RESEARCH LETTER

Multi-year assessment of the impact of ship-borne radiosonde observations on polar WRF forecasts in the Arctic

Yonghan Choi[1], Joo-Hong Kim[1*] , Sang-Yoon Jun[1], Taejin Choi[1] and Xiangdong Zhang[2]

Abstract

To compensate for the lack of conventional observations over the Arctic Ocean, ship-borne radiosonde observations have been regularly carried out during summer Arctic expeditions and the observed data have been broadcast via the global telecommunication system since 2017. With these data obtained over the data-sparse Arctic Ocean, observing system experiments were carried out using a polar-optimized version of the Weather Research and Forecasting (WRF) model and the WRF Data Assimilation (WRFDA) system to investigate their effects on analyses and forecasts over the Arctic. The results of verification against reanalysis data reveal: (1) DA effects on analyses and forecasts; (2) the reason for the year-to-year variability of DA effects; and (3) the possible role of upper-level potential vorticity in delayed DA effects. The overall assimilation effects of the extra data on the analyses and forecasts over the Arctic are positive. Initially, the DA effects are the most apparent in the temperature variables in the middle/lower troposphere, which spread to the wind variables in the upper troposphere. The effects decrease with time but reappear after approximately 120 h, even in the 240-h forecasts. The effects on forecasts vary depending on the proximity of the radiosonde observation locations to the high synoptic variability. The upper-level potential vorticity is known to play an important role in the development of Arctic cyclones, and it is suggested as a possible explanation for the delayed DA effects after about 120 h.

Keywords Ship-borne radiosonde observations, Data assimilation, Arctic forecasts, Year-to-year variability, Delayed effect

Introduction

Climate change casts new attention on the Arctic as the melting Arctic can lead to more frequent extreme weather events (e.g., cold surge, heat wave) in mid-latitudes (Cohen et al. 2014; Kim et al. 2014; Coumou et al. 2018) as well as increasing human activities of scientific and commercial purposes over the Arctic. Therefore, it

is important to improve the weather forecast skill over the Arctic. The accuracy of numerical weather prediction is limited by uncertainties in dynamics/physics of numerical models and initial/boundary conditions. The uncertainties in the initial conditions can be reduced by adopting a sophisticated data assimilation (DA) method and/or using more qualified observations. In the Arctic, however, the number of conventional observations (e.g., radiosonde) is rare, and the assimilation of radiance observations from polar-orbiting satellites is restricted due to surface conditions (i.e., sea-ice, snow) over the Arctic.

Recent studies have investigated the effects of Arctic observations on weather forecasts in the Arctic and the mid-latitudes. Lawrence et al. (2019) showed the positive

*Correspondence:
Joo-Hong Kim
joo-hong.kim@kopri.re.kr
[1] Division of Ocean and Atmosphere Sciences, Korea Polar Research Institute, 26 Songdomirae-ro, Yeonsu-gu, Incheon 21990, Republic of Korea
[2] North Carolina State University, Asheville, NC 28801, USA

impact of Arctic observations on forecasts in the Arctic and mid-latitudes by conducting observing system experiments (OSEs) and diagnosing forecast sensitivity to observation impact (FSOI) in the European Centre for Medium-range Weather Forecasts (ECMWF) system. They found that the key observing systems in summer and winter are microwave sounding and conventional observations, respectively. Based on both the OSE and FSOI methods, Laroche and Poan (2022) examined the impact of Arctic observing systems in the Canadian global forecast system and showed the importance of microwave sounding and radiosonde observing systems. They also indicated a positive impact of supplementary radiosondes during the first two special observing periods (SOPs) of the Year of Polar Prediction (YOPP) on short-range forecasts (less than 24 h). Naakka et al. (2019) demonstrated that high-quality radiosonde observations in data-sparse regions (e.g., the central Arctic Ocean) substantially improve the analyses over the Arctic. The impact of Arctic observations on regional weather forecasts was assessed using the Application of Research to Operations at Mesoscale-Arctic system by Randriamampianina et al. (2021), in which the relative impact of conventional and satellite observations was separated from the impact of lateral boundary conditions.

There have been studies that focused on the exclusive effect of additional Arctic radiosonde observations on weather predictability over the Arctic and beyond. First, a series of OSE studies was carried out using the global model and ensemble-based global DA system of the Japan Agency for Marine-Earth Science and Technology (e.g., Inoue et al. 2015; Yamazaki et al. 2015; Sato et al. 2017, 2018; Lee et al. 2019; and so on). Inoue et al. (2015) showed that additional radiosonde observations improve the forecasts of a strong wind event and sea-ice advection. Yamazaki et al. (2015) investigated the effects of assimilating radiosonde observations on the predictability of an Arctic cyclone. Additional radiosonde observations help to reproduce the upper tropospheric circulation, which contributes to an accurate prediction of the cyclone. Sato et al. (2017) revealed that additional Arctic radiosonde observations can lead to improved forecasts of winter weather extremes over the mid-latitudes; moreover, Sato et al. (2018) showed that additional radiosonde observations in the Arctic can improve the predictability of tropical cyclones. The positive effects of additional ship-borne radiosonde observations over the Chukchi Sea on weather forecasts over Alaska were shown by Lee et al. (2019). Finally, Hong et al. (2022) quantified the impact of extra radiosonde observations from the YOPP on forecasts over the Arctic using the U.S. Navy's numerical weather prediction system. They found that assimilation of additional YOPP radiosonde

observations improves the accuracy of short-term forecasts over the Arctic.

In this study, we investigated the effects of additional ship-borne radiosonde observations on regional/medium-range weather forecasts over the Arctic by conducting data denial experiments. In contrast to previous studies, we used the polar-optimized version of the Weather Research and Forecasting (polar WRF; Skamarock et al. 2019; Hines and Bromwich 2008; Bromwich et al. 2009; Hines et al. 2011) model and its DA system (WRFDA; Barker et al. 2012) and conducted a multi-year assessment to reveal the year-to-year variability of the effects. The remainder of this paper is organized as follows. "Observations and forecast experiments" section describes the ship-borne radiosonde observations and provides the model description/experimental design. The results and corresponding discussions are given in "Results and discussion" section. "Summary and conclusions" section closes the paper with a summary and conclusions.

Observations and forecast experiments

Ship-borne radiosonde observations over the Arctic Ocean

During the summertime Arctic expedition by the Korean ice-breaking research vessel (IBRV) *Araon*, ship-borne radiosonde observations (Model: RS41-SG, Vaisala) have been regularly conducted over the Arctic Ocean since 2015, except in 2020 due to the outbreak of COVID-19. Since 2017, radiosonde observations have been broadcast through the World Meteorological Organization's Global Telecommunication System in collaboration with the Korea Meteorological Administration. The observation period, interval (Table 1), and locations (Fig. 1) varied with the year depending on the observation plan of each year. Usually, the IBRV *Araon*'s

Table 1 Radiosonde observation period and interval during Arctic expedition of the Araon in 2017, 2018, 2019, and 2021

Year		Observation period	Observation interval
2017	First leg	08/07/2017–08/23/2017	6 h
	On berth	08/24/2017–08/27/2017	
	Second leg	08/28/2017–09/13/2017	
2018	First leg	08/05/2018–08/25/2018	6 h
	On berth	08/26/2018–08/29/2018	
	Second leg	08/30/2018–09/17/2018	
2019	First leg	08/05/2019–08/26/2019	6 h
	On berth	08/27/2019–08/30/2019	
	Second leg	08/31/2019–09/17/2019	
2021	First leg	07/19/2021–08/23/2021	12 h
	Second leg	08/24/2021–09/12/2021	

Choi *et al. Geoscience Letters* (2024) 11:8

Page 3 of 13
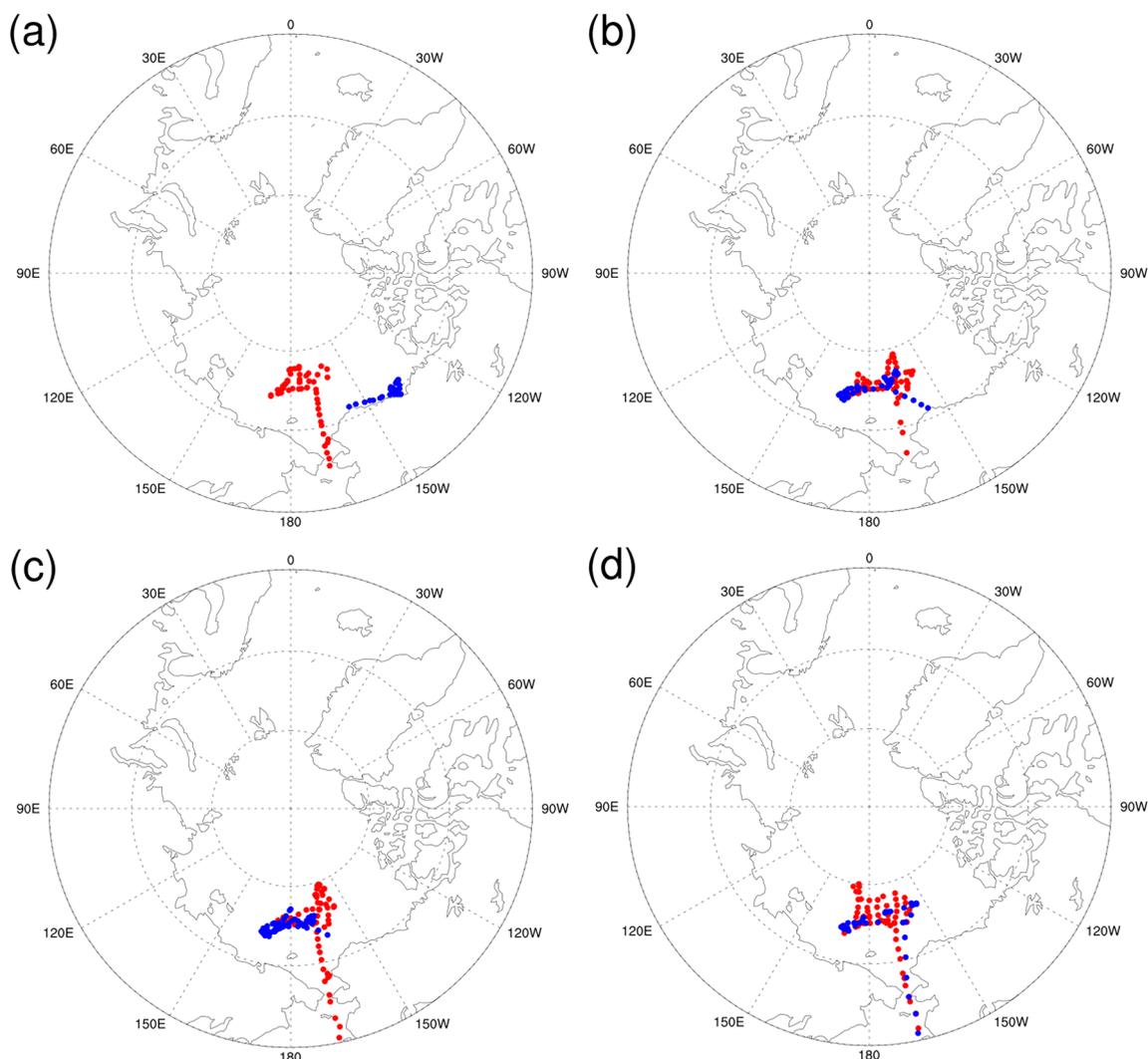
Fig. 1 Locations of radiosonde launches onboard the IBRV, Araon for **a** 2017, **b** 2018, **c** 2019, and **d** 2021. Red (blue) dots denote observation locations for the 1st (2nd) leg. Observing interval for 2017, 2018, and 2019 (2021) is 6 (12) h

Arctic cruise consists of two legs designed for oceanographic research purposes. In 2017, radiosonde observations were conducted over the Chukchi Sea during the first leg, whereas they were conducted over the Beaufort Sea during the second leg. In 2018, 2019, and 2021, radiosonde observations were mostly conducted over the Chukchi and East Siberian Seas during the first and second legs. Note that (1) radiosonde observations in 2018 contributed to YOPP's second Northern Hemisphere SOP (YOPP SOP-NH2; http://polarprediction.net/key-yopp-activities/special-observing-periods/), and that (2) there was no observational gap between the first and second legs in 2021 but radiosonde balloon launches were made less frequently at 12-h intervals.

Descriptions of model and forecast experiments

OSEs were conducted to investigate the effects of additional radiosonde observations from the IBRV *Araon* on analyses and forecasts over the Arctic. In the CTL experiment, all available observations, including the IBRV *Araon*'s radiosonde observations, were assimilated, whereas in the NoAraon experiment (i.e., data denial experiment), the IBRV *Araon*'s radiosonde observations were excluded. Continuous DA cycling was carried out during each year's cycling period, with an interval of 6 h (Table 2); i.e., a 6-h forecast from the previous cycle was used as a background when assimilating the observations (except the very first cycle) and a 10-day free forecast was conducted only at 00 UTC.

Choi *et al. Geoscience Letters*   (2024) 11:8

Page 4 of 13

Table 2 Cycling periods for OSEs of 2017, 2018, 2019, and 2021

Year	Cycling period
2017	08/07/2017–09/13/2017
2018	08/05/2018–09/18/2018
2019	08/05/2019–09/18/2019
2021	07/19/2021–09/12/2021

The WRF model optimized for polar regions (Polar WRF) version 4.3.3 was used to make both 6-h forecasts for cycling and 10-d free forecasts. The computational domain covers the Arctic Ocean and surrounding areas (Additional file 1: Fig. S1); its horizontal resolution is 27 km; the number of vertical levels is 44, with the model top at 10 hPa. Analyses and forecasts from the National Centers for Environmental Prediction Global Forecast System were used as the background for the first cycle and lateral boundary conditions, respectively. The physics schemes for running the WRF model are as follows: Kain–Fritsch cumulus (Kain 2004), Morrison double-moment microphysics (Morrison et al. 2009), Mellor–Yamada–Janjic boundary layer (Janjic 1994), Rapid Radiative Transfer Model for GCM (RRTMG) radiation (Iacono et al. 2008) schemes, and the Noah land surface model (Tewari et al. 2004).

Three-dimensional variational method included in the WRFDA system version 4.3.3 was used. The background error covariance was computed using the National Meteorological Center method (Parrish and Derber 1992), in which background error statistics were calculated based on the differences between 24-h and 12-h forecasts for the one-month period. Conventional (e.g., observations from radiosonde, aircraft, land surface, buoy, and ship), satellite-derived wind, Global Positioning System radio occultation, and satellite radiance observations were assimilated. Details of the radiance assimilation (e.g., observation operator, bias correction) are given in Table 3.

Results and discussion

Before analyzing and comparing the analyses/forecasts of the two experiments (i.e., CTL and NoAraon), O − B (observation minus background) and O − A (observation minus analysis) statistics for four variables (zonal/meridional wind, temperature, and water vapor) of radiosonde observations are examined. For all variables and cycles, O − A is smaller than O − B regardless of the year, which implies radiosonde observations are appropriately assimilated in all experiments (Additional file 1: Fig. S2; results for 2018 are shown as an example).

It is difficult to find a proper reference (or truth) when evaluating the quality of analyses or forecasts from DA cycling experiments, especially for regions, such as the Arctic Ocean, where independent observations (i.e., observations that are not assimilated) are rare. The ECMWF atmospheric reanalysis version 5 (ERA5) data (Hersbach et al. 2020) are known to perform better than other reanalyses in terms of temperature and wind fields in the Arctic (Graham et al. 2019) and present improved precipitation and snowfall over Arctic sea ice compared to ERA-Interim (Wang et al. 2019), although the ERA5 reanalysis suffers from cold season near-surface warm bias like other reanalyses (Batrak and Müller 2019). In this study, ERA5 is used to verify the analyses and forecasts from the experiments in terms of temperature, geopotential height, and wind fields, as ERA5's representation of these variables is considered reliable.

Differences in the root mean square error (RMSE) between the CTL and NoAraon (i.e., CTL minus NoAraon) experiments for the 900-hPa temperature analyses show that regardless of the year, the number of positive-effect cycles, where the RMSE of the CTL experiment is smaller than that of the NoAraon experiment, is larger than the number of negative-effect cycles (Additional file 1: Fig. S3). The ratios of the number of positive-effect cycles to the number of total cycles are larger than 50% for all years: about 64%, 66%, 59%, and 62% in 2017, 2018, 2019, and 2021, respectively. The averaged differences from all cycles are approximately − 0.022 K,

Table 3 Details of radiance data assimilation

Observation operator	Community Radiative Transfer Model (CRTM; Han et al. 2006)
Quality control	QC procedures in WRFDA (Liu et al. 2012)
Bias correction	Variational Bias Correction (VarBC; Dee 2004)
Thinning mesh	120 km
Sensors	
Infrared	AIRS from Aqua satellite IASI from MetOp-A, B satellites
Microwave	AMSU-A from NOAA-15, 18, 19, MetOp-A, B, Aqua satellites MHS from NOAA-19, MetOp-A, B satellites

Choi *et al. Geoscience Letters* (2024) 11:8

Page 5 of 13

− 0.036 K, − 0.026 K, and − 0.035 K in 2017, 2018, 2019, and 2021, respectively, and the differences for 2018, 2019, and 2021 are statistically significant at the 90% confidence level. Although the effects of additional radiosonde observations on the lower-tropospheric temperature analysis vary depending on the year and cycle, the overall effects are positive.

The number of 10-d forecasts in 2017, 2018, 2019, and 2021 is 38, 45, 45, and 56, respectively, and their sum is 184. Figure 2 shows the percentage differences in the RMSE (dRMSEs) between the CTL and NoAraon experiments (i.e., CTL minus NoAraon) for temperature, geopotential height, zonal wind, and meridional wind. The dRMSEs from 184 forecasts are averaged. The statistical significance for the difference is evaluated using a bootstrap resampling method with a 90% confidence level. Although radiosonde observations

include information on wind, temperature, and moisture in the troposphere and lower stratosphere, their effects on analyses and early forecasts are statistically significant only in the middle and lower tropospheric temperatures. The inclusion of IBRV *Araon*'s radiosonde observations improves temperature and geopotential height forecasts initially, and after 24 h, wind forecasts in the upper troposphere are improved. Bromwich et al. (2022) reported similar results from OSEs for Southern polar regions. The earlier effects of the additional radiosonde observations last for approximately 48 h. After approximately 120 h, the forecasts of all variables in the CTL experiment are better than those in the NoAraon experiment, and most improvements are statistically significant. These later effects of additional radiosonde observations are seen even at 240 h. Lee et al. (2019) also reported that the effects
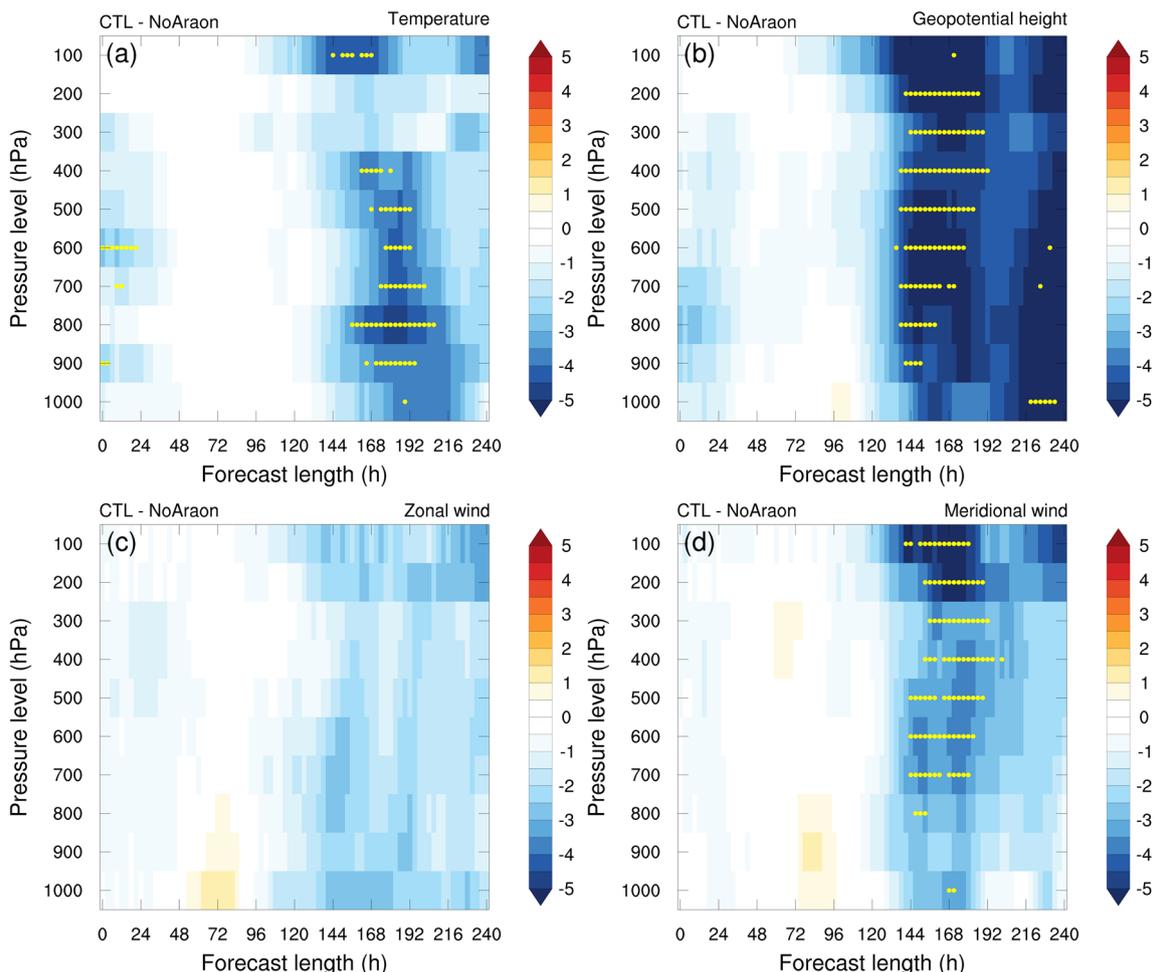
Fig. 2 Percentage differences in RMSE (%) between the CTL and NoAraon experiments (CTL minus NoAraon) for **a** temperature, **b** geopotential height, **c** zonal wind, and **d** meridional wind. RMSEs are calculated over areas north of 70°N using ERA5 reanalysis as a reference, and RMSEs from all cycles (i.e., sum of cycles of four years) are averaged. Statistical significance (denoted by yellow dots) for the difference is computed using a bootstrap resampling method and its confidence level is 90%

Choi *et al. Geoscience Letters*   (2024) 11:8

Page 6 of 13

of additional ship-borne radiosonde observations are remarkable in the later stage of the forecast.

To investigate year-to-year variability, the dRMSEs of geopotential height for individual years are plotted as functions of forecast lead time and pressure level (Fig. 3). The later effects (i.e., effects after about 120 h) of the additional radiosonde observations show large year-to-year variability. Consistent with the averaged dRMSE (Fig. 2b), for forecasts of lead times greater than 120 h, the geopotential height forecasts of the CTL experiment are notably improved compared to those of the NoAraon experiment in 2018 and 2019. The improvement at pressure levels between 900 and 200 hPa (500 and 300 hPa), and at lead times between 120 and 168 h (after 216 h) is statistically significant at the 90% confidence level in 2018 (2019). In contrast, in 2017 and 2021, although the geopotential height errors of the NoAraon experiment are

greater than those of the CTL experiment, the differences between the two experiments are not statistically significant. In particular, the difference is much smaller in 2021.

The temporal mean and standard deviation of the 500-hPa geopotential height over the cycling period are calculated for individual years (Fig. 4). The horizontal distributions of the mean 500-hPa geopotential height show different atmospheric waves in terms of wavenumber and trough/ridge locations. The standard deviation at a certain grid point indicates how large the temporal variation at that grid point is. As shown in Fig. 1, in 2018, 2019, and 2021, ship-borne radiosonde observations were conducted over the Chukchi and East Siberian Seas, while, in 2017, they were carried out over the Chukchi and Beaufort Seas. Areas of high temporal variation of geopotential height are very close to, or coincide with the locations of the IBRV *Araon*'s radiosonde observations in 2018,
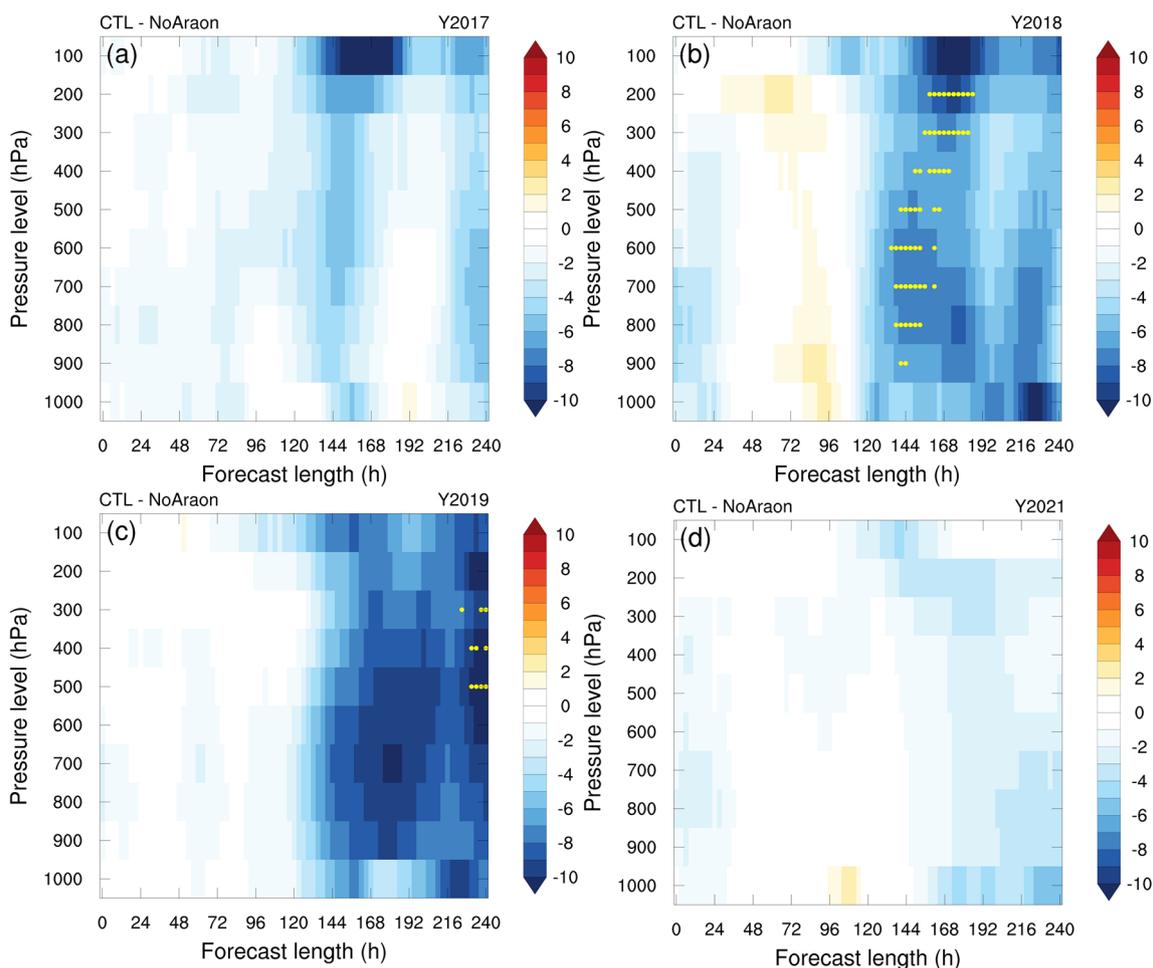
Fig. 3 Percentage differences in RMSE (%) between the CTL and NoAraon experiments (CTL minus NoAraon) for the year **a** 2017, **b** 2018, **c** 2019, and **d** 2021. RMSEs of geopotential height are calculated over areas north of 70°N using ERA5 reanalysis as a reference, and RMSEs from all cycles of the corresponding year are averaged. Statistical significance (denoted by yellow dots) for the difference is computed using a bootstrap resampling method and its confidence level is 90%

Choi *et al. Geoscience Letters*   (2024) 11:8

Page 7 of 13
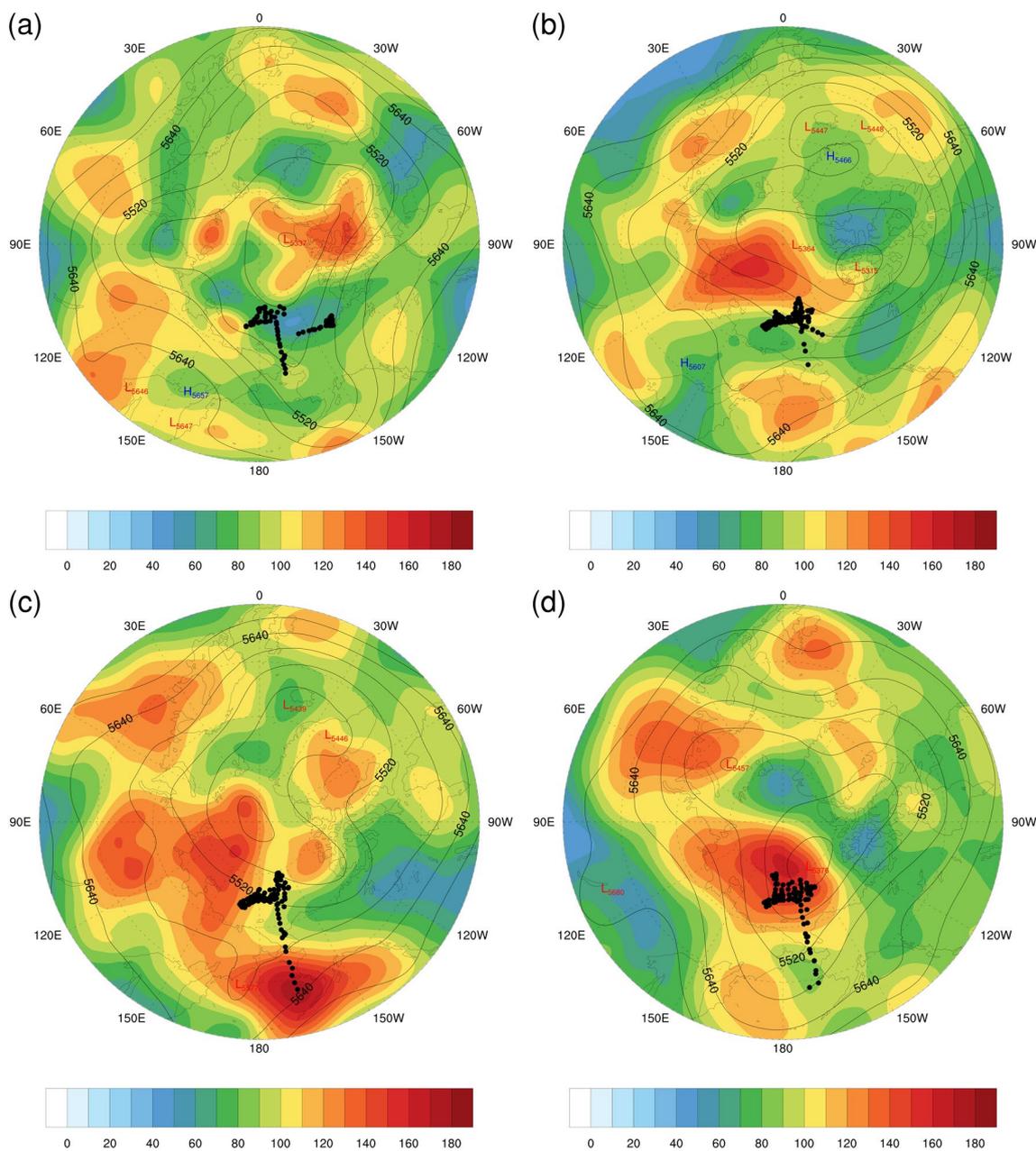
Fig. 4 Horizontal distributions of temporal mean (contour; contours from 5340 to 5700 by 60 gpm) and standard deviation (shading; gpm) of 500-hPa geopotential height for the year **a** 2017, **b** 2018, **c** 2019, and **d** 2021. Temporal mean and standard deviation are calculated over the cycling period of the corresponding year. Locations of radiosonde launches (black dots) are also shown

2019, and 2021. However, in 2017, the temporal variation of the geopotential height over the area of radiosonde observations is comparatively low.

In addition, the interannual variability of cyclone tracks and storm activity over the Arctic is analyzed. Cyclone tracks are computed using the cyclone tracking algorithm of Zhang et al. (2023), and the storm activity is measured by calculating eddy kinetic energy. Additional

file 1: Figure S4 shows horizontal distributions of cyclone tracks and storm activities for the years, 2017, 2018, 2019, and 2021. Tracks of cyclones that passed areas north of 70°N at least once during the cycling period of each year are shown. Synoptic eddy is defined as the deviation from a 2-week mean, and eddy kinetic energy attributed by 850-hPa zonal and meridional winds is averaged over the cycling period. Although there exists a large interannual

Choi *et al. Geoscience Letters*   (2024) 11:8

Page 8 of 13

variability in cyclone tracks, locations of additional radiosonde observations are close to areas where cyclone tracks are concentrated in 2018, 2019, and 2021, but in 2017, the density of cyclone tracks is relatively low near the radiosonde observation locations. Furthermore, in 2018, 2019, and 2021, locations of additional radiosonde observations are in the proximity of high storm activity.

To investigate the effects of the observing interval of additional radiosonde observations, an additional experiment, CTL_12h was conducted for 2018 and 2019. In the CTL_12h experiment, all observations except shipborne radiosonde observations were assimilated every 6 h but additional radiosonde observations from *Araon* were assimilated every 12 h. The CTL_12h experiments for 2018 and 2019 emulate the conditions of 2021, that locations of additional radiosonde observations are close to high synoptic variability but the observing interval of additional radiosonde observations is 12 h. Compared to Fig. 3b, c (i.e., 6-h observing interval), the improvements after 120 h, brought by additional radiosonde observations are significantly reduced both in 2018 and 2019, and no improvement in the CTL_12h experiment is statistically significant at the 90% confidence level (Additional file 1: Fig. S5). These results confirm the effects of the observing interval on analyses and forecasts in the Arctic.

Additional radiosonde observations in 2018 and 2019 reduce the analysis uncertainties over areas where synoptic variability is high. In 2021, a longer observing interval of 12 h limits the effects of additional radiosonde observations although radiosonde observations are conducted over areas of high synoptic variability, as in 2018 and 2019.

The differences in the absolute error of the 500-hPa geopotential height between the CTL and NoAraon experiments (i.e., CTL minus NoAraon) are shown in Fig. 5. The error is calculated against the ERA5 reanalysis, and the differences from all cycles of the years 2018 and 2019, when the later DA effects are significant, are averaged. Statistical significance for differences is computed using the bootstrap method with a 90% confidence level. At analysis time (i.e., $t=0$ h), errors of the CTL experiment are reduced compared to the NoAraon experiment over the Chukchi and East Siberian Seas, where additional radiosonde observations are assimilated. At 48 and 96 h, a mixture of positive and negative values appears, which is consistent with Fig. 3. At 144 h, the positive effects of additional radiosonde data are statistically significant over the East Siberian Sea, Siberia, and Greenland. Statistically significant positive effects appear over the Chukchi, Laptev, and Barents Seas, and even outside the Arctic Circle (i.e., south of 66.5°N) at 240 h. The temporal changes in the error difference between the two

experiments partly explain why later DA effects emerge after about 120 h.

To clarify reasons for the later DA effects, a total of four cases (one from each year) that show the largest later effects are selected. Figure 6 shows percentage differences in RMSE between the CTL and NoAraon experiments (CTL minus NoAraon) for temperature, geopotential height, zonal wind, and meridional wind using the selected four cases. Due to the small sample size, statistical significance for the difference is not tested. Initially, the effects of additional radiosonde observations are apparent in mass variables (i.e., temperature and geopotential height) in the middle and lower troposphere; after 24 h, the effects are transferred to wind variables in the upper troposphere. These are consistent with Fig. 2, results from all cycles of four years. Then, the effects spread upward to variables at levels higher than 300 hPa, resulting in the modification of potential vorticity in the lower stratosphere (through relative vorticity and static stability adjustments). Finally, the effects can spread downward to variables in the troposphere through the downward intrusion of potential vorticity after approximately 120 h (i.e., later effects). Note that all selected cases are associated with cyclone developments over the Arctic. The dominant role of lower stratospheric potential vorticity in the development of summer Arctic cyclones is explained in detail in Zhang et al. (2023; please see their Fig. 5).

We look into two cases to investigate the role of the upper-level potential vorticity in the development of Arctic cyclones. Figure 7 shows the horizontal distributions of the forecast errors of the NoAraon experiment and the differences between the CTL and NoAraon experiments (i.e., CTL minus NoAraon) for a 500-hPa geopotential height and 300-hPa potential vorticity. The errors are calculated against the ERA5 reanalysis, and 5-, 6-, 7-, and 8-d forecasts from the cycle at 00 UTC 11 August 2019 are displayed. In the reanalysis, an Arctic cyclone was stagnant and deepened over the Barents Sea from 00 UTC 16 to 00 UTC 19 August 2019, which corresponds to 5-to-8-d forecasts. However, the NoAraon experiment failed to simulate the development of the Arctic cyclone; a cyclone over the Barents Sea weakened after 00 UTC 16; instead, another cyclone falsely developed over the Beaufort Sea in the NoAraon experiment. In the CTL experiment, an Arctic cyclone was located over the Barents Sea from 00 UTC 16 to 00 UTC 19 August 2019, like the reanalysis, although its strength was not as strong as the reanalysis. The success (failure) of simulating the Arctic cyclone in the CTL (NoAraon) experiment can be explained by analyzing the upper-level potential vorticity. The role of upper-level potential vorticity in

Choi *et al. Geoscience Letters*   (2024) 11:8

Page 9 of 13
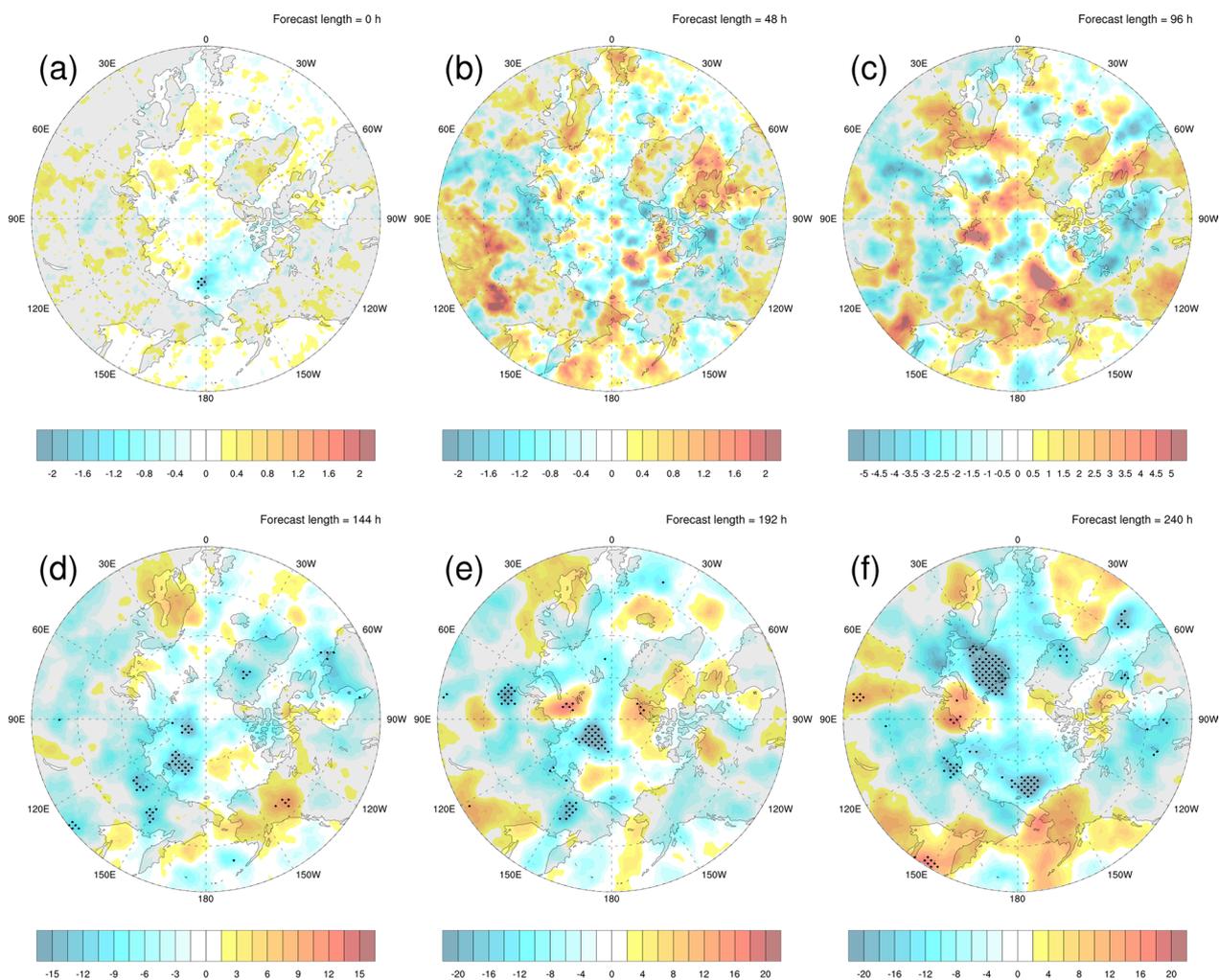
Fig. 5 Horizontal distributions of difference in absolute error of 500-hPa geopotential height (gpm) between the CTL and NoAraon experiments (CTL minus NoAraon) for forecast length of **a** 0 h, **b** 48 h, **c** 96 h, **d** 144 h, **e** 192 h, and **f** 240 h. Errors are calculated against ERA5 reanalysis, and differences in absolute error are averaged over all cycles from the years 2018 and 2019. Statistical significance for the difference at 90% confidence level is denoted by black dots

the development and maintenance of Arctic cyclones has been emphasized in many previous studies (e.g., Tao et al. 2017a, 2017b). In the NoAraon experiment, the 300-hPa potential vorticity was underestimated (overestimated) over the Barents (Beaufort) Sea and this underestimation (overestimation) led to improper weakening (strengthening) of the cyclone. However, the forecasted 300-hPa potential vorticity over both the Barents and Beaufort Seas was similar to the reanalysis in the CTL experiment, which resulted in improved forecasts of the Arctic cyclone. Case analysis from the specific cycle shows that the effects of additional radiosonde observations on forecasts reappear after approximately 120 h, possibly through the action of the upper-level potential vorticity.

The role of the upper-level potential vorticity in the later DA effects is confirmed by the analysis of another case, the cycle at 00 UTC 10 September 2018 (Additional file 1: Fig. S6). In the reanalysis, an Arctic cyclone moved from the central Arctic Ocean to the Canadian Arctic Archipelago during the period of 15–18 September 2018. However, in the NoAraon experiment, the Arctic cyclone remained near Svalbard and merged with another cyclone that developed from the East Siberian Sea. In the CTL experiment, the movement of the Arctic cyclone was successfully simulated although errors still occurred in the cyclone position. The improvement in the forecasts of the Arctic cyclone in the CTL experiment was attributed to better forecasts of the upper-level potential vorticity.
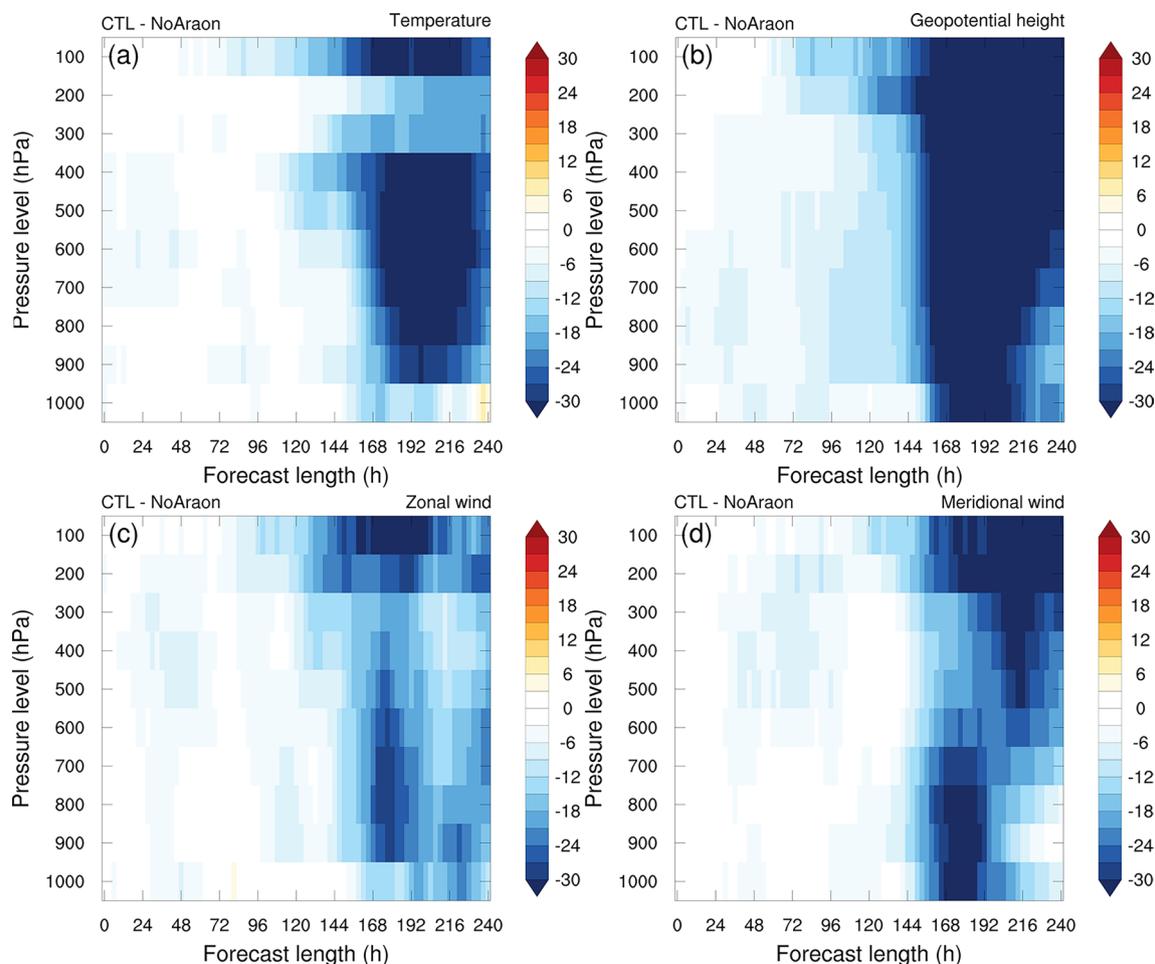
Fig. 6 Percentage differences in RMSE (%) between the CTL and NoAraon experiments (CTL minus NoAraon) for **a** temperature, **b** geopotential height, **c** zonal wind, and **d** meridional wind. RMSEs are calculated over areas north of 70°N using ERA5 reanalysis as a reference, and RMSEs from four selected cycles (i.e., cycles at 00 UTC on August 23, 2017, August 31, 2018, September 10, 2019, and August 23, 2021) are averaged

Summary and conclusions

The effects of ship-borne additional radiosonde observations from the IBRV *Araon* over the Chukchi, East Siberian, and Beaufort sectors of the Arctic Ocean on the analyses and forecasts over the Arctic were investigated using the Polar WRF model and WRFDA system. The locations and observing period/interval of the radiosonde observations varied, depending on each year's observation plan. For each year, two experiments were conducted: CTL (assimilating all available observations) and NoAraon (excluding *Araon*'s radiosonde observations). The effects of the additional radiosonde observations on the analyses and forecasts were explored by comparing the two experiments. The year-to-year variability of the DA effects and the reasons for later (delayed) effect were also examined.

The results of this study can be summarized as follows:

- Both analyses and forecasts over the Arctic are improved through assimilating additional radiosonde observations from the IBRV *Araon*.
- The effects of additional radiosonde observations are most noticeable in the temperature variable in the middle and lower troposphere initially; after 24 h, the DA effects spread to the wind variables in the upper troposphere. [Earlier effect]
- After approximately 120 h, the DA effects reappear, leading to improved forecasts in the CTL experiment. [Later effect]
- The DA effects can vary depending on 1) the proximity of the observation locations to high synoptic variability and 2) the observing interval.
- The action of the upper-level potential vorticity on the development of Arctic cyclones is suggested as a possible explanation of the later effect.

Fig. 7 Error of 500-hPa geopotential height in the NoAraon experiment (shading; gpm) and 500-hPa geopotential height from ERA5 reanalysis (contour; contours from 5280 to 5940 by 60 gpm) at forecast length of **a** 120 h, **b** 144 h, **c** 168 h, and **d** 192 h. Difference between the CTL and NoAraon experiments (CTL minus NoAraon) for 500-hPa geopotential height (gpm) at forecast length of **e** 120 h, **f** 144 h, **g** 168 h, and **h** 192 h. Error of 300-hPa potential vorticity in the NoAraon experiment (shading; PVU) and 300-hPa potential vorticity from ERA5 reanalysis (contour; contours from 0 to 15 by 5 PVU) at forecast length of **i** 120 h, **j** 144 h, **k** 168 h, and **l** 192 h. Difference between the CTL and NoAraon experiments (CTL minus NoAraon) for 300-hPa potential vorticity (PVU) at forecast length of **m** 120 h, **n** 144 h, **o** 168 h, and **p** 192 h. Forecasts of the CTL and NoAraon experiments from the cycle at 00 UTC 11 August 2019 and the corresponding ERA5 reanalysis data are used

The positive impact of additional radiosonde observations over the Arctic, which has been reported in previous studies, is also shown in this study. The earlier effect of the extra radiosonde observations is consistent with Bromwich et al. (2022), although they dealt with Antarctica. This study uniquely reveals the important role of upper-level potential vorticity in the later effect and the relationship between the observation location/interval and weather predictability.

To maximize the effects of assimilating additional radiosonde observations, a more sophisticated DA method (e.g., 4D-Var) and/or consideration of balloon drift can be taken into account in future studies. Furthermore, using a global model and the corresponding DA system, the relationship between the characteristics (e.g., location) of radiosonde observations and DA effects on forecasts over the mid-latitudes and Arctic can be explored. Finally, the effects of the IBRV *Araon*'s radiosonde observations will be monitored continuously, and the results will be used to determine the optimal observation locations.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40562-024-00326-w.

Additional file 1: Figure S1. Geographical area and topography height (m) of computational domain. Names of selected seas are indicated. **Figure S2**. Root mean square (RMS) of O-B (observation minus background, blue) and O-A (observation minus analysis, red) for **a** zonal wind (m s⁻¹), **b** meridional wind (m s⁻¹), **c** temperature (K), and **d** water vapor mixing ratio (g kg⁻¹) variables of radiosonde observations. O-B, O-A statistics from all DA cycles of the CTL experiment for the year of 2018 are shown. The number of assimilated observations (black) are also shown. **Figure S3**. Differences in RMSE (K) between the CTL and NoAraon experiments (CTL minus NoAraon) for the cycling period of **a** 2017, **b** 2018, **c** 2019, and **d** 2021. RMSEs of 900-hPa temperature analyses are calculated over areas north of 70°N and ERA5 reanalysis is used as a reference. Blue (red) bar denotes the cycle at which the RMSE of the NoAraon experiment is greater (less) than that of the CTL experiment. The black line denotes the average difference over all cycles, and the statistical significance at the 90% confidence level for the difference is indicated by a yellow star. **Figure S4**. Horizontal distributions of cyclone tracks (open black circles) and storm activities (shading) for **a** 2017, **b** 2018, **c** 2019, and **d** 2021. Cyclones that passed areas north of 70°N at least once during the cycling period of each year were considered. The cyclone tracking algorithm in Zhang et al. (2023) was used. Synoptic eddies are defined as deviations from a 2-week mean. Eddy kinetic energy is calculated using 850-hPa zonal and meridional winds for each cycle and eddy kinetic energy from all cycles is averaged. Locations of additional radiosonde observations (red filled circles) are also shown. **Figure S5**. Percentage differences in RMSE (%) between the CTL_12h and NoAraon experiments (CTL_12h minus NoAraon) for the year **a** 2018 and **b** 2019. RMSEs of geopotential height are calculated over areas north of 70°N using ERA5 reanalysis as a reference, and RMSEs from all cycles of the corresponding year are averaged. Statistical significance (denoted by yellow dots) for the difference is computed using a bootstrap resampling method and its confidence level is 90%. **Figure S6**. Error of 500-hPa geopotential height in the NoAraon experiment (shading; gpm) and 500-hPa geopotential height from ERA5 reanalysis (contour; contours from 5280 to 5940 by 60 gpm) at forecast length of **a** 120 h, **b** 144 h, **c** 168 h, and **d** 192 h. Difference between the CTL and NoAraon experiments (CTL minus NoAraon) for 500-hPa geopotential height (gpm) at forecast length of **e** 120 h, **f** 144 h, **g** 168 h, and **h** 192 h. Error of 300-hPa potential vorticity in

---

the NoAraon experiment (shading; PVU) and 300-hPa potential vorticity from ERA5 reanalysis (contour; contours from 0 to 15 by 5 PVU) at forecast length of **i** 120 h, **j** 144 h, **k** 168 h, and **l** 192 h. Difference between the CTL and NoAraon experiments (CTL minus NoAraon) for 300-hPa potential vorticity (PVU) at forecast length of **m** 120 h, **n** 144 h, **o** 168 h, and **p** 192 h. Forecasts of the CTL and NoAraon experiments from the cycle at 00 UTC 10 September 2018 and the corresponding ERA5 reanalysis data are used.

Author contributions
YC conceptualized the research, designed and performed numerical experiments, and carried out formal analysis and visualization. YC wrote the original manuscript (writing—original draft preparation). J-HK contributed to supervision, funding, and formal analysis. J-HK, S-YJ, TC, and XZ contributed to writing—review and editing. All authors read and approved the final manuscript.

Availability of data and materials
The WRF model version 4.3.3 and the WRFDA system version 4.3.3 can be downloaded from the WRF model's GitHub (https://github.com/wrf-model/WRF/releases) after a user registration. Source codes for the polar-optimized version of the WRF model are available from the BYRD Polar and Climate Research Center web page (https://polarmet.osu.edu/PWRF) after a user registration. The GFS analysis/forecast data can be downloaded from the NCEI/NOAA web page (https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast) upon request. The GDAS observational data including conventional, satellite-derived wind, GPS RO, and satellite radiance observations can be downloaded from the NCEI/NOAA web page (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00379) upon request. The ERA5 reanalysis data are available from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu). Data from the DA cycling experiments and radiosonde observations from IBRV, *Araon* are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

References
Barker D, Huang X-Y, Liu Z, Auligné T, Zhang X, Rugg S, Ajjaji R, Bourgeois A, Bray J, Chen Y (2012) The weather research and forecasting model's community variational/ensemble data assimilation system: WRFDA. Bull Am Meteor Soc 93(6):831–843

Batrak Y, Müller M (2019) On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice. Nat Commun 10(1):4170

Bromwich DH, Hines KM, Bai LS (2009) Development and testing of polar weather research and forecasting model: 2 Arctic Ocean. J Geophys Res. https://doi.org/10.1029/2008JD010300

Bromwich DH, Powers JG, Manning KW, Zou X (2022) Antarctic data impact experiments with Polar WRF during the YOPP-SH summer special observing period. Q J R Meteorol Soc 148(746):2194–2218

Cohen J, Screen JA, Furtado JC, Barlow M, Whittleston D, Coumou D, Francis J, Dethloff K, Entekhabi D, Overland J (2014) Recent Arctic amplification and extreme mid-latitude weather. Nat Geosci 7(9):627–637

Coumou D, Di Capua G, Vavrus S, Wang L, Wang S (2018) The influence of Arctic amplification on mid-latitude summer circulation. Nat Commun 9(1):2959

Dee DP (2004) Variational bias correction of radiance data in the ECMWF system. In: Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP, Reading, UK. pp 97–112.

Graham RM, Hudson SR, Maturilli M (2019) Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyses. Geophys Res Lett 46(11):6138–6147

Han Y, van Delst P, Liu Q, Weng F, Yan B, Treadon R, Derber J (2006) JCSDA community radiative transfer model (CRTM): Version 1. NOAA Technical Report. NOAA, Washington, D.C.

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049

Hines KM, Bromwich DH (2008) Development and testing of polar weather research and forecasting (WRF) model. Part I: greenland ice sheet meteorology. Monthly Weather Rev 136(6):1971–1989

Hines KM, Bromwich DH, Bai L-S, Barlage M, Slater AG (2011) Development and testing of Polar WRF Part III: Arctic land. J Clim 24(1):26–48

Hong X, Doyle JD, Tyndall DP (2022) Impact of assimilating the special radiosonde observations on coamps arctic forecasts during the year of polar prediction. Data assimilation for atmospheric, oceanic and hydrologic applications (Vol. IV):397–410. https://doi.org/10.1007/978-3-030-77722-7_14

Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008) Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. J GeophysRes. https://doi.org/10.1029/2008JD009944

Inoue J, Yamazaki A, Ono J, Dethloff K, Maturilli M, Neuber R, Edwards P, Yamaguchi H (2015) Additional Arctic observations improve weather and sea-ice forecasts for the Northern Sea Route. Sci Rep 5(1):16868

Janjić ZI (1994) The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon Weather Rev 122(5):927–945

Kain JS (2004) The Kain-Fritsch convective parameterization: an update. J Appl Meteorol 43(1):170–181

Kim B-M, Son S-W, Min S-K, Jeong J-H, Kim S-J, Zhang X, Shim T, Yoon J-H (2014) Weakening of the stratospheric polar vortex by Arctic sea-ice loss. Nat Commun 5(1):4646

Laroche S, Poan ED (2022) Impact of the Arctic observing systems on the ECCC global weather forecasts. Q J R Meteorol Soc 148(742):252–271

Lawrence H, Bormann N, Sandu I, Day J, Farnan J, Bauer P (2019) Use and impact of Arctic observations in the ECMWF Numerical Weather Prediction system. Q J R Meteorol Soc 145(725):3432–3454

Lee M-H, Kim J-H, Song H-J, Inoue J, Sato K, Yamazaki A (2019) Potential benefit of extra radiosonde observations around the Chukchi Sea for the Alaskan short-range weather forecast. Polar Sci 21:124–135

Liu Z, Schwartz CS, Snyder C, Ha S-Y (2012) Impact of assimilating AMSU-A radiances on forecasts of 2008 Atlantic tropical cyclones initialized with a limited-area ensemble Kalman filter. Mon Weather Rev 140(12):4017–4034

Morrison H, Thompson G, Tatarskii V (2009) Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one-and two-moment schemes. Mon Weather Rev 137(3):991–1007

Naakka T, Nygård T, Tjernström M, Vihma T, Pirazzini R, Brooks IM (2019) The impact of radiosounding observations on numerical weather prediction analyses in the Arctic. Geophys Res Lett 46(14):8527–8535

Parrish DF, Derber JC (1992) The National Meteorological Center's spectral statistical-interpolation analysis system. Mon Weather Rev 120(8):1747–1763

Randriamampianina R, Bormann N, Køltzow MA, Lawrence H, Sandu I, Wang ZQ (2021) Relative impact of observations on a regional Arctic numerical weather prediction system. Q J R Meteorol Soc 147(737):2212–2232

Sato K, Inoue J, Yamazaki A, Kim JH, Maturilli M, Dethloff K, Hudson SR, Granskog MA (2017) Improved forecasts of winter weather extremes over midlatitudes with extra Arctic observations. J Geophys Res Oceans 122(2):775–787

Sato K, Inoue J, Yamazaki A, Kim J, Makshtas A, Kustov V, Maturilli M, Dethloff K (2018) Impact on predictability of tropical and mid-latitude cyclones by extra Arctic observations. Sci Rep 8:12104

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Liu Z, Berner J, Wang W, Powers J, Duda M, Barker D (2019) A description of the advanced research WRF version 4. NCAR Tech. Note NCAR/TN-556+ STR 145.

Tao W, Zhang J, Zhang X (2017b) The role of stratosphere vortex downward intrusion in a long-lasting late-summer Arctic storm. Q J R Meteorol Soc 143(705):1953–1966

Tao W, Zhang J, Fu Y, Zhang X (2017a) Driving roles of tropospheric and stratospheric thermal anomalies in intensification and persistence of the Arctic Superstorm in 2012. Geophys Res Lett 44(19):10,017–10,025.

Tewari N, Tewari M, Chen F, Wang W, Dudhia J, LeMone M, Mitchell K, Ek M, Gayno G, Wegiel J (2004) Implementation and verification of the unified NOAH land surface model in the WRF model (Formerly Paper Number 17.5). In: Proceedings of the 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, Seattle, WA, USA.

Wang C, Graham RM, Wang K, Gerland S, Granskog MA (2019) Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution. Cryosphere 13(6):1661–1679

Yamazaki A, Inoue J, Dethloff K, Maturilli M, König-Langlo G (2015) Impact of radiosonde observations on forecasting summertime Arctic cyclone formation. J Geophys Res 120(8):3249–3273

Zhang X, Tang H, Zhang J, Walsh JE, Roesler EL, Hillman B, Ballinger TJ, Weijer W (2023) Arctic cyclones have become more intense and longer-lived over the past seven decades. Commun Earth Environ 4:348

Publisher's Note