environment in the Gulf of Mexico was sampled because the ensemble-based guidance suggested that this area was where the storm’s track forecast was sensitive to the environmental steering flow. The data collected during this mission was subsequently assimilated into the 1200 UTC 22 August global model runs.
3.2.1.2. Forecast Analysis

The 1200 UTC 22 August forecast is compared to the 0000 UTC 22 August forecast to investigate the impacts the dropsonde data had on the 1200 UTC 22 August ensemble forecasts. The 0000 UTC 22 August forecast is chosen as it is the final global model forecast initialization prior to the assimilation of the data collected during the 0530 UTC 22 August NOAA G-IV mission. Prior to understanding which processes contributed to the eastward shift in the 1200 UTC track forecast, it is necessary to analyze the track forecasts themselves. Figure 3.10 shows the ECMWF EPS and GEFS mean track forecast errors every 6-h beginning at 1200 UTC 22 August for both model runs. At the start of the forecast period, the GEFS and ECMWF EPS 0000 UTC model runs have a 180-km and 100-km position error, respectively, which rises to 375 km and 330 km by 0000 UTC 25 August, respectively. Meanwhile, the GEFS and ECMWF EPS 1200 UTC models runs have lower position errors at all times in the forecast, such that position errors remain below 125 km and 175 km, respectively. While the GEFS has larger errors in the 0000 UTC model run compared to the ECMWF EPS, the GEFS experiences a larger total reduction in error in the 1200 UTC model run compared to the ECMWF EPS. Nevertheless, both operational models display a reduction in track forecast error in the 1200 UTC model run at all times in the forecast period, suggesting that the inclusion of the reconnaissance dropsonde data in the models may have improved the track forecasts.

An alternative way to view TC position errors is in terms of its motion. Given that previous studies have identified the environmental steering flow as the main driver of TC motion (e.g., George and Gray 1976; Chan and Gray 1982; Pike 1985; Aberson and DeMaria 1994), TC position errors can be viewed as the integral of the environmental steering flow errors over time. For
example, an easterly 1 ms\(^{-1}\) error in the TC’s steering flow will result in an 86-km westerly position error after 24 h. Therefore, it is worth examining the differences in the environmental steering flow pattern between the forecasts initialized before and after dropsonde data is assimilated. This is evaluated by computing the ECMWF EPS ensemble mean \(u\) and \(v\) components of the environmental steering flow at every grid point at all times in each forecast, then comparing the resulting steering flow vectors at the same valid time in the forecast. Figure 3.11 shows the environmental steering flow vectors from both forecasts and the magnitude of the vector wind difference between the two initialization times before and after dropsonde assimilation. During the forecast period, there are minimal differences in the magnitude and direction of the steering flow near Marco’s position forecasts. In particular, the difference in the steering flow magnitude does not exceed 2 kt in the vicinity of Marco at any point in the forecast. Further, the steering flow direction within 2° of Marco’s forecasted positions does not differ by more than 5° at any point in the forecast. Thus, the eastward shift in Marco’s track forecast before and after dropsonde assimilation does not appear to be due to the dropsonde data altering the downstream steering flow.

The prior result indicates that another factor must be driving the position differences in Marco’s track forecast before and after dropsonde assimilation. Further assessment of the evolution of the steering flow shown in Figure 3.11 illustrates the presence of deformation in the Gulf of Mexico steering flow, with the axis of contraction extending from the southeast to the northwest in the Gulf of Mexico and into Louisiana starting at forecast hour 30 in the before forecast and forecast hour 18 in the after forecast (Figure 3.11b). The steering flow on the western side of the axis of contraction is southeasterly, while on the eastern side the flow is more south-southeasterly. As the forecast continues, the distance between the ensemble mean positions begins to increase, with the 1200 UTC 22 August forecast position taking a more eastward path than the
0000 UTC 22 August forecast (Figure 3.11c, d). This result suggests that the initial position correction in the 1200 UTC 22 August model run placed Marco on the opposite side of the deformation axis, leading to the eastward shift in the track forecast due to the TC being in more south-southeasterly flow on the eastern side of the deformation axis. Therefore, it appears that the critical factor behind the eastward shift in the track forecast after the reconnaissance mission is the change in the initial position in the 1200 UTC 22 August model run toward the observed TC position, not a change in the steering flow structure driving Marco’s motion.

It has been shown that there are negligible differences in the steering flow structure of the two forecasts; therefore, differences in the near-TC steering flow would support the hypothesis that the change in Marco’s initial position is the process behind the eastward shift in the entire track forecast. This finding is because changes in the TC initial position can place the TC in a different ensemble mean steering flow value, even if the overall steering flow regime did not change from the previous model run. The TC motion will be driven by a different steering flow than the previous model run, resulting in a change in the track forecast. Greater changes in the track forecast are the product of larger direction and speed changes in the near-TC steering flow. Therefore, it is worthwhile to quantifying changes in the near-TC steering flow and the corresponding shift of the forecasted TC position in the 1200 UTC 22 August model run compared to the 0000 UTC 22 August model run. Figure 3.12 shows the magnitude of the vector difference of the near-TC steering flow and the distance between the 0000 UTC and 1200 UTC ensemble mean position for all times. The difference in the steering flow is only analyzed through 0000 UTC 25 August, as the ensemble mean steering flow regime is established by forecast hour 72. The distance between the 0000 UTC and 1200 UTC ensemble mean positions is around 75 km at the start of the forecast period and remains relatively constant until 1200 UTC 23 August when the distance
begins to increase, reaching 200 km by the end of the forecast period. The magnitude of the vector difference of the 0000 UTC and 1200 UTC near-TC steering flow increases to just below 1.75 kt by 0000 UTC 23 August before subsequently dropping to near 0 kt 18 h later. However, the difference immediately begins to increase again and reaches 2 kt by the end of the forecast period. The second increase in the magnitude of the vector difference of the near-TC steering flow corresponds to the increase in the distance between the ensemble mean positions. This result indicates that the initial position change at the start of the 1200 UTC August 22 forecast placed Marco into a different near-TC steering flow pattern. As the forecast continues, the near-TC steering flow pattern remains different and results in the increasing distance between the two track forecasts. Furthermore, the timing of the increasing distance between the two track forecasts corresponds to when Marco enters a region characterized by greater deformation in the environmental steering flow, as displayed in Figure 3.11b, c, d. This result further supports that the eastward initial position change in the 1200 UTC 22 August model run is the critical factor influencing the eastward shift in the track forecast after the reconnaissance mission.

If the initial position change in the 1200 UTC 22 August model run is the main factor behind the eastward shift in Marco’s track forecast after the reconnaissance mission, then the position variability at later times in the forecast (e.g., 0000 UTC 25 August) should be strongly related to the position difference at earlier times in the forecast. One method of measuring this relationship is shown in Figure 3.13, which displays the ensemble correlation between Marco’s position along the major axis of the position ellipse at 0000 UTC 25 August to the distance along the major axis at earlier times in the forecast. The ellipses utilized as part of the MACC calculation are shown in Figure 3.7a, b, where the major axis of the ellipses for these two forecasts are mainly in the across-track direction (i.e., east–west direction). At 1200 UTC 22 August, the correlation
for the 1200 UTC model run is 0.31 which increases to 0.79 by 1200 UTC 23 August and maximizes near 1.00 by the end of the forecast period. For both initialization times, the MACC values are statistically significant at the 95% confidence level at all times. This result confirms the expected relationship between the position at later times in the forecast and the position at earlier times in the forecast when an initial position change leads to an eastward shift in the track forecast. Therefore, the significant eastward shift in Marco’s track forecast after the assimilation of dropsonde data collected during the 0530 UTC 22 August NOAA G-IV reconnaissance mission is likely due to the initial position change in the 1200 UTC 22 August ECMWF EPS model run.

3.2.2. Hurricane Zeta (2020)

3.2.2.1. Synoptic Overview

A multitude of factors contributed to the genesis of Zeta. On 20 October 2020, a weak surface trough had formed in the southwestern Caribbean Sea due to increased deep convection north of Panama, partially due to the influence of low-level southwesterly flow (Blake et al. 2021). Simultaneously, a tropical wave had propagated westward near the Yucatan Peninsula. This tropical wave was accompanied by a broad area of low pressure; however, high vertical wind shear inhibited further development of the wave. Instead, the southerly flow induced by this system caused the aforementioned surface trough to start drifting northward towards Jamaica. This disturbance became more organized over the next few days, and subsequently became a tropical depression by 1200 UTC 24 October. Despite misaligned low- and mid-level circulation centers,
strong central convection developed and the depression strengthened to Tropical Storm Zeta by 0000 UTC 25 October.

As high pressure built over the Gulf of Mexico, Zeta began to slowly move to the west or west-northwest and steadily strengthened due to low shear and warm water in the surrounding environment. By early morning on 26 October, Zeta strengthened into a hurricane southeast of Cozumel, Mexico. At the same time, Zeta began to move much faster as the high pressure over the Gulf of Mexico to the north of Zeta strengthened. At 0355 UTC 27 October, Zeta made its first landfall near Ciudad Chemuyil, Mexico with a minimum central pressure of 977 mb and an intensity of 75 kt. Later that day, Zeta emerged over the southern Gulf of Mexico as a tropical storm.

Meanwhile, the synoptic pattern ahead of Zeta began to rapidly change. The ridge over the Gulf of Mexico began to erode on its western side due to a deep-layer cutoff low over the southwestern United States moving eastward. Zeta continued its northwestward movement and re-intensified to a hurricane early on 28 October due to a favorable environment in the Gulf of Mexico. Zeta’s movement began to accelerate as the hurricane reached the south-central Gulf of Mexico. The cutoff low over western Texas and the eastward moving ridge near Florida led to Zeta accelerating even more to the north-northeast. Zeta continued to strengthen as it rapidly approached the Louisiana coast, and made landfall near Cocodrie, Louisiana around 2100 UTC 28 October with a minimum central pressure of 970 mb and an intensity of 100 kt. Early the next day, Zeta weakened to a tropical storm as it moved further inland before becoming post-tropical over central Virginia by 1800 UTC 29 October. Zeta dissipated over the western Atlantic Ocean early on 30 October after merging with a frontal zone.
NHC official forecast track errors for Zeta were lower than the five-year mean through 48 h (Blake et al. 2021). However, official forecast track errors were above the mean official track errors at forecast times 60–72 h. ECMWF EPS and GEFS forecasts initialized on 1200 UTC 25 October were characterized by mainly along-track variability, especially after 60 h (Figure 3.14a, c). Subsequent ensemble forecasts initialized on 0000 UTC 26 October continued to exhibit high along-track variability amongst ensemble members at later times in the forecast (Figure 3.14b, d). The high along-track variability exhibited in the ensemble forecasts aligns with the start of Zeta’s northwestward acceleration in the Gulf of Mexico on 28 October. This observation suggests that the variability in the ensembles could be related to the timing of the onset of Zeta’s acceleration, leading to the higher track errors at forecast times 60–72 h noted by the NHC.

In addition to the higher track errors at forecast times 60–72 h, CLP5 errors were approximately 50–75% greater than the five-year mean NHC official track errors at all forecast times, indicating that Zeta’s track was more difficult to forecast than a typical Atlantic tropical cyclone (Blake et al. 2021). Two factors may have contributed to the difficult track forecast: the previously mentioned disagreements in the timing of the onset of Zeta’s acceleration, as well as a northeastward bias in the ensemble track forecasts in the early parts of Zeta’s life. Figure 3.15 shows this northeastward bias in the 1200 UTC 25 October ECMWF EPS and GEFS forecasts as Zeta approaches the Yucatan Peninsula, particularly in the GEFS forecast. This bias decreases in the subsequent 0000 UTC 26 October forecast, alongside a reduction in Zeta’s forward motion (Figure 3.15). This result suggests that the forecasted environmental conditions changed in between these two forecasts, possibly due to the environmental data collected in the Gulf of Mexico during the 1730 UTC 25 October NOAA G-IV reconnaissance mission. This NOAA G-IV synoptic surveillance mission departed from Lakeland, Florida and proceeded to deploy 33
drops ondes at roughly 150–200-km spacing in the downstream environment in the eastern half of the Gulf of Mexico and the northwest quadrant of the storm (Figure 3.15). The downstream environment in the Gulf of Mexico was sampled during the mission because the ensemble-based guidance suggested that Zeta’s track forecast was sensitive to the environmental steering flow in this region. The dropsonde data collected during this mission was subsequently assimilated into the 0000 UTC 26 October global model runs. Therefore, this case study will focus on identifying the processes and features that contributed to the changes Zeta’s track forecast in the 0000 UTC 26 October ensemble forecasts.

3.2.2.2. Forecast Analysis

The 0000 UTC 26 October ensemble forecasts are compared to the 1200 UTC 25 October ensemble forecasts to investigate the impacts the NOAA G-IV dropsonde data had on the 0000 UTC 26 October ensemble forecasts. This case study follows the same methodology introduced for Hurricane Marco (2020). Figure 3.16 shows the ECMWF EPS and GEFS mean track forecast errors every 6 h beginning at 0000 UTC 25 October for both model runs. ECMWF EPS track errors were slightly (<60 km) lower than the GEFS track errors until 1800 UTC 26 October. Afterwards, GEFS track errors were lower compared to the ECMWF EPS, especially after 1200 UTC 28 October. There was a reduction in track forecast errors of up to 100 km in both ensembles until 0000 UTC 29 October. The up to 50-km reduction in error at the start of the forecast and the southwestward shift in the track forecasts shown in Figure 3.15 suggest that the dropsonde data led to the northeastward track bias in the 1200 UTC 25 October forecasts being corrected in the 0000 UTC 26 October forecasts. Furthermore, the greatest decreases in track forecast errors are
from 0000 UTC 27 October to 0000 UTC 29 October, corresponding to the onset of Zeta’s rapid acceleration. This result suggests that the inclusion of the reconnaissance dropsonde data in the models improved the environmental steering flow, leading to an improvement in the track forecast, especially at mid-ranges of the forecast.

The environmental steering flow is assessed to test the hypothesis that changes in the environmental steering flow led to position error reductions in the 0000 UTC 25 October model run. Figure 3.17 shows the environmental steering flow vectors from both forecasts and the magnitude of the vector difference between the two initialization times. At the start of the forecast (Figure 3.17a), the differences in the steering flow magnitude in the area surrounding the TC reach 5 kt, mainly due to shifts in the direction of the steering flow from southeasterly to southwesterly. The differences lower to 1.5–2 kt 18-hours later (Figure 3.17b) and the position forecasts are still proximate to each other. However, the magnitude of the vector difference of the steering flow starts to increase due to a 5–10-kt reduction in the steering flow speed in the near-TC and surrounding areas (Figure 3.17c). By forecast hour 72 in the before forecast and forecast hour 60 in the after forecast, the magnitude of the vector difference reaches 6 kt (Figure 3.17d). As the magnitude of the vector difference of the steering flow increases, the ensemble mean positions begin to separate, with the 1200 UTC 25 October forecast position accelerating northward at a faster rate. This separation suggests that the 5–10-kt decrease in the speed of the environmental steering flow in the 0000 UTC 26 October model run decreased Zeta’s forward motion compared to the previous model run, leading to a more accurate representation of the near-TC steering flow driving Zeta’s motion. Therefore, it appears that the dropsonde data impacted the steering flow structure driving Zeta’s motion, as opposed to Hurricane Marco where the dropsonde data changed the initial position.
If the dropsonde data impacted the steering flow driving Zeta’s motion, there should be a corresponding change in the near-TC steering flow. Figure 3.18 shows the magnitude of the vector difference of the near-TC steering flow (inner 100 km) and the distance between the 1200 UTC 25 October and 0000 UTC 26 October ensemble mean position for all times. Initially, the magnitude of the vector difference of the near-TC steering flow is around 3 kt; however, the difference drops below 1 kt by 1200 UTC 26 October. Starting at 0000 UTC 27 October, 24 h into the forecast period, the magnitude of the vector difference begins to rapidly increase, reaching a difference of 9 kt by 1200 UTC 28 October. This rapid increase in the magnitude of the vector difference of the near-TC steering flow corresponds to a rapid increase in the distance between the ensemble mean positions of the two forecasts, where the distance increases from 50 km to 250 km between 0000 UTC 27 October and 1200 UTC 28 October. These corresponding rapid increases indicate that there was a change in the steering flow structure in the 0000 UTC 26 October model run, starting at the first time in the forecast period. As the forecast continues, the near-TC steering flow pattern is much slower than in the previous forecast, leading to the distance between the two track forecasts increasing. The increasing distance between the two track forecasts may also be the result of the 1200 UTC 25 October model run position reaching the area of faster flow earlier than the 0000 UTC 26 October model run position. Furthermore, the onset of the largest differences in the near-TC steering flow and the ensemble mean positions at 0000 UTC 27 October corresponds to the greatest decreases in track forecast errors (Figure 3.15). This result further supports that the change in the steering flow structure in the 0000 UTC 26 October model run, particularly a decrease in the steering flow magnitude, is the factor influencing the slowdown in Zeta’s forward motion and the resulting improvement in the mid-ranges of the track forecast after the reconnaissance mission.
Since it is hypothesized that changes in the near-TC steering flow structure after 0000 UTC 27 October in the 0000 UTC 26 October model run led to a reduction in position errors, the position at later times in the forecast is expected to be dependent on the position after 0000 UTC 27 October. Figure 3.19 shows the correlation between Zeta’s position along the major axis at 1200 UTC 28 October to the distance along the major axis at earlier times in the forecast. The ellipses utilized as part of the MACC calculation are shown in Figure 3.14a, b, where the major axis of the ellipses for these two forecasts are in the along-track direction, particularly after the first 24 h of both forecasts. Between 0000 UTC 27 October and 1200 UTC 28 October, the major axes of the ellipses are oriented in the north-south direction. The MACC for the 0000 UTC 26 October forecast begins at about 0.28 and increases to about 0.85 at 0000 UTC 27 October. Afterwards, the MACC remains very high and approaches 1.0 by the end of the forecast period. The MACC values are statistically significant at the 95% confidence level at all times in the 0000 UTC 26 October forecast. There is a high correlation of Zeta’s position along the major axis at 1200 UTC 28 October to the position at all times after 0000 UTC 27 October, suggesting that Zeta’s position at 0000 UTC 27 October is determining the position later in the forecast. This result supports the hypothesis that the change in the steering flow structure in the 0000 UTC 26 October model run is the factor influencing the improvement in the mid-ranges of Zeta’s track forecast after the NOAA G-IV reconnaissance mission. Therefore, the decrease in Zeta’s forward motion in the mid-ranges of the forecast and resulting reduction of track forecast error after the assimilation of dropsonde data collected during the 1730 UTC 25 October NOAA G-IV reconnaissance mission can be attributed to a change in the environmental steering flow in the 0000 UTC 26 October ECMWF EPS model run.
3.2.3. Hurricane Dorian (2019)

3.2.3.1. Synoptic Overview

Hurricane Dorian was the strongest hurricane to hit the northwestern Bahamas in modern history, devastating Great Abaco and Grand Bahama Islands as a powerful category 5 hurricane. On 19 August 2019, a tropical wave moved westward from the coast of Africa, propagating westward over the tropical Atlantic Ocean for a few days with minimal convection (Avila et al. 2020). By 22 August, the wave reached the central Atlantic Ocean and developed into a small low-pressure area. Despite southeasterly deep tropospheric vertical wind shear, the low continued gaining convective organization whilst moving westward and by 0600 UTC 24 August the low became a tropical depression while located about 700 n mi east-southeast of Barbados. The depression quickly became better organized and strengthened into Tropical Storm Dorian at 1800 UTC 24 August.

Over the next few days, Dorian struggled to strengthen and stay well-organized due to intrusions of dry air and made landfall over Barbados around 0130 UTC 27 August with an intensity of 45 kt. Dorian continued to move across the Windward Islands, passing over St. Lucia around 1100 UTC 27 August. Dorian’s interactions with the terrain on St. Lucia disrupted its low-level circulation, causing the center to reform to the north. Afterwards, Dorian developed an inner core and intensified to a category 1 hurricane as its center passed over eastern St. Croix in the U.S. Virgin Islands around 1530 UTC 27 August. Dorian continued to strengthen as it travelled northwestward between an upper-level low over the Straits of Florida and the Atlantic subtropical ridge. During this time, the upper-level low moved southward and the subtropical ridge expanded.
to the west, resulting in a shift in Dorian’s motion to the west-northwest. This change in the steering flow guiding Dorian’s motion moved the storm into a region of low vertical wind shear, abundant atmospheric moisture, and warm ocean temperatures. This favorable environment allowed Dorian to rapidly intensify and become a category 5 hurricane just prior to making landfall at Elbow Cay, Great Abaco, in the northwestern Bahamas at 1640 UTC 1 September with a minimum central pressure of 910 mb and an intensity of 160 kt.

Dorian’s westward motion drastically slowed over the next day as the high pressure to its north weakened, causing the hurricane to batter Great Abaco for several hours before making landfall near South Riding Point on Grand Bahama around 0215 UTC 2 September with an intensity of 155 kt. While Dorian exited along the north coast of Grand Bahama later that day, a large mid-level trough positioned over the eastern U.S. swung eastward, generating a steering flow pattern that shifted Dorian’s motion from the west to the north. This change in Dorian’s motion prevented it from making landfall over Florida. Instead, Dorian travelled northward then northeastward along the East Coast of the United States. Dorian made its final landfall as a tropical cyclone over Cape Hatteras, North Carolina at 1230 UTC 6 September with an intensity of 85 kt. Afterwards, Dorian became embedded within the mid-latitude flow and accelerated northeastward before becoming post-tropical at 1800 UTC 7 September, then fully extratropical at 0600 UTC 8 September.

NHC official forecast track errors for Dorian were lower than the five-year mean NHC official track errors (Aviles et al. 2020). However, prior to Dorian’s center reformation after crossing over the terrain of St. Lucia on 27 August, models incorrectly forecasted Dorian to stay to the south of Puerto Rico and cross over eastern Hispaniola. After Dorian’s center reformed north of its previous location around 1200 UTC 27 August, there was a significant shift in subsequent
official and ensemble track forecasts on 0000 UTC 28 August (Figure 3.20). There is a noticeable northeastward shift in both the GEFS and ECMWF EPS position forecasts in the 0000 UTC 28 August forecast, which predicted Dorian to make landfall over Puerto Rico and stay offshore of Hispaniola. This shift suggests that some set of observations led to the global models picking up on Dorian’s position change due to the center reformation. Meanwhile, the near-TC and surrounding environments were sampled using dropsondes deployed by the 1730 UTC 27 August NOAA G-IV reconnaissance mission. The aircraft departed from St. Croix, United States Virgin Islands and proceeded to deploy 30 dropsondes at roughly 150–200-km spacing as the aircraft circumnavigated the storm and in the surrounding environment in the eastern Caribbean Sea (Figure 3.20). The environment to the northeast of Dorian was sampled because the ensemble-based guidance suggested that this area was where the storm’s track forecast was sensitive to the environmental steering flow. The in-situ observations collected during this mission were assimilated into the 0000 UTC 28 August global model runs, coinciding with the northeastward shift in the GEFS and ECMWF EPS position forecasts. This result suggests that the observational data collected during the G-IV mission may be a factor that contributed to the global models resolving Dorian’s northward center reformation and the subsequent northeastward shift in track in the 0000 UTC 28 August model run. Therefore, this case study will focus on identifying the processes and features that contributed to the northeastward shift in the 0000 UTC 28 August position forecasts.
3.2.3.2. Forecast Analysis

The 0000 UTC 28 August ensemble forecasts are compared to the 1200 UTC 27 August ensemble forecasts to investigate the impacts the NOAA G-IV dropsonde data had on the 0000 UTC 28 August ensemble forecasts. ECMWF EPS and GEFS forecasts initialized at 1200 UTC 27 August were characterized by across-track variability amongst ensemble members in the first 60 h of the forecast and along-track variability in later times in the forecast, particularly in the ECMWF EPS (Figure 3.21a, c). The subsequent ensemble forecast initialized on 0000 UTC 28 August continued to exhibit across-track variability amongst ensemble members in the first hours of the forecast (Figure 3.21b, d); however, there was a noticeable northeastward shift in the position forecasts with all ensemble members now forecasting Dorian to make landfall on Puerto Rico and miss Hispaniola to the north. Figure 3.22 shows the ECMWF EPS and GEFS mean track forecast errors every 6 h beginning at 0000 UTC 28 August for both model runs. Both models experience a reduction in error in the 0000 UTC model run, ranging from 50 km at the start of the forecast period up to 200 km by 0000 UTC 30 August. At 0000 UTC 28 August, the error in both models reduces to near 0 km, indicating that both models resolved Dorian’s center reformation. The rate of error reduction for both models is greatest in the first 48 h of the forecast period, with minimal changes in the amount of error reduction despite the westward shift in the 0000 UTC 28 August track forecast. This result suggests that the critical factor driving the improvements in the GEFS and ECMWF EPS track forecasts comes earlier in the forecast period. Overall, the ECMWF EPS has lower errors compared to the GEFS in both forecasts. In the 0000 UTC 28 August forecast, the ECMWF EPS has track errors around 100 km from 0000 UTC 29 August to 1200 UTC 31 August, while the GEFS has track errors increasing to a peak of 200 km by 0000 UTC 30 August.
before lowering to 100 km by 1200 UTC 31 August. After 1200 UTC 31 August, the GEFS has track errors around 75 km lower than the ECMWF EPS. Nevertheless, both operational models display a reduction in track forecast error in the 0000 UTC 28 August model run at all times in the forecast period, although the greatest reductions in error are in the first 48 h of the forecast. This result suggests that the inclusion of the reconnaissance dropsonde data in the models improved the track forecasts, particularly at earlier times in the forecast.

The environmental steering flow is evaluated to determine if the steering flow structure changed in the 0000 UTC 28 August model run. Figure 3.23 shows the environmental steering flow vectors from both forecasts and the magnitude of the vector difference between the two initialization times. During the forecast period, Dorian is embedded in a complex flow environment. Dorian is located in a steering flow saddle point of a break in the subtropical ridge at the start of the forecast (Figure 3.23a). Throughout the forecast, there are 5°–15° differences in the direction and 0–5-kt differences in the speed of the environmental steering flow on the western side of the subtropical ridge, resulting in up to 4-kt differences in the steering flow magnitude. Near Dorian, however, the difference in the steering flow magnitude does not exceed 2 kt at any point in the forecast, indicating that the steering flow around Dorian exhibits minimal directional and/or speed changes in the steering flow. Instead, it appears that the approximately 50-km northeastern shift in the TC initial position after the reconnaissance mission placed Dorian closer to the subtropical ridge and away from the saddle point in the break of the subtropical ridge, leading the storm to be steered by the subtropical ridge to the northeast more directly than in the 1200 UTC model run. Thus, the northeastward shift in Dorian’s track forecast after dropsondes are assimilated is not due to the dropsonde data altering the steering flow regime near Dorian. Instead, these results suggest that the critical factor behind the northeastward shift in the track forecast after the
reconnaissance mission is the change in the initial position in the 0000 UTC 28 August model run toward the observed TC position, not a change in the steering flow structure driving Dorian’s motion.

The near-TC steering flow is examined to quantitatively determine whether the 50-km northeastern shift in the TC initial position after the reconnaissance mission placed Dorian in a different near-TC steering flow pattern. Figure 3.24 shows the magnitude of the vector difference of the near-TC steering flow and the distance between the 1200 UTC 27 August and 0000 UTC 28 August ensemble mean position for all times. The distance between the 1200 UTC and 0000 UTC ensemble mean positions is around 60 km at the start of the forecast period. As the forecast continues, this distance increases, reaching a peak of around 190 km at 0000 UTC 30 August before decreasing to 125 km 12 h later. The magnitude of the vector difference of the 1200 UTC and 0000 UTC near-TC steering flow is around 0.55 kts at the start of the forecast period and increases to 1.0 kt 18 h later. Overall, there are minimal differences in the magnitude of the vector difference of the 1200 UTC and 0000 UTC near-TC steering flow. Nevertheless, while the differences in the magnitude of the vector difference of the near-TC steering flow are small, the difference increase in the first half of the forecast corresponds to the increase in the distance between the ensemble mean positions. This result indicates that the initial position change at the start of the 0000 UTC forecast placed Dorian into a different near-TC steering flow pattern, leading to a northeastward shift in the track forecast toward the observed TC position.

As shown with Hurricane Marco, if changes in the initial TC position led to a shift in the track forecast, the position at the end of the forecast will be dependent on the position at the start of the forecast. Figure 3.25 shows the correlation between Dorian’s position along the major axis at 1200 UTC 30 August to the distance along the major axis at earlier times in the forecast. The
ellipses utilized as part of the MACC calculation are shown in Figure 3.21a, b, where the major axis of the ellipses for these two forecasts are in the across-track direction until 1200 UTC 30 August in both forecasts. The correlation for the first 12 h of the 0000 UTC model run is lower than the correlation of the 1200 UTC model run, which is likely due the ECMWF EPS spinning up the TC (i.e., the perturbations are still balancing). At 0000 UTC 28 August, the correlation for the 0000 UTC model run is 0.12 which increases to 0.65 by 1800 UTC 28 August and maximizes near 1.0 by the end of the forecast period. The 1200 UTC 27 August model run MACC values are statistically significant at the 95% confidence level at all times in the forecast. The 0000 UTC 28 August model run MACC values are statistically significant at the 95% confidence level from 1800 UTC 28 August to the end of the forecast period. There is a high correlation of Dorian’s position along the major axis at 1200 UTC 30 August to the position along the major axis at all times after 1800 UTC 28 August in the 0000 UTC forecast. Therefore, the position variability later in the forecast (e.g., 1200 UTC 30 August) is strongly related to the position at earlier times in the forecast. The large correlation early in the forecast supports the hypothesis that the TC initial position change in the 0000 UTC 28 August model run is the factor influencing the northeastward shift of Dorian’s track forecast after the NOAA G-IV reconnaissance mission. This initial position change in the 0000 UTC 28 August model run placed Dorian closer to the subtropical ridge and into a different near-TC steering flow pattern. This placement resulted in a northeastward shift in Dorian’s track, bringing the storm closer to the observed track, and a corresponding reduction of error in the track forecast.
3.2.4. Hurricane Jerry (2019)

3.2.4.1. Synoptic Overview

Hurricane Jerry was a category 2 hurricane that ultimately did not directly affect any land areas as a tropical cyclone. On 11 September 2019, a tropical wave moved westward from the coast of Africa, propagating westward over the eastern tropical Atlantic for a few days with poorly organized convection (Brown 2019). As the wave reached the central tropical Atlantic on 15 September, it generated an area of low pressure within the Intertropical Convergence Zone. Over the next 24–36 h, the area of low pressure gradually became better defined as it slowly moved west-northwestward. By 0600 UTC 17 September, the low formed into a tropical depression about 950 n mi east of the Windward Islands and slowly strengthened as it continued to move west-northwestward through a region of moderate northeasterly vertical wind shear to the south of a strong deep-layer ridge in the central Atlantic. By 0600 UTC 18 September, the convection became more organized and the depression strengthened into Tropical Storm Jerry. Shortly after becoming a tropical storm, Jerry began a period of rapid intensification that continued until it reached its peak intensity of a 90 kt hurricane at 0000 UTC 20 September. Over the next 24 h, Jerry moved quickly west-northwestward through a region of strong northwesterly upper-level winds and dry mid-level air located to the south of the strong ridge. These unfavorable conditions resulted in Jerry weakening to a tropical storm by 0000 UTC 21 September while the storm passed about 120 n mi north of the Leeward Islands and began to move northwestward around the western edge of the ridge. By early 22 September, a mid-level trough to Jerry’s north associated with recurving Hurricane Humberto produced a pronounced weakness in the ridge, resulting in Jerry turning
north-northwestward and slowing its forward motion. Over the next two days, Jerry continued its north-northwestward motion whilst its center became increasingly separated from the deep convection due to westerly shear, causing it to gradually weaken. By 1800 UTC 24 September, Jerry’s deep convection dissipated and Jerry degenerated into a post-tropical cyclone while being located about 245 n mi west-southwest of Bermuda.

NHC official forecast track errors for Jerry were generally close to the five-year mean NHC official track errors (Brown 2019). However, there were disagreements in the ensemble guidance on 20 September as Jerry’s motion began slow down and to shift from west-northwestward to northwestward as the storm approached a break in the subtropical ridge created by a deep-layer trough moving across the northwestern Atlantic. ECMWF EPS and GEFS forecasts initialized on 1200 UTC 20 September were characterized by large across-track variability amongst ensemble members, particularly at lead times greater than 48-h (Figure 3.26a, c). Both ensembles exhibited large variability in the timing of Jerry’s northwestward shift in motion as the storm travelled around the southwestern edge of the subtropical ridge. Subsequent ensemble forecasts initialized on 0000 UTC 21 September had reduced variability amongst ensemble members (Figure 3.26b, d), as well as a northeastward shift in the track forecasts toward the observed TC track. This result suggests that some set of observations led to the global models becoming more confident in the timing of Jerry’s north-northwestward shift in motion as the storm reached the break in the subtropical ridge. Meanwhile, the near-TC and surrounding environments were sampled using dropsondes deployed by the 1730 UTC 20 September NOAA G-IV reconnaissance mission. The aircraft departed from Barbados and proceeded to deploy 33 dropsondes at roughly 150–200-km spacing as the aircraft circumnavigated the storm to the north of the Leeward Islands and in the surrounding environment north of the storm in the Atlantic Ocean (Figure 3.27). The environment
north of Jerry in the Atlantic Ocean was sampled because the ensemble-based guidance suggested that this area was where the storm’s track forecast was sensitive to the environmental steering flow. The in-situ observations collected during this mission were assimilated into the 0000 UTC 21 September global model runs, coinciding with the northeastward shift in the GEFS and ECMWF EPS position forecasts.

3.2.4.2. Forecast Analysis

The 0000 UTC 21 September ensemble forecasts are compared to the 1200 UTC 20 September ensemble forecasts to investigate the impacts the NOAA G-IV dropsonde data had on the 0000 UTC 21 September ensemble forecasts. Figure 3.28 shows the ECMWF EPS and GEFS mean track forecast errors every 6 h beginning at 0000 UTC 21 September for both model runs. Both models have similarly large errors in the 1200 UTC 20 September model run until 1200 UTC 23 September. On 0000 UTC 21 September, both 1200 UTC 20 September model runs have errors around 40 km. These errors continue to grow before reaching their peaks of 260 km (ECMWF EPS) and 225 km (GEFS) at 1800 UTC 22 September. After 1200 UTC 23 September, the GEFS error increases to 330 km while the ECMWF EPS error decreases to 140 km by 1200 UTC 24 September. Nevertheless, both models display a reduction in track forecast error in the 0000 UTC 21 September model run compared to the 1200 UTC 20 September. At 0000 UTC 21 September, both 0000 UTC 21 September models have errors around 25 km, only 10–20 km lower than the 1200 UTC model runs. While the 0000 UTC ECMWF EPS and GEFS track errors increase steadily to over 200 km by 1200 UTC 22 September, the 1200 UTC track errors remain below 100 km. The 1200 UTC ECMWF EPS and GEFS track errors continue to increase throughout the remainder
of the forecast, reaching 175 km and 210 km, respectively, by 1200 UTC 24 September. Overall, the greatest reduction in error in both models occurs 24–60 h after the initialization of the 0000 UTC 21 September model runs, coinciding with the change in Jerry’s motion due to the weakness in the subtropical ridge on 22 September (Brown 2019). This result suggests that the inclusion of the reconnaissance dropsonde data in the models altered the steering flow driving Jerry’s motion and improved the subsequent track forecasts.

The environmental steering flow is examined to determine if there were changes in the flow driving Jerry’s motion as the storm rounded the subtropical ridge. Figure 3.29 shows the environmental steering flow vectors from both forecasts and the magnitude of the vector difference between the two initialization times. During the forecast period, Jerry is embedded in a complex flow environment due to the interaction of the subtropical ridge with a mid-level trough that resulted in a pronounced weakness in the ridge. At the start of the forecast, Jerry is rounding the southwestern edge of the subtropical ridge and then proceeds to move poleward along the western edge of the weakening ridge (Figure 3.29). Throughout the forecast, there are differences in the direction and speed of the environmental steering flow both near the TC and in the environment ahead of the TC. At 0000 UTC 21 September, the 0000 UTC model run steering flow direction shifts from easterly to east-southeasterly and the magnitude decreases 5 kt near the 0000 UTC TC position (Figure 3.29a). There continue to be differences in the steering flow 12 h later and the TC position forecasts begin to separate, with the 0000 UTC model run position beginning to curve around to the western side of the subtropical ridge (Figure 3.29b). The 0000 UTC model run steering flow direction shifts from southeasterly to south-southeasterly in a broad area around the TC while the magnitude generally remains the same except for an area directly east of the 0000 UTC position where the steering flow magnitude is 5 kt higher (Figure 3.29b). During the
remainder of the forecast period, there continue to be 30°–45° shifts in the steering flow associated with the northern half of the subtropical ridge in the 0000 UTC 21 September model run (Figure 3.29c, d). Accordingly, the 0000 UTC model run position continues to be upstream of the 1200 UTC model run position. These results suggest that the directional changes in the environmental steering flow in the 0000 UTC 21 September model run led to Jerry curving around the southwestern edge of the ridge earlier than in the previous model run, leading to a more accurate representation of the near-TC steering flow driving Jerry’s motion. Therefore, it appears that the dropsonde data impacted the steering flow structure driving Jerry’s motion as opposed to a change in the initial position of the TC, similar to the case study of Hurricane Zeta in section 3.2.2.

If the changes in the environmental steering flow led to the position changes in the 0000 UTC 21 September model run, there should be corresponding differences in the near-TC steering flow between the two model runs. Figure 3.30 shows the magnitude of the vector difference of the near-TC steering flow and the distance between the 1200 UTC 20 September and 0000 UTC 21 September ensemble mean position for all times. The distance between the 1200 UTC and 0000 UTC ensemble mean positions is around 40 km at the start of the forecast period. As the forecast continues, this distance steadily increases to 230 km by 0000 UTC 23 September and remains around 230 km through the end of the forecast period. Initially, the magnitude of the vector difference of the near-TC steering flow is around 1.3 kt; however, the difference sharply rises to 2.7 kt by 0600 UTC 21 September and remains there until 0600 22 September. Afterwards, the difference decreases to 1.0 kt by 0000 UTC 23 September and remains between 1.0 and 1.5 kt until the end of the forecast period. The greatest differences in the magnitude of the vector difference of the near-TC steering flow in the first half of the forecast correspond to the steady increase of the distance between the ensemble mean positions of the two forecasts. This result suggests that
the near-TC steering flow differences played a role in the increasing distance between the two position forecasts. Overall, this result further supports that the directional changes in the environmental steering flow in the 0000 UTC 21 September model run are the critical factor influencing the northeastward shift in the track forecast towards the observed TC position after the reconnaissance mission.

Since the greatest differences in the near-TC steering flow occurred in the first 30 h of the forecast period, it is expected that the position at 1200 UTC 23 September will be dependent on the position in the first 30 h of the forecast. Figure 3.31 shows the correlation between Jerry’s position along the major axis at 1200 UTC 23 September to the distance along the major axis at earlier times in the forecast. The correlation curves for both model runs have a similar shape. At 0000 UTC 21 September, the correlation for the 0000 UTC model run is 0.12 which briefly decreases to 0.08 at 0600 UTC 21 September. Afterwards, the correlation quickly increases to 0.63 by 0000 UTC 22 September and steadily increases up to around 1.00 by the end of the forecast period. The MACC values for both model runs are statistically significant at the 95% confidence level from 1800 UTC 21 September to the end of the forecast period. There is a high correlation of Jerry’s position along the major axis at 1200 UTC 23 September to the position along the major axis at all times after 0000 UTC 22 September in the 0000 UTC forecast. Therefore, the position variability later in the forecast is strongly related to the 1800 UTC 21 September position and its location relative to the subtropical ridge. This result supports the hypothesis that the directional changes in the near-TC steering flow structure in the 0000 UTC 21 September model run are the factors influencing the northeastward shift in Jerry’s track forecast and the resulting reduction in track forecast error after the NOAA G-IV reconnaissance mission.
Figure 3.1. The mean ECMWF EPS (blue), GEFS (magenta), and CLP5 (orange) 72-h track forecast error (km) from the same forecast initialization times in each CLP5 error percentile bin.
Figure 3.2. The mean track forecast error (km) for (a) ECMWF EPS (blue) and GEFS (magenta) and (b) CLP5 (orange) forecasts initialized with dropsondes (circle markers) and without dropsondes (x markers). The number of forecasts within each average error is listed in black text at the top of each panel. The percentage of randomly sampled forecasts from climatology that have lower average track forecast error than the forecasts with dropsondes is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.
Figure 3.3. The average skill relative to CLP5 (%) for ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized with dropsondes (circle markers) and forecasts initialized without dropsondes (x markers). The black text denotes the number of forecasts within the average skill at each forecast hour. The blue (ECMWF EPS) and magenta (GEFS) text denote the percentage of randomly sampled forecasts from climatology that have less skill relative to CLP5 than the forecasts initialized with dropsondes.
Figure 3.4. The mean track forecast error (km) for (a) ECMWF EPS (blue) and GEFS (magenta) and (c) CLP5 (orange) forecasts initialized with dropsondes (circle markers) and 12 h before dropsondes (left-facing triangle markers). The difference in mean track forecast error (km) for (b) ECMWF EPS (blue) and GEFS (magenta) and (d) CLP5 (orange) forecasts initialized with dropsondes and 12 h before dropsondes (circle markers) and the difference between all pairs of forecast initializations in the climatology (diamond markers). The number of forecasts within each average error is listed in black text at the top of each panel. The percentage of randomly sampled pairs of forecasts from climatology that have a lower reduction in track forecast error than the forecasts with dropsondes is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.
Figure 3.5. The frequency of superior performance (FSP; %) for (a) ECMWF EPS, (b) GEFS, and (c) CLP5 forecasts initialized after the assimilation of dropsondes (circle marker) and all “after” forecast initializations in the climatology (diamond marker). The percentage of randomly sampled “after” forecasts from climatology that have a lower FSP than the after dropsondes forecast initializations is in blue (ECMWF EPS), magenta (GEFS), and orange (CLP5) text at the top of the corresponding panel.
Figure 3.6. The (a) average skill relative to CLP5 (%) for ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized 12 h before dropsondes (left-facing triangle marker) and with dropsondes (circle marker) and (b) the difference in average skill relative to CLP5 (%) between ECMWF EPS (blue) and GEFS (magenta) track forecasts initialized 12 h before and with dropsondes (circle markers) and all pairs of forecast initializations in the climatology (diamond marker). The black text denotes the number of forecasts within the average skill at each forecast hour. The blue (ECMWF EPS) and magenta (GEFS) text denote the percentage of randomly sampled pairs of forecasts from climatology that have a greater reduction in average skill than forecasts with dropsondes.
Figure 3.7. (a), (b) ECMWF EPS and GEFS (c), (d) forecasts of Hurricane Marco initialized at (a), (c) 0000 UTC 22 August 2020 and (b), (d) 1200 UTC 22 August 2020 (gray lines). The dots indicate the location of each ensemble member at 24-h intervals. The colored ellipses show the bivariate normal distribution of the forecast positions at 24-h intervals, as in Hamill et al. (2011). Purple denotes the 24-h locations, cyan denotes 48-h locations, green denotes 72-h locations, red denotes 96-h locations, and magenta denotes 120-h locations.
Figure 3.8. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Marco beginning at 1200 UTC 22 August 2020 and ending at 0000 UTC 25 August 2020. The forecasts initialized at 0000 UTC 22 August 2020 are denoted by left-facing triangle markers, while the forecasts initialized at 1200 UTC 22 August 2020 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 0600 UTC 22 August 2020 and by black squares for the position starting at 1200 UTC 22 August 2020. The dropsonde deployment locations during the 0530 UTC 22 August 2020 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.
Tropical Storm Marco Discussion Number 10
NWS National Hurricane Center Miami FL AL142020
400 PM CDT Sat Aug 22 2020

Cutting to the chase, there have been some big changes among the model guidance, and subsequently the NHC forecast, for Marco this afternoon. While at this point it's a little speculative, the data collected by this morning's NOAA G-IV flight in the environment around Marco and across the Gulf of Mexico may have played a key role in the significant eastward shift seen in nearly all the 12z models. This isn't to say that the uncertainty in the eventual track has diminished. In fact, various ensemble members from some of the global models still show a potential risk to the coast anywhere from Texas to Alabama, and it's entirely possible that the volatile shifts seen in the models could continue. That being said, the new NHC track forecast has been shifted significantly eastward and now shows the center of Marco reaching southeastern Louisiana in about 2 days, which is the scenario currently shown by the GFS, ECMWF, HCCA, Florida State Superensemble, and the TVCN multi-model consensus. After Marco reaches the coast, the western Atlantic ridge is expected to build westward and should cause the cyclone to move more slowly toward the west-northwest across southern portions of Louisiana.
Figure 3.10. ECMWF EPS (blue) and GEFS (magenta) mean track error every 6 h for forecasts initialized at 0000 UTC 22 August 2020 (left-facing triangle marker) and at 1200 UTC 22 August 2020 (circle marker) for Hurricane Marco.
Figure 3.11. The ensemble mean environmental steering flow vectors (kt) from the 0000 UTC 22 August 2020 (black barbs) and the 1200 UTC 22 August 2020 (blue barbs) ECMWF EPS forecast initializations for Hurricane Marco valid (a) 1200 UTC 22 August 2020, (b) 0600 UTC 23 August 2020, (c) 0000 UTC 24 August 2020, and (d) 1200 UTC 24 August 2020. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 0000 UTC 22 August 2020 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 1200 UTC 22 August 2020 forecast is denoted by the blue circle marker.
Figure 3.12. The magnitude of the vector difference in the near-TC steering flow (i.e., within 100 km of the TC position) (kts; red) and the distance between the ensemble mean position (km; black) for 6-h forecasts initialized at 0000 UTC 22 August 2020 and 1200 UTC 22 August 2020 for Hurricane Marco. Forecasts are valid from 1200 UTC 22 August 2020 to 0000 UTC 25 August.
Figure 3.13. The major axis correlation coefficient (MACC) for 6-h position forecasts initialized for Hurricane Marco at 0000 UTC 22 August 2020 (left-facing triangle marker) and 1200 UTC 22 August 2020 (circle marker). Forecasts are valid from 1200 UTC 22 August 2020 to 0000 UTC 25 August 2020.
Figure 3.14. As in Figure 3.7, but for Hurricane Zeta initialized at (a), (c) 1200 UTC 25 October 2020 and (b), (d) 0000 UTC 26 October 2020.
Figure 3.15. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Zeta beginning at 0000 UTC 26 October 2020 and ending at 1200 UTC 29 October 2020. The forecasts initialized at 1200 UTC 25 October 2020 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 26 October 2020 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 25 October 2020 and by black squares for the position starting at 0000 UTC 26 October 2020. The dropsonde deployment locations during the 1730 UTC 25 October 2020 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.
Figure 3.16. As in Figure 3.10, but for Hurricane Zeta initialized at 1200 UTC 25 October 2020 (left-facing triangle marker) and at 0000 UTC 26 October 2020 (circle marker).
Figure 3.17. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 25 October 2020 (black barbs) and the 0000 UTC 26 October 2020 (blue barbs) ECMWF EPS forecast initializations for Hurricane Zeta valid (a) 0000 UTC 26 October 2020, (b) 1800 UTC 26 October 2020, (c) 1800 UTC 27 October 2020, and (d) 1200 UTC 28 October 2020. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 25 October 2020 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 26 October 2020 forecast is denoted by the blue circle marker.
Figure 3.18. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 25 October 2020 and 0000 UTC 26 October 2020 for Hurricane Zeta. Forecasts are valid from 0000 UTC 25 October 2020 to 1200 UTC 28 October 2020.
Figure 3.19. The MACC 6-h position forecasts initialized for Hurricane Zeta at 1200 UTC 25 October 2020 (left-facing triangle marker) and 0000 UTC 26 October 2020 (circle marker). Forecasts are valid from 0000 UTC 26 October 2020 to 1200 UTC 28 October 2020.
Figure 3.20. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Dorian beginning at 0000 UTC 28 August 2019 and ending at 0000 UTC 02 September 2019. The forecasts initialized at 1200 UTC 27 August 2019 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 28 August 2019 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 27 August 2019 and by black squares for the position starting at 0000 UTC 28 August 2019. The dropsonde deployment locations during the 1730 UTC 27 August 2019 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.
Figure 3.21. As in Figure 3.7, but for Hurricane Dorian initialized at (a), (c) 1200 UTC 27 August 2019 and (b), (d) 0000 UTC 28 August 2019.
Figure 3.22. As in Figure 3.10, but for Hurricane Dorian initialized at 1200 UTC 27 August 2019 (left-facing triangle marker) and at 0000 UTC 28 August 2019 (circle marker).
Figure 3.23. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 27 August 2019 (black barbs) and the 0000 UTC 28 August 2019 (blue barbs) ECMWF EPS forecast initializations for Hurricane Dorian valid (a) 0000 UTC 28 August 2019, (b) 1800 UTC 28 August 2019, (c) 1800 UTC 29 August 2019, and (d) 1200 UTC 30 August 2019. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 27 August 2019 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 28 August 2019 forecast is denoted by the blue circle marker.
Figure 3.24. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 27 August 2019 and 0000 UTC 28 August 2019 for Hurricane Dorian. Forecasts are valid from 0000 UTC 28 August 2019 to 1200 UTC 30 August 2019.
Figure 3.25. The MACC 6-h position forecasts initialized for Hurricane Dorian at 1200 UTC 27 August 2019 (left-facing triangle marker) and 0000 UTC 28 August 2019 (circle marker). Forecasts are valid from 0000 UTC 28 August 2019 to 1200 UTC 30 August 2019.
Figure 3.26. As in Figure 3.7, but for Hurricane Jerry initialized at (a), (c) 1200 UTC 20 September 2019 and (b), (d) 0000 UTC 21 September 2019.
Figure 3.27. ECMWF EPS (blue) and GEFS (magenta) 6-h track forecasts for Hurricane Jerry beginning at 0000 UTC 21 September 2019 and ending at 1200 UTC 24 September 2019. The forecasts initialized at 1200 UTC 20 September 2019 are denoted by left-facing triangle markers, while the forecasts initialized at 0000 UTC 21 September 2019 are denoted by circle markers. The NHC best-track position every 6 h is denoted by a red square for the position at 1800 UTC 20 September 2019 and by black squares for the position starting at 0000 UTC 21 September 2019. The dropsonde deployment locations during the 1730 UTC 20 September 2019 NOAA G-IV mission are denoted by gray circle markers. The order of the dropsonde deployment locations is indicated by the black, numbered text within the gray circle markers.
Figure 3.28. As in Figure 3.10, but for Hurricane Jerry initialized at 1200 UTC 20 September 2019 (left-facing triangle marker) and at 0000 UTC 21 September 2019 (circle marker).
Figure 3.29. The ensemble mean environmental steering flow vectors (kt) from the 1200 UTC 20 September 2019 (black barbs) and the 0000 UTC 21 September 2019 (blue barbs) ECMWF EPS forecast initializations for Hurricane Jerry valid (a) 0000 UTC 21 September 2019, (b) 1200 UTC 21 September 2019, (c) 0600 UTC 22 September 2019, and (d) 0000 UTC 23 September 2019. The shaded contours represent the magnitude of the vector difference (kt) between the two initializations times at the given valid time. The ensemble mean position for the 1200 UTC 20 September 2019 forecast is denoted by the blue left-facing triangle marker, while the ensemble mean position for the 0000 UTC 21 September 2019 forecast is denoted by the blue circle marker.
Figure 3.30. As in Figure 3.12, but for 6-h forecasts initialized at 1200 UTC 20 September 2019 and 0000 UTC 21 September 2019 for Hurricane Jerry. Forecasts are valid from 0000 UTC 21 September 2019 to 1200 UTC 23 September 2019.
Figure 3.31. The MACC 6-h position forecasts initialized for Hurricane Jerry at 1200 UTC 20 September 2019 (left-facing triangle marker) and 0000 UTC 21 September 2019 (circle marker). Forecasts are valid from 0000 UTC 21 September 2019 to 1200 UTC 23 September 2019.
4. Conclusions

The impacts of early operational NOAA G-IV synoptic surveillance missions on TC track forecasts have been well documented (e.g., Aberson and Franklin 1999; Aberson 2002, 2008, 2010). While studies have since been published exploring the impacts of all NHC reconnaissance missions, including the G-IV missions, the most recent study solely on the impacts of G-IV missions was published in 2010 and included missions through the 2006 Atlantic hurricane season. Since 2006, multiple studies have been conducted analyzing the optimal targeting strategies of the G-IV missions (e.g., Majumdar et al. 2006; Peng and Reynolds 2006; Wu et al. 2007; Torn and Hakim 2008; Wu et al. 2009). More recently, the results of Ryan et al. (2018) found that dropsondes deployed in the near-TC environment led to the greatest improvements in position forecasts, leading to the permanent implementation of a circumnavigation of the TC core with dropsonde releases on all G-IV missions. Therefore, due to the lack of recent published analyses on the impacts of G-IV missions and the update to the G-IV targeting strategy, it is worthwhile to assess the recent impacts of the assimilation of G-IV dropsonde data on track forecasts.

This research quantifies the impacts of dropsonde data collected during NOAA G-IV synoptic surveillance missions on the position variability of Atlantic basin TCs from 2018–2021. ECMWF EPS, GEFS, and CLP5 position forecasts were utilized from 675 forecast initialization times, in which 56 had dropsonde data assimilated at the time of initialization. Two subsets of forecast initialization times were analyzed: those initialized with (56 initializations) and without (619 initializations) the assimilation of G-IV dropsondes, and those initialized 12 h before (56 initializations) and directly after (56 initializations) the assimilation of G-IV dropsondes. The average track forecast error and skill is computed and compared within each subset of forecast.
initialization times to quantify the impacts of the dropsonde data. In addition, the difference in average track forecast error and skill between the before and after dropsonde forecast initialization times is computed to identify any impacts on the forecasts following the assimilation of the dropsondes with respect to previous forecasts for the same TC. The FSP is also computed for the before and after dropsonde forecast initialization times to assess the frequency in which the forecasts after dropsondes have lower position errors than the forecasts 12 h before dropsondes, regardless of the amount of error reduction. Finally, a bootstrap resampling without replacement is utilized to determine the statistical significance of the results.

GEFS and ECMWF EPS track forecasts initialized with the assimilation of dropsonde data had lower average track error at all lead times (0–120 h), which increases with increasing lead time (Figure 3.2a). For GEFS forecasts at 12, 24, and 36 h, and ECMWF EPS forecasts at 12, 96, 108, and 120 h, fewer than 5% of randomly sampled forecast initializations from climatology have lower position errors at these forecast hours compared to forecast initializations with dropsondes. Although the statistical significance of the decreased average position errors in the forecast initialization times with dropsondes differs between the two operational models, this result ultimately supports the proposed hypothesis that adding dropsondes improves TC track forecasts. In contrast, the differences in average track forecast error between forecasts initialized 12 h before and with the assimilation of dropsondes are similar to those found in randomly selected pairs of forecast initializations from climatology (Figure 3.4b). Therefore, it appears that the impacts of dropsondes on position forecasts are most significant when compared to forecasts initialized without dropsondes from all TCs, rather than the forecast initialized 12 h before the dropsonde data is assimilated. This result suggests that the dropsonde data is most impactful to track forecasts.
when analyzed across multiple forecasts or seasons, as opposed to individual forecasts for the same TC, which is consistent with Aberson (2008).

The impacts of the dropsonde data on forecast skill relative to CLP5 are similar to those for average track error, such that there is an improvement in forecasts with dropsondes compared to those without dropsondes. From 2018–2021, the ECMWF EPS and GEFS both exhibit a 10% increase in track forecast skill relative to CLP5 for forecasts initialized with dropsondes (Figure 3.3). This increase in skill for forecasts initialized with dropsondes is statistically significant at all forecast hours for GEFS forecasts and at 48, 96, 108, and 120 h for ECMWF EPS forecasts. This result suggests that the dropsonde data from NOAA G-IV missions are likely adding skill to position forecasts, particularly in GEFS forecasts, and supports the hypothesis that forecast initializations with dropsondes have a higher degree of skill than those without dropsondes.

Four case studies were conducted in an effort to diagnose the potential sources of significant position error reductions in individual track forecasts due to the assimilation of dropsondes. The four cases studies, described by mission take-off data and time, were: 0530 UTC 22 August 2020 for Hurricane Marco, 1730 UTC 25 October 2020 for Hurricane Zeta, 1730 UTC 27 August 2019 for Hurricane Dorian, and 1730 UTC 20 September 2019 for Hurricane Jerry. For each case study, the track and environmental steering flow of the ECMWF EPS track forecast 12 h before and immediately following a NOAA G-IV synoptic surveillance mission were compared. It was hypothesized that large error reductions in individual track forecasts are the result of changes in the environmental steering flow driving the TC’s motion, a change in the initial position of the TC that placed the TC into a different steering flow regime, or a combination of both.

The results of the case studies can be divided into two categories based on which aspect of the hypothesis the results aligned with. First, the reduction in track error seen in Hurricane Marco
and Hurricane Dorian in the forecasts initialized with dropsondes is likely to be the result of an initial position change in the ECMWF EPS forecast initialization with dropsondes. The initial position change in the 12000 UTC 22 August ECMWF EPS model run for Hurricane Marco placed the storm on the opposite side of the deformation axis in the Gulf of Mexico steering flow, resulting in a different near-TC steering flow pattern and the subsequent eastward shift in Marco’s track toward the observed TC position. Similar results were seen for Hurricane Dorian. The initial position change in the 0000 UTC 28 August ECMWF EPS model run placed Dorian closer to the subtropical ridge and to the observed location of the storm after its center reformation that occurred at the same time. The northwestward shift in Dorian’s initial position led to the storm experiencing a different near-TC steering flow and a corresponding reduction of error in the position forecast.

The second category of case study results pertains to the reduction of track error in Hurricane Zeta and Hurricane Jerry. The reduction in track error seen in Hurricane Zeta and Hurricane Jerry is likely the result of a change in the environmental steering flow structure driving the motion of the TC in the ECMWF EPS forecast initialization with dropsondes with respect to the forecast initialization 12 h before dropsondes. The 0000 UTC 26 October ECMWF EPS model run for Hurricane Zeta exhibited a decrease in in the speed of the environmental steering flow, resulting in a decrease in the storm’s motion in after 24 h. This decrease in forward motion brought the forecast track closer to the observed TC track, resulting in a reduction in position error in the forecast initialization with dropsondes compared to the forecast initialization 12 h before dropsondes. Similarly, the 0000 UTC 21 September ECMWF EPS model run for Hurricane Jerry exhibited changes in the direction of the environmental steering flow related to the nearby subtropical high. The directional changes in the environmental steering flow in the forecast initialization with dropsondes led to Jerry curving around the southwestern edge of the subtropical
ridge earlier than in the previous model run, as well as bringing the forecasted track closer to the observed track.

There are a few caveats related to the methodology employed in this thesis which are worth considering in future work. The first, and likely most important caveat relates to the decision to compared forecasts initialized with dropsondes to those initialized without or 12 h before dropsondes. Previous studies on the impacts of G-IV dropsonde data on TC track forecasts utilized two sets of forecasts: one set of forecasts initialized with dropsonde data, and the same set of forecasts but without the assimilation of the dropsonde data (e.g., Aberson and Franklin 1999; Aberson 2002, 2008, 2010). While this approach allows for changes in the track forecasts to be directly attributed to the assimilation of dropsonde data, it requires having the resources to rerun operational models without the assimilation of the dropsonde data. This approach is computationally expensive and was outside the scope of this thesis. That being said, repeating the multi-season analysis and the case studies using forecasts generated by rerunning the operational models without the assimilation of dropsondes would allow for a more robust set of conclusions of the impacts of G-IV dropsondes and would be of interest for future work on this subject. In addition, this thesis did not address other potential sources of differences in forecasts initialized 12 h before dropsondes and with dropsondes, as other types of observations could have also contributed to the differences in average error and skill. Finally, this thesis did not consider the operational models themselves when conducting the multi-seasonal analysis. This caveat is particularly applicable for the GEFS, as there was a major update to the model in September 2020, halfway through the analysis period, which increased the number of ensemble members and replaced the Global Spectral Model with the global FV3 dynamical core (NOAA 2020). Future work may consider separating the GEFS forecast initializations before and after this update to
address changes in the model. Furthermore, it may be worth investigating whether the dropsonde data has a greater impact on the ECMWF EPS or GEFS forecasts, and why that may be.
REFERENCES


Magnusson, L., and Coauthors, 2021: Tropical cyclone activities at ECMWF. ECMWF Technical Memo. 888, ECMWF, 140 pp., Reading, United Kingdom.


