A Comparison of Arctic and Atlantic Cyclone Predictability

PEYTON K. CAPUTEa AND RYAN D. TORNa

*Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, New York

ABSTRACT: Arctic cyclones (ACs) are synoptic-scale features that can be associated with strong, intense winds over the Arctic Ocean region for long periods of time, which in turn lead to rapid declines of sea ice during the summer. As a consequence, sea ice predictions may rely on the predictability of cyclone-related wind speed and direction, which critically depends on the cyclone track and intensity. Despite this, there are relatively few studies that have documented the predictability of ACs during the summer, beyond a few case studies, nor has there been an extensive comparison of whether these cyclones are more or less predictable relative to comparable midlatitude cyclones, which have been studied in greater detail. The goal of this study is to document the practical predictability of AC position and intensity forecasts over 100 cases and compare it with 89 Atlantic Ocean basin midlatitude cyclones using the Global Ensemble Forecast System (GEFS) Reforecast V2. This dataset contains 11-member ensemble forecasts initialized daily from 1985 to the present using a fixed model. In this study, forecasts initialized 1 and 3 days prior to the cyclone development time are compared, where predictability is defined as the ensemble mean root-mean-square error and ensemble standard deviation (SD). Although Atlantic basin cyclone tracks are characterized by higher predictability relative to comparable ACs, intensity predictability is higher for ACs. In addition, storms characterized by low ensemble SD and predictability are found in regions of higher baroclinic instability than storms characterized by high predictability. There appears to be little, if any, relationship between latent heat release and precipitable water and predictability.

SIGNIFICANCE STATEMENT: The purpose of this study is to compare the position and intensity uncertainty of Arctic and Atlantic basin cyclones, since Arctic cyclone uncertainty has been less studied. A comparable number of long-duration, intense, Arctic and Atlantic cyclones were compared to find that on average, Atlantic cyclone intensity is more uncertain than that of Arctic cyclones, whereas Arctic cyclone position is more uncertain than that of Atlantic cyclones. It was also found that the most uncertain cyclones in both basins are in regions associated with larger horizontal temperature differences. Future studies will examine what synoptic-scale features limit Arctic cyclone position uncertainty.

KEYWORDS: Arctic; Extratropical cyclones; Ensembles; Numerical weather prediction/forecasting

1. Introduction

Arctic cyclones (ACs) are synoptic-scale features that can be associated with strong, intense winds over the Arctic Ocean region for long periods of time, which in turn lead to rapid declines of sea ice (Simmonds and Rudeva 2012; Boisvert et al. 2016; Woods and Caballero 2016; Binder et al. 2017). Over the past 40 years, Arctic sea ice extent has decreased in coverage by over 40% during the summer months (Stroeve et al. 2012; Serreze and Stroeve 2015; Simmonds 2015; Koyama et al. 2017); however, during the summer months these cyclones originate most frequently over Eurasia (Serreze 1995). The high frequency of ACs over Eurasia is possibly linked to a reduction in static stability associated with turbulence and radiative heat fluxes from heated landmass, in addition to northern Eurasia topography (Serreze 1995). AC duration and intensity has been found to be highest during the summer months (Zhang et al. 2004), which can play a significant role in sea ice reduction. One example of a long-lived event is the great AC of August 2012 (AC12; Simmonds and Rudeva 2012). This system resulted from a midlatitude cyclone that formed over Siberia, merged with an AC that originated over the East Siberian Sea, and subsequently intensified in the Arctic region (Yamagami et al. 2018b). AC12 had a duration of approximately 10 days and is believed to have been responsible for widespread sea ice loss (Parkinson and Comiso 2013; Crawford and Serreze 2016; Petty 2018; Yamagami et al. 2018b).

The predictability of AC12 has been examined in-depth, notably by Yamagami et al. (2018b) who used operational medium-range ensemble forecasts to determine that position variability is much larger than intensity variability. They concluded that the predictability of this event is particularly low,
such that accurate forecasts of the cyclone position and intensity were not available until 2–3 days prior to its maximum intensity (Yamagami et al. 2018b). The lack of predictability for this case does not seem to be unique; medium-range forecast skill for 10 “extraordinary” ACs during the summer months from 2008 to 2016 was found to be lower in multiple global modeling systems, measured via cyclone position and intensity, than that for midlatitude cyclones in the Northern Hemisphere (Yamagami et al. 2018a). Furthermore, Sandu and Bauer (2018) suggest that forecast skill for medium-range forecasts in the Arctic is lower than in midlatitudes based on 6-day 500-hPa geopotential height anomaly correlation scores. Although there have been case studies for AC predictability, there has yet to be a comprehensive climatological study that evaluates AC track and intensity predictability against cyclones in other parts of the globe.

Even though there are relatively few studies that have focused on AC predictability, there is a comparatively large body of literature on midlatitude cyclone predictability. In addition to numerous case studies (e.g., Sanders et al. 2000; Hacker et al. 2003; Zhu and Thorpe 2006), there have been several studies on the predictability of midlatitude cyclones (e.g., Leary 1971; Froude 2010; Zhang et al. 2019). An early-winter midlatitude storm analysis found that the National Meteorological Center model did not forecast intensity deep enough over the ocean, Rocky Mountain lee cyclones were forecast to be too intense and too warm, and very intense cyclones were forecast to the right of the observed track (Leary 1971). Similar results were found by Silberberg and Bosart (1982) using the NMC Limited Area Fine Mesh Model. A more recent analysis of Northern Hemisphere extratropical cyclone predictability found that there is little variation between nine ensemble prediction systems for cyclone position as compared with intensity (Froude 2010). In addition, there have been multiple studies that have sought to understand how uncertainty in synoptic features may impact midlatitude cyclone predictability. For example, midlatitude cyclone forecasts have been found to be sensitive to the motion of upper tropospheric features and Rossby wave packets (e.g., Langland et al. 2002; Chang et al. 2013; Grazzini and Vitart 2015), potential vorticity gradients, (e.g., Hakim 2005), lower tropospheric baroclinicity and rate of latent heat release (LHR) (e.g., Anthes et al. 1983; Kuo and Reed 1988; Joos and Wernli 2012; Binder 2016), as well as resulting in a reduction in downstream predictability (e.g., Torn 2017). As a consequence, Atlantic cyclone predictability is fairly well-understood relative to ACs, which in turn makes them a good comparison to help to better understand AC predictability.

Although Atlantic basin cyclones make a good comparison for AC predictability, there are many known differences between Arctic and Atlantic basin midlatitude cyclone environments and dynamics. A main difference is that ocean waters in the Atlantic basin are much warmer than those in the Arctic, which allows for extraction of latent and sensible heat and hence allows storms to intensify more rapidly. On the contrary, ACs are able to sustain themselves for longer period of times compared to midlatitude cyclones due to the ample supply of vorticity via the upper-level polar vortex (Tanaka et al. 2012). Such ACs were found to originate from the merging of small-scale cyclones, rather than through baroclinic instability for midlatitude cyclones. These small-scale cyclones can form from preexisting tropospheric waves that develop along the Arctic frontal zone (AFZ) during the summer (Serreze and Barrett 2008). In addition, the tracks of the ACs studied in Tanaka et al. (2012) were found to follow the movement of the upper level polar vortex at 300 hPa, resulting in less predictable motion, unlike Atlantic basin cyclones that generally follow the North Atlantic jet stream at 500 hPa during the boreal winter. These ACs were found to have cold cores associated with upward motion from the surface to the lower stratosphere and a warm core associated with downward motion above the lower stratosphere, another unique feature in comparison with midlatitude cyclones.

The purpose of this study is to compare the predictability of long-lived, intense midlatitude cyclones within the western Atlantic basin with cyclones in the Arctic Ocean. Here, practical predictability will be measured by evaluating the ensemble mean error and ensemble standard deviation (SD) from the Global Ensemble Forecasting System Reforecast, version 2. Practical predictability uses realistic uncertainties in both the forecast model and initial and boundary conditions when assessing predictability (Lorenz 1969; Zhang et al. 2006). The Atlantic basin could have more dynamical factors that might promote position variability and fast cyclone intensification; however, the Arctic basin is not as well observed as the Atlantic, so there is likely more analysis uncertainty in the Arctic (Bromwich and Wang 2005). Therefore, the hypothesis tested in this study is that the predictability of AC track and intensity is indistinguishable from comparable midlatitude cyclones. In addition, it is expected that cyclones characterized by large SD are found in regions of large baroclinic instability and LHR, as compared with cyclones characterized by low SD. The remainder of this study is outlined as follows. Section 2 covers the data and methods used to create a climatology of intense, long-lived cyclones, cyclone tracking, and statistical calculations performed in this study. Section 3 compares Atlantic and AC predictability and examines a few large-scale parameters for cases characterized by large and small SD. Section 4 provides a summary of the results and conclusions.

2. Data and methods

This study compares the predictability of Arctic and Atlantic basin cyclones during two different seasons over a long period of time. For the Arctic, the focus is on summer months (June–August) so as to capture cyclones that may be responsible for sea ice break up (e.g., Crawford and Serreze 2016). By contrast, extended winter month (November–March) Atlantic cyclones are examined, as cyclones are most frequent in this region during this time of the year, and because they have been extensively studied (e.g., McCabe et al. 2001). Comparing cyclones for these two seasons is justified because baroclinic environments are similar on average during the Arctic summer and Atlantic basin winter, measured by the mean Eady growth rate (EGR) (not shown). This occurs as baroclinic instability shifts poleward in the Northern Hemisphere from winter to...
summer, associated with the seasonal shift of the polar front jet stream and storm tracks (Hoskins and Hodges 2002, 2019). In addition, the presence of the AFZ during summer months contributes to enhanced baroclinic instability in the Arctic region (Crawford and Serreze 2016).

A subset of cyclones in both regions are selected from the Sprenger et al. (2017) cyclone climatology, which makes use of 6-hourly ERA-Interim data from 1985 to 2016. The Atlantic basin was chosen to represent midlatitude cyclones as they are well studied and allow for similar case numbers to compare. Atlantic cyclones are selected if the cyclones’ reference time occurred within the region characterized by the highest storm track density in the basin (denoted in Fig. 2 by the black polygon). The reference time, denoted as \( t_N \), is defined as the time step when the difference between the environmental pressure and cyclone minimum sea level pressure (SLP), referred to as pressure depression, first exceeds 12 hPa. This value represents the largest depression that occurred for all cyclones in both the Atlantic and Arctic basins; therefore, it represents a uniform benchmark for cyclones in both basins, since the environmental pressure varies between basins. The reference time occurred most commonly during the developing stage of the selected cyclones. All forecasts from 5 days prior to the reference time up to the reference time were computed and show consistent results to what is presented here, except for -4-day position variability that shows slightly higher SD for the Atlantic (not shown). However, for brevity, only forecasts initialized 1 and 3 days prior to \( t_N \) are shown in this study. For the Arctic, a cyclone was considered here if at least 80% of the 6-hourly positions lie north of 70°N. This condition was used in order to capture cyclones that develop in or near the Arctic region and spend a majority of their life cycle in the Arctic Ocean, which could in turn impact Arctic sea ice. In addition, this condition was used to eliminate cyclones that originate in the midlatitudes and subsequently move into the Arctic region, often during cycloysis (AC tracks are shown in Fig. 2).

The focus of this study is on long-duration, intense cyclones, which are more likely to impact Arctic sea ice. As a consequence, the cyclone must have a track within the Sprenger dataset that lasts at least 3 days (twelve 6-h time steps). Here, the most intense cyclones in each basin are identified using the accumulated pressure depression (APD), which is calculated as

\[
APD = \sum_{t=1}^{N_{\text{max}}} (P_{\text{environment}} - P_{\text{central}}),
\]

where \( P_{\text{central}} \) is the pressure center and \( P_{\text{environment}} \) is the pressure of the last closed contour on a 1-hPa interval. The sum is calculated every 6 h over the duration of each storm in the Sprenger dataset. This metric is similar to the accumulated cyclone energy, or ACE, metric, which is used to measure the time-integrated tropical cyclone activity over a storm and/or one season (e.g., Zhan and Wang 2016; Davis and Zeng 2019). For this study, only cyclones that exceed the 85th percentile of APD for each respective basin are used (Fig. 1). The 85th percentile was chosen as a similar standard to capture the strongest storms in each basin without having to account for the more intense cyclones in the Atlantic basin. For the Arctic the 85th percentile is equivalent to an APD of 250 hPa, yielding 100 cyclones, whereas in the Atlantic the 85th percentile is equivalent to an APD of 450 hPa, yielding 89 cyclones (tracks shown in Fig. 2).

The predictability of Arctic and Atlantic basin cyclones is evaluated by averaging ensemble mean errors and ensemble forecast SD as a function of lead time for a comparable number of storms. Specifically, the Global Ensemble Forecasting System (GEFS) Reforecast, version 2, 1° 6-hourly dataset (Hamil11 et al. 2013) is employed, which consists of 11 ensemble members (control + 10 perturbed members) initialized each day at 0000 UTC during 1985–2016. The GEFS reforecast employs the GEFS ensemble prediction system that went into operations in February 2012. Until 2011, the ensemble mean initial condition came from the Climate Forecast System Reanalysis (Saha et al. 2010). After that, the ensemble mean analysis is the operational GFS analysis. Initial condition perturbations determined using the ensemble transform technique with rescaling (Wei et al. 2008). The interested reader is directed to Hamill et al. (2013) for additional details about this system.

The GEFS reforecast dataset was chosen because it provides 89 Atlantic and 100 Arctic cases from which to test this hypothesis about AC predictability. Furthermore, unlike a typical operational forecasting system, there have been minimal changes to the model over this time period, which eliminates the possibility that the results could be impacted by the continuous operational forecast model improvement that have occurred over this 30-yr period. However, note that increases in observations over time may influence forecast skill. In addition, observations are sparser in the Arctic than in the Atlantic basin (Bromwich and Wang 2005).

Cyclones are tracked within each ensemble member and each initialization time that occurs 5 days prior to and up to the reference time by finding the maximum in 925-hPa
area-averaged vorticity. GEFS reforecasts are only initialized at 0000 UTC; therefore, cyclones with a reference time at 0600 or 1200 UTC are shifted to the closest initialization time prior and cyclones with a reference time of 1800 UTC are shifted to the closest initialization time after, similar to Torn (2017).

In this study, Arctic and Atlantic cyclone predictability is evaluated by considering both the ensemble mean root-mean-square error (RMSE) and ensemble SD in cyclone position and minimum SLP, with higher error and larger SD indicating lower predictability. Average ensemble RMSE and SD is calculated using the number of available storms at each time step (numbers shown along the top of Fig. 3). The statistical significance of the difference in the SD and RMSE is determined via the t test every 6 h. The GEFS position and minimum ensemble SD is found to be smaller than the RMSE, signifying that the ensemble is underdispersive (Fig. 3). Nevertheless, the GEFS ensemble SD appears to be able to distinguish between low- and high-error cases. One way to demonstrate this is to employ an error–spread Pearson correlation skill score, which was introduced by Hopson (2014). This metric compares the correlation between individual forecast ensemble mean errors and SD with what is expected for a perfect ensemble prediction system, given the variation in the forecast SD present in the ensemble system. This metric is defined as

$$SS_j = \frac{r_{fores}}{r_{perf}} \tag{2}$$

where \(r_{fores}\) is the correlation between the ensemble mean error and SD for all cases at a given forecast hour, and \(r_{perf}\) is the correlation that would come from a perfect EPS system, given the variability in the ensemble SD [see Hopson (2014) for a derivation of this]. Here, a value of 1 indicates that the correlation between ensemble mean error and SD is equivalent to what one would expect for a perfect ensemble prediction system with a given range of SD.

Figure 4 shows this skill score for Atlantic and Arctic cyclones as a function of hours since \(t_o\) for forecasts initialized 1 and 3 days prior to \(t_o\). The position forecast skill scores appear to exceed 0.6 for most times, except for after 36 h for both Atlantic and Arctic cyclones for the −1-day forecasts (Figs. 4a,b). Furthermore, the

\[\text{FIG. 2. Cyclone tracks each 6 h from the Sprenger et al. (2017) ERA-Interim cyclone climatology for the 100 Arctic (blue) and 89 Atlantic basin (red) cyclones from 1985 to 2016 used in this study. The first time step for each cyclone is denoted by a white circle. Longitude lines are plotted every 30°, and latitude is plotted every 10°. The black-outlined polygon represents the region of maximum Atlantic cyclones.}\]
The cyclone minimum SLP skill score is also generally above 0.6 for all lead times, except for Atlantic basin cyclones prior to 18 h (Figs. 4c,d). Regardless, this calculation suggests that the ensemble SD does provide a relatively skillful estimate of the error at most lead times and hence can be used as a complementary measure of predictability in this study along with the ensemble mean error.

Cyclone position and minimum SLP increase with lead time and have different units; therefore, a standardized SD is calculated to determine track and intensity ensemble variability in both basins. It is given by

$$\text{standardized SD} = \frac{SD_{\text{storm}} - SD}{SD}$$  \hspace{1cm} (3)

Here, $SD_{\text{storm}}$ is the SD in either position or minimum SLP for each individual storm and $SD$ is the average SD in position or minimum SLP each basin. From this calculation, the top quartile of storms can be identified, which in turn represents the most variable storms, and vice versa. The statistical significance of the difference between these quartiles is determined using bootstrap resampling without replacement. This method proceeds by randomly selecting two subsets of unique cases from each basin and running 10,000 times to determine the probability of obtaining the difference between the Arctic and Atlantic RMSE and SD cyclone values, respectively, via the $t$ test.

There are several potential mechanisms that could explain the difference between Arctic and Atlantic cyclones that are relatively more or less predictable. Based on previous studies, these processes could include both dry and moist dynamics, which can be measured using baroclinic instability and LHR. To evaluate the potential role baroclinic instability and LHR play in the predictability of these storms, area-averaged EGR and a proxy of column-integrated LHR, respectively, are evaluated from ERA-Interim reanalysis data, which is the same dataset used in the Sprenger cyclone climatology. Using the 6 hourly ERA-Interim dataset that was converted to a 1° fixed grid, these parameters are computed for the most and least predictable storms in each basin. Here, EGR is used as a measure of baroclinity and is computed via

$$EGR = 0.31 \left( \frac{f}{N} \right) \frac{\partial \mathbf{v}}{\partial z}$$  \hspace{1cm} (4)

where $f$ is planetary vorticity, $N$ is the Brunt–Väisälä frequency, $z$ is the vertical coordinate, and $\mathbf{v}$ is the vector
horizontal wind calculated between 850 and 400 hPa (Hoskins and Valdes 1990). Area-averaged EGR has been used to examine the role of surface baroclinicity for an Arctic storm case study by Tao et al. (2017). Baroclinic instability, or EGR, can contribute to the maintenance of a cyclones’ intensity. Therefore, it is important to determine the role EGR plays in the predictability of these cyclones.

Unfortunately, the ERA dataset does not explicitly contain an estimate of LHR; therefore, the residual of the column-integrated water budget equation is used as a proxy for the column-integrated LHR. Specifically, the column-integrated LHR is computed via

$$C - E = \text{LHF}_{\text{ave}} - \frac{1}{g} \frac{\partial}{\partial t} \int_S q \, dp - \int_S \left( \frac{\partial u q}{\partial x} + \frac{\partial v q}{\partial y} \right) dp,$$

where LHF_{ave} is surface latent heat flux, the second term is the Eulerian derivative of vertically integrated water vapor, and the third term is vertically integrated horizontal moisture flux convergence. The left-hand side of this equation is condensation C minus evaporation E, which is used as a proxy for column-integrated LHR. For this study, area-averaged EGR and LHR are computed for each time step within a 700-km radius. Cyclones are characterized by upward and downward motion and hence latent heat releases in different locations within the cyclone, which may not be consistent from across all storms. As a consequence, an area average over a circle centered on the storm may not represent the mean LHR. Furthermore, ACs are often characterized by erratic motion; therefore, there is no reasonable method for computing area-averaged LHR, either in an cyclone-relative framework, or even in a framework aligned with the motion of the cyclone. As a consequence, the area-averaged LHR metric only considers points that are above the 50th percentile of LHR within 700 km of the cyclone center, which in turn limits the metric to grid points where positive LHR might take place. For each cyclone, the 50th percentile is determined by only considering the grid points within 700 km for that particular case. For consistency, this method was applied to both Arctic and Atlantic basin cyclones. Similarly, this method is used to calculate the amount of moisture associated with Arctic and Atlantic basin cyclones by area-averaging precipitable water (PW).

3. Results

The aforementioned criteria yields 100 summer Arctic and 89 winter Atlantic cyclones, whose tracks are shown in Fig. 2. A majority of summer ACs originate over Eurasia north of 60°N, consistent with Serreze (1995), and subsequently spend most of their existence over the Arctic region, often exhibiting erratic motion. By contrast, the Atlantic cyclones selected for this study originate over North America or off the east coast of the

![Fig. 4. Average (a),(b) position and (c),(d) minimum SLP skill score SSr for all 100 Arctic (blue) and 89 Atlantic (red) cyclones as a function of time relative to \( t_0 \), for forecasts initialized (left) 1 day and (right) 3 days prior to \( t_0 \).]
United States, generally move to the northeast to the east of Greenland, and often enter the Arctic at the end of their life when they undergo cyclolysis. These Atlantic storm tracks are similar to those found in previous extratropical cyclone climatologies off the east coast of North America such as Grise et al. (2013) and Bentley et al. (2019).

For this study, the average ensemble mean RMSEs and ensemble SDs of minimum SLP and position are used to test the hypothesis that AC track and intensity predictability is indistinguishable from comparable Atlantic basin cyclones. To be expected, the RMSE and ensemble SD in cyclone position and minimum SLP increases with forecast lead time for forecasts initialized 1 and 3 days prior to the reference time (Fig. 3). For ACs, forecasts initialized 1 day prior to $t_o$ are characterized by 40-km-higher position errors and 25-km-higher position SD on average relative to the Atlantic basin; however, for most times the differences are not statistically significant (Fig. 3a). In addition, AC minimum SLP SD is initially approximately 0.25-hPa higher than Atlantic cyclones; however, beginning 12 h after $t_o$, Atlantic basin intensity SD increases at a faster rate through the remainder of the forecast, such that the Atlantic basin cyclones have a 0.7-hPa-higher SD than ACs by 30 h after $t_o$ (Fig. 3c). In addition, the RMSE for Atlantic basin cyclones is nearly identical to the Arctic at $t_o$ but increases to 1 hPa by 18 h. Forecasts initialized 3 days prior to $t_o$ show qualitatively similar results, but with larger magnitude differences between the basins (Figs. 3c,d). AC position SD remains on average 52 km higher than the Atlantic basin cyclones for all forecast lead times while the RMSE is approximately 100 km greater (Fig. 3b). Moreover, Atlantic basin cyclone intensity RMSE and SD increases to 10 and 7.5 hPa, respectively, by 36 h, while the corresponding maximum AC intensity error and SD are 7 and 4 hPa, respectively (Fig. 3d). The combination of these results suggest that, on average, Atlantic basin cyclone tracks are more predictable than comparable AC tracks, particularly for longer lead times, and vice versa for intensity. This finding is consistent with Yamagami et al. (2018b) who found that AC position variability is higher than intensity variability.

One potential explanation for these results is that cyclone motion and intensity are distinct between these two basins. This hypothesis is evaluated by computing the average intensity of each cyclone at each time step relative to $t_o$ within the Sprenger database. Furthermore, the average speed is calculated via the distance each cyclone traveled between each 6 h time step within the Sprenger database. During the 48 h following $t_o$, the average minimum SLP decreases from approximately 995 to 968 hPa in the Atlantic basin and from 989 to 985 hPa in the Arctic, indicating that Atlantic basin cyclones become 18 hPa deeper than ACs (Fig. 5a). As a consequence, Atlantic basin cyclones are more likely to intensify, which in turn could provide more potential variability in cyclone minimum SLP in comparison with ACs, which tend to maintain a similar intensity throughout their life. Despite this relationship, there does not seem to be a strong linear relationship between the intensification rate and the ensemble SD on a case-by-case basis. The correlation between the 0–48-h change in mean sea level pressure (MSLP) and the 48-h ensemble SD is $-0.28$ and $-0.34$ for the Arctic and Atlantic basins, respectively, meaning that cyclones with larger decreases in MSLP are associated with higher ensemble SD. Further, the speed of Atlantic basin cyclones is approximately 2 times as fast as comparable ACs (Fig. 5b), with the average speed for Atlantic cyclones being 52.5 km h$^{-1}$, as compared with 27.6 km h$^{-1}$ for ACs. Despite this, Atlantic basin cyclone position may be easier to predict, and hence have lower position variability, based on the speed of the cyclone and location of being within the Atlantic storm track, whereas there is no comparable storm track that regulates the steering flow in the Arctic region.

The remainder of this section examines the environmental conditions that are associated with cases characterized by high variability in cyclone intensity in each basin. For each basin, the least and most predictable storms are identified based on the time average standardized SD [Eq. (3)]. Here, the most-predictable cases are those that fall within the lowest quartile of standardized minimum SLP for that particular initialization time relative to $t_o$, whereas the least predictable are those that fall into the upper quartile of standardized minimum SLP (Fig. 6). For forecasts initialized 1 day prior to $t_o$, average SD for the 25th-percentile cases vary slightly, such that the minimum SLP SD increases from 1 hPa to 1.5 and 2 hPa over the forecast for Arctic and Atlantic cyclones, respectively (Fig. 6a). By contrast, the average SD for the 75th-percentile cases exhibit higher SD for AC than the Atlantic from $t_o$ to 12 h (Fig. 6a). After 12 h, the SD for the Atlantic cases increases to above 5-hPa SD, while the Arctic SD levels off just above 4 hPa (Fig. 6a). As a consequence, the SD is 3 hPa higher for the 75th percentile than for the 25th percentile in the Atlantic, whereas the difference is 2.5 hPa in the Arctic. A similar trend is present.
in the forecasts initialized 3 days prior to $t_m$, with approximately a 1.7-hPa SD difference between the Arctic and Atlantic 25th-percentile cases at 48 after $t_m$, and a 4.5-hPa SD difference between the 75th-percentile cases starting at 24 after $t_m$ (Fig. 6b). These results, combined with the spread–error skill score presented earlier, indicate that the standardized SD method can be used to distinguish between most and least variable cases in both basins based on minimum SLP (Fig. 6).

Furthermore, the results from the 75th-percentile ACs are comparable to errors found by Yamagami et al. (2019), who examined 26 “extraordinary” ACs. A comparison of the average minimum SLP and speed using Sprenger et al. (2017) data for the most and least variable storms in each basin shows that the most variable storms tend to have a lower minimum SLP and faster speeds for forecasts initialized at both 1 and 3 days prior to $t_m$ (Table 1). In addition, least predictable storms are characterized by a larger minimum SLP tendency on average in both basins, where tendency is defined as the difference in minimum SLP between the first time step and the minimum value and divided over the time each cyclone took to reach maximum intensity (not shown).

To determine whether baroclinic instability and LHR play a role in the predictability of these storms, area-averaged EGR and column-integrated $C^2P$, respectively, are calculated for the most- and least-predictable storms shown in Fig. 7. In both the Arctic and Atlantic basin, less-predictable cyclones are characterized by higher EGR for forecasts initialized 1 day prior to $t_m$ (Fig. 7a). The differences between the most- and least-variable storms are statistically significant (99% level) in both basins, particularly for the first 12 h after $t_m$ in the Arctic region and between 24 and 36 h after $t_m$ in the Atlantic basin for forecasts initialized 1 day prior to $t_m$ (Fig. 7a) and for the first 12 h after $t_m$ in both basins for forecasts initialized 3 days prior to $t_m$ (Fig. 7b). For the least-predictable storms, average EGR decreases from 0.91 to 0.60 day$^{-1}$ for ACs and from 1.04 to 0.8 day$^{-1}$ for Atlantic cyclones (Fig. 7a). For most-predictable storms, average EGR decreases from 0.73 to 0.6 day$^{-1}$ for ACs and from 0.96 to 0.69 day$^{-1}$ for Atlantic cyclones (Fig. 7a). A qualitatively similar result is obtained when repeating the calculation for forecasts initialized 3 days prior to $t_m$ (Fig. 7b).

As a consequence, this result suggests that less-predictable cyclones typically occur in environments characterized by greater baroclinic instability.

Whereas less-predictable cyclones are characterized by a clear difference in baroclinicity, the result for LHR is less clear. In the Atlantic basin, the less-predictable cyclones are characterized by higher LHR for forecasts initialized 1 day prior to $t_m$ (Fig. 7c). By contrast, the 75th-percentile ACs have higher LHR values only through 36 h after $t_m$, and the differences are not statistically significant for a consistent period of time. Forecasts initialized 3 days prior to $t_m$ show similar results, but with smaller discrepancies between the 75th- and 25th-percentile cases (Fig. 7d). These results indicate that higher LHR values are not associated with systematic differences in intensity predictability beyond the Atlantic cyclones.

Results from PW are similar to LHR such that forecasts initialized 1 day prior to $t_m$ show larger amounts of PW, particularly in the first 24 h since $t_m$ with statistically significant differences in the Atlantic basin in the first 12 h (Fig. 7e). After 24 h, PW is larger for the 25th-percentile cases in both basins. Similarly, −3-day forecasts demonstrate larger PW for 75th-percentile cases in the first 12 h in the Atlantic basin and 6 h in

![FIG. 6. Mean ensemble SD in minimum SLP for least-predictable (75th percentile; long dashed) and most-predictable (25th percentile; short dashed) storms for forecasts initialized (a) 1 day and (b) 3 days prior to $t_m$.](unavailable)

<table>
<thead>
<tr>
<th>Basin/forecast</th>
<th>75th-percentile min SLP (hPa)</th>
<th>25th-percentile min SLP (hPa)</th>
<th>75th-percentile speed (km h$^{-1}$)</th>
<th>25th-percentile speed (km h$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arctic/1 day</td>
<td>984.4</td>
<td>988.5</td>
<td>29.7</td>
<td>29.5</td>
</tr>
<tr>
<td>Atlantic/1 day</td>
<td>973.7</td>
<td>979.9</td>
<td>64.1</td>
<td>48.1</td>
</tr>
<tr>
<td>Arctic/3 days</td>
<td>983.0</td>
<td>988.2</td>
<td>63.3</td>
<td>47.6</td>
</tr>
<tr>
<td>Atlantic/3 days</td>
<td>972.0</td>
<td>983.9</td>
<td>26.9</td>
<td></td>
</tr>
</tbody>
</table>
the Arctic basin but remain lower than the 25th-percentile cases for the remainder of the forecast (Fig. 7f); therefore, it appears that higher PW values are not associated with systematic differences in intensity predictability.

The same standardized SD analysis was done to identify the most- and least-variable cases based on forecast position variability (Fig. 8). The 75th- and 25th-percentile cases for the 1-day forecast show similar position SD for Atlantic and ACs, with approximately a 100-km SD difference between these two percentiles (Fig. 8a). However, the 3-day forecast exhibit large SD differences between the 75th-percentile Arctic and AC basins and similar SD for the 25th-percentile cases (Fig. 8b). At 48 h since \( t_0 \), Atlantic position SD has a maximum difference of about 360 km, whereas Arctic position SD has a maximum difference of about 550 km at 36 h since \( t_0 \) (Fig. 8b).

Unlike the most- and least-predictable cases based on minimum SLP forecasts, most- and least-predictable cases based on position do not exhibit substantial differences in
cyclones’ environmental properties. In particular, the 75th- and 25th-percentile cases for a −1-day forecast have similar minimum SLP and speed values (Table 2). In addition, the 75th-percentile storms in both basins tend to move slower than the 25th-percentile storms for −1-day forecasts. Furthermore, the 75th-percentile ACs for −3-day forecasts reach a stronger intensity than the 25th-percentile storms, similar to what was found when the cases are sorted by minimum SLP variability. Therefore, this result suggests that although the least-predictable storms based on minimum SLP are characterized by higher intensity and faster speeds, least-predictable storms based on position are less distinguishable.

4. Summary and conclusions

This study examines the hypothesis that the practical predictability of cyclone track and intensity forecasts for long-duration, intense summer ACs is indistinguishable from comparable cyclones in the Atlantic primary storm track. This hypothesis is tested using forecasts of 100 Arctic and 89 Atlantic cyclones are selected from Sprenger et al. (2017) cyclone climatology during 1985–2016. The predictability of cyclone track and intensity in each basin is approximated using the ensemble mean error and ensemble SD computed from the GEFS Reforecast dataset. For the purpose of this paper, only forecasts initialized −1 and −3 days prior to the reference time \( t_0 \) are shown, although all forecasts out to −5 days have been evaluated and demonstrate similar results.

On average, Atlantic basin cyclone position is found to be more predictable than ACs, as indicated by the larger AC position error and SD for all initialization times, despite the fact that Atlantic basin cyclones move about 2 times as fast as ACs. One possible explanation is that the position of Atlantic basin cyclones may be easier to predict due to the consistent presence of a jet stream over the Atlantic Ocean, opposed to the more variable Arctic frontal zone that develops along the temperature gradient between the land and ocean during the summer (Crawford and Serreze 2016). As a consequence, ACs tend to have a more erratic track, which also might limit cyclone position predictability in this region (not shown).

Simmonds and Rudeva (2014) examined five of the most intense ACs for each calendar month over a 30-yr time period, suggesting that long-duration ACs may result in the merging of cyclones, resulting in one long track. Upon closer examination, this was found to be true for multiple ACs included in this study, which also might make it difficult to obtain accurate forecast for AC tracks. Another possibility is that Atlantic cyclones are better observed, which in turn would yield lower errors.

In contrast to cyclone position, AC minimum SLP is found to be more predictable than Atlantic basin cyclones. The intensity of Atlantic basin cyclones may be more difficult to predict because cyclones in this basin are on average stronger, which in turn yields a larger range of minimum pressure center values relative to summer ACs. On average, summer ACs intensify to 985 hPa, whereas the mean intensity of winter Atlantic basin cyclones are below 970 hPa. To determine if there is a relationship that exists between intensity variability and intensification of cyclones, the correlation between the change in cyclone MSLP over the 48 h after \( t_0 \) and the 48-h cyclone MSLP SD was calculated for each cyclone independently for each basin. For the −1-day forecasts, a correlation coefficient of −0.28 was found for the Arctic and −0.34 for the Atlantic, meaning that cyclones that have a greater decrease in

<table>
<thead>
<tr>
<th>Basin/forecast</th>
<th>75th-percentile min SLP (hPa)</th>
<th>25th-percentile min SLP (hPa)</th>
<th>75th-percentile speed (km h(^{-1}))</th>
<th>25th-percentile speed (km h(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arctic/−1 day</td>
<td>984.9</td>
<td>985.2</td>
<td>26.1</td>
<td>26.7</td>
</tr>
<tr>
<td>Atlantic/−1 day</td>
<td>980.3</td>
<td>975.6</td>
<td>53.1</td>
<td>54.8</td>
</tr>
<tr>
<td>Arctic/−3 days</td>
<td>985.5</td>
<td>987.5</td>
<td>31.4</td>
<td>23.1</td>
</tr>
<tr>
<td>Atlantic/−3 days</td>
<td>977.4</td>
<td>980.5</td>
<td>59.7</td>
<td>49.5</td>
</tr>
</tbody>
</table>
pressure tended to be associated with higher intensity forecast SD, but this is not the only predictor. Nevertheless, there are likely other factors that could lead to the difference in Atlantic and Arctic cyclone predictability including differences in the observation network and the nature of the large-scale environment and cyclones in the basin, as discussed in the results section. Furthermore, the environmental pressure in the extratropics is lower than over the Arctic, which in turn yields an even larger pressure depression within the midlatitudes. In addition, Arctic cyclones examined in this study are found in regions of higher baroclinic instability relative to that observed for ACs, also limiting their predictability (e.g., Torn 2017; Zhang et al. 2019).

The most- and least-variable storms based on SD are examined to understand the role of the large-scale environment on predictability in both basins. Here, the most- and least-predictable storms for position and intensity are identified via the standardized SD over the 48-h period, independently. On average, the least predictable storms for position and intensity are characterized by statistically lower minimum SLP and faster cyclone speed relative to the more predictable quartile. Past literature has suggested that cyclone predictability is limited in regions of high baroclinic instability and LHR in both the Arctic (e.g., Tao et al. 2017) and midlatitude regions (e.g., Torn 2017). Therefore, area-averaged EGR and column-integrated moisture flux convergence is examined for the 75th- and 25th-percentile cases to determine whether less-predictable storms are characterized by higher baroclinic instability and/or potential for LHR. For most lead times, the less-predictable storms are characterized by higher EGR relative to the more-predictable cases. By contrast, the LHR of each quartile is similar, except for ~3-day forecasts in the Arctic basin, but they are not statistically significant. In the Arctic, there is strong baroclinicity between the sea ice edge and continents, which can play an important role in the cyclogenesis of ACs, similar to midlatitude cyclones (e.g., Inoue and Hori 2011). By contrast, stratifying the cyclone cases based on the position forecast SD, indicates that the least predictable position cases are characterized by similar intensity and motion speed; therefore, the bulk cyclone properties do not appear to be tied to the position predictability.

Research within the Arctic is still relatively recent, especially in comparison with midlatitude cyclone research, which allows Atlantic cyclones to be a good assessment for exploring AC predictability. This study suggests that, although AC tracks are less predictable relative to their midlatitude counterparts, intensity is found to be more predictable, potentially because forecasts of intensity persistence is less challenging for models. In addition, AC storms characterized by lower intensity predictability are located in regions of higher baroclinic instability. However, further work is required to determine what features and processes limit AC position predictability in particular. Future work will further examine AC case studies that are characterized by lower predictability to determine what synoptic-scale features may limit AC predictability.

Acknowledgments. The authors thank Tomer Berg for providing cyclone-tracking code. This research also benefited from insightful suggestions and conversations from Dr. Ryan Torn’s research group, Dr. Lance Bosart, and Dr. Daniel Keyser. This paper benefited from the insightful comments of three anonymous reviewers. This research is supported by Office of Naval Research Grant 1145883-1-81016.

Data availability statement. The Global Ensemble Forecast System (GEFS) Reforecast V2 used in this study is openly available from the Physical Sciences Laboratory (https://psl.noaa.gov/forecasts/reforecast2). The cyclone climatology from Sprenger et al. (2017) uses ERA-Interim data that are openly available from the European Centre for Medium-Range Weather Forecasts (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim), as cited by Berrisford et al. (2011).

REFERENCES


