

Extraordinary Arctic Cyclones in Summer and Their Predictability on Medium-range Timescales

January 2019

Akio YAMAKAMI

Extraordinary Arctic Cyclones in Summer and Their Predictability on Medium-range Timescales

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Science

Akio YAMAKAMI

Abstract

Arctic cyclones (ACs) are atmospheric phenomena that severely affect the Arctic environment. Accurate prediction of ACs is therefore necessary to address associated environmental and societal concerns. This study investigated the predictability of extraordinary ACs during summer (June–August) using operational medium-range ensemble forecasts and the second-generation Global Ensemble Forecast System (GEFS) reforecast.

Ten extraordinary ACs were identified for summer seasons during 2008–2016. Because of the sparse observational network, larger uncertainties in assimilated atmospheric state are expected over the Arctic than at lower latitudes. The central pressure and position of individual mature ACs operationally analyzed by numerical weather prediction (NWP) centers are similar to those of ERA-Interim reanalysis. The average differences between each NWP center's analysis of central pressure and position for the 10 ACs and those of ERA-Interim are less than 0.5 hPa and 37.6 km, respectively.

We found that accurate predictions of an upper-level trough and tropopause polar vortices (TPVs) were important for prediction of extraordinary ACs. Predicted development of ACs is affected by merging prediction between pre-existing ACs over the Arctic and cyclones migrating from lower latitude (mid-latitude cyclones). Predictions of development of ACs are also influenced by the prediction of upper-level trough and ridge through baroclinic growth of mid-latitude cyclones. Predicted positions of ACs are primarily influenced by predicted positions of TPVs because of coupling between surface cyclone of pre-existing ACs and TPVs. Predictions of upper-level trough and ridge also have an influence on predicted position of ACs position through steering winds. In addition, the predicted upper-level trough and ridge affect the prediction of cyclogenesis of the mid-latitude cyclones. Therefore, when both the upper-level trough and TPV are correctly predicted, the merging of a mid-latitude cyclone with a pre-existing AC is well predicted, resulting in an accurate prediction of the development and

position of the AC. In some cases, the accurate predictions of the cyclone merging are more important for predictions of the AC development than those of the AC location. In other cases, the accurate predictions of cyclone merging are more important to prediction of the AC location.

The forecast skill for 10 extraordinary ACs that occurred in summer during 2008–2016, based on five leading operational medium-range ensemble forecasts, indicates that 90 % ensemble member predict the existence of the ACs at lead times of ≤ 3.5 days. The average central position errors of the predicted ACs are less than half the mean radius of the 10 ACs (469.1 km) at lead times of 2.5–4.5 days. Average central pressure errors of the predicted ACs are 5.5–10.7 hPa at such lead times. Overall, the European Centre for Medium-Range Weather Forecasts (ECMWF) shows a 1.0- to 1.5-day advantage in predicting the existence probability, central pressure, and central position of extraordinary ACs compared with other centers.

The forecast skills of existence probability, central pressure, and central position for 26 extraordinary ACs in summer during 1986–2016 from the GEFS reforecast are similar to those for the 10 ACs from the operational medium-range ensemble forecasts. No improvements appeared in the forecast skills of the extraordinary ACs at all lead times. From the verification of the strike probability for ACs within a 400-km radius, the probability information is useful compared with a reference forecast in 1.0- to 5.0-day forecasts, although the forecast probabilities were overconfident for lead times more than 1.0 days.

Keywords: Arctic cyclone, medium-range ensemble forecast, cyclone merging, tropopause polar vortex, forecast skill, predictability

Contents

Abstract	i
List of Tables	v
List of Figures	1
CHAPTER 1 Introduction	8
1.1 Arctic climate and cyclonic activity in the Arctic	8
1.2 Arctic cyclones	9
1.3 Predictability of Arctic atmosphere	10
1.4 Purpose of this study	11
CHAPTER 2 Data and Methods	13
2.1 Data	13
2.2 Methodology	15
2.2.1 Detection of extraordinary ACs	15
2.2.2 Verification of the AC prediction	16
2.2.3 Strike probability and its verification using Brier Skill Score and reliability diagram	17
CHAPTER 3 Results	22
3.1 Observed extraordinary ACs in summer during 2008–2016 and their analysis uncertainties	22
3.2 Case studies of extraordinary AC predictions	28
3.2.1 AC in August 2012	28
3.2.2 AC in August 2013	32
3.2.3 AC in August 2016	35
3.2.4 AC in June 2015	38
3.2.5 Summary of case studies	41
3.3 Average predictability of Arctic cyclones	60
3.3.1 Forecast skill for summer ACs during 2008–2016 by TIGGE	60
3.3.2 Forecast skill for summer ACs during 1986–2016 by GEFS reforecast	63
3.3.3 Summary of average predictability	69
CHAPTER 4 Discussion	81

4.1	<i>Predictions of individual extraordinary ACs</i>	81
4.2	<i>Forecast skill for extraordinary ACs</i>	82
CHAPTER 5	Conclusions	85
	Acknowledgements	89
	References	90

List of Tables

Table 2.1	Configurations for operational medium-range ensemble prediction systems from TIGGE project (as of August 2016).....	19
Table 2.2	Forecast length, and ensemble size, forecast model resolution, and data assimilation method for the GEFS reforecast.	20
Table 3.1	Date, central pressure, central pressure anomaly from climatology, radius, and central position during the mature stage of each AC event.	26
Table 3.2	Date, central pressure, and radius during the mature stage of each AC event in Summer during 1986–2016.....	75

List of Figures

- Figure 2.1 (a) Observed SLP (contour, 2 hPa interval) and all cyclone centers (red circle, the values are given on the right of each center). (b) Observed SLP (contour, 3 hPa interval) and temperature anomaly at 250 hPa (shading). The cyclone center of extraordinary AC are given by a white cross.21
- Figure 3.1 Observed sea-level pressure (SLP) fields for the 10 extraordinary ACs at their mature stage during summer 2008–2016 (ERA-Interim). The date of the mature stage is given above each panel. Shading represents SLP <1000 hPa. The white cross indicates the AC center.25
- Figure 3.2 Central positions of the analyzed ACs in ERA-Interim (black circle) and control analyses (cross) of CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), and (b) the difference of central position between ERA-Interim and the control analysis. (c) Central pressures of the analyzed ACs for ERA-Interim (black horizontal line; the value is given at the top-right corner of each frame) and each control analysis (colored bars; differences >2.0 hPa between ERA-Interim and each analysis are given above or below each bar). The letters (a–j) for each symbol in (a) correspond to the AC events shown in (b), (c), Fig. 3.1, and Table 3.1.27
- Figure 3.3 Observed sea-level pressure (SLP) fields for the 10 extraordinary ACs at their mature stage during summer 2008–2016 (ERA-Interim). The date of the mature stage is given above each panel. Shading represents SLP < 1000 hPa. The white cross indicates the AC center.43
- Figure 3.4 Predicted tracks of AC12 in forecasts initialized at 1200 UTC on (a) 1, (b) 2, and (c) 3 August 2012 for CMC, ECMWF, JMA, NCEP, and UKMO (colored lines). The circle represents the positions at 1800 UTC on 2 August 2012. The color of the squares indicates the predicted minimum SLP at 1200 UTC on 6 August 2012. The

	right blue lines represent the analyzed track and the contours show the analyzed SLP at 1200 UTC on 6 August (ERA-Interim).	44
Figure 3.5	Predicted minimum SLP at the center of AC12, valid at 1200 UTC on 6 August 2012 for (a) CMC, (b) ECMWF, (c) JMA, (d) NCEP, and (e) UKMO, for different forecast initialization dates. Each ensemble member is represented by a symbol following the legend (predicted central pressure (P) and distance (D) between the predicted and observed cyclone center at 1200 UTC on 6 August 2012). The red dashed line (ANL) represents the minimum SLP of AC12 in ERA-Interim at 1200 UTC on 6 August 2012. The number of all ensemble members is given at the top-right corner of each frame, the number of ensemble members that predicted the AC existence is given at the bottom of each frame at each initial date.	45
Figure 3.6	Predicted SLP (contour, intervals are 4 hPa) and temperature anomaly at 250 hPa (T250 anomaly, shading) of the five (a) best- and (c) worst-performing ensemble members for each NWP center. Forecasts are initialized at 1200 UTC on 2 August 2012 and valid at 1200 UTC on 6 August. Hatching in (a) and (c) represents areas with a significant (95% level) difference in T250 anomaly between the best and worst members. (b) Observed SLP at 1200 UTC on 6 August 2012 from ERA-Interim. The cross marks the analyzed AC center and the contour interval is 4 hPa.	46
Figure 3.7	Predicted geopotential height at 300 hPa for composites of (a) higher-skill, (b) lower-skill, and (c) only-position ECMWF members initialized at 1200 UTC on 2 August 2012, valid at 0000 UTC on 4 to 6 August 2012 at 12-hourly intervals. The shading in (b) and (c) represents the difference between (a) and (b) (lower- minus higher-skill members), and between (a) and (c) (only-position minus higher-skill members), respectively, where where the significance is at the 90% confidence level. The cross represents the AC12 center in ERA-Interim at that time step.	47
Figure 3.8	Same as Figure 3.7, but for predicted temperature anomaly at 250 hPa.	48

Figure 3.9	Predicted (a, b) geopotential height at 300 hPa and (c, d) temperature anomaly at 250 hPa for composites of (a, c) higher-skill and (b, d) only-development ECMWF members initialized at 1200 UTC on 3 August 2012, valid at 0000 UTC on 4 to 6 August 2012 at 12-hourly intervals. The shading in (b) and (d) represents the difference (only-development minus higher-skill members, where the significance is at the 90% confidence level) between (a) and (b), and between (c) and (d), respectively. The cross represents the AC12 center in ERA-Interim at that time step.	48
Figure 3.10	Same as Fig. 3.3, but for AC13.	49
Figure 3.11	Same as Fig. 3.5, but for AC13.	49
Figure 3.12	Predicted Z300 for composites of members (a, c) with and (b) without AC initialized at 1200 UTC on (b, c) 2 August and (a) 3 August 2012, valid at 0000 UTC on 5 August 2013 for CMC, ECMWF, JMA, NCEP, and UKMO (left to right columns). The shading in (b) represents the difference between (a) and (b), and that in (c) represents the difference between (a) and (c). Bold colors indicate a significant difference at the 95% confidence level. The cross represents the AC13 center in ERA-Interim.	50
Figure 3.13	Same as Fig. 3.9a and b, but for ECMWF members (a) with AC initialized on 3 August and (b) without AC initialized on 2 August, valid from 1200 UTC on 3 August to 1200 UTC on 5 August. Bold color indicates the region in which all NWP centers agree on the sign of the difference between members with and without AC13.	51
Figure 3.14	Same as Fig. 3.3, but for AC16.	52
Figure 3.15	Same as Fig. 3.5, but for AC16.	52
Figure 3.16	(a) Analyzed SLP (contour) from 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns) in ERA-Interim. The cross represents the AC16 center in ERA-Interim. (b, c) Predicted SLP (contour) and its error (shading) for composites of (b)	

only-position members initialized at 1200 UTC on 11 August and (c) lower-skill ECMWF members initialized at 1200 UTC on 10 August, valid at 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). The shading in (b) represents the error between (a) and (b), and that in (c) represents the error between (a) and (c). Bold color indicates the region in which all NWP centers agree on the sign of the errors.....53

Figure 3.17 (a) Analyzed geopotential height at 300 hPa (contour) and temperature anomaly at 250 hPa (shading) from 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). (b, c) Predicted geopotential height at 300 hPa (b, contour) and temperature anomaly at 250 hPa (c, contour) for composites of lower-skill ECMWF members initialized at 1200 UTC on 10 August 2016, valid at 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). The shading in (b) represents the error between (a) and (b), and that in (c) represents the error between (a) and (c). Bold color indicates the region in which all NWP centers agree on the sign of the errors. The cross represents the AC16 center in ERA-Interim.54

Figure 3.18 Same as Fig. 3.3, but for AC15.....55

Figure 3.19 Same as Fig. 3.5, but for AC15.....55

Figure 3.20 Same as Fig. 3.4, but for AC15.....56

Figure 3.21 Predicted (a, b) geopotential height at 300 hPa and (c, d) temperature anomaly at 250 hPa for composites of (a, c) only-development and (b, d) lower-skill ECMWF members initialized at 1200 UTC on 1 June 2015, valid at 1200 UTC on 2 to 6 June 2015 at 24-hourly intervals. The shading in (b) and (d) represents the difference (lower-skill minus only-development members, where the significance is at the 95% confidence level) between (a) and (b), and between (c) and (d), respectively. The cross represents the AC15 center in ERA-Interim.57

Figure 3.22 Predicted geopotential height at 300 hPa for composites of (a) higher-skill and only-development ECMWF members with the predicted AC position of (b) $< 90^{\circ}\text{E}$ and (c)

> 90°E members initialized at 1200 UTC on 2 June 2015, valid at 1200 UTC on 2 to 6 June 2015 at 24-hourly intervals. The shading in (b) and (c) represents the difference between (a) and (b) and between (a) and (c) (only-position minus higher-skill members), respectively. Bold colors indicate a significant difference at the 95% confidence level. The cross represents the AC12 center in ERA-Interim at that time step. 58

Figure 3.23 Schematic diagram for classification of predicted ACs. Top panels indicate the key factors for ACs prediction. “Accurate” (“inaccurate”) indicate a forecast accurately (inaccurately) predicts the key factor described at top panels. The colored arrows indicate the higher-skill (blue), only-development (orange), only-position (green), lower-skill (red) members’ forecasts. The gray arrows indicate the members without ACs. 59

Figure 3.24 Average (a) existence probability, (b) central pressure error, and (c) central position error for the mature stages of 10 extraordinary ACs in summer of 2008-2016, predicted by CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), as a function of forecast lead time. The forecasts were verified against ERA-Interim. 71

Figure 3.25 As in Fig. 3.24a, but for each AC. 72

Figure 3.26 As in Fig. 3.24b, but for each AC. 72

Figure 3.27 As in Fig. 3.24c, but for each AC. 73

Figure 3.28 Observed sea-level pressure (SLP) fields for the 26 extraordinary ACs at their mature stage during summer 1986–2016 (GEFS control analysis). The date of the mature stage is given above each panel. Shading represents SLP < 1000 hPa. The white cross indicates the AC center. 74

- Figure 3.29 (a) Observed central pressure (solid line; left vertical axis) and radius (broken line; right vertical axis) for the 26 ACs. (b) Time series of the number of extraordinary ACs in each summer from 1986 to 2016. 76
- Figure 3.30 Average (a) existence probability, (b) central pressure error, and (c) central position error during the mature stage of 26 ACs in summer during 1985–2016 (black) and 10 ACs in summer during 2008–2016 (brown) from the GEFS reforecast, as a function of forecast lead time. Colored broken lines are the same as Figure 3.24, but verified against own control analysis. 77
- Figure 3.31 Existence probability for individual mature ACs in summer during 1985–2016 from the GEFS reforecast. The vertical axis is forecast lead time, and colored bars indicate existence probabilities of $\geq 6/11$ (red), $\geq 9/11$ (green), and $11/11$ (blue). (b, c) Ensemble mean central (b) pressure error and (c) position error for individual mature ACs at lead times of 1 (blue), 3 (green), 5 (yellow), and 7 (red) days, predicted by the GEFS reforecast. Error bars indicate one standard deviation and are given for forecasts with existence probabilities of $\geq 6/11$. These forecasts were verified against own control analyses. 78
- Figure 3.32 Strike probability map for the AC in August 2016 from (a) 1-, (b) 3-, (c) 5-, and (d) 7-day forecasts of the GEFS reforecast. The strike probability is defined as the probability that the AC is located within 400 km of a given location during its mature stage. This probability is calculated based on the number of ensemble members which predicted ACs. Colored shading indicates probabilities of 5%–25% (blue), 25%–50% (green), 50%–75% (yellow), and 75%–100% (red). Black shading indicates a location within 400 km of the center of an observed AC during its mature stage. ... 79
- Figure 3.33 Reliability diagrams for forecast probabilities that an AC exists within a 400-km radius during its mature stage from (a) 1-, (b) 3-, (c) 5-, and (d) 7-day forecast of the GEFS reforecast. Reliabilities and their regression lines are indicated by gray and black lines, respectively (left vertical axis). The number of samples (grid points)

predicting an event is indicated by blue bars, and the number of samples for which the event actually occurred is indicated by dashed red bars (right vertical axis). The Brier Skill Score (BSS) are given at the upper-left corner of each panel. Forecast probabilities were verified for 26 ACs during 1986–2010..... 80

CHAPTER 1 Introduction

1.1 Arctic climate and cyclonic activity in the Arctic

A radiative cooling forms a global low-pressure area and maintains the year-round polar vortex over the Arctic. Cold air mass over the Arctic influences the mid-latitude weather when the cold air blows into mid-latitude by the meandering of the westerly (Osborn, 2011; Shepherd, 2016). The accumulation and blowing out of the cold air are known as the natural variations like Arctic Oscillation (Thompson and Wallace, 1998, Ogi et al. 2004) and North Atlantic Oscillation (Ambaum et al. 2001). These natural variations become pronounced in winter.

In recent years, the effects on high- and mid-latitude environments not only by natural variability but also by reduction of sea ice caused by Arctic Amplification (Serreze and Francis, 2006; Nagato and Tanaka, 2012; Cohen et al. 2014) attract a lot of attention. In particular, the relationship between reduction of sea ice and the change of atmospheric circulation is studied actively (e.g. Coumou et al., 2015; Knudsen et al., 2015; Lynch et al. 2016; Mori et al. 2014; Nakamura et al., 2016; Ogi et al. 2016; Semmler et al. 2016; Serreze et al., 2016; Yang et al., 2016; Screen 2017a, b). However, the effects of sea ice reduction and the interaction between high- and mid-latitude are very complex (Overland et al. 2016), and the studies of those relations and interactions are progressing (Screen, 2017c).

Cyclonic activity over the Arctic affects the Arctic climate systems on a variety of timescales from long-term to short- and medium-term ranges. As long-term effects, cyclones migrating from mid-latitudes bring large amounts of heat and water vapor (Graversen and Wang, 2009). Luo et al. (2017) and Johansson et al. (2017) indicated that the combination of cyclonic activity over the North Atlantic and blocking anticyclone over the Ural Mountain caused intrusion of a large amount of water vapor into the Arctic. Cyclones over the Arctic also influence the ice-albedo feedback through changes in snow, sea ice, and cloud cover (Yoshimori et al. 2014; Letterly et al. 2015). These influences contribute to the progressing of Arctic

Amplification. As short- to medium-range impacts, cyclones over the Arctic Ocean decreased sea water temperature at 0–20 m depth (Inoue and Hori, 2011). Besides, intense cyclones caused reduction of sea ice not only in summer (Parkinson and Comiso, 2013; Zhang et al., 2013a) but also in winter (Boisvert et al. 2016). However, the statistical relationship between cyclone activity and reduction of sea ice is unclear. Screen et al (2011) represented that the fewer cyclones over the Arctic during May to July caused to the lower September sea ice area. Simmonds and Keay (2009) showed the September sea ice extent reduction with increasing cyclone strength and size over the Arctic. In contrast, Koyama et al. (2017) indicated that the potential of cyclogenesis and an amount of moisture over the Arctic increase in the following year with less September sea ice, but frequency and intensity of individual cyclones show no consistent increase with the reduction of sea ice. In general, the number (strength) of cyclones in summer is greater (weaker) than in winter over the Arctic (Zhang et al. 2004; Serreze and Barrett, 2008, Sepp and Jaagus, 2011, Koyama et al. 2017).

1.2 Arctic cyclones

Among cyclones over the Arctic, a unique type of cyclone sometimes occurs, referred to as an Arctic Cyclone (AC) (Tanaka et al., 2012; Aizawa and Tanaka 2016). ACs have a different structure from mid-latitude cyclones during their mature stage. ACs are characterized by a warm core and a descending motion at the upper troposphere and the lower stratosphere and a cold core and an ascending motion at the lower troposphere. In addition, ACs show a barotropic structure of vorticity from a surface to lower stratosphere due to the establishment of a connection between a surface (lower-level) cyclone and an upper-level polar vortex. The connection of vortices causes a long lifetime of ACs (Simmonds and Rudeva, 2014), mainly due to the downward intrusion of positive potential vorticity from lower stratosphere to the upper troposphere (Tao et al., 2017). By contrast, Crawford and Serreze (2015, 2016) and Day

and Hodges (2017) reported that the cyclones migrating into the Arctic in summer were intensified by the baroclinicity of the Arctic frontal zone. Furthermore, Aizawa et al. (2014) also displayed that ACs intensified through merging with mesoscale cyclones generated over the Arctic frontal zone by numerical simulations using a cloud-resolving global model, called Nonhydrostatic ICosahedral Atmospheric Model (NICAM). For an extreme AC occurred in August 2016, Yamagami et al. (2017) indicated that both baroclinic instability and vertical connection of the vortices associated with the merging of surface cyclones contribute to the development of ACs.

1.3 Predictability of Arctic atmosphere

The human activities over the Arctic, like shipping on the Northern Sea Route (Eguíluz et al., 2016; Melia et al., 2016) and extraction of natural gas and oil resources, becomes accessible due to the recent reduction in sea ice reduction. Meanwhile, the cyclone activity over the Arctic becomes high in the summer and autumn. Therefore, accurate short- to medium-range weather forecasts are required to ensure safe human activities.

However, limited studies were conducted on predictability for Arctic atmosphere on short- to medium-range timescales. Bauer et al. (2015) demonstrated that medium-range forecasts showed a similar trend of forecast improvement over the Arctic compared with improvements at lower latitude using numerical weather prediction (NWP) models from the European Centre for Medium-Range Weather Forecasts (ECMWF) medium-range forecast. They also showed that the difference in initial analyses among NWP centers was large over the Arctic, especially in surface and lower-level variables. Jung and Matsueda (2016) showed that the predictability of the geopotential height at 500 hPa over the Arctic region in winter was comparable to that over mid-latitude regions using operational medium-range ensemble weather forecast systems. As demonstrated by Bauer et al. (2015), they also showed that the analysis uncertainty—

especially in near-surface variables—for the Arctic was larger than for mid-latitude regions. Since cyclones over the Arctic is affected by the baroclinicity in the Arctic frontal zone located between continent, ocean, and sea ice (Crawford and Serreze, 2016; Inoue and Hori, 2011; Serreze et al. 2001), the analysis uncertainties of near-surface variables can influence their prediction.

In addition to the lower-level fields, upper-level fields also have an impact on the predictability of the Arctic atmosphere. Data assimilation experiments for radiosonde observations at ice-free ocean by the Atmospheric general circulation model For the Earth Simulator (AFES) showed that these observations improved analyses and forecasts of upper-level temperature over the Arctic (Inoue et al., 2012), predictions of cold weather events at mid-latitude (Sato et al., 2017), and tropical and mid-latitude cyclones predictions (Sato et al., 2018). In particular, Yamazaki et al. (2015) investigated the improvement of prediction for the AC in August 2012 (AC12). Their results showed that the predictions assimilated with the additional radiosonde observations were improved compared with the predictions without the radiosonde observations, mainly due to an improvement of initial error for the upper polar vortex. Furthermore, forecasts of sea-ice cover over the Northern Sea Route were dependent on predicted AC12 (Ono et al., 2016). Therefore, accurate prediction of ACs is vital for the safety of human activity over the Arctic.

1.4 Purpose of this study

Ensemble forecasts with perturbed initial and/or boundary conditions were conducted by numerous operational NWP centers routinely to provide probabilistic information regarding high-impact weather and to estimate weather forecast uncertainties. Previous studies have used medium-range ensemble forecasts to investigate the forecast skill for various atmospheric phenomena in the tropics and mid-latitudes (e.g. Frame et al., 2011, 2013; Froude, 2010, 2011;

Matsueda, 2009, 2011; Matsueda and Endo, 2011; Matsueda and Kyouda, 2016; Matsueda and Palmer, 2018; Yamaguchi et al. 2012; Zhang et al., 2013b). Although extraordinary ACs in summer have significant impacts on the Arctic climate and human activities over the Arctic, their forecast skills have not yet been assessed.

In this study, extraordinary ACs in summer (June–August) and their predictability were investigated using medium-range ensemble forecasts, focusing on existence probability, central pressure, and central position at the mature stage.

Data and methods for cyclone detection and verification of ensemble forecasts are described in Chapter 2. Chapter 3 depicts the results of analyses for observed and predicted extraordinary ACs in summer. Discussion and conclusion are described in the Chapters 4 and 5, respectively.

CHAPTER 2 Data and Methods

2.1 Data

Operational medium-range ensemble forecast data from five leading NWP centers, available at the data portal of The Interactive Grand Global Ensemble (TIGGE, Swinbank et al., 2016), are used in this study. The NWP centers used in this study are the Canadian Meteorological Centre (CMC), ECMWF, the Japan Meteorological Agency (JMA), the US National Centers for Environmental Prediction (NCEP), and the United Kingdom Met Office (UKMO). These five NWP centers show higher forecast skill over the Northern Hemisphere (Matsueda and Tanaka, 2008; Swinbank et al. 2016) and the Arctic (Jung and Matsueda, 2016) than the remaining NWP centers.

Configurations of the five operational ensemble prediction systems (EPSs) are summarized in Table 2.1. In summary, JMA (UKMO) has the shortest forecast length of 9.0 (7.25) days until (after) November 2014. CMC and NCEP have the longest forecast length (16 days). The ensemble sizes of CMC and NCEP are the smallest (21 members), and ECMWF and JMA are the largest (51 members). After November 2014, UKMO has the smallest ensemble size of 12 members. All NWP centers changed the horizontal and vertical resolutions of forecast models more than once during 2008–2016. Initial perturbations for ECMWF and JMA are made by singular vector. On the other hand, initial perturbations for CMC, NCEP, and UKMO are made by the ensemble kalman filter, ensemble transform rescaling, and ensemble transform kalman filter, respectively. Only the ECMWF atmospheric model coupled with a dynamical ocean model. All EPSs uncouple with dynamical sea-ice model. The data assimilation method for CMC and NCEP are the Ensemble Kalman filter and the Grid-Point Statistical Interpolation (GSI, Klest et al., 2009). The other centers use the four-dimensional variational data assimilation method. Matsueda and Nakazawa (2015), Jung and Matsueda (2015), and Swinbank et al. (2016) give additional details of these EPSs. We use the ensemble forecast data

initialized at 1200 UTC on each day in summer (June–August) during 2008–2016. The ensemble forecast data are interpolated into a grid spacing of 2.5° and have a 6-hourly temporal resolution. Note that NCEP has some data gaps in 2016.

The second-generation Global Ensemble Forecast System (GEFS) reforecast data (Hamill et al., 2013) provided by National Ocean and Atmosphere Administration (NOAA) are also used in this study. This dataset is created using NCEP EPS with an approximately fixed NWP model and data assimilation system. The GEFS reforecast is a dataset consisting of medium-range ensemble forecasts initialized at 0000 UTC on each day from 1 December 1984 to the present. Configurations of EPS for the GEFS reforecast are summarized in Figure 2.1. The GEFS reforecast has the forecast length of 16 days and 11 ensemble members throughout the entire period. The GEFS reforecast was created by the NWP model with a resolution of T190L28 from 1 December 1985 to 19 February 2012. The resolution changed to T254L42 for day 0–8 and T190L42 for day 8–16 after 20 February 2012. The method of data assimilation was also upgraded twice during 1986–2016. The NCEP climate forecast system reanalysis (CFSR) was used as a control analysis (defined as an initial analysis of a control forecast) from 1 December 1984 to 19 February 2011. The operational GSI analysis was used as the control analysis of the GEFS reforecast from 20 February 2011 to 21 May 2012. After 22 May 2012, the control analysis was estimated by the Hybrid ensemble Kalman filter-variational analysis (Hamill et al. 2011). In contrast to operational EPSs, the GEFS reforecast data is created using an approximately fixed EPS. Therefore, an improvement of forecast skill in the Northern Hemisphere (Hamill et al., 2013) and the Arctic (Daily scores for GEFS reforecasts available at the TIGGE museum, (http://gpvjma.ccs.hpcc.jp/TIGGE/GEFS_score_all.html)) during 1986–2016 is expected to be primarily due to an increase in the number of observations used for data assimilation. The GEFS reforecast data were interpolated into a grid spacing of 1.0° , and a 12-hourly temporal resolution in all the period.

The ECMWF reanalysis (ERA-) Interim (Dee et al., 2011) is used for the detection of extraordinary ACs. The ERA-Interim are interpolated into a horizontal grid spacing of 2.5° and a 6-hourly temporal resolution. Forecasts of the five leading NWP centers in TIGGE was verified against the ERA-Interim. In contrast, the verification of the GEFS reforecast was conducted against own control analysis.

2.2 Methodology

2.2.1 Detection of extraordinary ACs

Detection of extraordinary ACs consists of two steps (Figure 2.1). First, all cyclone centers at each time step were detected by applying a method developed by Aizawa and Tanaka (2016) (Figure 2.1a). In this method, sea level pressure (SLP) of a common longitude–latitude grid is interpolated into an equal-distance x – y grid (40 km) centered on the North Pole. The SLP at the target grid point is compared with SLP averaged over all grid points between 500 and 550 km from the target grid point. When SLP at the target grid point is 2 hPa lower than the averaged SLP, and the derivative at the target grid point shows a local minimum for both x and y directions, the grid is regarded as a candidate for a cyclone center. Second, extraordinary ACs were selected from the identified cyclone centers in the first step. The cyclones satisfying the following three criteria were regarded as the extraordinary ACs (Figure 2.1b): the central pressure was less than 980 hPa, the central position was north of 70°N , and the area-averaged (within 800 km of the cyclone center) temperature anomaly at 250 hPa was higher than 5 K. The mature stage of each AC was defined as the time when the AC showed the lowest central pressure. Because all the NWP centers in TIGGE provides control analyses at 0000 and 1200 UTC on each day, the time of the mature stage of the ACs was selected at those times.

The cyclone tracking was conducted based on the nearest-neighbor method (Murray and Simmonds, 1991; Pezza et al., 2012), and the radius of each AC was measured based on an area of an outermost closed SLP contour with 1 hPa intervals including single cyclone center (Yamagami et al., 2017).

The centers of both ACs and mid-latitude cyclones can be detected using the method employed in this study. Here, ACs and mid-latitude cyclones were separated further based on their temperature in the upper troposphere at 250 hPa. A cyclone center was identified as an AC (mid-latitude cyclone) when the cyclone exhibited positive temperature anomalies aloft (in the rear of) the surface cyclone. In addition, an upper-level polar vortex with a warm core is referred as the tropopause polar vortex (TPV) in Cavallo and Hakim (2010).

2.2.2 Verification of the AC prediction

The same cyclone center detection method was applied to the ensemble forecast data. The existence probability of the ACs was assessed for each NWP center, at each forecast lead time. The existence of an AC at its mature stage was defined based on the cyclone center less than 995 hPa enclosed by at least two closed contours (circular or elliptical) with 5 hPa contour interval.

In order to investigate the forecasts for individual ACs, predicted ACs were classified based on thresholds for the great-circle distance and the difference in the central pressure between the predicted and observed ACs. The threshold for the difference in the central positions (central pressure) was set at 450 km (8 hPa). Ensemble members were classified into four groups: both the development and position of ACs were predicted well (higher-skill members); only the development was predicted well (only-development members); only the position was predicted well (only-position members); or both the development and the position were predicted poorly (lower-skill members). Welch's test was used to examine the significance of differences between the groups (Welch, 1947).

Existence probabilities and forecast skills for ACs at their mature stage was assessed at each lead time. Existence probabilities were defined as the ratio of ensemble members predicting the existence of ACs to the total number of ensemble members. Forecast skills were measured as the root mean square error of central pressure and the great-circle distance. The forecast skills for each AC were measured by the ensemble mean of the errors for the members that predicted the existence of the ACs. The average

central pressure and position errors for extraordinary ACs were calculated based on a weighted average of the errors by the number of ensemble members that predict the existence of each AC.

2.2.3 Strike probability and its verification using Brier Skill Score and reliability diagram

The strike probability was used as probabilistic forecasts for ACs existence. The strike probability for tropical cyclones (TCs) is defined as the probability that a TC will pass within a certain radius of a given location within a certain lead-time window (Van der Grijn et al., 2004; Yamaguchi et al., 2013). In this study, a radius using the calculation of strike probability for ACs was set to 400 km. The strike probability was calculated at each equal-distance x–y grid point north of 65°N (15,209 grid points). The probability was calculated based on the number of ensemble members that predicted the AC existence within 400 km radius.

Brier Skill Score (BSS) and reliability diagrams illustrate the performance and characteristics of a probabilistic forecast system. BSS was defined as the ratio of Brier Score (BS) calculated from ensemble forecasts to BS for a reference forecast (BS_{ref}). BS and BSS were calculated as follows:

$$BS = \frac{1}{N} \sum_i^N (p_i - o_i)^2,$$

$$BSS = \frac{BS_{ref} - BS}{BS_{ref}} = 1 - \frac{BS}{BS_{ref}},$$

where, p_i is the strike probability and o_i is the existence of observed AC within 400 km (0: none, 1: exist) at i -th grid point. N is the total number of samples (i.e., the number of grid points \times the number of AC events). The reference probability was calculated at each grid point and defined as the average existence over all AC events at their mature stage. When BSS is equal to 1.0 (i.e., BS is 0), the probabilistic forecast is perfect. On the other hand, when BSS is less than and equal to 0.0 (i.e., BS is greater than or equal to BS_{ref}), the probabilistic forecast provides useless information.

Reliability diagrams are graphical means to assess the reliability of probabilistic forecasts. The horizontal axis is the predicted probability and the vertical axis is the observed frequency. When the forecast probability agrees with the observed frequency, the distributions lie along the diagonal. In such a case, the forecast is considered to be reliable. When low-occurrence frequency is underestimated and high-occurrence frequency is overestimated, the distribution that is flatter than 45°. In this case, the

forecast is considered ‘overconfident’. In contrast, when low-occurrence frequency is overestimated and high-occurrence frequency is underestimated, a distribution that is steeper than 45°. In this case, the forecast is considered ‘underconfident’.

Table 2.1 Configurations for operational medium-range ensemble prediction systems from TIGGE project (as of August 2016). “TxLy” in model resolution indicates the horizontal total wavenumber truncation (x) and the number of vertical levels (y). DA, EnKF, SV, ETR, ETKF, 4D-Var, and GSI represent Data assimilation, Ensemble Kalman Filter, Singular Vector, Ensemble Transform with Rescaling, Ensemble Transform Kalman Filter, Four-dimensional Variational data assimilation, and Grid-point Statistical Interpolation, respectively.

NWP center	Forecast model resolution	Forecast length (days)	Ensemble size	Initial perturbation	DA	Sea surface temperature	Sea ice
CMC	0.9degL28 (– 16 Aug. 2011) 0.9degL40 (17 Aug. 2011 –)	16	21	EnKF	EnKF	uncoupled	uncoupled
ECMWF	TL399L62 (day 0-10), TL255L62 (day 11-15) (– 31 Oct. 2013) TL639L91 (day 0-10), TL319L91 (day 11-15) (1 Nov. 2013 –)	15	51	SV	4D-Var	uncoupled (– 31 Oct. 2013) coupled (1 Nov. 2013 –)	uncoupled
JMA	TL319L60 (– 25 Feb. 2014) TL479L60 (26 Feb. 2014 –)	9 (– 25 Feb. 2014) 11 (26 Feb. 2014 –)	51 (– 26 Feb. 2014) 27 (26 Feb. 2014 –)	SV	4D-Var	uncoupled	uncoupled
NCEP	T126L28 (– 22 Feb. 2010) T190L28 (23 Feb. 2010 – 13 Feb. 2012) T254L42 (day 0-8), T190L42 (day 9-16) (14 Feb. 2012 – 1 Dec. 2012) TL574L64 (day 0-8), TL372L64 (day 9-16) (2 Dec. 2012 –)	16	21	ETR	GSI	uncoupled	uncoupled
UKMO	N216L70 (– 29 Apr. 2013) N216L85 (30 Apr. 2013 – 5 Nov. 2014) N400L70 (6 Nov. 2014 –)	15 (– 5 Nov. 2014) 7.25 (6 Nov. 2014 –)	24 (– 5 Nov. 2014) 12 (6 Nov. 2014 –)	ETKF	4D-Var	uncoupled	uncoupled

Table 2.2 Forecast length, and ensemble size, forecast model resolution, and data assimilation method for the GEFS reforecast.

Period	Model resolution	Fcst. Length (days)	Ens. Size	Initial perturbation	Data assimilation	Sea surface temperature	Sea ice
01 Dec. 1984 – 19 Feb. 2011	T190L28	16	11	ETR	GSI from NCEP-CFSR (Saha et al., 2010) at T382L64	uncoupled	uncoupled
20 Feb. 2011 – 14 Feb. 2012					Operational GSI analysis at T574L64		
14 Feb. 2012 – 21 May. 2012	T254L42 (day 0 – 8)				Hybrid EnKF-variational analysis (Hamill et al., 2011)		
22 May. 2012 – 31 Dec. 2016	T190L42 (day 9 – 16)						

Cyclone centers & T250 anomaly (2008.06.23 0000UTC, ERA-Interim)

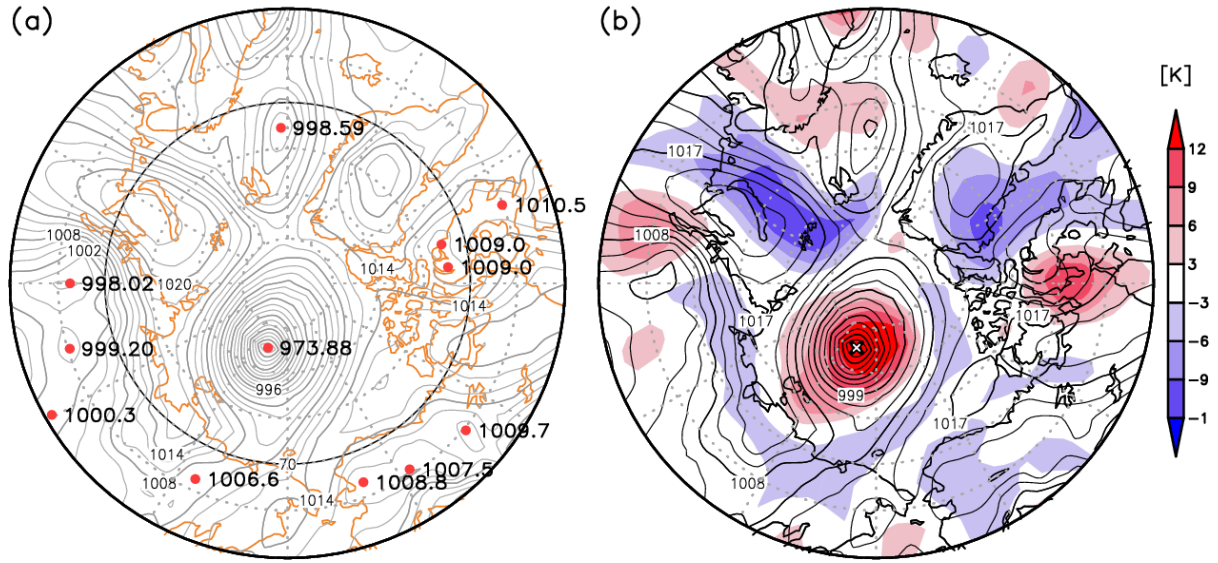


Figure 2.1 (a) Observed SLP (contour, 2 hPa interval) and all cyclone centers (red circle, the values are given on the right of each center). (b) Observed SLP (contour, 3 hPa interval) and temperature anomaly at 250 hPa (shading). The cyclone center of extraordinary AC are given by a white cross.

CHAPTER 3 Results

3.1 Observed extraordinary ACs in summer during 2008–2016 and their analysis uncertainties

Ten extraordinary ACs in summer during 2008–2016 were detected using ERA-Interim (Figure 3.1 and Table 3.1). Although an average of one or two extraordinary ACs occurred during these summers, no extraordinary ACs occurred in the summers of 2009, 2011, or 2014. Thus, the occurrence frequency of extraordinary ACs is dependent on year, owing to the strength of the upper-level polar vortex (e.g., Tao et al., 2017) and baroclinicity over the AFZ (e.g., Crawford and Serreze, 2016).

Of the 10 ACs, the AC12 had the lowest central pressure at its mature stage (965.3 hPa, Fig. 1e), followed by the AC in August 2016 (967.3 hPa, Fig. 3.1j). The AC12 also had the lowest central pressure anomaly of -43.7 hPa. Interestingly, the AC in June 2015 (Fig. 3.1h) showed the third lowest central pressure but the largest size (radius of 1370.5 km) of all the ACs and the second lowest central pressure anomaly (-42.3 hPa). A correlation between the central pressure and radius of the ACs was -0.65 , suggesting that the stronger AC tended to have the larger size. Except for the AC in August 2010, the ACs reached their mature stage over the Pacific sector of the Arctic Ocean, where maximum cyclone deepening occurs frequently (Serreze and Barrett, 2008). Besides the ACs in August 2012 and 2016 (Parkinson and Comiso, 2013; Petty et al., 2018), the other ACs, especially long-lived ACs, also influenced sea-ice and snow (Robinson, 2016).

As observations are limited over the Arctic (Jung et al., 2016), uncertainties in related analyses are larger over the Arctic than over the Northern mid-latitudes (Jung and Matsueda, 2016). Therefore, it is important to assess the difference between control analyses and ERA-Interim at the mature stage of the ACs.

Overall, the central position of each AC at its mature stage in the control analyses was comparable to that in the ERA-Interim (Figure 3.2a, b). The difference in central position between each control analysis and ERA-Interim was less than 90 km in the majority of events. The average differences for all AC events between ERA-Interim and the CMC, ECMWF, JMA, NCEP, and UKMO analyses were 36.7, 36.2, 42.6, 28.7, and 43.7 km, respectively. The largest difference (144.1 km) occurred between the ECMWF analysis and ERA-Interim for the AC in August 2016 (event j in Figure 3.2b), followed by 126.4 km between the CMC analysis and ERA-Interim for the AC in June 2008 (event a in Figure 3.2b), and 113.0 km between the UKMO analysis and ERA-interim for the AC in August 2016 (event j in Figure 3.2b). Although differences were larger for these ACs than for other ACs, differences were less than 15% of the radius of the individual ACs.

Regarding the analyzed central pressure (Figure 3.2c), the differences between each analysis and ERA-Interim were generally larger for the ACs from June 2008 to August 2012 (events a–e in Figure 3.2c) than for the ACs from July 2013 to 2016 (events f–i in Figure 3.2c). The differences for the AC in August 2016 (event j in Figure 3.2c) were, however, similar to those for the first five events. The central pressures in the CMC and UKMO analyses were generally higher than those in ERA-Interim. In contrast, the central pressures in the ECMWF analysis were lower than those in ERA-Interim, except for the ACs in June 2008 and August 2016. The average differences for all the events between ERA-Interim and the CMC, ECMWF, JMA, NCEP, and UKMO analyses were 2.1, –0.9, 0.2, 0.0, and 1.2 hPa, respectively. For individual ACs, the largest difference (10.3 hPa) occurred between the CMC analysis and ERA-Interim for the AC in June 2008 (event a in Figure 3.2c), followed by 4.1 hPa between the UKMO analysis and ERA-Interim for the same AC. Excluding the AC in June 2008 reduced the average difference for the remaining nine ACs between the CMC analysis and ERA-Interim to 1.2 hPa.

These comparisons indicate that there were insignificant differences in both the central pressure and position of the ACs at their mature stage. Therefore, ERA-Interim was treated as the observed conditions for verifying the forecasts in the following subsections.

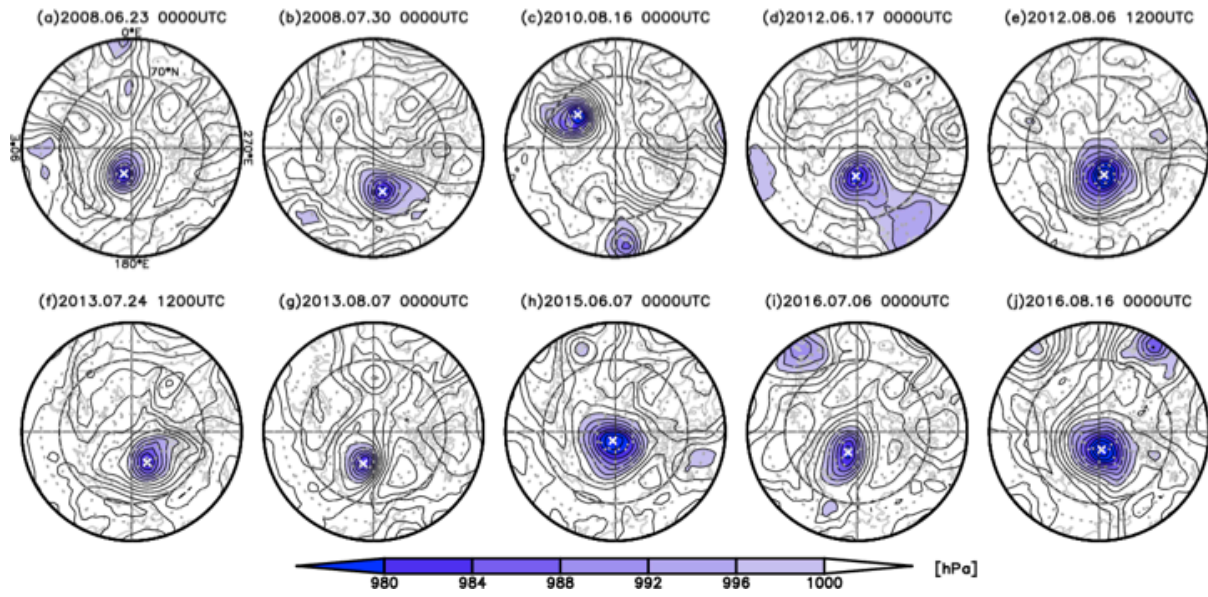


Figure 3.1 Observed sea-level pressure (SLP) fields for the 10 extraordinary ACs at their mature stage during summer 2008–2016 (ERA-Interim). The date of the mature stage is given above each panel. Shading represents SLP < 1000 hPa. The white cross indicates the AC center.

Table 3.1 Date, central pressure, central pressure anomaly from climatology, radius, and central position during the mature stage of each AC event.

Event (mature time)	(a) 0000UTC 23 Jun. 2008	(b) 0000 UTC 30 Jul. 2008	(c) 0000 UTC 16 Aug. 2010	(d) 0000 UTC 17 Jun. 2012	(e) 1200 UTC 06 Aug. 2012
Central pressure (hPa)	973.9	976.0	974.2	976.5	965.3
Central pressure anomaly (hPa)	-36.6	-34.9	-37.7	-35.5	-43.7
Radius (km)	797.4	849.3	1056.5	748.2	1107.8
Central position (lat, lon)	82.5°N, 163.3°E	77.5°N, 191.6°E	76.0°N, 48.1°E	82.1°N, 177.4°E	82.3°N, 190.8°E
Event (mature time)	(f) 1200 UTC 24 Jul. 2013	(g) 0000 UTC 07 Aug. 2013	(h) 0000 UTC 07 Jun. 2015	(i) 0000 UTC 06 Jul. 2016	(j) 0000 UTC 16 Aug. 2016
Central pressure (hPa)	978.1	979.1	971.0	978.9	967.6
Central pressure anomaly (hPa)	-32.3	-30.6	-42.3	-30.7	-41.2
Radius (km)	886.0	678.2	1370.5	878.5	1010.4
Central position (lat, lon)	80.4°N, 206.6°E	80.6°N, 162.3°E	87.4°N, 164.1°E	83.7°N, 156.4°E	84.9°N, 188.1°E

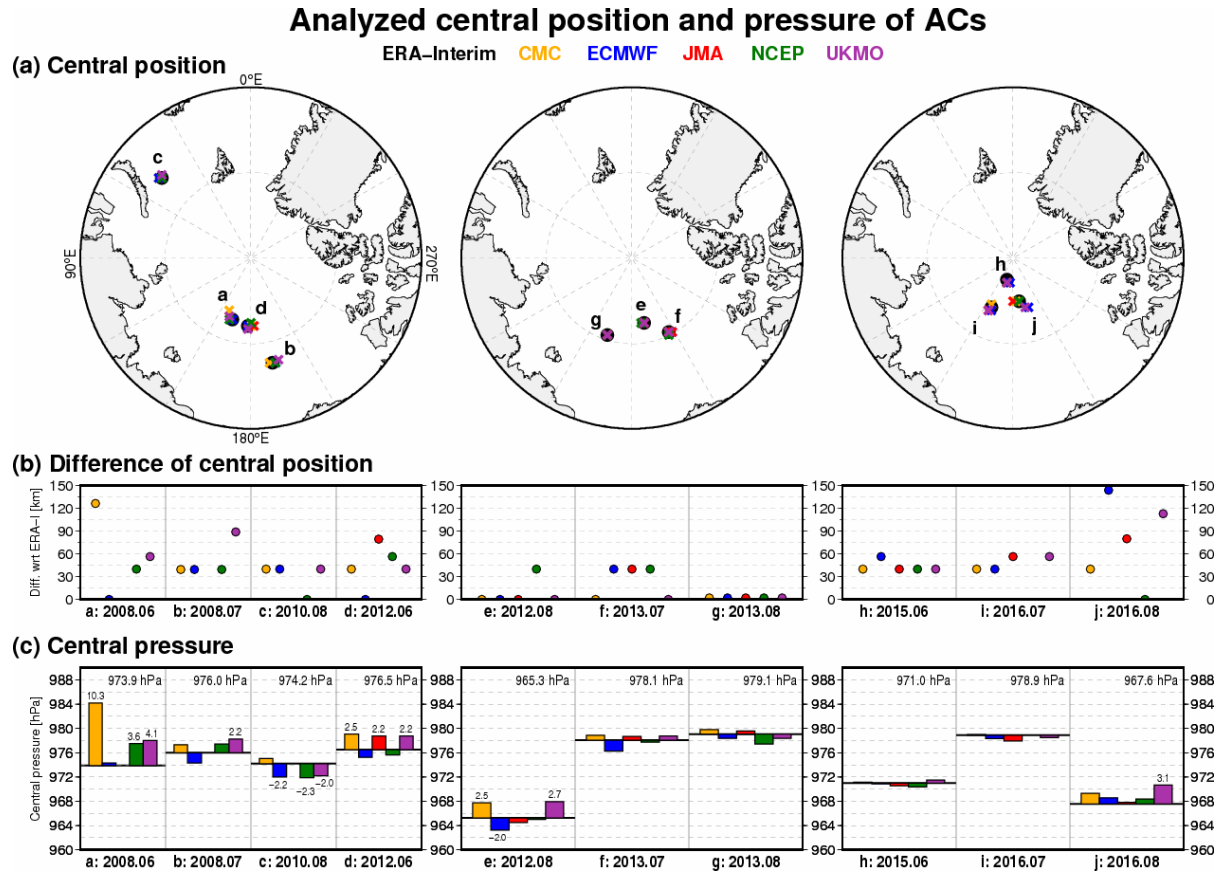


Figure 3.2 Central positions of the analyzed ACs in ERA-Interim (black circle) and control analyses (cross) of CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), and (b) the difference of central position between ERA-Interim and the control analysis. (c) Central pressures of the analyzed ACs for ERA-Interim (black horizontal line; the value is given at the top-right corner of each frame) and each control analysis (colored bars; differences >2.0 hPa between ERA-Interim and each analysis are given above or below each bar). The letters (a–j) for each symbol in (a) correspond to the AC events shown in (b), (c), Fig. 3.1, and Table 3.1.

3.2 Case studies of extraordinary AC predictions

3.2.1 AC in August 2012 – the strongest AC – (Yamagami et al., 2018a)

AC in August 2012 (AC12) had the lowest central pressure of all the extraordinary ACs in summer during 2008–2016 (Table 3.1). A cyclone center first detected to the east of an upper-level trough over Siberia at 1800 UTC on 2 August (Figure 3.3a and f). The cyclone moved northeastward and migrated to the north of 70°N from the coast of the East Siberian Sea on 4 August (Figure 3.3b and c). The cyclone was classified as the mid-latitude cyclone, because the upper-level trough accompanied by a warm core was located over the west of the surface cyclone (Figure 3.3g and h). Meanwhile, a cyclone coupling with the TPV and having an upper-level warm core existed over the East Siberian Sea on 3 August (Figure 3.3b and g). Given its structure, the cyclone over the East Siberian Sea was classified as an AC. On 4–5 August (Figure 3.3c, d, and m), the AC and the mid-latitude cyclone merged over the East Siberian Sea, forming a multicenter cyclone (Hanley and Caballero, 2012). The upper-level warm cores associated with the trough over Siberia and the TPV also merged during the merging of the surface cyclones, and the merged warm core broadened and intensified (Figure 3.3h and i). The merged AC12 reached its mature stage at 1200 UTC on 6 August (Figure 3.3e, j, and o). AC12 wandered over the Arctic Ocean for 8 days after its mature stage, and dissipated over the Canadian Arctic Archipelago on 14 August (Simmonds and Rudeva, 2012).

Some ensemble members initialized on 1 August failed to predict the cyclogenesis position of the mid-latitude at 1800 UTC on 2 August (Figure 3.4a). In addition, almost all members predicted a slower moving cyclone or a location that was farther south, except for some ECMWF and NCEP members. The cyclogenesis prediction was improved in the members initialized on 2 August, in which the cyclogenesis and the subsequent northeastward movement were predicted well (Figure 3.4b). The cyclone track spread, however, became large when the predicted mid-latitude cyclones passed over the East Siberian Sea, resulted in the different position at their mature time. In the forecast initialized on 3 August (Figure

3.4c), all the members show an improvement in predicting the track and development, but with a slight southward shift in the predicted position at the mature time.

For the prediction of AC12, several CMC, ECMWF, and NCEP members initialized at the end of July predicted the development and position of AC12 with high skill (Figure 3.5). Members that predicted the correct position existed in all five centers' forecasts, and the number of members was much larger than the number of ECMWF and NCEP members that predicted the correct central pressure. This suggests that the position of AC12 is more predictable than its development. However, most members predicted the development and position of AC12 poorly. Forecasts initialized on 1–4 August gradually improved in predicting AC12. Predictions initialized on 2 and 3 August include a mix of higher-skill, only-development, only-position, and lower-skill members for all NWP centers, suggesting that the spread in the predicted central pressure of AC12 was the largest on these days. The 9 (25) ECMWF members initialized on 2 (3) August predicted the development and position accurately. The CMC, NCEP, and UKMO ensembles initialized on 3 (4) August showed large improvement in predicting the development (position) of AC12. Almost all members initialized on 4 August predicted the position and development of AC12 accurately. Similarly, the JMA ensembles initialized on 4 August showed a large improvement. However, some JMA members initialized on 4 August still failed to predicted the development of AC12. Thus, the JMA ensemble may have difficulty in predicting the development of AC12. These results suggest that predictability of AC12 in operational EPSs is no more than 2–3 days.

Composites of predicted SLP and temperature anomaly at 250 hPa (T250 anomaly) for the five best- and worst-performing members at 1200 UTC on 6 August are shown in Figure 3.6. The best five members were defined as forecasts with a position error of ≤ 450 km, and a lower central pressure than that of other members. The worst five members had the largest position error, as the position error was proportional to the central pressure error of the predicted AC12 (Figure 3.5a). Along with deepening of the central pressure of AC12, the size of AC12 and the strong and broad warm core over AC12 was correctly predicted by the best five members, except CMC (note that all the CMC members failed to

predict the development and position of AC12). The strength of the predicted warm core by the best performing ECMWF, JMA, NCEP, and UKMO members was similar to that of the analyzed warm core (Figure 3.5b), but the northward expansion of the warm core was narrower. In contrast, the T250 anomaly (Figure 3.5c) reveals that CMC, ECMWF, JMA, and UKMO predicted two warm cores over the Arctic region. One of the warm cores was predicted over the Pacific sector of the Arctic Ocean, and the other south or southwest of that. NCEP predicted a weak warm core with two peaks over the Arctic Ocean. The significant difference in T250 anomaly between the five best- and worst-performing members corresponds to the center of the warm core in the best five members, and the two warm cores in the worst five members. Therefore, merging of the warm cores accompanying the mid-latitude cyclone from Siberia and the AC over the Arctic caused the development of the warm core of AC12. This implies that an accurate prediction of the merging of the warm cores led to better prediction of the development of AC12.

The difference of the upper-level predicted warm core at the mature stage is caused by the difference in predicted upper-level geopotential height. The comparison of geopotential height at 300 hPa (Z300) between higher-skill and lower-skill ECMWF members initialized at 1200 UTC on 2 August (Figure 3.7a and b) shows that the significant difference appeared at the trough over Siberia at 1200 UTC on 4 August. The trough in lower-skill members was deep compare with that in higher-skill members. At the same time, positive and negative differences in Z300 appeared at the western and eastern parts of TPV, respectively. These differences developed exponentially with increasing lead-time, resulted in the deeper trough and the northward shift of TPV at 0000 UTC on 6 August in lower-skill members. The difference in predicted T250 anomaly was consistent with that in Z300 (Figure 3.8b). A significant negative difference appeared around the boundary between the two warm cores at 0000 UTC on 5 August. At 1200 UTC on 5 August, the warm core accompanied by the trough merged with that in TPV in higher-skill members, while the warm cores did not merge in lower-skill members due to both

the southward displacement of trough and the different position of TPV. As a result, both the development and position of AC12 failed to predict in the lower-skill members.

The comparison between higher-skill and only-position members (Figure 3.7c and Figure 3.8c) shows that a significant negative difference in Z300 appeared on the ridge over the Laptev Sea at 0000 and 1200 UTC on 4 August. In addition to the negative difference, significant positive difference appeared on the west of the trough over Siberia at 1200 UTC on 4 August, as with the difference between higher- and lower-skill members. At 0000 and 1200 UTC on 5 August, positive and negative differences appeared in the western part and at the center of the polar vortex, respectively, suggesting that the polar vortex in the only-position members was farther eastward than in the higher-skill members. The polar vortex in the higher-skill members was round in shape at 0000 UTC on 6 August, whereas the vortex in the only-position members was oval shaped, with the long axis stretching from the East Siberian Sea to the central Arctic Ocean. The central position of the polar vortex was similar in both categories. At 1200 UTC on 5 August, two warm cores joined and formed a merged warm core with a center over the coast of northern Siberia in the higher-skill members, whereas the merging of two warm cores was underway in the only-position members (Figure 3.8c). Concurrently, a significant negative difference appeared at the boundary between the two peaks. At 0000 UTC on 6 August, the higher-skill members had a circular warm core, while the only-position members had the warm core extending from the coast of the East Siberian Sea to the central Arctic Ocean. Overall, the difference in predictions of AC12 between the higher-skill and only-position members is predominantly associated with the difference in timing of the merging of warm cores. Interestingly, some only-position members predicted that the lowest central pressure of AC12 occurred after 1800 UTC on 6 August (not shown). This delay suggests that the main difference between the only-position and other members was the timing of merging of upper-level warm cores. Moreover, this delay highlights the importance of upper-level warm cores to the development of the AC12.

In contrast to the previous two categories, the difference in Z300 between higher-skill and only-development ECMWF members, initialized on 3 August, was not significant at the troughs over Siberia at on 4 August (Figure 3.9a and b). However, a significant positive difference appeared around TPV on 4 August, and the difference developed with increasing lead-time. The positive difference indicates that the TPV in only-development members was shifted south compared with that in higher-skill members. Besides, the significant negative difference appeared at the trough over the Siberia at 0000 UTC on 5 August, indicating a larger meandering of the trough and ridge in only-development members. The positive difference persisted until 0000 UTC on 6 August. The T250 anomaly shows significant negative and positive differences at eastern and western parts of warm core over the East Siberian Sea, respectively, at 0000 UTC on 5 to 6 August (Figure 3.9c and d). Thus, the merging position of the warm cores in only-development members was farther westward compared with that in higher-skill members due to the deeper trough and the southwestward-shifted TPV. Although the predicted position of TPV was different, the merging of two warm cores and consequent development were accurately predicted in only-development members.

3.2.2 AC in August 2013 – short period from the generation to mature –

The AC in August 2013 (AC13; Figure 3.10g) is characterized by a short period from its cyclogenesis to mature stage. The cyclone center was first identified in the northern part of central Siberia at 1800 UTC on 3 August, where it remained until 1200 UTC on 4 August (Figure 3.10a). An upper-level trough was located west of the cyclone center, indicating that this cyclone can be classified as a mid-latitude cyclone (Figure 3.10f). Concurrently, a weak cyclone with an AC structure existed over the Pacific side of the Arctic Ocean (Figure 3.10a, f, and k). The mid-latitude cyclone moved along the north coast of Siberia and intensified gradually (~ 11.6 hPa/day) at 1200 UTC on 4–6 August, associated with eastward movement of the upper-level trough (Figure 3.10 b–d, and g–i). Although the surface AC disappeared at 0000 UTC on 6 August (Figure 3.10c), the upper-level warm core associated with the

trough over Siberia merged with the warm core associated with TPV during the eastward movement of the mid-latitude cyclone. At 0000 UTC on 7 August, the mid-latitude cyclone moved over the central Arctic Ocean and attained its lowest central pressure of 979.1 hPa. The cyclone became coupled with TPV, indicating that it developed to the AC structure during its mature stage (Figure 3.10e, j, and o).

The central pressure predictions initialized on 31 July to 6 August show that operational EPSs have difficulty predicting the existence of AC13 during its mature stage (Figure 3.11). In forecasts initialized from 31 July to 2 August, all NWP centers show a low existence probability (~ 0.3). In particular, only two JMA members initialized on 1 August predicted the existence of AC13 at its mature stage (Figure 3.11c). Thus, both the cyclogenesis and maintenance of AC13 were difficult to predict in 4.5–6.5-day forecasts. The existence probability improved significantly in the forecasts initialized on 3 August in all NWP centers (>0.76). In addition, more than half of the ECMWF and UKMO members performs well in predicting the development and location of AC13 (blue symbols in Figure 3.11). In forecasts initialized on 4 August, almost all of the CMC, ECMWF, and UKMO members show better skill in predicting the development and location of AC13, whereas half of JMA and NCEP members did not predict its development. In forecasts initialized on 5 and 6 August, all the members predicted the development and location of AC13 correctly, except for a few CMC, JMA, and NCEP members initialized on 5 August. Therefore, the predictability of AC13 was largely similar to that of AC12, although the difficulty in predicting AC13 was related to its generation and maintenance rather than the merging.

The SLP and Z300 fields (Figure 3.12) indicate that cyclogenesis of the mid-latitude cyclone over central Siberia was induced by baroclinicity. Thus, the upper-level trough played an important role in the generation and maintenance (development) of the surface cyclone. Composited Z300 fields for members with and without AC13 are shown in Figure 3.12. For members that show AC13 initializing on 3 August, the trough was located west of the center of the cyclone (Figure 3.12a), indicating a developing stage for the cyclone. For these members, the meandering of the trough over the Kara Sea

(~9030 m contour) is consistent with that over western Siberia (~9310 m contour). As a result, the deep trough with an axis stretching from western Siberia to the central Arctic Ocean formed over the west of the surface cyclone center. In contrast, for members that did not show the AC initializing on 2 August, the trough over the Kara Sea did not meander (Figure 3.12b). Accordingly, there exists a significant positive difference in Z300 over the Kara Sea in all NWP centers between members with and without AC13. Consequently, the trough over western Siberia was separated from the trough over the Kara Sea and did not form the broad trough seen in Figure 3.12a. This significant positive difference did not appear in the comparison between members for which AC13 initialized on 2 and 3 August (Figure 3.12c), suggesting that the positive difference is a key factor in the prediction of cyclogenesis and maintenance of AC13. In addition, a significant negative difference is apparent over the Canadian Arctic Archipelago (CAA) in both Fig 3.11b and c, indicating that it is insignificant in predicting the existence of AC13.

The comparison between ECMWF members with and without AC for each lead time shows that the significant positive difference in Z300 at 1200 UTC on 5 August originated from the positive difference over the Barents Sea at 1200 UTC on 3 August (Figure 3.13). The positive difference at 1200 UTC on 3 August developed over 24 hours. This result suggests that the small trough over the Barents Sea at 1200 UTC on 3 August was difficult to predict correctly in the ECMWF forecast initialized on 2 August. These results also indicate that prediction of the depth of the trough was difficult in the forecasts initialized from 31 July to 2 August because these forecasts showed a low probability of existence. In the members without AC13, this growth in the positive error over the Barents Sea was seen not only in ECMWF but also in the other NWP centers. Therefore, all NWP centers had difficulty predicting the depth of the trough over the Barents Sea. In addition, the lower-level geopotential and temperature fields between members with and without AC13 show small region in which all NWP centers agree on the sign of the difference (not shown). In conclusion, the difficulty in predicting the cyclogenesis and maintenance of the mid-latitude cyclone was due to the difficulty in predicting the upper-level trough in the AC13 case.

3.2.3 AC in August 2016 – the AC with multiple merges –

AC in August 2016 (AC16) was maintained for more than one month due to multiple merges with cyclones both generated in the Arctic and migrating from lower latitudes (Yamagami et al. 2017). AC16 was similar to the AC12 not only in central pressure but also in radius and central position at mature stage (Figure 3.1 and Table 3.1).

At 0000 UTC on 10 August, two cyclones existed over the Scandinavian Peninsula and the central Arctic Ocean (Figure 3.14a). The upper-level TPV accompanied by a warm core was located above the cyclone over the Arctic Ocean (Figure 3.14f), indicating that the cyclone had an AC structure. In contrast, an upper-level trough appeared over the North Atlantic, suggesting that the cyclone over the Scandinavian Peninsula was classified as a mid-latitude cyclone. The mid-latitude cyclone traveled along the north coast of Siberia, and formed a multicenter cyclone at 0000 UTC on 12 August (Figure 3.14b). While the merged AC was over the central Arctic Ocean, another mid-latitude cyclone migrated to the Arctic from the Scandinavian Peninsula at 1800 UTC on 13 August (Figure 3.14c, h, and m). The mid-latitude cyclone intensified significantly due to baroclinicity on 13–14 August, and it formed a multicenter cyclone together with the AC at 1800 UTC on 14 August (Figure 3.14d and n). At 0000 UTC on 16 August, the merged AC recorded the lowest central pressure (Figure 3.14e and o). Upper-level warm cores associated with the mid-latitude cyclone and the AC also merged on 14–16 August (Figure 3.14i and j).

Almost all members initialized on 9 August failed to predict the development and/or position of AC16 (Figure 3.15). The ECMWF forecast was improved in predicting central pressure at 10 August, except for two ensemble members with central pressure of ~990 hPa. The central position forecasts initialized on 11 August showed a vast improvement, particularly in the CMC, ECMWF and UKMO forecasts. The CMC and UKMO members initialized on the same date also showed significant improvement in prediction of central pressure. In addition, 16 NCEP members initialized on 11 August

accurately predicted the central position of AC16. JMA shows improvements for both central pressure and position in the forecasts initialized on 11 August and on 13 August. These results suggest that the predictions of central pressure and position were improved simultaneously with decreasing lead-time. Predicted central pressures in CMC, ECMWF, and JMA members initialized on 13–15 August are similar to the analyzed central pressure in ERA-Interim. In contrast, those in the NCEP and UKMO members were weaker than analyzed central pressure. The central pressure in the UKMO control analysis was ~ 3.1 hPa higher than that in ERA-Interim (Figure 3.2c), which is consistent with the weaker predicted AC in the UKMO forecasts. These results suggest that the predictability of AC16 is about 3–4 days on average.

The forecast error of SLP in the only-position ECMWF members initialized on 11 August were small around the multicenter cyclone over the central Arctic Ocean at 0000 UTC on 12 August (Figure 3.16b). A positive error appeared around the center of merged AC on 13 August, which was seen in all NWP centers. Positive and negative errors in the south and north of the migrating mid-latitude cyclone in the only-position ECMWF members at 0000 UTC on 14 August, implying poleward shifted track of migrating cyclone. The positive and negative errors persisted on 14–15 August, and the positive error was common to all NWP centers. In addition to the poleward shift of cyclone track, the positive error indicates that the predicted mid-latitude cyclone was weak compared with the analyzed mid-latitude cyclone. The cyclone merging between the mid-latitude cyclone and the AC was, however, predicted correctly in only-position members at 0000 UTC on 15 August. Since merged AC shows the correct position in only-position members, the cyclone merging contributes mainly to not the development but the position of AC16.

Lower-skill ECMWF members initialized on 10 August also show the positive error around the merged AC center on 13 August (Figure 3.16c). On 14 August, positive error of more than 7 hPa appeared around the mid-latitude cyclone, indicating that the development of the cyclone was weak compared to the analyzed cyclone. In addition to the positive error, a negative error appeared over the

East Siberian Sea at 1200 UTC on 14 August. These positive and negative errors developed and persisted until 1200 UTC on 15 August. These errors indicate that the predicted location of the merged AC was correct on 13 August, although its development was weak compared with the analyzed AC. After 14 August, the predicted mid-latitude cyclone was shifted southeast compared with the analyzed position. As a result, the merging between the mid-latitude cyclone and AC was not predicted correctly.

The predicted trough over the mid-latitude cyclone shows a positive error at 1200 UTC on 10 August in lower-skill ECMWF members initialized on 10 August (Figure 3.17). In consistent with the positive error, a negative error appeared in the T250 anomaly around the trough (Figure 3.17c). These errors indicate that the shallower trough with the weaker warm core lead to the baroclinic development of the mid-latitude cyclone in the lower-skill members (Figure 3.16c). In addition, some lower-skill members did not predict the cyclogenesis of the mid-latitude cyclone on 12 August, as with AC13 prediction (not shown). Concurrently, a negative error appeared around the ridge and trough over the central Siberia (Figure 3.17b). These positive and negative errors developed at 1200 UTC on 13 August, indicating that the predicted meandering of steering wind is smaller in lower-skill members than analyzed meandering. In T250 anomaly, a negative error appeared at the warm core over the Arctic Ocean on 13 August (Figure 3.17c), suggesting that the lower-skill members failed to predict the merging of warm cores on 12–13 August. The positive and negative errors in Z300 developed around the TPV on 14 August, and the positive error combined with a positive error over the CAA (Figure 3.17b). These positive and negative errors indicate that the predicted TPV was displacement to the Pacific side compared to the analyzed TPV. In the T250 anomaly, the positive error appeared over the Pacific side of the warm core, and the negative error appeared at the center of the warm core on 14–15 August (Figure 3.17c). Therefore, mid-latitude cyclone migrated to the southeast of analyzed position due to the weak meandering of upper-level trough and ridge, resulted in the failure of merging of upper-level warm cores. The positive and negative errors in Z300 and T250 anomaly appeared in the lower-skill members in all NWP centers. Note that the common errors among all NWP centers' lower-skill

members corresponded to the significant difference (95 % significance level) between higher-skill and lower-skill ECMWF members.

These results indicate that the cyclone merging on 12–13 August (first merging), baroclinic growth, and the cyclone merging on 14–15 August (second merging) are a key factor in prediction of AC16. When the three processes are correctly predicted, the predicted AC would have accurate central pressure and position (higher-skill members). When only the baroclinic growth of the mid-latitude cyclone were correctly predicted, the predicted AC would be categorized as the only-development members. When only the second merging was predicted well, the position of predicted AC would be correct (only-position members). When a forecast failed to predict all the processes, the predicted AC would be categorized as the lower-skill members.

3.2.4 AC in June 2015 –the largest AC without merging –

The AC in June 2015 (AC15) has the largest size of all ACs (Figure 3.1 and Table 3.1). AC15 was generated over the North America, and it reached over the North Atlantic at 0000 UTC on 1 June (Figure 3.18a). AC15 migrated in the Arctic from the north of the Scandinavian Peninsula on 3–4 June (Figure 3.18b, c). The structure of AC15 was classified as the mid-latitude cyclone because the upper-level trough existed over the west of surface cyclone on 4 August (Figure 3.18c, h). The upper-level geostrophic wind oriented to the central Arctic Ocean and AC15 traveled further northward on 5 August (Figure 3.18d and i). Thus, the traveling of AC15 was attributed to the upper-level steering wind. AC15 reached its lowest central pressure (971.0 hPa) near the North Pole at 0000 UTC on 7 June (Figure 3.18e). Concurrently, TPV was located above the surface cyclone, indicating that the structure of AC15 became that of ACs (Figure 3.18j). In contrast to the previous three ACs, there was no ACs over the Arctic during the development of AC15. Although AC15 formed a multicenter cyclone on 2–3 June (Figure 3.18a, b, k, and l), a period from the cyclone merging to the mature stage of AC15 was ~3 days. Therefore, the merging might have a little contribution to the development of AC15.

Almost all ensemble members initialized from 31 May to 2 June failed to predict the position of AC15 (Figure 3.19). In contrast, the development of AC15 was accurately predicted by some members in all NWP centers at these lead times. The number of only-development members increased in CMC, ECMWF, and JMA predictions initialized on 1–2 June. The NCEP forecast initialized on 2 June shows an improvement in predicting the central pressure of AC15. The central position prediction shows vast improvement in the CMC, ECMWF, NCEP, and UKMO forecasts (the JMA forecast) initialized on 3 (4) August. Almost all members initialized on 4–6 June predicted the central pressure and position of AC15 correctly. These results indicate that the development of AC15 was more predictable than its central position. In addition, existence of AC15 at its mature stage was accurately predicted in all ensemble members throughout the entire lead times.

All ensemble members initialized on 1 June accurately predicted the initial position and subsequent northeastward traveling of AC15 (Figure 3.20a). The predicted cyclone tracks, however, began to differ between members when predicted AC15 passed over the northern part of the Scandinavian Peninsula. Some members predicted a slower moving of AC15. In particular, all NCEP members predicted the location of AC15 over Svalbard at 0000 UTC on 7 June. Others predicted the cyclone tracks with an eastward shift, especially the ECMWF members. In forecasts initialized on 2 June (Figure 3.20b), the predicted AC15 was also located further south from the observed position at the mature stage. Almost all the CMC, ECMWF, JMA, and NCEP members initialized on 2 June predicted central pressures <980 hPa. The slower traveling speed of predicted AC15 was improved in the forecasts initialized on 3 August, except for some CMC and JMA members (Figure 3.20c).

Since no ACs existed over the Arctic during the development of AC15, it is expected that the upper-level trough and ridge play an important role in its development and position. The difference in Z300 between only-development and lower-skill ECMWF members initialized at 1200 UTC on 1 June shows significant differences around trough and ridge over the surface cyclone (Figure 3.21a and b). A significant positive difference appeared over the Barents at 1200 UTC on 2 August, and the difference

developed on 3–5 June. Meanwhile, another positive difference appeared over the North Atlantic. The positive difference around the trough over the surface cyclone indicates the westward shift of steering wind in lower-skill members. In addition, the exit region of the steering wind over the central Siberia oriented westward in lower-skill members compared with that in higher-skill members. As a result, the surface cyclone traveled further west. Hence, predicted AC15 was coupled with TPV at lower latitude and did not developed due to baroclinic instability in lower-skill members.

The significant difference in the upper-level temperature anomaly between higher- and lower-skill members did not appear (not shown.). In contrast, lower-level temperature (temperature at 850 hPa) shows significant differences on 4–6 June (Figure 3.20c and d). Positive and negative difference appeared around the positive temperature anomaly, indicating that the warm advection due to cyclonic wind of AC15 was shifted westward in lower-skill members. This westward shift of warm anomaly corresponds to that of the upper-level steering wind. Similar significant difference appeared in the comparisons between only-development and lower skill in the other NWP centers (not shown).

The comparison between higher-skill and only-development ECMWF members initialized on 2 June was shown in Figure 3.22. In this comparison, only-development members were divided into two groups. One group contains the members with AC's central position of $< 90^{\circ}\text{E}$ (western only-development members; Figure 3.22b). The other contains the members with AC's central position of $>90^{\circ}\text{E}$ (eastern only-development members; Figure 3.22c). The comparison between higher-skill and western only-development members shows that a significant positive difference appeared over the central Arctic at 1200 UTC on 2 August (Fig. 3.22b). This positive difference developed during 2–4 June, and the difference was located at the exit of upper-level steering wind at 1200 UTC on 5 June. At 1200 UTC on 6 August, positive and negative differences appeared in the TPV over the Barents Sea, though the negative difference is not significant. These differences indicated southward shift of TPV in western only-development members compared with that in higher-skill members. In contrast, the comparison between higher-skill and eastern only-development members shows that the significant

negative difference appeared around the exit of steering wind at 1200 UTC on 6 August (Fig. 3.22c). The origin of the negative difference in eastern only-development members was an insignificant negative difference over the northern part of the Scandinavian Peninsula at 1200 UTC on 2 June. As a result of the significant negative difference, the upper-level steering wind in eastern only-development members displaced southeast compared to that in the higher-skill members. Central position of TPV in eastern only-development members was almost similar to that in higher-skill members at 1200 UTC on 6 June. These comparisons indicate that the prediction of an upper-level trough and ridge has a large influence on the prediction of central position of AC15. These results also revealed that the difference in upper-level fields over the central Arctic at initial time influenced to position forecasts of AC15, although there is no pre-existing AC over the Arctic.

3.2.5 Summary of case studies

Accurate predictions of cyclogenesis, cyclone traveling, and cyclone merging lead to accurate predictions of development and location of extraordinary ACs. Upper-level fields associated with trough, ridge, and TPV have significant impacts on these three processes.

For ACs with merging event (Figure 3.23a), first key process for accurate prediction of development and location of ACs is the cyclogenesis of migrating (mid-latitude) cyclone. The cyclogenesis is predicted accurately in higher-skill, only-development, and only-position members. In contrast, when a forecast failed to predict cyclogenesis, it is classified as lower-skill members (red; e.g., AC16) or non-existent members (gray; e.g., AC13 and AC16). When the cyclogenesis is predicted correctly, forecast skills of ACs depend on predicted upper-level fields. When upper-level trough and ridge, and position of TPV predicted correctly, merges between mid-latitude cyclones and ACs were predicted accurately. As a result, predicted AC shows correct central pressure and position. These forecasts are classified as higher-skill members (blue). When predictions of TPV position are accurate and predictions of trough and

ridge are poor, those predictions did not predict development of AC due to a lack of cyclone merging. In this case, the predicted AC shows an accurate position because the location of TPV was predicted well. These forecasts were classified as only-position members (green; e.g., AC12). When forecasts are classified as only-development members (orange), prediction skills for trough and ridge and positions of TPV shows several combinations. First, the upper-level trough and ridge was predicted accurately, but position of TPV was not predicted accurately. Second, both the upper-level trough and ridge and position of TPV was not predicted correctly. In these two cases, merges of upper-level warm cores associated with merges of surface cyclones are predicted correctly at different location from observation. The merging of upper-level warm cores and subsequent development of merged warm core lead to the development of ACs. Hence these predictions are classified as only-development members (e.g., AC12). When baroclinic instability is the main contributor to the development of ACs, the development is predicted accurately without cyclone merging prediction (e.g., AC16).

The combination of forecast skills for individual processes in each classification for ACs without merging event is simple compared with that for ACs with merging (Figure 3.23b; e.g., AC15). When a forecast fail to predict cyclogenesis of ACs, it is classified as non-existent members (gray). When cyclogenesis, an upper-level trough and ridge, and position of TPV were predicted accurately, the forecast classified as higher-skill members (blue). When upper-level trough and ridge are not predicted correctly, the forecast is classified as only-development (orange), only-position (green), or lower-skill members (red). When a relation of locations between a surface cyclone and trough are favorable to the baroclinic development, the forecast is classified as only-development members. When predicted upper-level steering winds are accurate, the forecast is classified as the only-development members. When the baroclinic growth and steering wind was not predicted correctly, the forecast was classified as the lower-skill members.

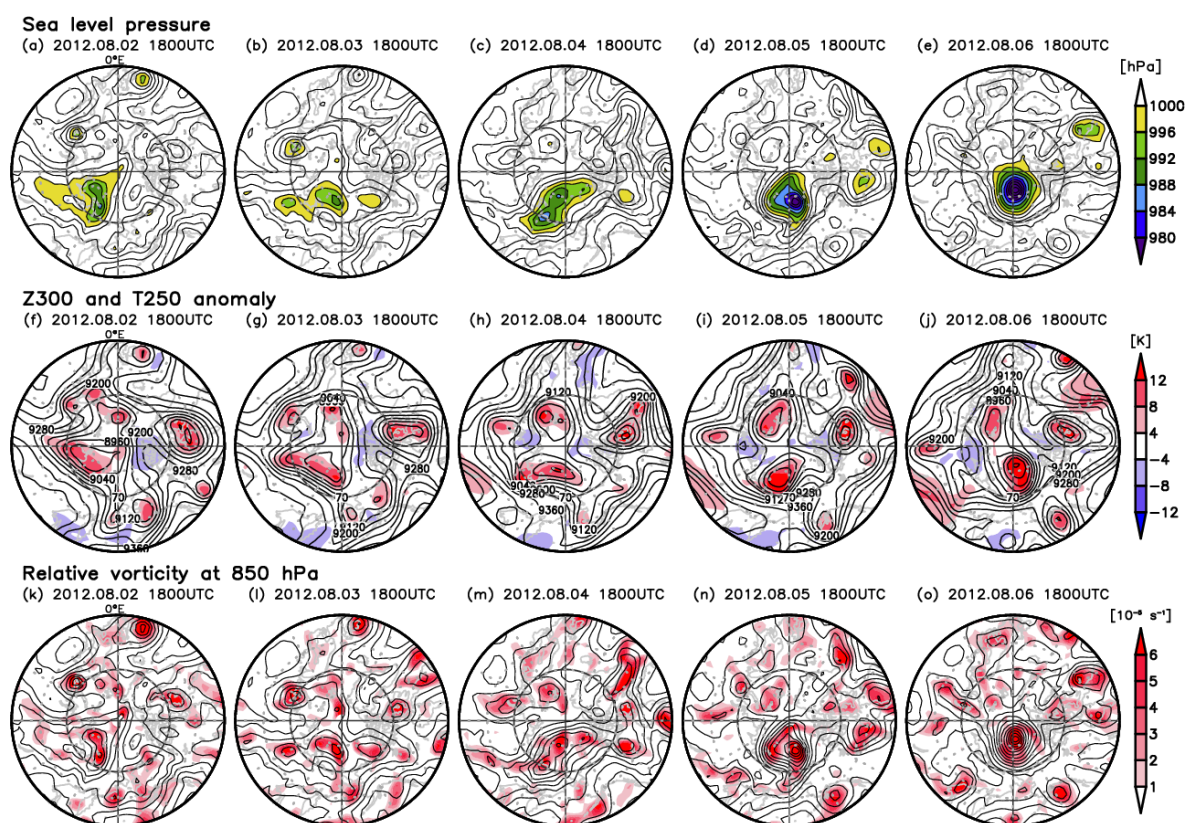


Figure 3.3 Analyzed (a–e) SLP, (f–j) geopotential height at 300 hPa (contours) and temperature anomaly at 250 hPa (shading), and (k–o) SLP (contour) and relative vorticity at 850 hPa (shading) at 1800 UTC on 2 to 6 August 2012. Contour intervals are 4 hPa for SLP and 80m for geopotential height.

Cyclone track forecasts

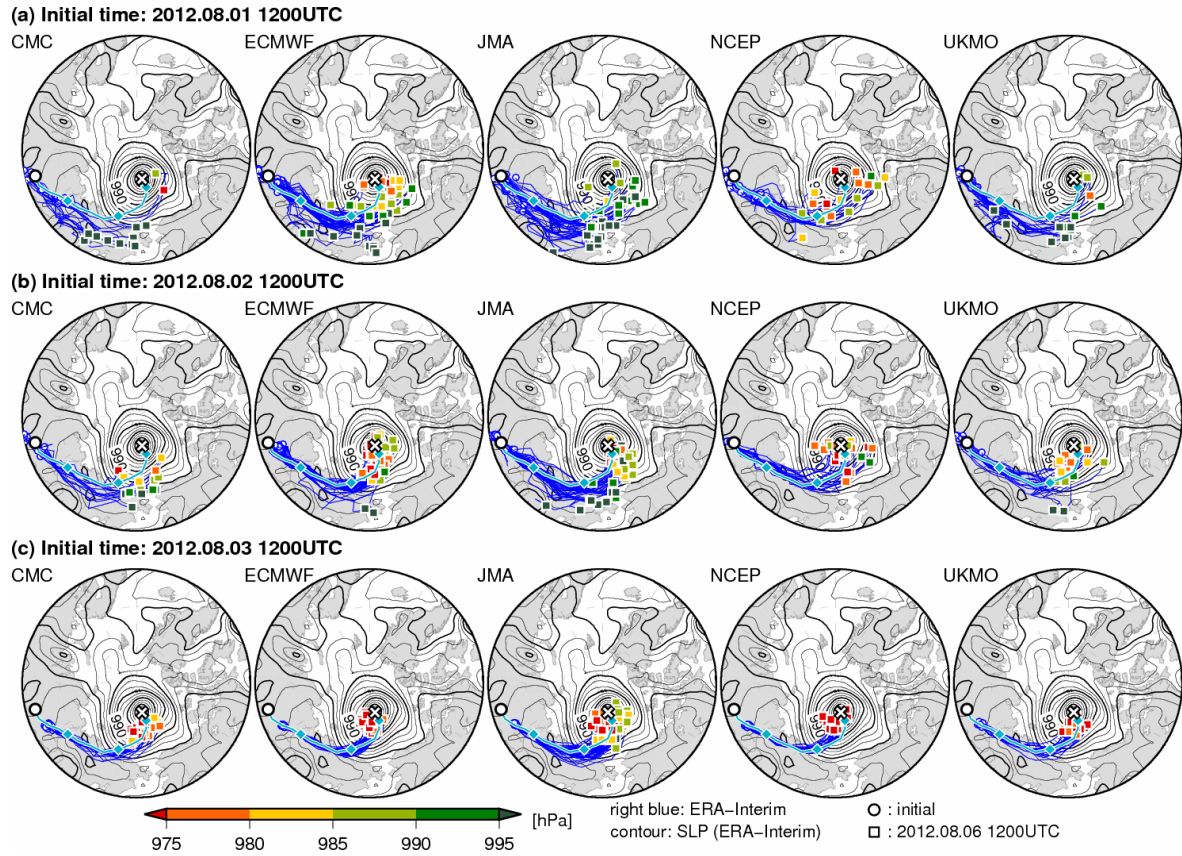


Figure 3.4 Predicted tracks of AC12 in forecasts initialized at 1200 UTC on (a) 1, (b) 2, and (c) 3 August 2012 for CMC, ECMWF, JMA, NCEP, and UKMO (colored lines). The circle represents the positions at 1800 UTC on 2 August 2012. The color of the squares indicates the predicted minimum SLP at 1200 UTC on 6 August 2012. The right blue lines represent the analyzed track and the contours show the analyzed SLP at 1200 UTC on 6 August (ERA-Interim).

Minimum Sea-level Pressure

Valid: 2012.08.06 1200UTC

red line: ERA-Interim (2012.08.06 1200UTC)

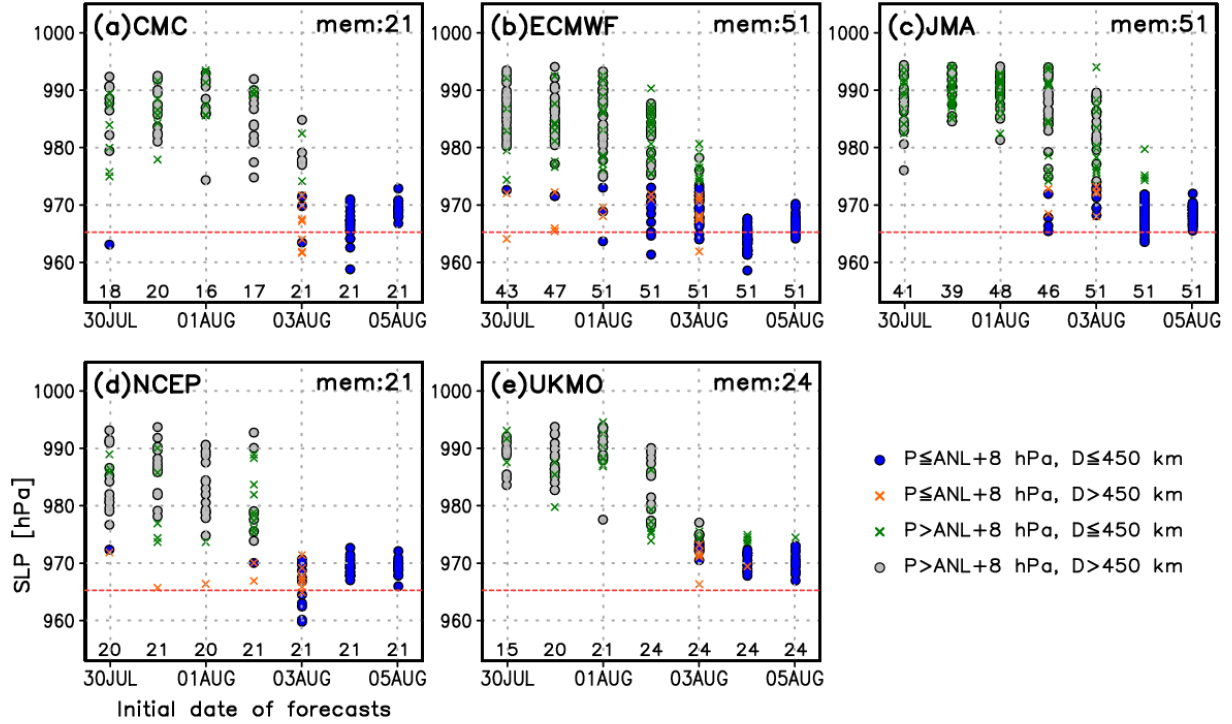


Figure 3.5 Predicted minimum SLP at the center of AC12, valid at 1200 UTC on 6 August 2012 for (a) CMC, (b) ECMWF, (c) JMA, (d) NCEP, and (e) UKMO, for different forecast initialization dates. Each ensemble member is represented by a symbol following the legend (predicted central pressure (P) and distance (D) between the predicted and observed cyclone center at 1200 UTC on 6 August 2012). The red dashed line (ANL) represents the minimum SLP of AC12 in ERA-Interim at 1200 UTC on 6 August 2012. The number of all ensemble members is given at the top-right corner of each frame, the number of ensemble members that predicted the AC existence is given at the bottom of each frame at each initial date.

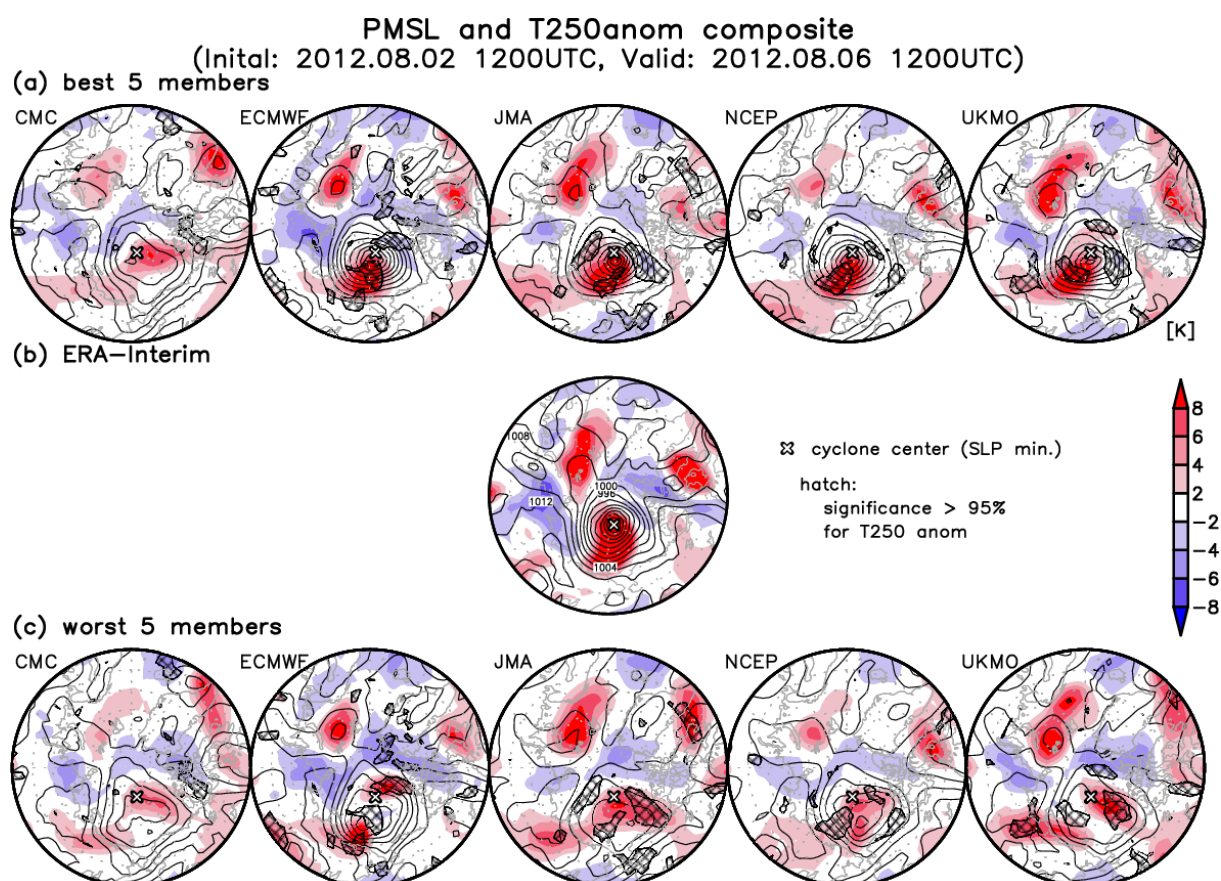


Figure 3.6 Predicted SLP (contour, intervals are 4 hPa) and temperature anomaly at 250 hPa (T250 anomaly, shading) of the five (a) best- and (c) worst-performing ensemble members for each NWP center. Forecasts are initialized at 1200 UTC on 2 August 2012 and valid at 1200 UTC on 6 August. Hatching in (a) and (c) represents areas with a significant (95% level) difference in T250 anomaly between the best and worst members. (b) Observed SLP at 1200 UTC on 6 August 2012 from ERA-Interim. The cross marks the analyzed AC center and the contour interval is 4 hPa.

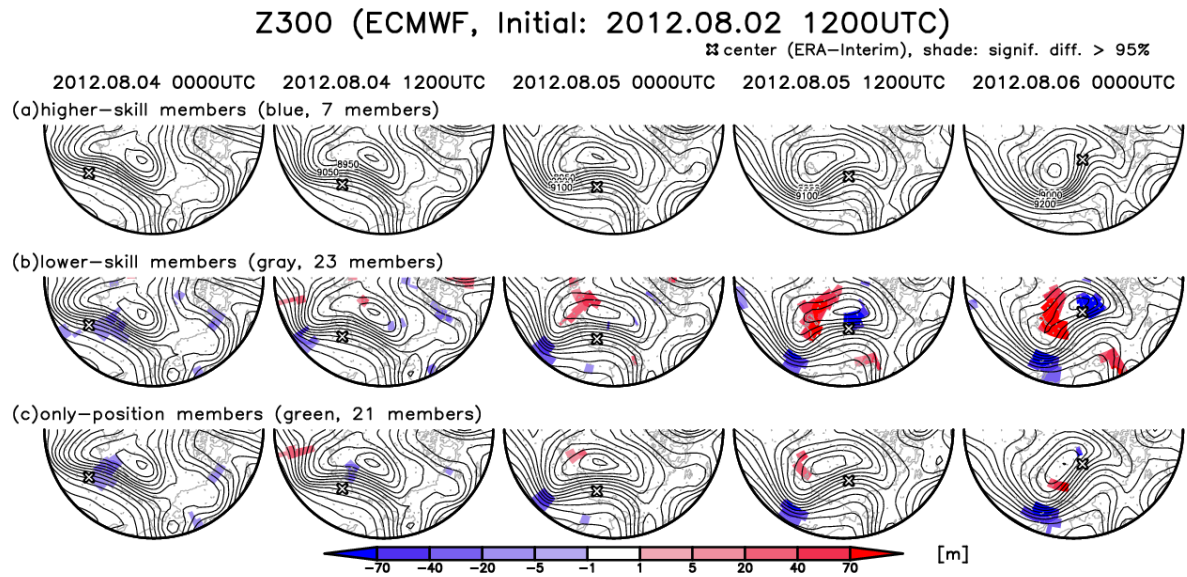


Figure 3.7 Predicted geopotential height at 300 hPa for composites of (a) higher-skill, (b) lower-skill, and (c) only-position ECMWF members initialized at 1200 UTC on 2 August 2012, valid at 0000 UTC on 4 to 6 August 2012 at 12-hourly intervals. The shading in (b) and (c) represents the difference between (a) and (b) (lower- minus higher-skill members), and between (a) and (c) (only-position minus higher-skill members), respectively, where the significance is at the 90% confidence level. The cross represents the AC12 center in ERA-Interim at that time step.

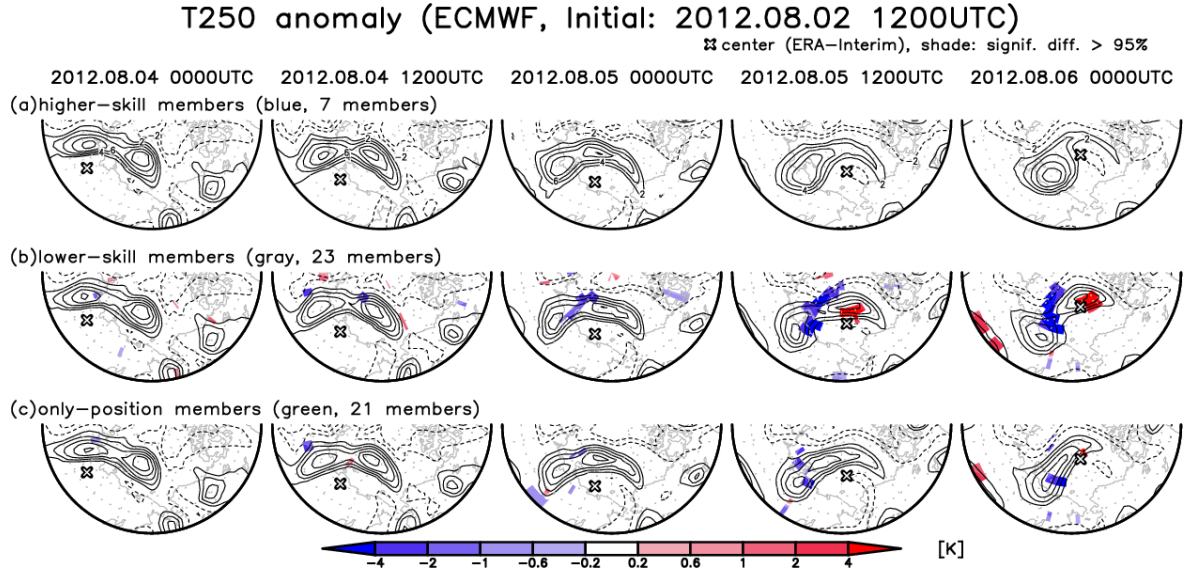


Figure 3.8 Same as Figure 3.7, but for predicted temperature anomaly at 250 hPa.

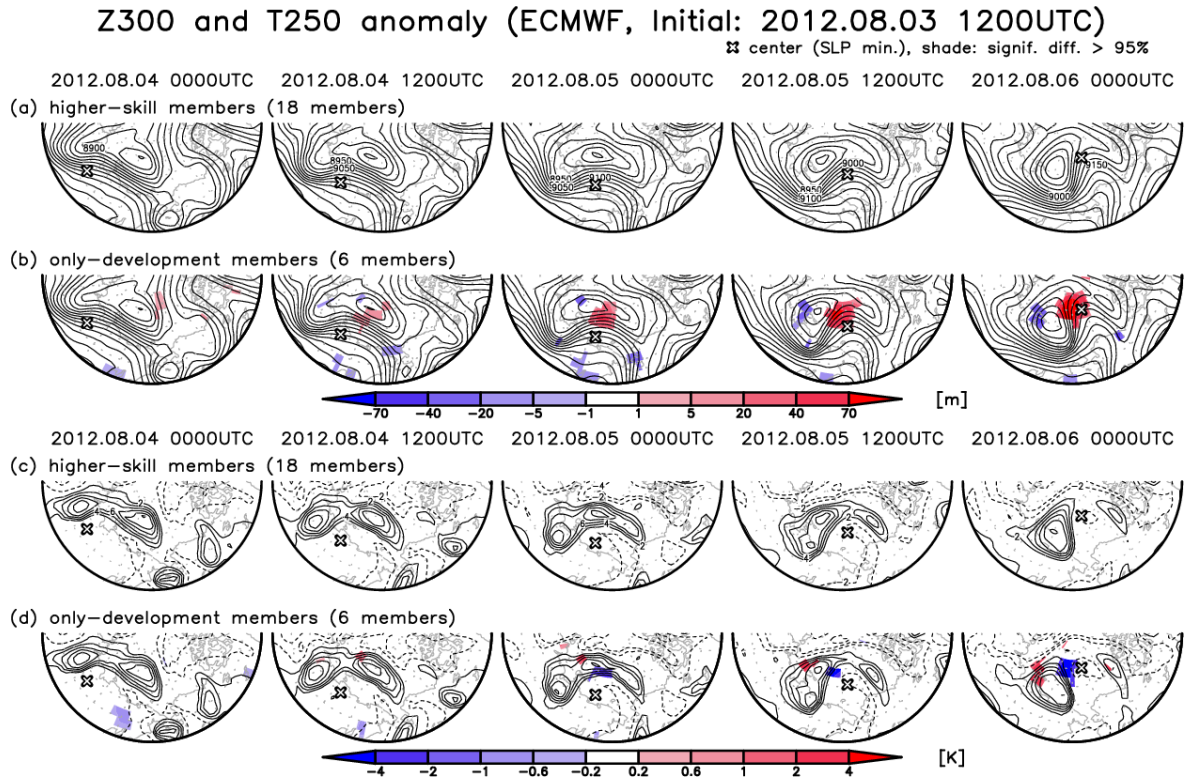


Figure 3.9 Predicted (a, b) geopotential height at 300 hPa and (c, d) temperature anomaly at 250 hPa for composites of (a, c) higher-skill and (b, d) only-development ECMWF members initialized at 1200 UTC on 3 August 2012, valid at 0000 UTC on 4 to 6 August 2012 at 12-hourly intervals. The shading in (b) and (d) represents the difference (only-development minus higher-skill members, where the significance is at the 90% confidence level) between (a) and (b), and between (c) and (d), respectively. The cross represents the AC12 center in ERA-Interim at that time step.

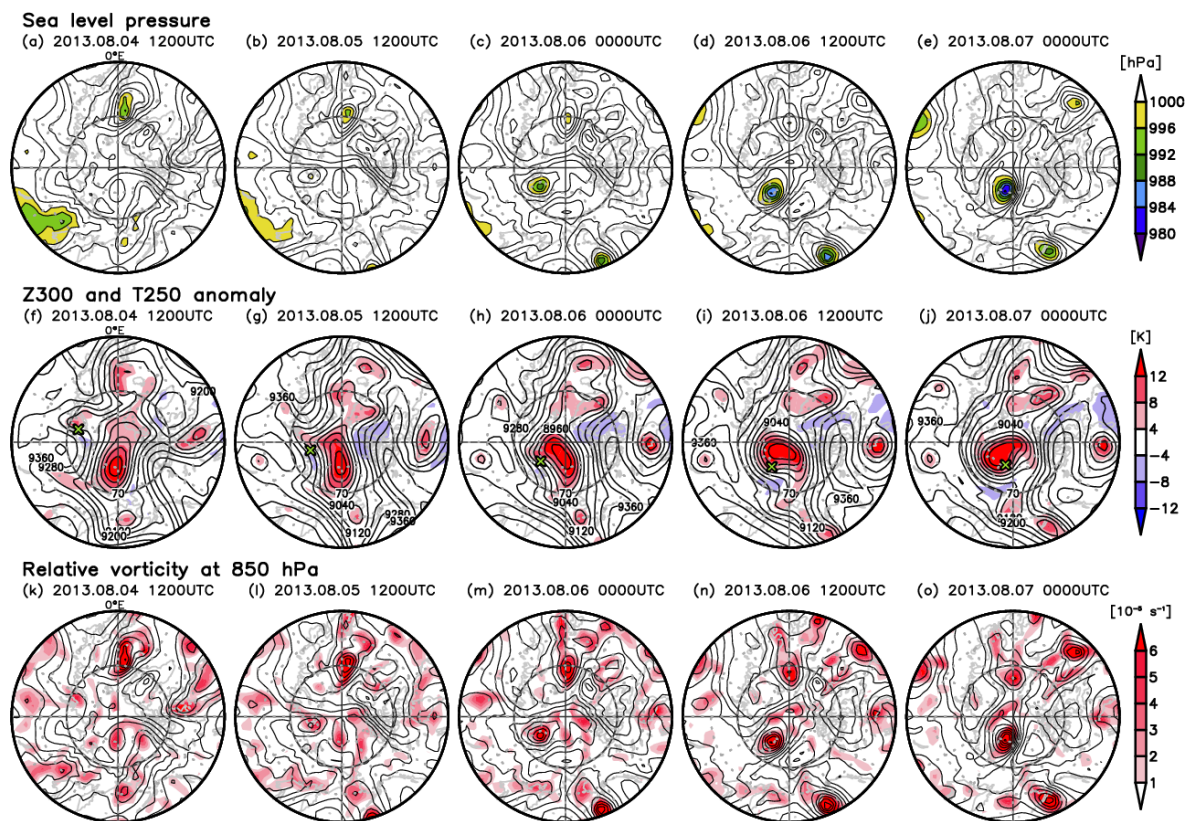


Figure 3.10 Same as Fig. 3.3, but for AC13.

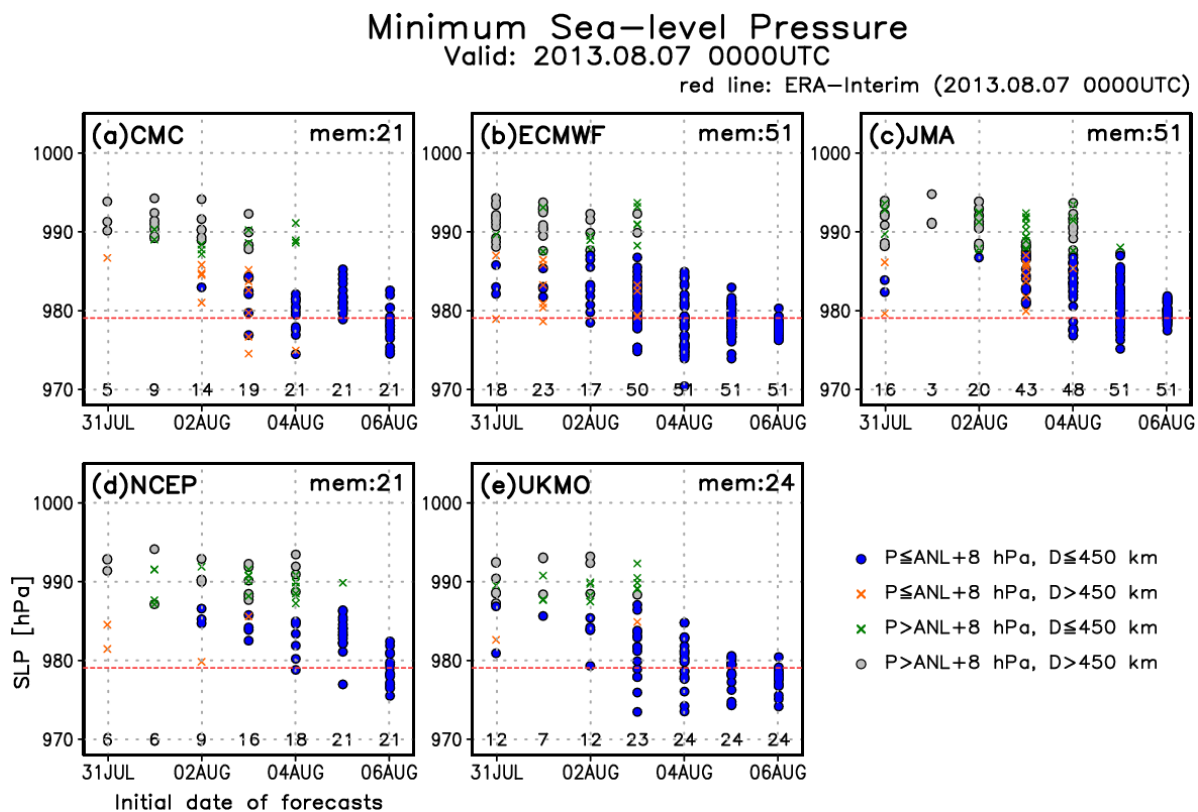


Figure 3.11 Same as Fig. 3.5, but for AC13.

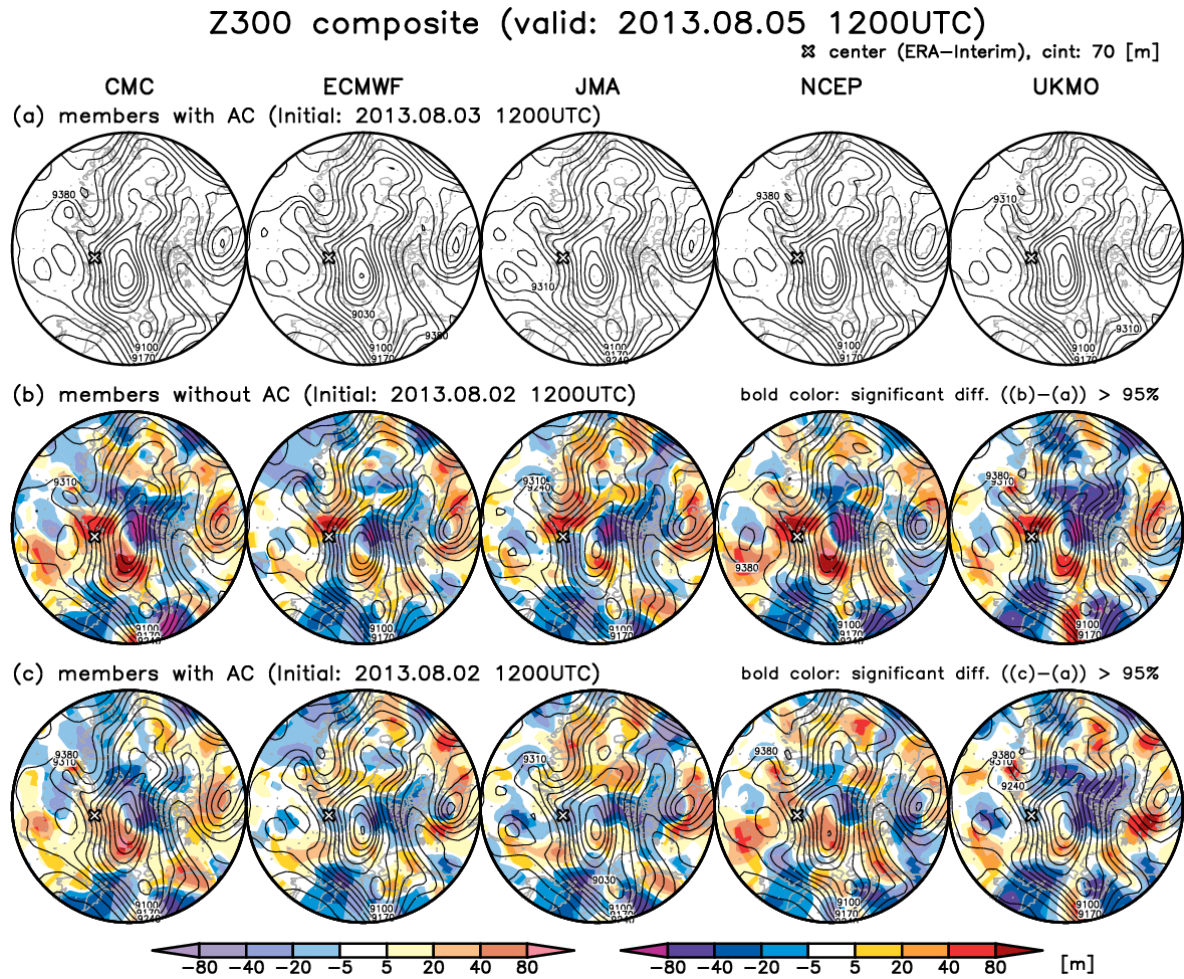


Figure 3.12 Predicted Z300 for composites of members (a, c) with and (b) without AC initialized at 1200 UTC on (b, c) 2 August and (a) 3 August 2012, valid at 0000 UTC on 5 August 2013 for CMC, ECMWF, JMA, NCEP, and UKMO (left to right columns). The shading in (b) represents the difference between (a) and (b), and that in (c) represents the difference between (a) and (c). Bold colors indicate a significant difference at the 95% confidence level. The cross represents the AC13 center in ERA-Interim.

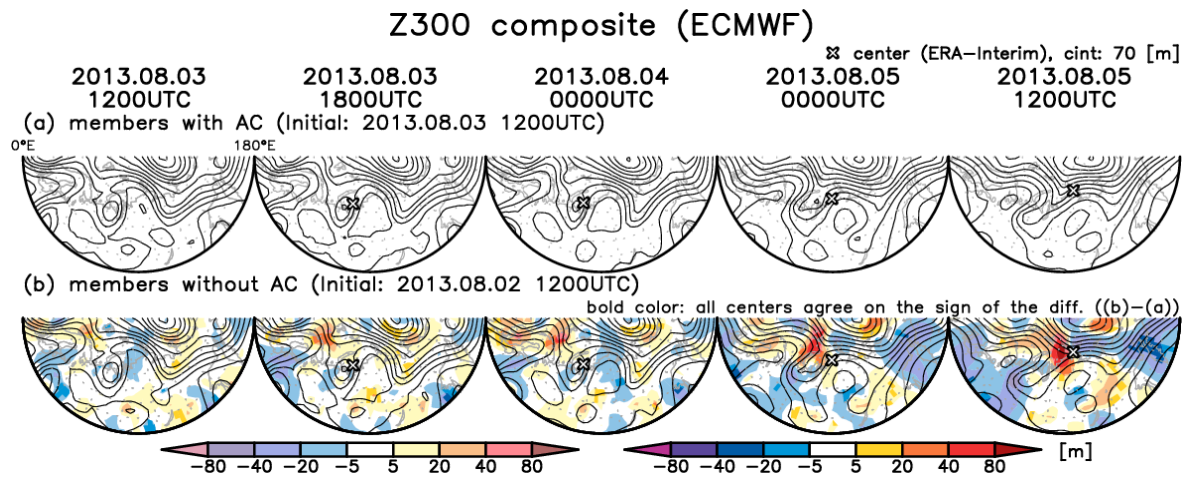


Figure 3.13 Same as Fig. 3.9a and b, but for ECMWF members (a) with AC initialized on 3 August and (b) without AC initialized on 2 August, valid from 1200 UTC on 3 August to 1200 UTC on 5 August. Bold color indicates the region in which all NWP centers agree on the sign of the difference between members with and without AC13.

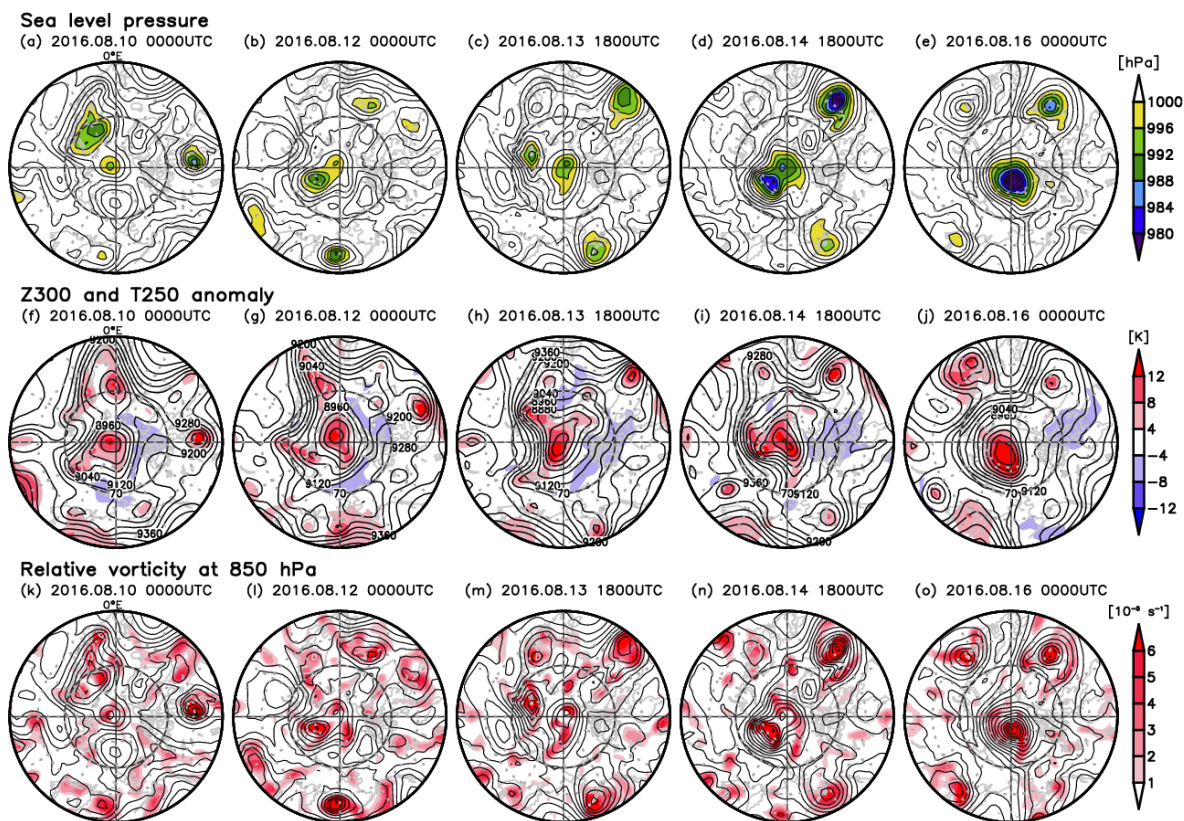


Figure 3.14 Same as Fig. 3.3, but for AC16.

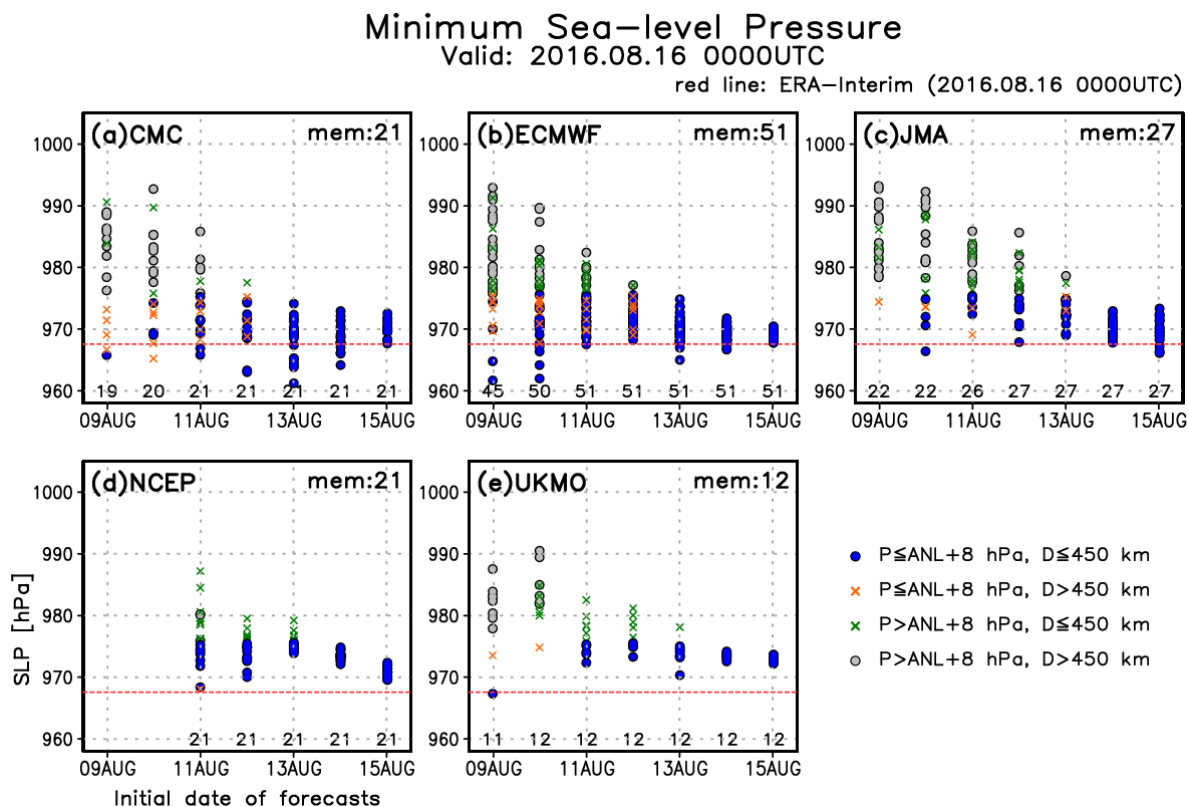


Figure 3.15 Same as Fig. 3.5, but for AC16.

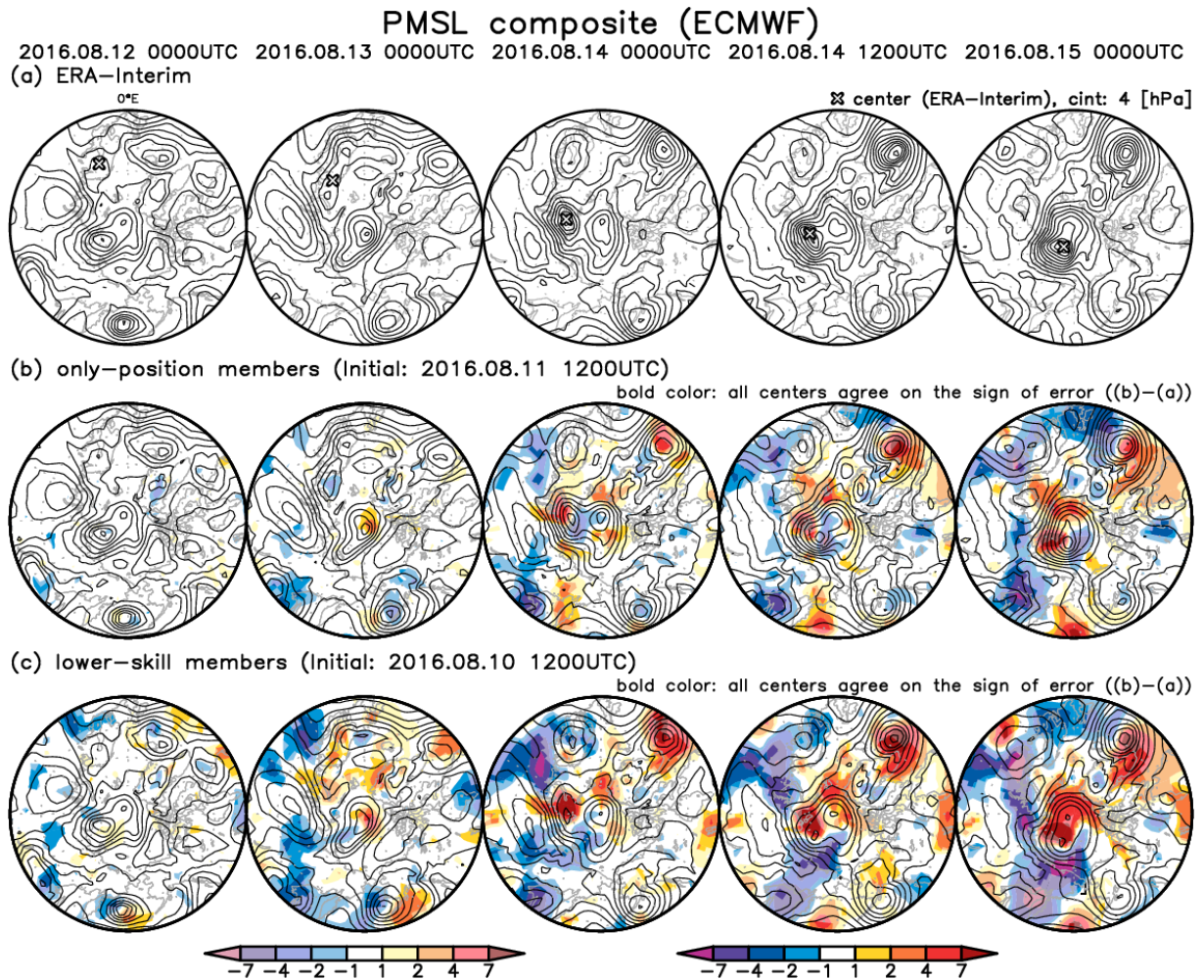


Figure 3.16 (a) Analyzed SLP (contour) from 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns) in ERA-Interim. The cross represents the AC16 center in ERA-Interim. (b, c) Predicted SLP (contour) and its error (shading) for composites of (b) only-position members initialized at 1200 UTC on 11 August and (c) lower-skill ECMWF members initialized at 1200 UTC on 10 August, valid at 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). The shading in (b) represents the error between (a) and (b), and that in (c) represents the error between (a) and (c). Bold color indicates the region in which all NWP centers agree on the sign of the errors.

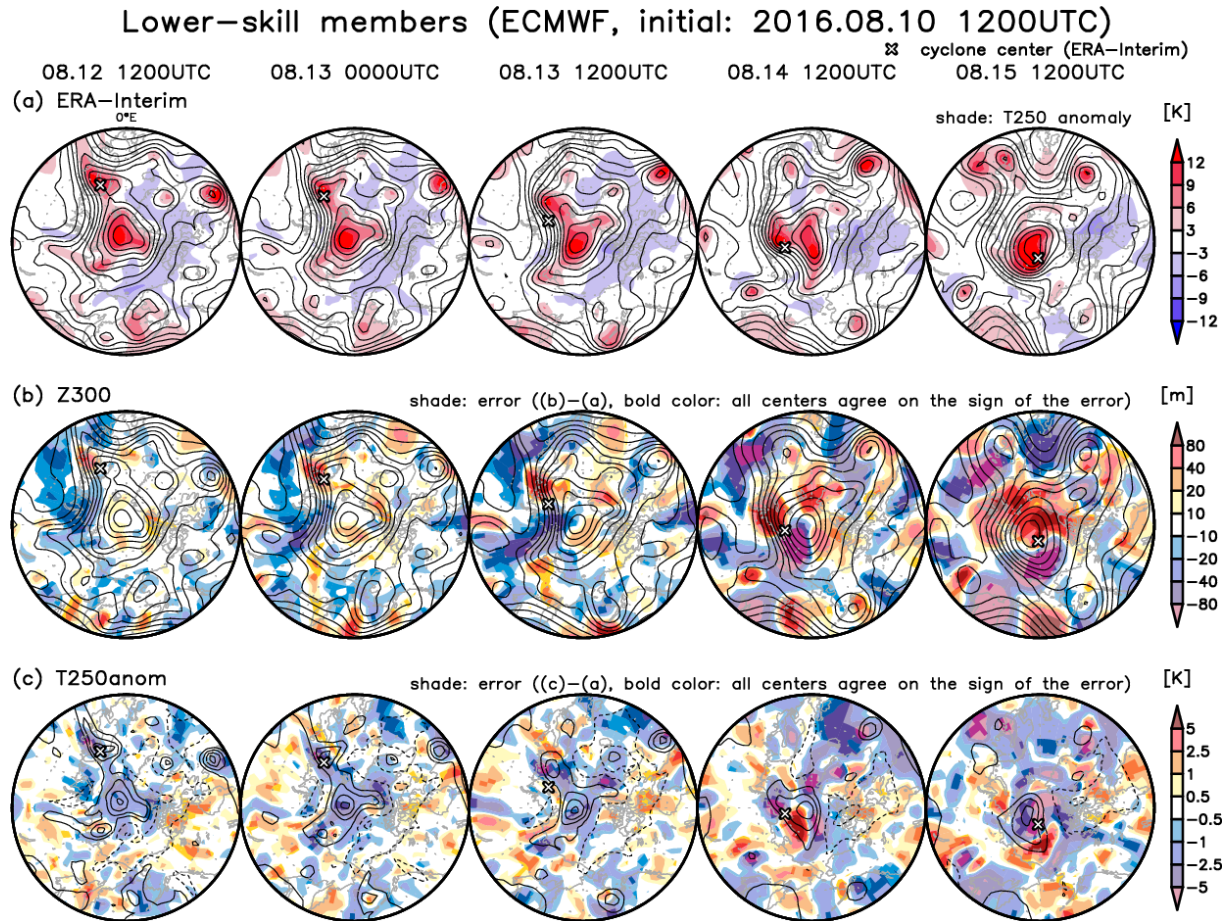


Figure 3.17 (a) Analyzed geopotential height at 300 hPa (contour) and temperature anomaly at 250 hPa (shading) from 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). (b, c) Predicted geopotential height at 300 hPa (b, contour) and temperature anomaly at 250 hPa (c, contour) for composites of lower-skill ECMWF members initialized at 1200 UTC on 10 August 2016, valid at 1200 UTC on 12 August to 1200 UTC on 15 August (left to right columns). The shading in (b) represents the error between (a) and (b), and that in (c) represents the error between (a) and (c). Bold color indicates the region in which all NWP centers agree on the sign of the errors. The cross represents the AC16 center in ERA-Interim.

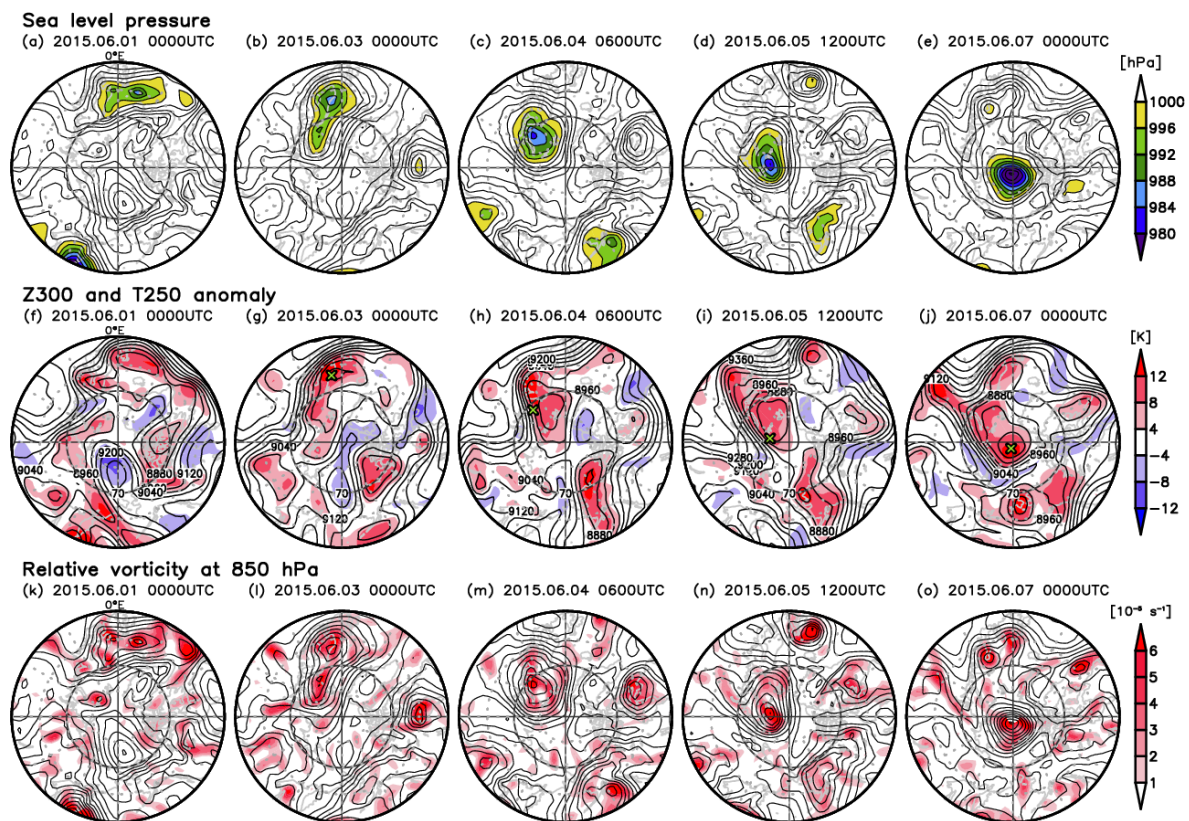


Figure 3.18 Same as Fig. 3.3, but for AC15.

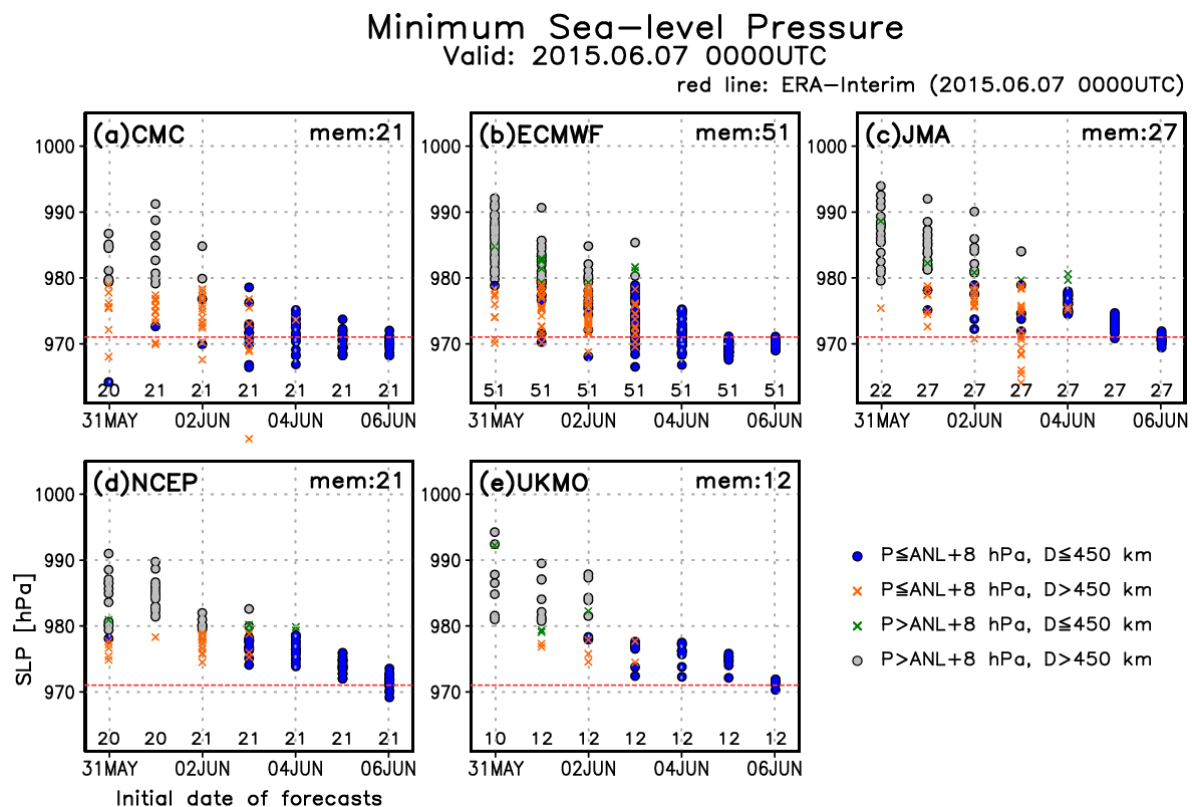
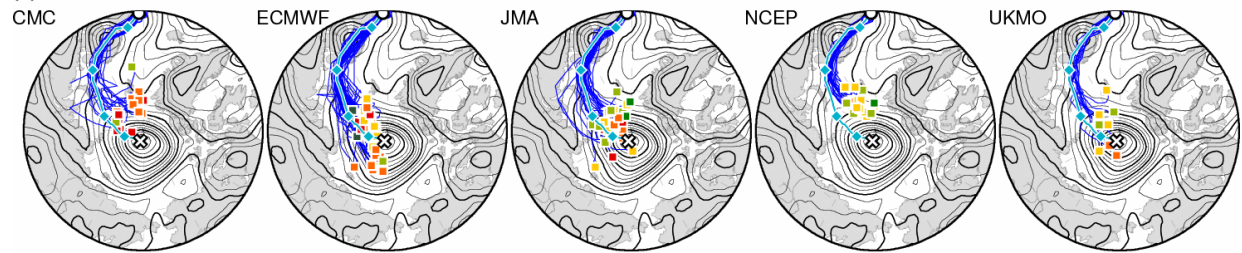


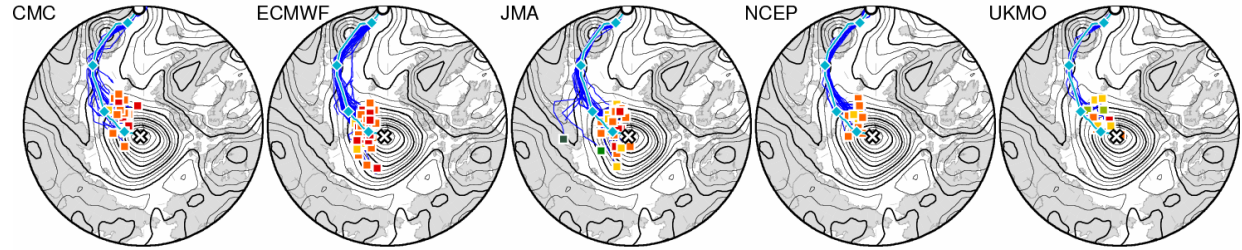
Figure 3.19 Same as Fig. 3.5, but for AC15.

Cyclone track forecasts

(a) Initial time: 2015.06.01 1200UTC



(b) Initial time: 2015.06.02 1200UTC



(c) Initial time: 2015.06.03 1200UTC

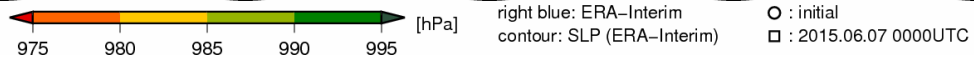
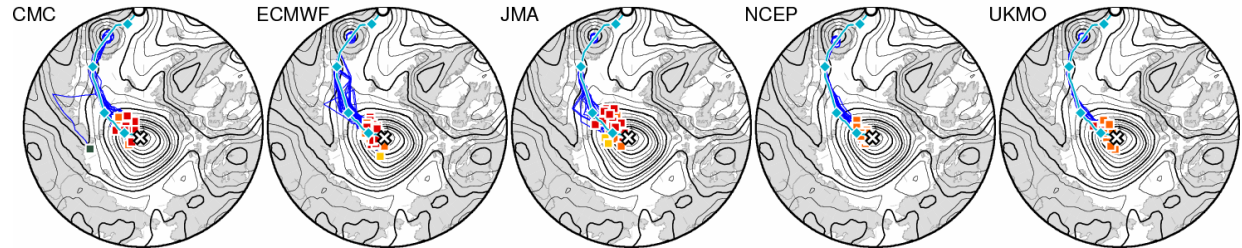


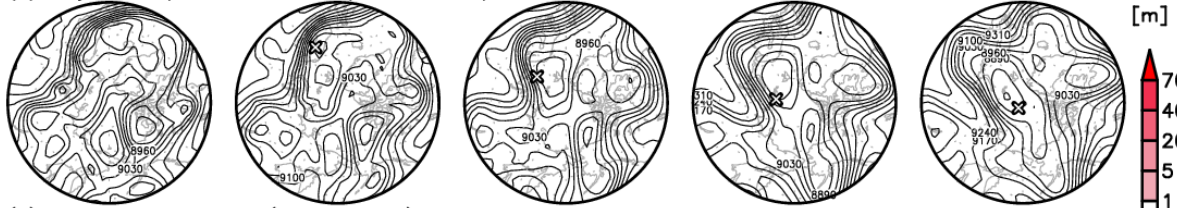
Figure 3.20 Same as Fig. 3.4, but for AC15.

Z300 and T850 anomaly (ECMWF, Initial: 2015.06.01 1200 UTC)

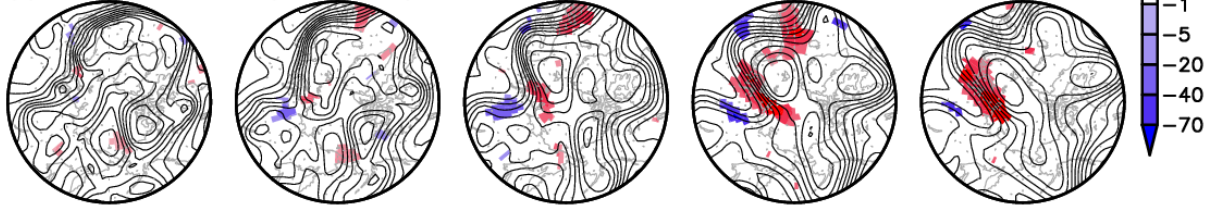
✕ center (SLP min.), shade: signif. diff. > 95%

2015.06.02 1200UTC 2015.06.03 1200UTC 2015.06.04 1200UTC 2015.06.05 1200UTC 2015.06.06 1200UTC

(a) only-development members (12 members)

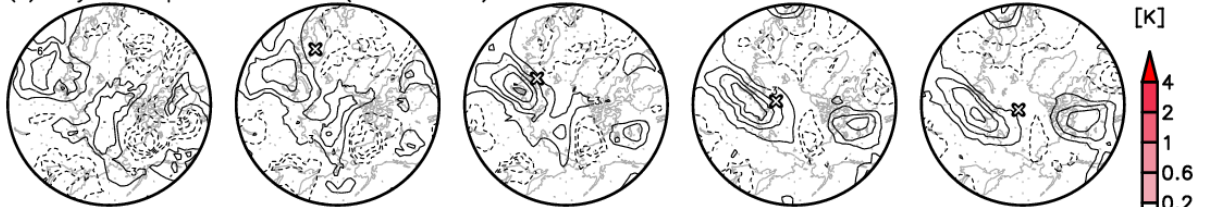


(b) lower-skill members (28 members)



2015.06.02 1200UTC 2015.06.03 1200UTC 2015.06.04 1200UTC 2015.06.05 1200UTC 2015.06.06 1200UTC

(c) only-development members (12 members)



(d) lower-skill members (28 members)

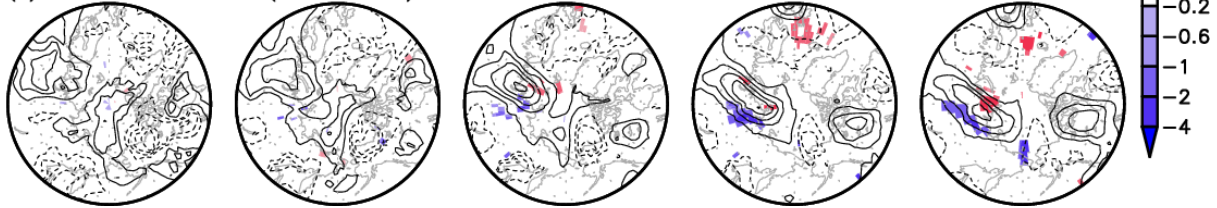


Figure 3.21 Predicted (a, b) geopotential height at 300 hPa and (c, d) temperature anomaly at 250 hPa for composites of (a, c) only-development and (b, d) lower-skill ECMWF members initialized at 1200 UTC on 1 June 2015, valid at 1200 UTC on 2 to 6 June 2015 at 24-hourly intervals. The shading in (b) and (d) represents the difference (lower-skill minus only-development members, where the significance is at the 95% confidence level) between (a) and (b), and between (c) and (d), respectively. The cross represents the AC15 center in ERA-Interim.

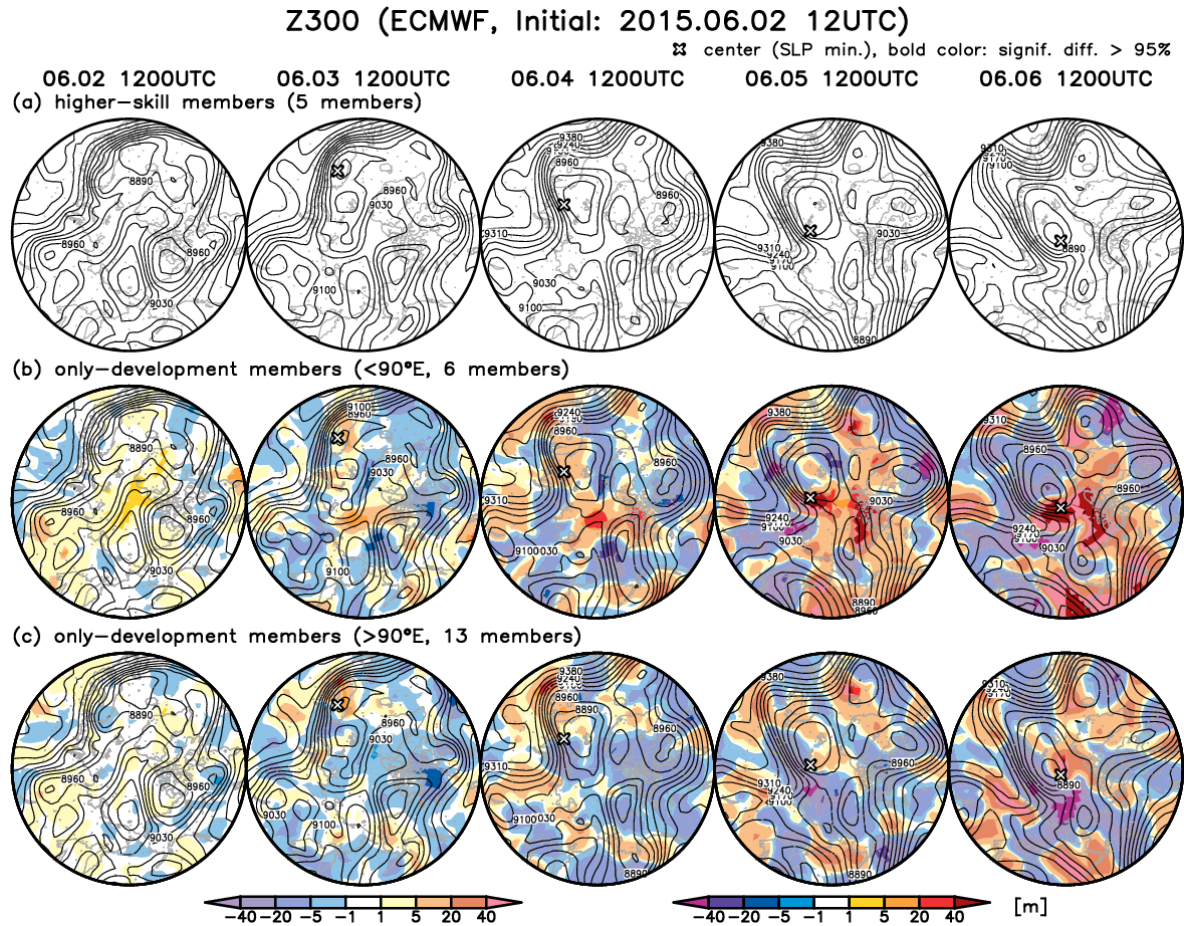


Figure 3.22 Predicted geopotential height at 300 hPa for composites of (a) higher-skill and only-development ECMWF members with the predicted AC position of (b) < 90°E and (c) > 90°E members initialized at 1200 UTC on 2 June 2015, valid at 1200 UTC on 2 to 6 June 2015 at 24-hourly intervals. The shading in (b) and (c) represents the difference between (a) and (b) and between (a) and (c) (only-position minus higher-skill members), respectively. Bold colors indicate a significant difference at the 95% confidence level. The cross represents the AC12 center in ERA-Interim at that time step.

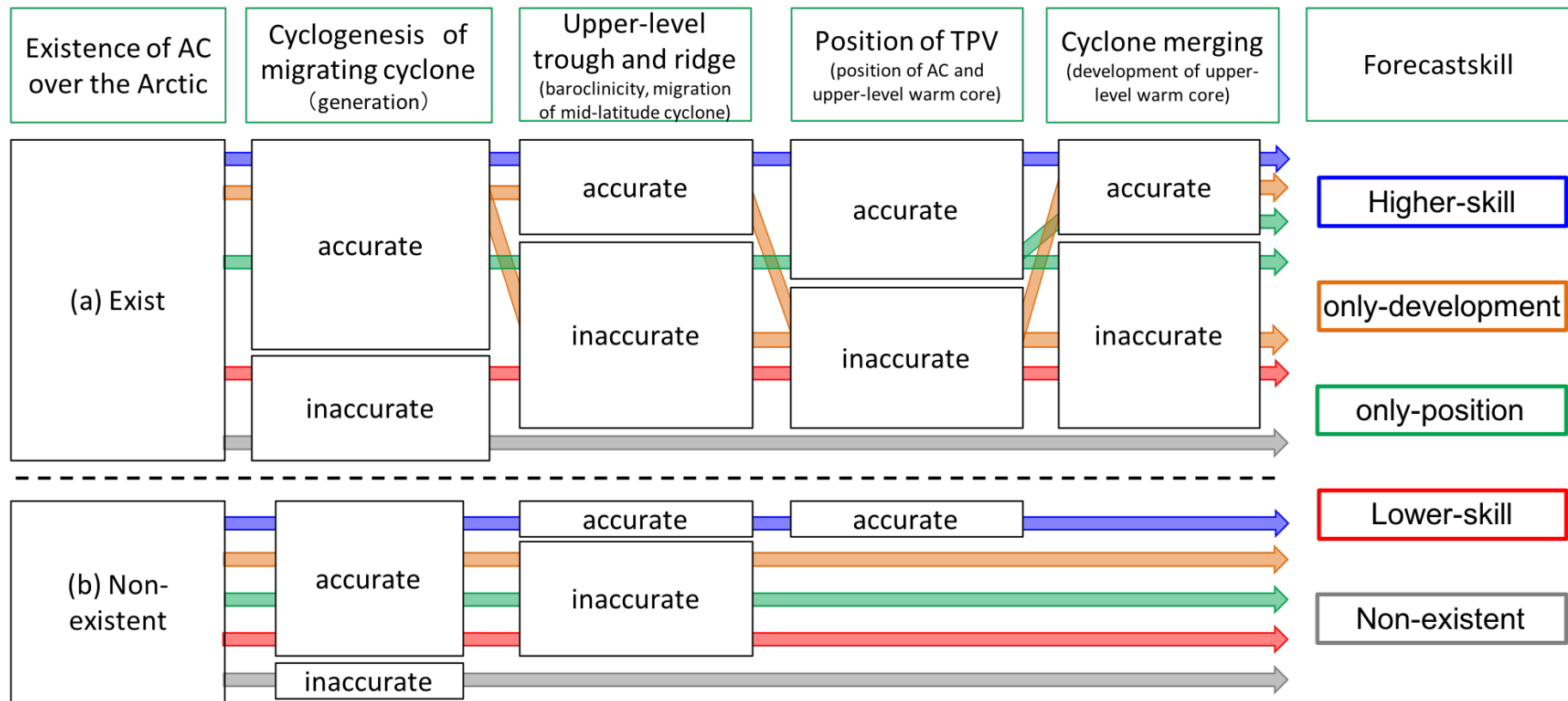


Figure 3.23 Schematic diagram for classification of predicted ACs. Top panels indicate the key factors for ACs prediction. “Accurate” (“inaccurate”) indicate a forecast accurately (inaccurately) predicts the key factor described at top panels. The colored arrows indicate the higher-skill (blue), only-development (orange), only-position (green), lower-skill (red) members’ forecasts. The gray arrows indicate the members without ACs.

3.3 Average predictability of Arctic cyclones

3.3.1 Forecast skill for summer ACs during 2008–2016 by TIGGE (Yamagami et al. 2018b)

The existence probability of ACs is mainly related to the forecast performance for the generation (but also maintenance) of the ACs (Figure 3.24a). For 10 extraordinary ACs, ECMWF exhibited the highest probability at lead times of 2.5–6.5 days, with probabilities of ~ 0.77 and ~ 1.0 on average, at lead times of 6.5 and 2.5 days, respectively. Both CMC and NCEP were the second-best performing centers at lead times of 4.5–6.5 days. The probability differences between ECMWF and these two centers were ~ 0.10 and ~ 0.05 at lead times of 6.5 and 4.5 days, respectively. Overall, JMA and UKMO showed the lowest probability at equivalent lead times, and their differences in existence probability from ECMWF were ~ 0.18 and ~ 0.08 at lead times of 6.5 and 4.5 days, respectively. Thus, ECMWF showed 1- to 1.5-day advantage in predicting AC generation and maintenance at these lead times, compared with the other centers, this is, the probability for ECMWF at a lead time of 6.5 days was similar to the probability for the other centers at lead times of 5.0–5.5 days. At lead times of 2.5–3.5 days, JMA and UKMO displayed vast improvement whereby probabilities were > 0.95 , similar to ECMWF and CMC. Although NCEP showed the lowest existence probabilities of 0.93–0.98 at these lead times, all centers had a probability of 1.0 at lead times of ≤ 1.5 days.

The existence probability verified for each AC were shown in Fig. 3.25. For example, the existence probability for the AC in June 2015 (Figure 3.25h) was ~ 1.0 at almost all lead times for all NWP centers. The probability was also higher for the ACs in July 2013 (except at a lead time of 7 days, Figure 3.25f) and August 2016 (Figure 3.25j) than for all remaining ACs. Conversely, the existence probability of ~ 0.3 for the AC in August 2013 (Figure 3.10 and Figure 3.25g) at lead times of 4.5–6.5 days was the lowest out of the 10 ACs, but significantly improved at a lead time of 3.5 days. As seen in the Section 3.1.2, the generation of the AC over the Eurasian continent

was difficult to predict at lead times ≥ 4.5 days, due to the failure of the upper-level trough prediction. For these ten events, the long-lived ACs are likely to show the high existence probability, especially in its developing stage.

Regarding the central pressure forecast, on the whole, ECMWF exhibited the highest forecast skill in predicting central pressures at all lead times (Figure 3.24b). The central pressure errors were 13.3 hPa and 1.7 hPa at lead times of 6.5 and 0.5 days, respectively. Overall, ECMWF had 1- to 1.5-day advantage compared with the other centers, as well as the existence probability. Further, ECMWF had the higher skill in predicting central pressures compared with the other NWP centers for most AC events (Figure 3.26). At lead times of 4.5–6.5 days (Figure 3.24b), CMC and NCEP were the second-best performing centers. The differences in central pressure errors between ECMWF and these two centers were 1.0 and 2.8 hPa at lead times of 6.5 and 4.5 days, respectively. For these lead times, JMA and UKMO exhibited the lowest performance, and their differences from ECMWF were 2.5 and 4.3 hPa at lead times of 6.5 and 4.5 days, respectively. CMC was also the second-best performing center at lead times of 1.5–3.5 days, whereas NCEP had the lowest forecast skill at these shorter lead times. The performance of JMA improved considerably at lead times ≤ 3.5 days and its central pressure errors were similar to those for CMC at lead times of 3.5 and 1.5 days. Furthermore, JMA had the second lowest central pressure error of 2.6 hPa at a lead time of 0.5 days. For individual ACs (Figure 3.26), the forecast for the AC in August 2012 showed particularly low skill compared with the other ACs (Figure 3.26e). As seen in the Section 3.1.1, the development of an upper-level warm core resulting from a merging cyclone was difficult to predict in this AC event. In contrast, the central pressure errors for the ACs in July 2013 (Figure 3.26f) and 2016 (Figure 3.26i) were apparently lower than those for the other ACs.

For predictions of the central position (Figure 3.24c), the average errors in the 6.5-day forecasts were 650–800 km for all NWP centers. These average position errors were

approximately equal to the radius of the ACs (Table 3.1), suggesting that all NWP centers failed to predict the positions of the ACs at this lead time. At lead times of 1.5–5.5 days, ECMWF showed the highest performance in predicting the central position and a 1-day advantage compared with the other centers, as well as the existence probability and forecast skill of central pressure. The difference in the average errors between ECMWF and the other centers was ~100 km. At lead times of 3.5–5.5 days, the average position error was larger for CMC than for the other centers. However, the CMC forecast improved significantly at a lead time of 2.5 days and the error was similar to that for JMA, NCEP, and UKMO at lead times of 1.5–2.5 days. At a lead time of 0.5 days, all NWP centers exhibited average errors <100 km. The central position errors for individual ACs (Figure 3.27) showed that the AC in June 2015 (Figure 3.27h) had the largest error of all the ACs, especially in the longer forecast lead time range. As seen in the Section 3.1.4, the northward movement of predicted AC was short compared with that of the analyzed AC in this event. Conversely, the errors for the ACs in July 2013 (Figure 3.27f) and 2016 (Figure 3.27i) were lower than those for the remaining ACs. In these ACs, the position of an upper-level polar vortex was predicted well compared with the other ACs. Unlike the central pressure predictions, the best performing center was dependent on the AC event, particularly at lead times of 4.5–6.5 days.

The average central position error for ECMWF at a lead time of 4.5 days reduced to less than half of the mean observed radius (469.1 km) for the 10 ACs, with the average pressure error of 8.8 hPa. In addition, JMA, NCEP, and UKMO had average position errors <496.1 km at a lead time of 3.5 days and their average central pressure errors were 9.2, 10.7, and 10.4 hPa, respectively. Predictions of CMC had an average central position error <496.1 km at a lead time of 2.5 days, with average central pressure error of 5.5 hPa.

3.3.2 Forecast skill for summer ACs during 1986–2016 by GEFS reforecast

Twenty-six extraordinary ACs were detected in summer during 1986–2016 using ERA-Interim. At the mature stages for these 26 ACs in the GEFS control analysis (Figure 3.28; Table 3.2), the ACs on 7 August 1995 (event I) and 2012 (event U) indicate the lowest central pressure (964.7 hPa), followed by ACs on 7 August 1991 (event F; 967.8 hPa) and 16 August 2016 (event Z; 968.3 hPa). Overall, these AC observations agree with the results of Simmonds and Rudeva (2012), although their results suggest that the AC on 7 August 2012 had a lower central pressure than that on 7 August 1995. This inconsistency may be attributed to differences in the datasets and cyclone detection methods. The AC on 27 June 1988 (event C) indicates the largest size (radius of 1512.8 km), followed by the ACs on 25 July 2013 (event V; 1356.5 km) and 7 June 2015 (event X; 1161.9 km). The average radius of the 26 ACs is 866.3 km. Excluding the ACs on 27 June 1988 and 25 July 2013, the radii of the ACs are inversely correlated with their central pressures over the period 1986–2016 (correlation coefficient of -0.60). This relationship is consistent with the result of ERA-Interim for 10 ACs during 2008–2016. The ACs that occurred during the most recent 10 years (events Q–Z) reached their mature stage over the Pacific side of the Arctic Ocean, except for the AC in August 2010 (event S). However, five ACs (events E, I, J, O, and P) reached their mature stage over the Barents or Kara Sea during 1986–2006, as for the AC in August 2010. No long-term trend is found in their strength, size, or position (Figure 3.29a). The year-to-year variation in the number of extraordinary ACs was larger after 2002 than before 1998 (Figure 3.29b). Namely, the Arctic environment in some summers becomes a more favorable condition for the development of ACs, but that in other summers becomes a more adverse condition.

Forecast performances of the GEFS reforecast for cyclogenesis and maintenance for the 26 ACs are shown as an average existence probability in Figure 3.30a. The average existence probabilities of the GEFS reforecast (hereafter referred to as “GEFS”) for the 26 ACs during

1986–2016 (black) increase with decreasing lead times for 4.0–7.0 days. The probabilities of GEFS for the 26 ACs reached more than 0.9 up to a lead time of 3.0 days. The existence probabilities of GEFS for the 10 ACs during 2008–2016 (brown), which is equivalent verification to the operational EPSs' forecasts, were similar to those for the 26 ACs after lead times of 4.0 days, except at the lead time of 7.0 days. This similarity indicates no improvement in predictions of AC existence with lead times of 4.0–6.0 days. The probabilities become more than 0.9 with lead times of 1.0–3.0 days for both the 10 and 26 ACs. The existence probability for the 10 ACs indicates that GEFS has similar skill to JMA (red) and UKMO (purple) in predicting AC existence.

The existence probabilities for 16 out of 26 ACs are more than 50% with a lead time of 7 days (Figure 3.30a), which is consistent with the average existence probability (i.e., GEFS failed to predict the existence of ~40% of the extraordinary ACs one week in advance). In particular, ACs on 4 July 2002 (event L) and 25 July 2013 (event V) are predicted by all ensemble members with this lead time. The maintenance (generation) of the AC on 4 July 2002 (25 July 2013) is well predicted. Up to a lead time of 4.0 days, the existence of all 26 ACs was predicted by more than half of the members. The existence probabilities with lead times less than 4.0 days are 100% for all AC events, except for those on 9 July 1986 (event D), 25 July 2002 (event M), and 18 August 2006 (event P). Particularly, the existence probabilities of ACs on 9 July 1986 and 25 July 2002 were less than 80% even with lead times of 2.0–3.0 days. The cyclogenesis for these two ACs is well predicted by almost all members, but their maintenances are difficult to predict. These two events contributed to the lower average existence probability for the 26 ACs compared with the 10 ACs with lead times of 2–3 days (Figure 3.30a). With a lead time of 1.0 day, the existence probability is 100% for all 26 ACs. The existence probabilities for individual ACs also indicate no interannual improvement in predicting AC existence throughout entire lead times.

The average central pressure errors of GEFS for the 26 ACs (black) are more than 10 hPa with lead times of 5.0–7.0 days (Figure 3.30b). The average central pressure error is almost similar to

those for the 10 ACs (brown) with these lead times. The central pressure errors for GEFS becomes less than 9.0 hPa at lead times of 4.0 days for both the 26 and 10 ACs. GEFS shows the second largest central pressure for the 10 AC with lead times of 1.0–2.0 days. The differences in the average central pressure errors between for the 26 and 10 ACs are small throughout entire lead times, as with existence probabilities. Thus, the forecast skill for central pressure of AC suggests no improvement with all lead times. Although the average central pressure errors of GEFS for the 10 ACs are equivalent to those of CMC (yellow) and ECMWF (blue) with lead times of 5.0–7.0 days, the existence probabilities of GEFS are lower than those for these two NWP centers (Figure 3.30a). This result indicates that GEFS has difficulty in predicting the cyclogenesis. However, after the cyclogenesis was correctly predicted, GEFS predicts the development of AC as well as CMC and ECMWF. The central pressure errors for NCEP (green) are larger than those for GEFS for the 10 ACs throughout entire lead times. The average central pressure error for GEFS for 9 ACs without the AC on 6 July 2016 (event Y in Figure 3.31b) are almost the same that for the 10 ACs (the AC is not included in the calculation of average errors for NCEP due to data gaps). One of the causes might be the difference in control analyses, used for the verification, between GEFS and NCEP. The largest difference in analyzed central pressures between NCEP and GEFS is ~8.8 hPa for AC in June 2008, followed by ~7.4 hPa for AC in August 2010 (not shown). These results suggest that the differences in central pressure errors between GEFS and NCEP are attributed not to the difference in forecast models but to the difference in control analysis, especially in 2008–2010.

From the central pressure errors for individual ACs (Figure 3.31b), forecast skill of central pressure significantly depends on AC events in 1.0- to 7.0-day forecasts. For a lead time of 1.0 day (blue symbols in Figure 3.31b), the largest central pressure error (7.94 hPa) is for the AC on 7 August 1991 (event F). The central pressure error for the AC on 7 August 1995 (event I), in addition to that for the AC on 7 August 1991, is large compared with those for the remaining ACs

for a lead time of 3.0 days (green symbols in Figure 3.31b). The AC on 7 August 2012 (event U) also has a larger central pressure error than the other ACs for lead times of 5.0 and 7.0 days (yellow and red symbols in Figure 3.31b). The observed central pressures of these three ACs are the first lowest and third lowest of all ACs (Table 3.2). These results suggest that the central pressure errors correlate with the observed central pressure, especially after a lead time of 5.0 days. The correlation coefficient between the observed central pressure and the central pressure error for lead times of 7.0 days is -0.75 , revealing a strong inverse relationship. As with tropical and extratropical cyclones, the development of ACs, particularly stronger ACs, is difficult to predict.

Average central position errors of GEFS for the 26 ACs (black) monotonically decrease with decreasing lead times (Figure 3.30c). The average central position errors are 812.0 km and 108.9 km for lead times of 7.0 and 1.0 days, respectively. The average position errors for the 26 ACs decrease to less than half of the mean observed radius for the 26 ACs (433.1 km) at lead time of 3.0 days. Average central position errors of GEFS for the 10 ACs (brown) are similar to those for the 26 ACs for all lead times, except for lead times of 7.0 and 4.0 days. GEFS for the 10 ACs shows smaller position error than NCEP (green) with lead times of 5.0–7.0 days. Given that the difference in central position between the GEFS control analyses and ERA-Interim is small for the 10 ACs (the average difference in analyzed position for the 10 ACs is 39.3 km), the differences in control analyses used for the verification presumably have little impact on these differences in central pressure errors. The difference in existence probabilities (i.e., weights used for the calculation of average errors) between NCEP and GEFS can lead to the difference in the average central position errors. For example, the central position error of 7.0-day GEFS forecast is weighted with an existence probability of 0.45 for the AC on 17 June 2012 (event T in Figure 3.31c), whereas the central position error of 6.5-day NCEP forecast is weighted with an existence probability of 0.76 for the same AC (Figure 3.27d). Namely, ACs which have larger central

position error are excluded in the calculation of average errors of GEFS because their existences are not predicted. With the lead time of 4.0 days, the average position error of GEFS for the 10 ACs is similar to that in CMC (yellow). The position error for the AC on 6 July 2016 (event Y in Figure 3.31c) might be a main contributor to this larger position error in the 4.0-day forecast. Excluding AC on 6 July 2016 reduced the average position error about 60 km in the 4.0-day forecast in GEFS. Overall, GEFS has an intermediate forecast skill in predicting location of ACs among the all NWP centers.

The central position errors for individual ACs with a lead time of 1.0 day (blue symbols in Figure 3.31c) are less than a quarter of the mean radius for the 26 ACs (216.6 km), indicating that the AC positions are correctly predicted. The AC on 6 July 2016 (event Y) has the largest central position error (205.7 km) with this lead time. With a lead time of 3.0 days (green symbols in Figure 3.31c), the central position errors for all the ACs are less than half the mean radius of the 26 ACs (433.1 km), except for the four ACs (event A, C, I, and T). The locations of these four ACs were not predicted correctly even in a 3.0-day forecast. The prediction for the AC on 25 July 2002 (event M) shows the largest position error (1175.9 km) with a lead time of 5.0 days (yellow symbols in Figure 3.31c), which are larger than the average radius of the 26 ACs. However, the central position error for the 8 ACs (events A, F, H, K, L, Q, R, and U) is smaller than half the mean radius of the 26 ACs at this lead time. Besides, the central position errors of three ACs (event V, T, and Q) are smaller than half of the average radius of the 26 ACs even in a 7.0-day forecast (red symbols in Figure 3.31c). These results indicate that the forecast skill of AC location significantly depends on AC events, as with that of AC development (Figure 3.31b).

Figure 3.32 shows an example of the strike probability map for the AC on 16 August 2016 (AC16) during its mature stage. In the 1.0-day forecast, the 75–100% probability area is located over the Pacific side of the Arctic Ocean (Figure 3.32a). The large area of the probability of 75–100% indicates a small ensemble spread of AC location. The high-probability area agrees well

with the observed AC position for this lead time. The area of high strike probability (75–100%) is smaller in the 3.0-day forecast than in the 1.0-day forecast (Figure 3.32b). In addition, the areas of lower probability (less than 75%) are broad in the north–south direction. The areas of 75%–100% probability agree with the half of observed areas. The remaining half of observed areas overlaps with the areas of 25–75% probability. Two-thirds of the 50–75% probability area overlap with the observed strike area, indicating that the strike probability in the 3.0-day forecast roughly agrees with the occurrence frequency of AC16. In the 5.0-day forecast, the area more than 10% probability shifts southwestward (Figure 3.32c). Fewer grid points are included in the 75–100% probability area, and these grid points do not overlap with the observed strike area. The area more than 10% strike probability does not correspond to the eastern part of the observed location. In the 7.0-day forecast (Figure 3.32d), the 75%–100% and 50%–75% probability areas are not present, indicating a large ensemble spread for AC location. Low existence probabilities (compared with shorter lead time forecasts) also contribute to this low strike probability (Figure 3.31a). A probability more than 10% appears over the central Arctic Ocean and the Laptev Sea. The separation suggests that some ensemble members predict the location of the mid-latitude cyclone migrating from lower latitude and others predict the location of the AC after cyclone merging (section 3.2.3).

The BSS and reliability diagram for probability of AC existence within 400-km radius from a given location, calculated based on the strike probability, are shown in Figure 3.33. In the 1.0-day forecast, the regressed line (bold black) of reliability follows the diagonal line for the entire forecast probability (Figure 3.33a), which is consistent with the small position error and high existence probability for each AC (Figure 3.33a and c). The GEFS reforecast, however, overestimates the high-occurrence area in the 3.0-day forecast (Figure 3.33b), indicating that the GEFS reforecast is overconfident in predicting AC existence within 400-km radius. As indicated by the raw value line (thin gray), the high-probability range (probability more than 50%) is the

primary contributor to the overconfidence of the GEFS reforecast. Overestimation in the high-probability range is larger in the 5.0-day forecast than in the 3.0-day forecast (Figure 3.33c). Furthermore, a 100% probability area is not present with this lead time, and the GEFS reforecast underestimates the AC existence in the low-probability range (probability less than 40%). The number of missed grid points (0%) is more than ten times that of the 1.0-day forecast. However, the BSS are positive up to the 5.0-day forecast. Therefore, the strike probability is useful for estimation of the AC existence up to a lead time of 5.0 days. The 7.0-day forecast (Figure 3.33d), is even more overconfident and has negative BSS. Overall, although the strike probabilities are overconfident with lead time of more than 1.0 day, they provide useful information for existence of extraordinary ACs up to a lead time of 5.0 days.

3.3.3 Summary of average predictability

Overall, ECMWF showed 1- to 1.5-day advantage in predicting the existence, central pressure, and central position of the ACs, compared with the other centers. The second-best performing center was dependent on the forecast lead time and the AC event. The average central position error for ECMWF (CMC) decreased to ≤ 469.1 km, which is half of the average radius for the 10 ACs in 2008–2016, at a lead time of 4.5 (2.5) days, and its average central pressure error was 8.8 (5.5) hPa. Further, JMA, NCEP, and UKMO had an average position error ≤ 469.1 km at a lead time of 3.5 days, and their average central pressure errors were 9.2, 10.7, and 10.4 hPa, respectively. The results suggest that the operational EPSs generally predict the position of the ACs within 469.1 km at 2.5–4.5 days before the mature stage, with a central pressure error of 5.5–10.7 hPa.

The forecast skills in the existence, central pressure, and central position of the 10 ACs predicted by the GEFS reforecast were similar to those predicted by the five leading NWP centers in TIGGE. The GEFS reforecast exhibited the existence probability of >0.9 at lead times of 1.0–

3.0 days. The existence probability was similar to that in JMA and UKMO throughout entire lead times, except at the lead time of 7.0 days. The forecast skill of central position was similar to that in JMA, NCEP, and UKMO, at all lead times, except at lead time of 4.0 days. At the lead time of 4.0 days, the GEFS reforecast showed the similar skill to the CMC forecast. Therefore, the average central position error for the GEFS reforecast decreased to ≤ 433.1 km at a lead time of 3.0 days, and its average central pressure error was 9.0 hPa. Furthermore, the existence probability, central pressure error, and position error for the 26 ACs in the GEFS reforecast showed almost similar to those in the five leading NWP centers in TIGGE. The average central position error decreased to ≤ 433.1 km, which is half of the average radius for the 10 ACs in 1986–2016, at lead time of 3.0–3.5 days. The existence probability and forecast skills of central pressure and position estimated from the 26 ACs in the GEFS reforecast supported the generality of these estimated from the 10 ACs in the five leading NWP center in TIGGE.

Neither existence probability nor forecast errors for the individual ACs show any improvements in the forecast skill. Therefore, the forecast skills of extraordinary ACs significantly depend on events, regardless year-by-year improvement of forecast skills in Z500 in the Northern Hemisphere (Hamill et al. 2013) and the Arctic (Daily scores for Z500 predicted by the GEFS reforecast are available at the TIGGE Museum,). These results indicate that the variability of forecast skills among the AC events is so large that an improvement by increasing the number of observations does not appeared.

Probabilistic forecasts of the GEFS reforecast for the AC existence within 400-km radius, based on the strike probability, are reliable at a lead time of 1.0 day. On the other hand, the probabilistic forecasts are overconfident at lead times of 3.0, 5.0, and 7.0 days, as with other severe weather events (Matsueda and Nakazawa, 2015). However, Brier Skill Scores are positive up to a lead time of 5.0 days, indicating that the strike probability forecast provides useful information for approaching of extraordinary ACs up to the 5.0-day forecast.

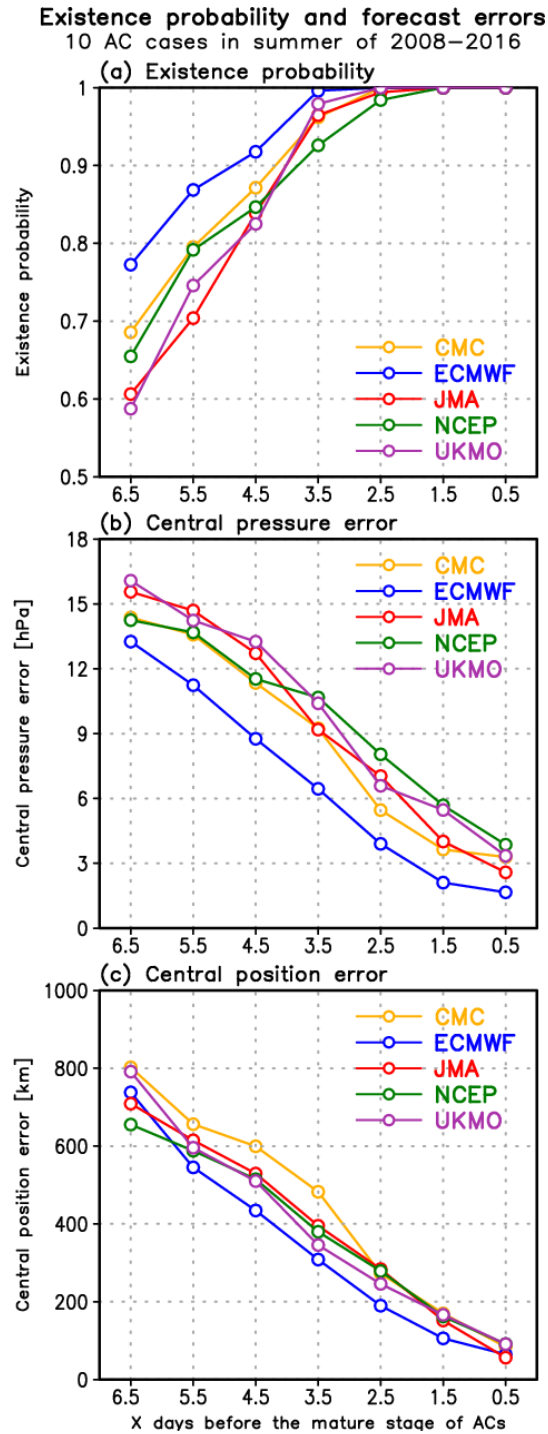


Figure 3.24 Average (a) existence probability, (b) central pressure error, and (c) central position error for the mature stages of 10 extraordinary ACs in summer of 2008-2016, predicted by CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), as a function of forecast lead time. The forecasts were verified against ERA-Interim.

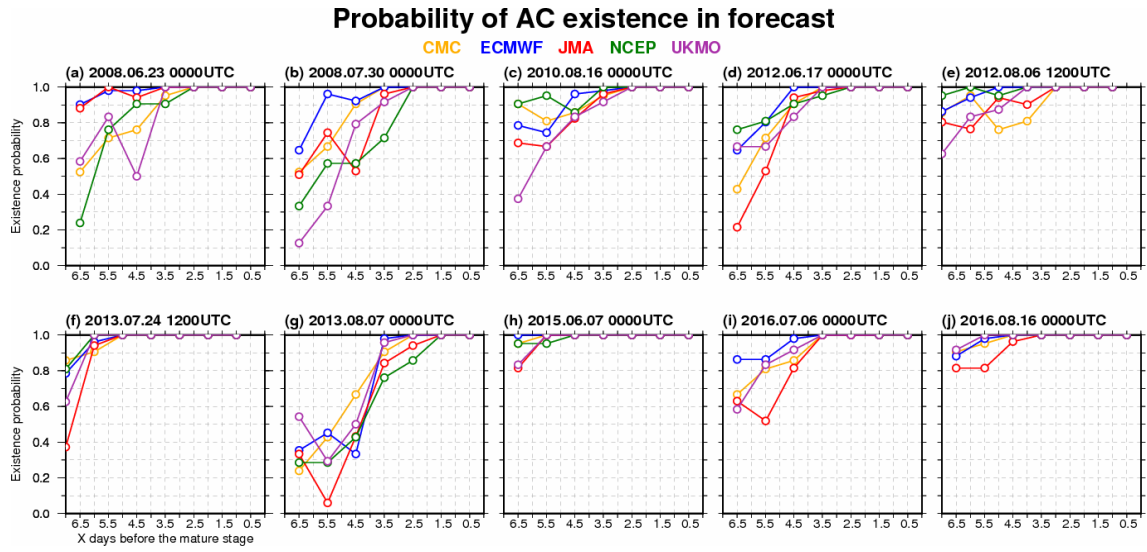


Figure 3.25 As in Fig. 3.24a, but for each AC.

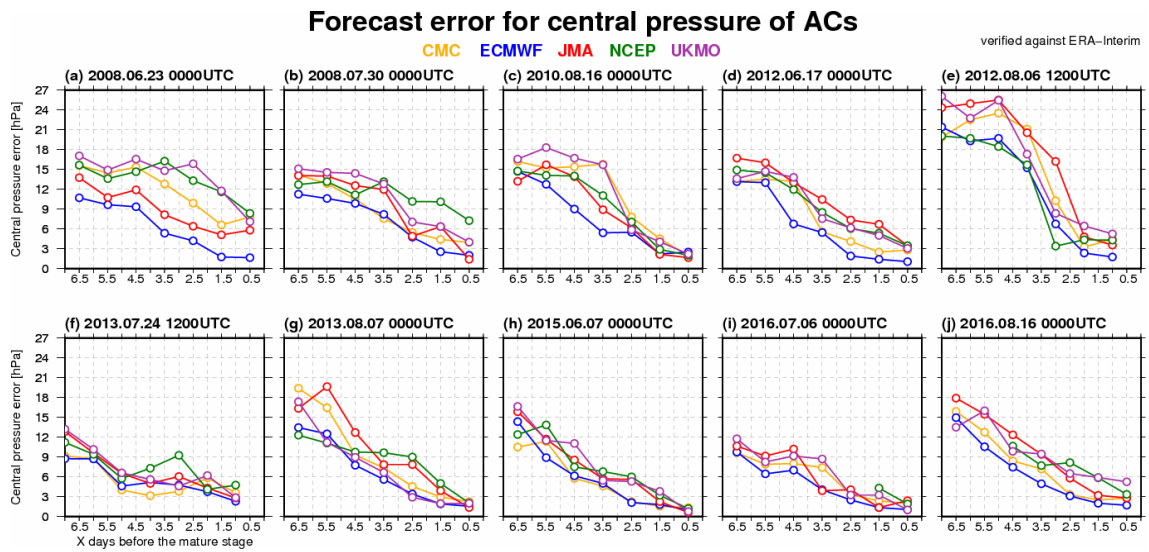


Figure 3.26 As in Fig. 3.24b, but for each AC.

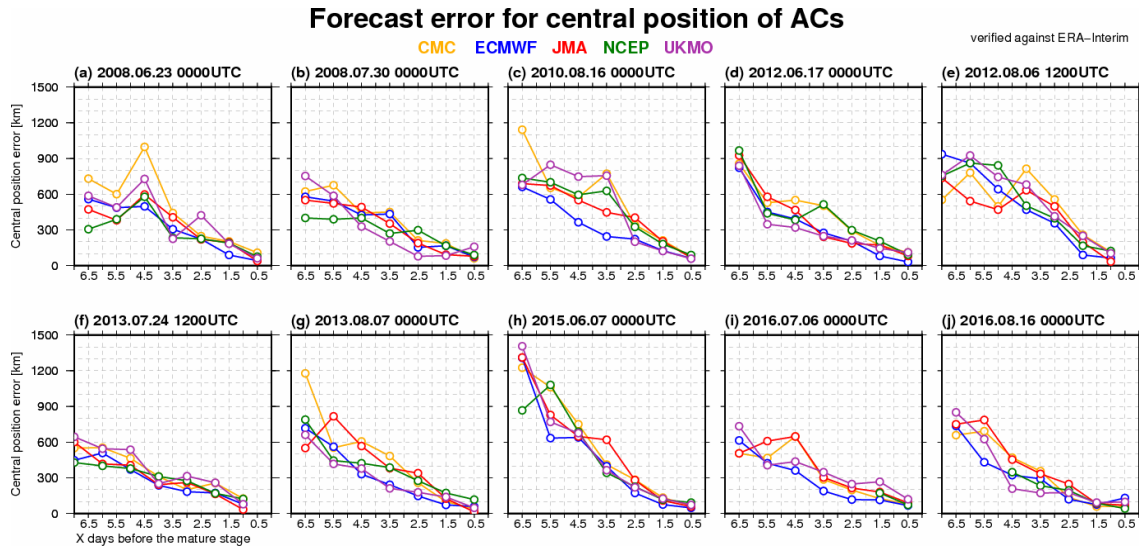


Figure 3.27 As in Fig. 3.24c, but for each AC.

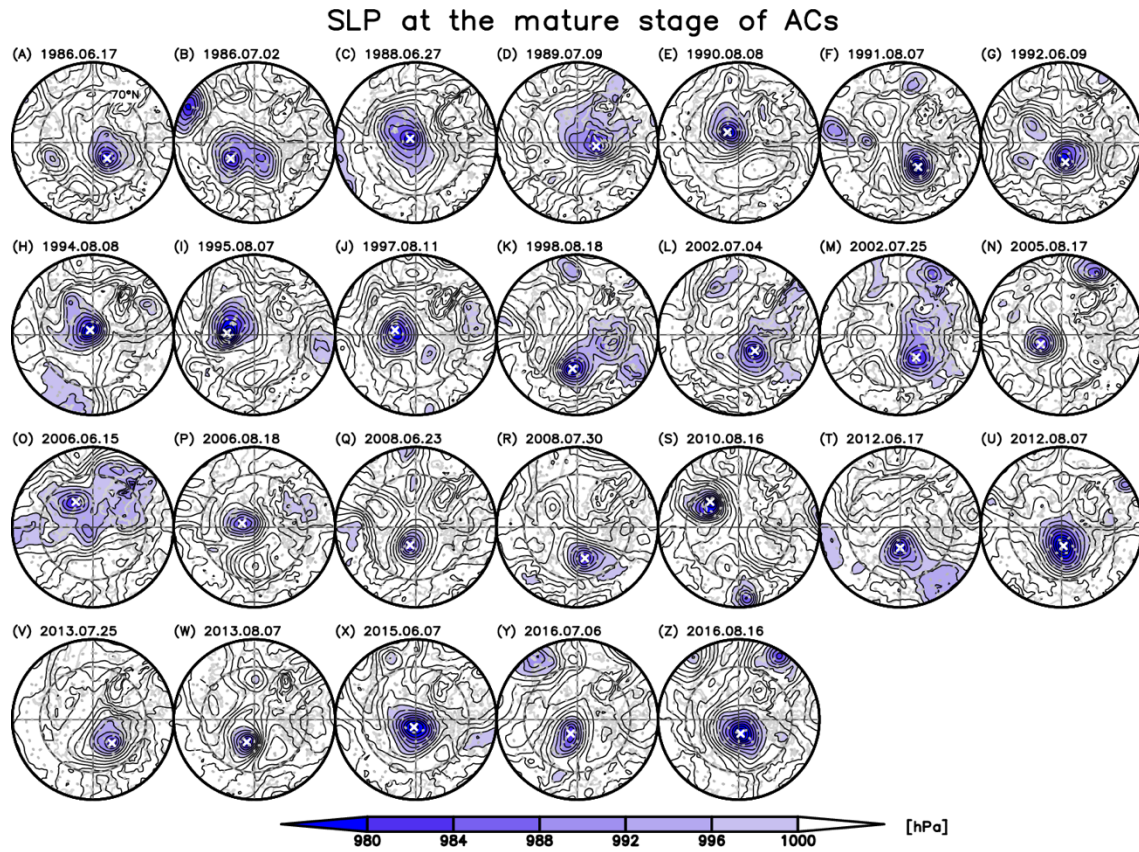


Figure 3.28 Observed sea-level pressure (SLP) fields for the 26 extraordinary ACs at their mature stage during summer 1986–2016 (GEFS control analysis). The date of the mature stage is given above each panel. Shading represents SLP < 1000 hPa. The white cross indicates the AC center.

Table 3.2 Date, central pressure, and radius during the mature stage of each AC event in Summer during 1986–2016.

Event (mature date)	(A) 17 Jun. 1986	(B) 02 Jul. 1986	(C) 27 Jun. 1988	(D) 09 Jul. 1989	(E) 08 Aug. 1990	(F) 07 Aug. 1991
Central pressure (hPa)	975.2	976.3	976.8	977.1	972.8	967.8
Radius (km)	541.3	542.3	1512.8	811.0	1015.0	919.2
Event (mature date)	(G) 09 Jun. 1992	(H) 08 Aug. 1994	(I) 08 Aug. 1995	(J) 11 Aug. 1997	(K) 18 Aug. 1998	(L) 04 Jul. 2002
Central pressure (hPa)	972.8	971.0	964.7	975.2	972.2	973.1
Radius (km)	834.6	660.9	1072.3	910.5	700.7	612.6
Event (mature date)	(M) 25 Jul. 2002	(N) 17 Aug. 2005	(O) 15 Jun. 2006	(P) 18 Aug. 2006	(Q) 23 Jun. 2008	(R) 30 Jul. 2008
Central pressure (hPa)	979.1	979.5	978.9	980.9	979.1	978.5
Radius (km)	651.3	728.4	590.1	885.4	783.2	739.1
Event (mature date)	(S) 16 Aug. 2010	(T) 17 Jun. 2012	(U) 07 Aug. 2012	(V) 25 Jul. 2013	(W) 07 Aug. 2013	(X) 07 Jun. 2015
Central pressure (hPa)	969.1	974.7	964.7	979.7	975.5	970.4
Radius (km)	1032.5	749.8	1115.0	1356.5	718.8	1161.9
Event (mature date)	(Y) 06 Jul. 2016	(Z) 16 Aug. 2016				
Central pressure (hPa)	980.1	968.3				
Radius (km)	866.9	1011.7				

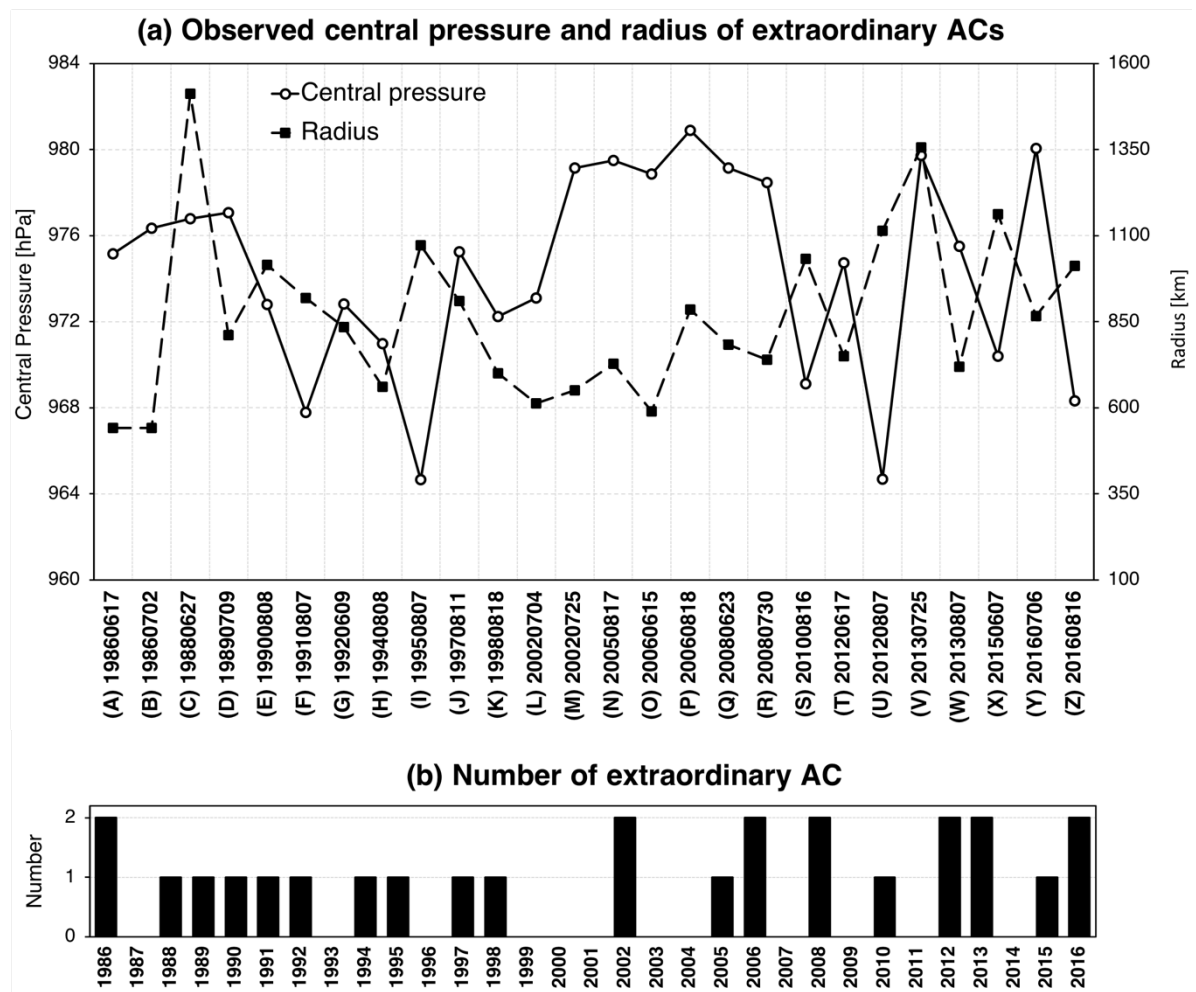


Figure 3.29 (a) Observed central pressure (solid line; left vertical axis) and radius (broken line; right vertical axis) for the 26 ACs designated by (A) to (Z). (b) Time series of the number of extraordinary ACs in each summer from 1986 to 2016.

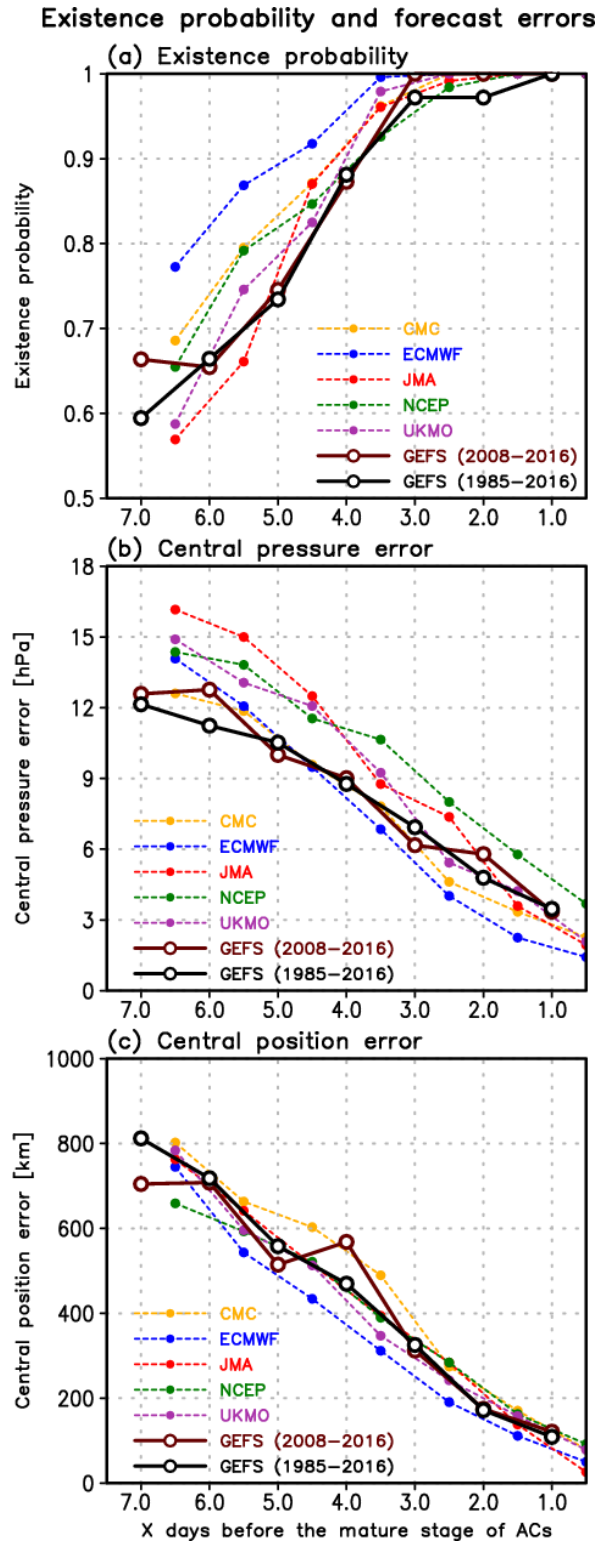


Figure 3.30 Average (a) existence probability, (b) central pressure error, and (c) central position error during the mature stage of 26 ACs in summer during 1985–2016 (black) and 10 ACs in summer during 2008–2016 (brown) from the GEFS reforecast, as a function of forecast lead time. Colored broken lines are the same as Figure 3.24, but verified against own control analysis.

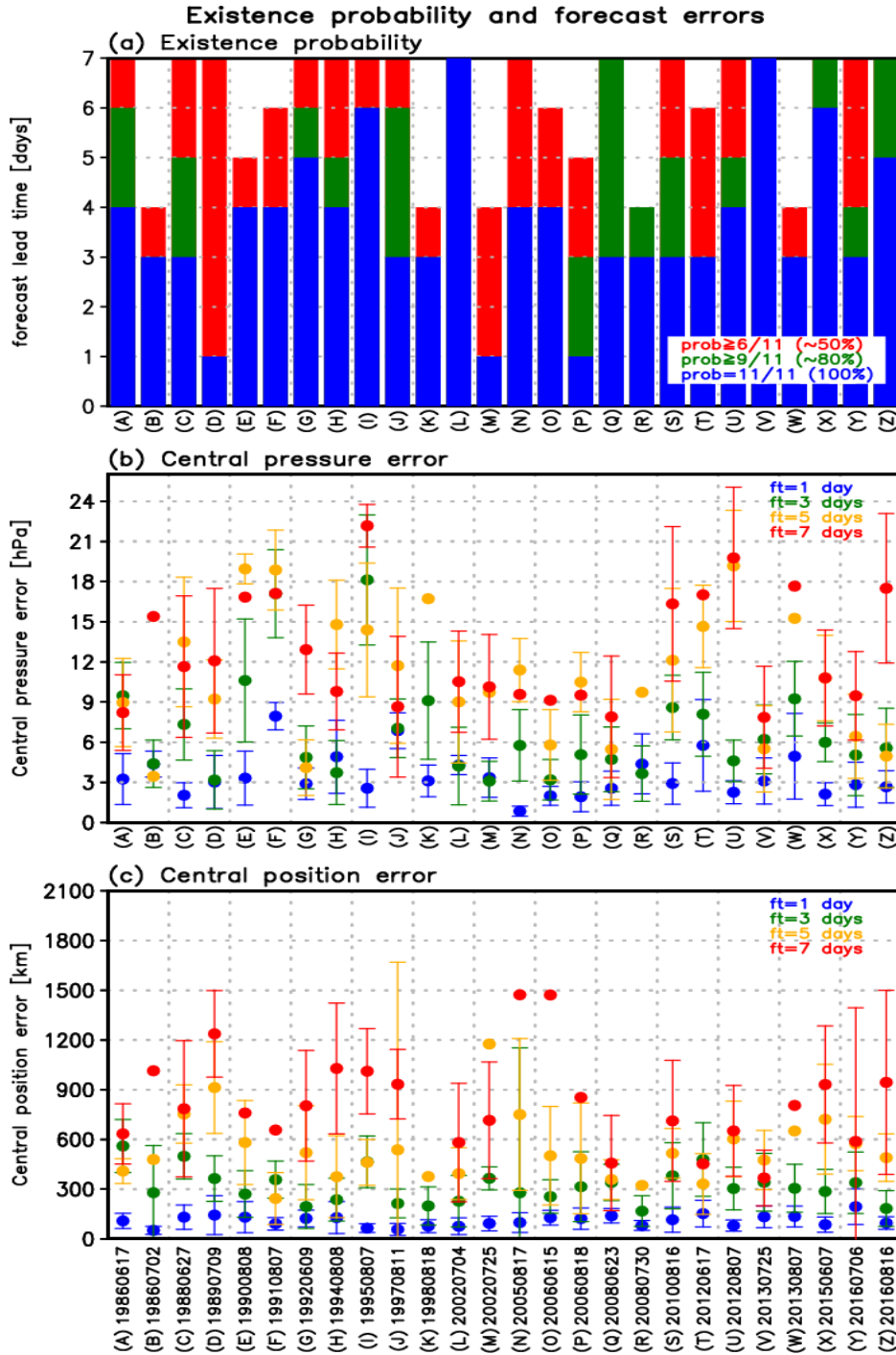


Figure 3.31 Existence probability for individual mature ACs in summer during 1985–2016 from the GEFS reforecast. The vertical axis is forecast lead time, and colored bars indicate existence probabilities of $\geq 6/11$ (red), $\geq 9/11$ (green), and 11/11 (blue). (b, c) Ensemble mean central (b) pressure error and (c) position error for individual mature ACs at lead times of 1 (blue), 3 (green), 5 (yellow), and 7 (red) days, predicted by the GEFS reforecast. Error bars indicate one standard deviation and are given for forecasts with existence probabilities of $\geq 6/11$. These forecasts were verified against own control analyses.

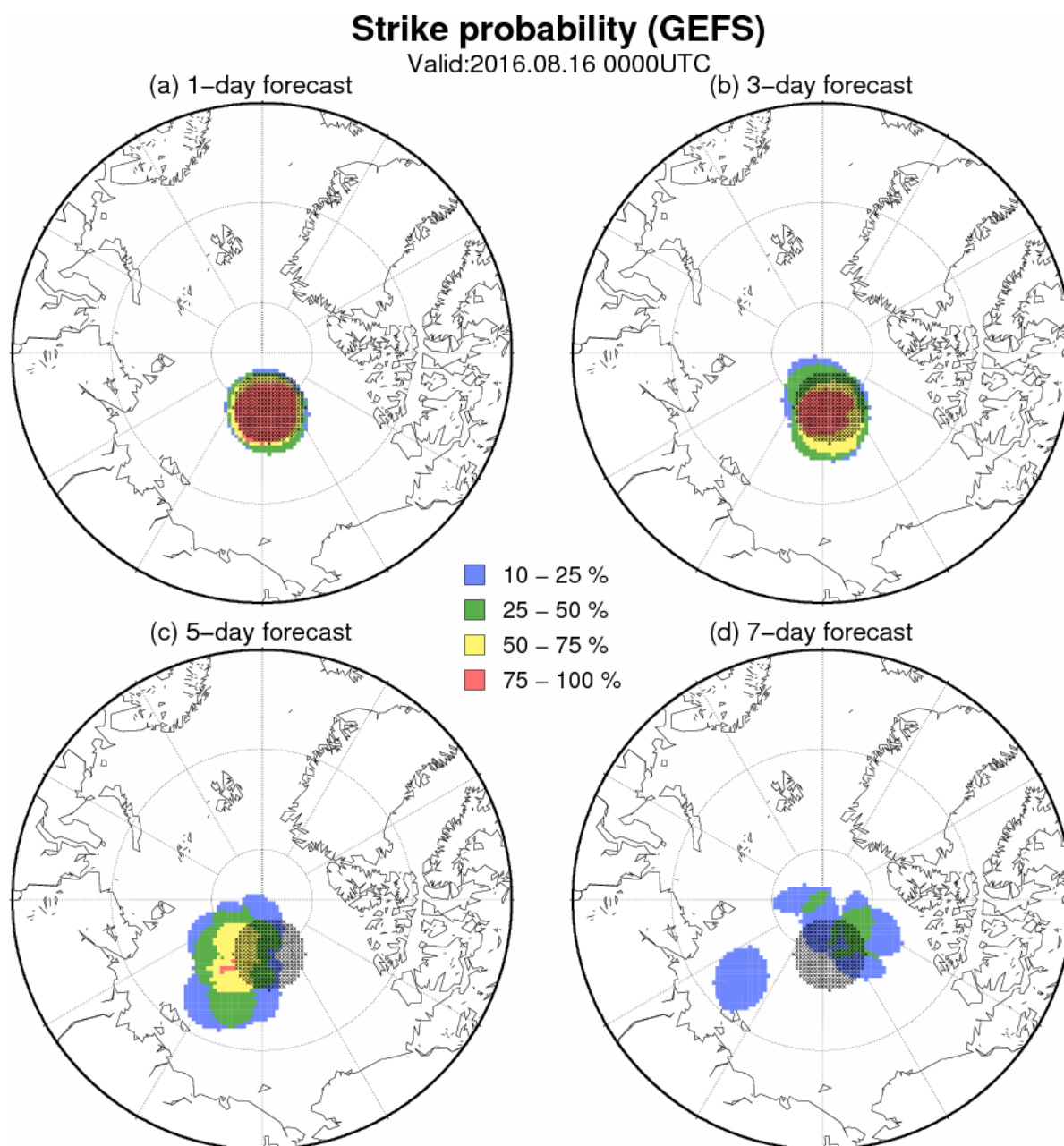


Figure 3.32 Strike probability map for the AC in August 2016 from (a) 1-, (b) 3-, (c) 5-, and (d) 7-day forecasts of the GEFS reforecast. The strike probability is defined as the probability that the AC is located within 400 km of a given location during its mature stage. This probability is calculated based on the number of ensemble members which predicted ACs. Colored shading indicates probabilities of 5%–25% (blue), 25%–50% (green), 50%–75% (yellow), and 75%–100% (red). Black shading indicates a location within 400 km of the center of an observed AC during its mature stage.

Reliability diagram for position forecasts of AC

1986–2016 (26 events), radius:400km

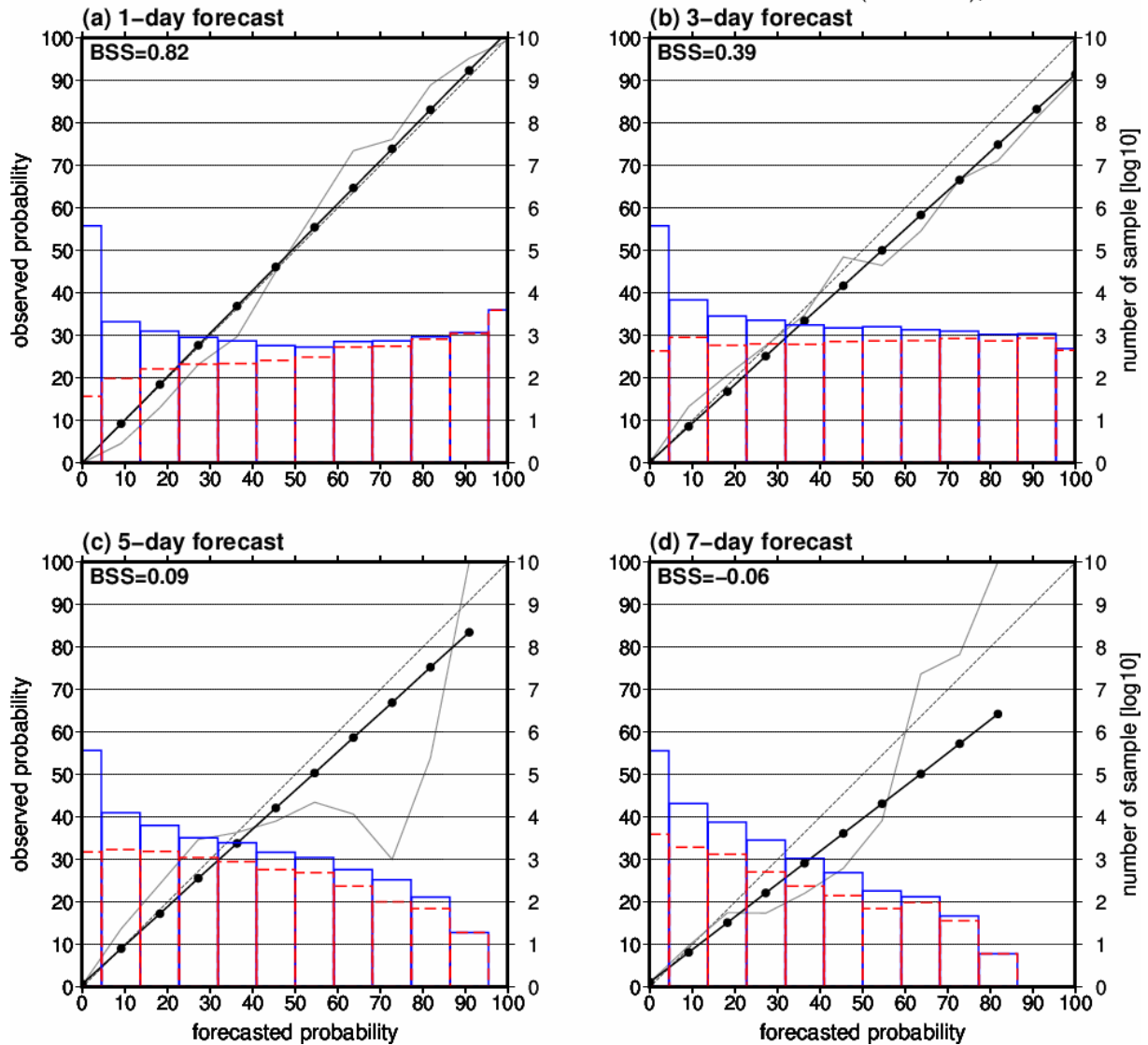


Figure 3.33 Reliability diagrams for forecast probabilities that an AC exists within a 400-km radius during its mature stage from (a) 1-, (b) 3-, (c) 5-, and (d) 7-day forecast of the GEFS reforecast. Reliabilities and their regression lines are indicated by gray and black lines, respectively (left vertical axis). The number of samples (grid points) predicting an event is indicated by blue bars, and the number of samples for which the event actually occurred is indicated by dashed red bars (right vertical axis). The Brier Skill Score (BSS) are given at the upper-left corner of each panel. Forecast probabilities were verified for 26 ACs during 1986–2010.

CHAPTER 4 Discussion

4.1 *Predictions of individual extraordinary ACs*

Upper-level fields have significant influences on the development and location of extraordinary ACs. In particular, accurate predictions of upper-level troughs and ridges are necessary for predicting cyclogenesis and development via baroclinic instabilities during the developing stage for cyclones migrating from lower latitudes. The trough and ridge also influence the track of the migrating cyclones. Zheng et al. (2013) performed a sensitivity analysis of a snowstorm (a mid-latitude cyclone) that developed over the east coast of North America. They found that the predictability of its development was influenced by initial errors associated with the upper-level trough and ridge of cyclones over the northeastern Pacific and central United States, while the predictability of its position was influenced by initial errors associated with the upper-level short-wave troughs over the southern Great Plains. Present study also showed that the initial difference in the upper-level trough and ridge between higher- and lower-skill members propagates northeastward and increases during the developing stage of migrating cyclones in the AC12, AC15, and AC16 events. In addition, AC13 predictions revealed that initial errors around upper-level troughs lead to failures in predicting cyclogenesis for mid-latitude cyclones. Therefore, the development and position of predicted ACs depend on predictions of upper-level troughs and ridges.

In addition to upper-level troughs and ridge, TPVs are a sensitivity source in AC predictions. The location of a TPV is correlated with that of surface AC during the developing stage through vertical coupling in the AC12 event. Hence, accurate predictions of TPVs are also required to correctly predict the development and location of ACs. Radiative forcing can contribute to the intensification and lifetime of TPVs (Cavallo and Hakim, 2009, 2010). Cavallo and Hakim (2012) showed that radiative forcing leads to a shift in TPV due to changes in large-scale atmospheric circulation over the Arctic, as inferred from a comparison of simulations with and without radiative forcing. In the present study, given that the forecast model in each NWP center uses the same radiation scheme for all ensemble members, the

difference in its impact on the spread of TPV locations would be small among ensemble members. However, differences in large-scale circulation between higher-skill members and lower-skill members can cause difference in TPV location. Therefore, not only initial perturbations around the TPV but also those at lower latitudes can lead to the difference in TPV location.

This study also suggests that both baroclinic instability and vertical coupling between TPVs and surface cyclones are required for the development of extraordinary ACs. Vertical coupling plays an essential role in the development of some ACs, whereas baroclinic instability is more important in the development of other ACs. For example, the vertical coupling and consequent merging of upper-level warm cores were as important as the baroclinic instability in the development of AC12. Tao et al. (2017b) used the Weather Research and Forecasting model to show that temperatures at 300–100 hPa caused an intensification of AC12 and also affected its location during the mature stage. In contrast, baroclinic instability was more important in the development of AC16 than the vertical coupling of vortices. The influence of the vertical coupling is larger on prediction of AC16 location than on prediction of AC16 development. The findings of the present study are consistent with the results in Tao et al. (2017b).

4.2 Forecast skill for extraordinary ACs

This study also investigated the average forecast skill for extraordinary ACs during summer. Ten (twenty-six) extraordinary ACs were detected in summer during 2008–2016 (1986–2016). In contrast to previous studies, the 26 extraordinary ACs did not show a long-term trend in their number, central pressure (strength), or radius (Figure 3.29). Simmonds and Keay (2009) reported that the number of cyclones over the Arctic in September displayed no significant trend during 1979–2009, whereas their strengths and radii followed a significant increasing trend. In contrast, Koyama et al. (2016) showed that the number of cyclones over the Arctic during summer indicated a significant decreasing trend. They also reported that the number of extreme cyclones over the Arctic shows no significant temporal trend in the period of 1979–2014. However, Rinke et al. (2017) determined that the number of extreme

cyclones increased during 1979–2015, using station data from Ny-Ålesund. These discrepancies in the number, strength, and radius of cyclones over the Arctic may be attributed to differences in data, methods of cyclone detection, and period (Neu et al., 2013; Simmonds et al., 2008).

The average central position errors (Figure 3.24) indicate that the average forecast skill for extraordinary ACs is lower than that for mid-latitude cyclones in the Northern Hemisphere (Froude, 2010), although the position error for mid-latitude cyclones in Froude (2010) was estimated from both extraordinary and ordinary mid-latitude cyclones. The forecast skill for mid-latitude cyclones in the Southern Hemisphere is lower than that for the Northern Hemisphere (Froude, 2011) presumably due to the sparse network of observations. The network of observations is also sparse over the Arctic (Jung et al., 2016), resulting in uncertainties in control analysis (Inoue et al., 2015; Jung and Matsueda, 2016). Given that the forecast skill for extraordinary ACs is similar to that for mid-latitude cyclones in the Southern Hemisphere, the sparse observation network may contribute to the lower forecast skills for both extraordinary ACs and mid-latitude cyclones in the Southern Hemisphere.

Mechanisms associated with the development and maintenance of extraordinary ACs may also contribute to the lower forecast skill. Yamazaki et al. (2015) showed that radiosonde observations over the Arctic improved predictions of the development and position of AC12, primary due to improvements in TPV forecasts. Thus, both the upper-level trough and the TPV can contribute to forecast uncertainty. In addition, the cyclone merging might be an additional source of uncertainties. For the AC12 prediction, the ensemble spread significantly increased during the merging between AC and mid-latitude cyclone. Prieto et al. (2003) showed that the interaction of two barotropic vorticities significantly depends on the ratios of their strengths and sizes, and the distance between the vortices. Since a barotropic vortex (AC) merged with a baroclinic vortex (mid-latitude cyclone) in AC cases, the interaction between these two types of vortices might be highly nonlinear and sensitive to their strengths and sizes, and the distance compared with two barotropic vortices.

The difference in average forecast skills between ECMWF and JMA suggests a minimal influence of ensemble size on forecast skill. Given that the model resolution has little impact on forecast skill of Z500 over the Arctic (Bauer et al., 2016) and the representation of extraordinary mid-latitude cyclones (Jung et al., 2006), and the ECMWF control analysis is the closest to the mean of all the NWP centers' control analysis (Park et al., 2008; Wei et al., 2010), the higher quality of the ECMWF control analysis may be the primary contributor to its higher forecast skill in the comparison among the NWP centers for the same AC event. In contrast, neither existence probability nor forecast errors for the individual ACs show any improvements in the forecast skill of the GEFS reforecast, regardless year-by-year improvement of forecast skills in Z500 in the Northern Hemisphere (Hamill et al. 2013) and the Arctic (Daily scores for Z500 predicted by the GEFS reforecast are available at the TIGGE Museum). The variability of forecast skills among the AC events is so large that an improvement by increasing the number of observations is not apparent. Therefore, accurate prediction of upper-level fields due to increase of observation improves predictions for some ACs. For the other ACs, accurate prediction of other processes, like surface temperature, might have large impacts on improvement of AC prediction.

The best-performing NWP center in terms of forecast skill in predicting AC central positions depends on the AC event, as well as other severe events (Matsueda and Nakazawa, 2015), particularly at lead times of 4.0–7.0 days. This suggests that an estimate of central-position forecast uncertainties using a multi-center grand ensemble (MCGE) approach (Matsueda and Tanaka, 2008) would be useful to ensure the safety of human activities over the Arctic, such as shipping on the Northern Sea Route and aviation on the Polar Route. In addition, the MCGE approach would reduce the overconfidence of position predictions for extraordinary ACs, as for tropical cyclones (Yamaguchi et al., 2012). To demonstrate the usefulness of the MCGE approach for AC forecasts, future work should focus on extending the period of analysis and including ordinary AC events due to small sample size of extraordinary ACs.

CHAPTER 5 **Conclusions**

This study investigated the forecast skill for extraordinary ACs in summer and key factors for their accurate prediction, using five leading operational medium-range ensemble forecasts and the second-generation GEFS reforecast. In particular, this work evaluated predictions of the probabilities of existence, central pressure, and central position of extraordinary ACs.

First, predictions of individual extraordinary ACs were investigated using the operational medium-range ensemble forecast. Ten extraordinary ACs were detected in summer during 2008–2016 based on criteria of central pressure, central position, and area-averaged upper-level temperature anomalies. Because of the sparse network of observations, larger analysis uncertainties are expected over the Arctic than at lower latitudes. The central pressure and position of individual mature ACs from control analyses of each NWP center are similar to those of ERA-Interim data. The average differences in the central pressure and position for the 10 ACs between each center's control analysis and ERA-Interim are less than 0.5 hPa and 37.6 km, respectively. The AC in August 2012 (AC12) had the lowest central pressure of all the ACs, followed by the AC in August 2016 (AC16), whereas the AC in June 2015 (AC15) had the largest size. In addition to predictions for these three ACs, the forecast for an AC that occurred in August 2013 (AC13) was analyzed in detail, because the period between its cyclogenesis and mature stage was short.

After classifying ensemble members based on errors in their predictions for the development and position of each AC, a comparison of the composite fields of geopotential height at 300 hPa and temperature anomalies at 250 hPa revealed that accurate prediction of the upper-level trough is one of the key factors in accurate predictions of extraordinary ACs. Differences in the predicted upper-level troughs between ensemble members lead to differences in the development and location of surface cyclones migrating into the Arctic through baroclinic instability and the direction of the upper-level steering flow, respectively (AC12, AC15, and AC16). The predicted upper-level trough and ridge also influence the cyclogenesis of migrating cyclones (AC13). Tropopause polar vortices (TPVs) are another

key factor in accurate predictions of extraordinary ACs. As evidenced by a developing-stage surface AC over the Arctic that was vertically coupled with a TPV, TPV locations are correlated with surface AC locations. For AC12, predictions of cyclone merging and the consequent development of an upper-level warm core were essential for accurate predictions of intensification and position. However, forecasts for cyclone merging were important mainly for position predictions in the case of AC16. Therefore, both TPV and upper-level trough and ridge predictions have an impact on predictions of the development and position of extraordinary ACs.

Second, average forecast skills of extraordinary ACs were investigated. On average, ECMWF exhibits a 1.0- to 1.5-day advantage in predicting the existence, central pressure, and central position of extraordinary ACs, compared with the other centers. The second-best performing center depends on the forecast lead time and the AC event. The average central position error for ECMWF (CMC) decreases to ≤ 469.1 km (half the average radius for the 10 ACs) at a lead time of 4.5 (2.5) days, and its average central pressure error is 8.8 (5.5) hPa. Furthermore, JMA, NCEP, and UKMO have an average position error of ≤ 469.1 km at a lead time of 3.5 days, and their average central pressure errors are 9.2, 10.7, and 10.4 hPa, respectively. These results indicate that the operational EPSs generally predict the position of the ACs within 469.1 km at 2.5–4.5 days before the mature stage, with a central pressure error of 5.5–10.7 hPa. Results also suggest that a higher quality of the control analysis, rather than higher model resolution or larger ensemble size, is a main contributor to improve forecast skill for extraordinary ACs.

The mean forecast skill for extraordinary ACs is lower than that for mid-latitude cyclones in the Northern Hemisphere (Froude, 2010), but similar to that in the Southern Hemisphere (Froude, 2011). The sparse network of observations over both the Arctic and Southern Hemisphere (Jung et al., 2016) and the consequent analysis uncertainties in initial conditions (Inoue et al., 2015; Jung and Matsueda, 2016) are presumably important reasons for the observed similarity in forecast skills. However, mechanisms associated with the development and maintenance of ACs during summer, particularly a coupling between upper- and lower-level vorticities, also contribute to the lower predictability

(Yamazaki et al., 2015). Furthermore, the best-performing center in predicting the central position depends on the AC event, along with other severe events (Matsueda and Nakazawa, 2015). This suggests that an estimate of uncertainties in the central position forecast using a multi-center grand ensemble approach would provide useful information for shipping on the Northern Sea Route and aviation on the Polar Route.

In addition, this study also assessed average forecast skills for 26 extraordinary ACs occurred in summer during 1986–2016. More than 90% ensemble members of the GEFS reforecast predicted existence for the 26 ACs up to a lead time of 3.0 days. The average existence probabilities of the GEFS reforecast for the 10 ACs during 2008–2016 were similar to those for the 26 ACs. The average central position errors of the GEFS reforecast for the 26 and 10 ACs decreases to less than 433.1 km (half of average radius for observed 26 ACs at their mature stage) with a lead time of 3.0 days. Average central pressure errors for both the 26 and 10 ACs are 6.9 hPa and 6.1 hPa at such a lead time. This result indicates that the GEFS reforecast has similar forecast skills in predicting central pressure and position to the operational EPSs' forecasts. Besides, probabilistic forecasts of the GEFS reforecast for the AC existence within 400-km radius, based on the strike probability, are reliable at a lead time of 1.0 day. On the other hand, the probabilistic forecasts of the GEFS reforecast are overconfident at lead times of 3.0, 5.0, and 7.0 days, as with other severe weather events (Matsueda and Nakazawa, 2015). However, BSS are positive up to a lead time of 5.0 days, indicating that the strike probability forecast provides useful information for approaching of extraordinary ACs up to the 5.0-day forecast.

This study focused on predictabilities for extraordinary ACs; however, accurate predictions for ACs of smaller scale and weaker intensity are also important to ensure the safety of human activities in the Arctic. These ACs are generated by baroclinic instabilities due to temperature differences over land, the open ocean, and sea ice (Inoue and Hori, 2011). Therefore, lower-boundary conditions (e.g., SST, SIC, and soil moisture) and analysis uncertainties in surface variables (Bauer et al., 2016; Jung and Matsueda, 2016) will have significant influences on the predictability of these ACs. In addition, cyclonic activity

in summer is different each year. For example, ACs occurred frequently during the summer of 2016, but no extraordinary ACs occurred during the summer of 2014. Hence, predicting cyclonic activity during upcoming summers on sub-seasonal to seasonal timescales will also be important for decisions related to human activity in the Arctic. Further AC predictability studies are therefore required at various time scales.

Acknowledgements

First of all, I would like to express special appreciation to Prof. Hiroshi L. Tanaka, Center for Computational Sciences, University of Tsukuba, for his valuable comments and encouragements. I am grateful to Prof. Mio Matsueda, Center for Computational Sciences, University of Tsukuba, for his useful advices and encouragements. I am also thankful to Profs. H. Ueda and H. Kusaka for accepting the committee of my doctoral thesis and their helpful comments. I also grateful to Profs. K. Ueno and Y. Kamae for their constructive comments and suggestions. I express my appreciation to Prof. M. Ishii, the Meteorological Research Institute in Japan, for accepting the committee of my doctoral thesis. I would like to thank Dr. T. Aizawa of the Meteorological Research Institute in Japan / the University of Tokyo for his constructive suggestions and providing the cyclone center detection algorithm. I am grateful to all other students and staff of the Climatology and Meteorology Group, the University of Tsukuba, for their comments and supports. The authors also thank ECMWF for providing ERA-Interim and TIGGE data and NOAA for providing GEFS reforecast data. This study was supported by the Arctic Challenge for Sustainability (ArCS) Project. Finally, I am most thankful to my family for their support and understanding.

References

- Aizawa, T., and H.L. Tanaka, 2016: Axisymmetric structure of the long lasting summer arctic cyclones. *Polar Science*, **10**, 192–198.
- Aizawa, T., H.L. Tanaka, and M. Satoh, 2014: Rapid development of arctic cyclone in June 2008 simulated by the cloud resolving model NICAM. *Meteorol. Atmos. Phys.*, **137(660)**, 105–117.
- Ambaum, M.H.P., B.J. Hoskins, and D.B. Stephenson, 2001: Arctic Oscillation or North Atlantic Oscillation? *J. Clim.*, **14**, 3495–3507.
- Bauer, P., L. Magnusson, J.-N. Thépaut, and T.M. Hamill, 2016: Aspects of ECMWF model performance in polar areas. *Quart. J. Royal Meteor. Soc.*, **142**, 583–596.
- Boisvert, L.N., A.A. Petty, and J.C. Stroeve, 2016: The impact of the extreme winter 2015/16 arctic cyclone on the Barents–Kara Seas. *Mon. Wea. Rev.*, **144**, 4279–4287.
- Cavallo, S.M., and G.J. Hakim, 2009: Potential vorticity diagnosis of a tropopause polar cyclone. *Mon. Wea. Rev.*, **138**, 1358–1371.
- Cavallo, S.M., and G.J. Hakim, 2010: Composite structure of tropopause polar cyclones. *Mon. Wea. Rev.*, **138**, 3840–3857.
- Cavallo, S.M., and G.J. Hakim, 2012: Radiative impact on tropopause polar vortices over the Arctic. *Mon. Wea. Rev.*, **140**, 1683–1702.
- Cohen, J., J. A. Screen, J.C. Furtado, M. Barlow, D. Whittleston, D. Coumou, J. Francis, K. Dethloff, D. Entekhabi, J. Overland, and J. Jones, 2014: Recent Arctic amplification and extreme mid-latitude weather. *Nature Geosci.*, **7**, 627–637.
- Coumou, D., J. Lehmann, and J. Beckmann, 2015: The weakening summer circulation in the Northern Hemisphere mid-latitudes. *Mon. Wea. Rev.*, **140**, 1683–1702.
- Crawford, A.D., and M.C. Serreze, 2015: A new look at the summer Arctic frontal zone. *J. Clim.*, **28**, 737–754.
- Crawford, A.D., and M.C. Serreze, 2016: Does the summer arctic frontal zone influence arctic ocean cyclone activity? *J. Clim.*, **29**, 4977–4993.
- Dee, D., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A.C.M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A.J. Geer, L. Haimberger, S.B. Healy, H. Hersbach, E.V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A.P. McNally, B.M. Monge-Sanz, J.J. Morcrette, B.K. Park, C. Peubey, P. de Rosnay, C. Tavalato, J.N. Thépaut, and F. Vitart, 2011: The ERA-Interim

- reanalysis: configuration and performance of the data assimilation system. *Quart. J. Royal Meteor. Soc.*, **137**, 553–597.
- Eguíluz, V.M., J. Fernández-Gracia, X. Irigoien, and C.M. Duarte, 2016: A quantitative assessment of Arctic shipping in 2010–2014. *Nature Scientific Reports*, **6**, 30682, doi:10.1038/srep30682.
- Frame, T.H.A., M.H.P. Ambaum, S.L. Gray, and J. Methven, 2011: Ensemble prediction of transitions of the north atlantic eddy-driven jet. *Quart. J. Royal Meteor. Soc.*, **137**, 1288–1297.
- Frame, T.H.A., J. Methven, S.L. Gray, and M.H.P. Ambaum, 2013: Flow-dependent predictability of the North Atlantic jet. *Geophys. Res. Lett.*, **40**, 2411–2416.
- Froude, L.S.R., 2010: TIGGE: Comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **25**, 819–836.
- Froude, L.S.R., 2011: TIGGE: Comparison of the prediction of Southern Hemisphere extratropical cyclones by different ensemble prediction systems. *Wea. Forecasting*, **26**, 388–398.
- Graversen, R.G., and M. Wang, 2009: Polar amplification in a coupled climate model with locked albedo. *Clim. Dyn.*, **33**, 629–643.
- Hamill, T.M., G.T. Bates, J.S. Whitaker, D.R. Murray, M. Fiorino, T.J. Galarneau, Y. Zhu, and W. Lapenta, 2011: Predictions of 2010's Tropical cyclones using the GFS and ensemble-based data assimilation methods. *Bull. American Meteor. Soc.*, **94**, 1553–1565.
- Hamill, T.M. G.T. Bates, J.S. Whitaker, D.R. Murray, M. Fiorino, T.J. Galarneau, Y. Zhu, and W. Lapenta, 2013: NOAA's second-generation global medium-range ensemble reforecast dataset. *Mon. Wea. Rev.*, **139**, 3243–3247.
- Hanley, J, and R. Caballero, 2011: Objective identification and tracking of multicenter cyclones in the ERA-Interim reanalysis dataset. *Quart. J. Royal Meteor. Soc.*, **138**, 612–625.
- Inoue, J, T. Enomoto, and M.E. Hori, 2013: The impact of radiosonde data over the ice-free Arctic Ocean on the atmospheric circulation in the Northern Hemisphere. *Geophys. Res. Lett.*, **40**, 1–6, L12502, doi:10.1029/2011GL047696.
- Inoue, J, and M.E. Hori, 2012: Arctic cyclogenesis at the marginal ice zone: A contributory mechanism for the temperature amplification? *Geophys. Res. Lett.*, **38**, 1–6, doi:10.1002/grl.50207.
- Inoue, J., A. Yamazaki, J. Ono, K. Dethloff, and M. Maturilli, 2015: Additional Arctic observations improve weather and sea-ice forecasts for the Northern Sea Route. *Nature Scientific Reports*, **5**, 16868, doi:10.1038/srep16868.
- Johansson, E., A. Devasthale, M. Tjernström, A.M.L. Ekman, and T. L'Ecuyer, 2017: Response of the lower troposphere to moisture intrusions into the Arctic. *Geophys. Res. Lett.*, **44**, 2527–2536, doi:10.1002/2017GL072687.

- Jung, T., S.K. Gulev, I. Rudeva, and V. Soloviev, 2006: Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Quart. J. Royal Meteor. Soc.*, **132**(619), 1839–1857.
- Jung, T., N.D. Gordon, P. Bauer, D.H. Bromwich, M. Chevallier, and J.J. Day, K. Dawson, F. Doblas-Reyes, C. Fairall, H.F. Goessling, M. Holland, J. Inoue, T. Iversen, S. Klebe, P. Lemke, M. Losch, A. Makshtas, B. Mills, P. Nurmi, D. Perovich, P. Reid, I.A. Renfrew, G. Smith, G. Svensson, M. Tolstykh, and Q. Yang, 2016: Advancing polar prediction capabilities on daily to seasonal time scales. *Bull. American Meteor. Soc.*, **97**(9), 1631–1647.
- Jung, T., and M. Matsueda, 2016: Verification of global numerical weather forecasting systems in polar regions using TIGGE data. *Quart. J. Royal Meteor. Soc.*, **142**, 574–582.
- Kleist, D.T., D.F. Parrish, J.C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Wea. Forecasting*, **24**, 1691–1705.
- Knudsen, E., Y. Orsolini, T. Furevik, and K. Hodges, 2015: Observed anomalous atmospheric circulation in summers of unusual Arctic sea ice reduction. *J. Geophys. Res. Atmos.*, **120**, 2595–2611.
- Koyama, T., J. Stroeve, J. Cassano, and A. Crawford, 2017: Sea ice loss and arctic cyclone activity from 1979 to 2014. *J. Clim.*, **30**, 4735–4754.
- Letterly, A., J. Key, and Y. Liu, 2016: The influence of winter cloud on summer sea ice in the Arctic, 1983–2013. *J. Geophys. Res. Atmos.*, **121**, doi:10.1002/2015JD024316.
- Luo, B., D. Luo, L. Wu, L. Zhong, and I. Simmonds, 2017: Atmospheric circulation patterns which promote winter Arctic sea ice decline. *Env. Res. Lett.*, **12**, 054017, doi: 10.1088/1748-9326/aa69d0.
- Matsueda, M., 2009: Blocking predictability in operational medium-range ensemble forecasts. *SOLA*, **5**, 113–116.
- Matsueda, M., 2011: Predictability of Euro-Russian blocking in summer of 2010. *Geophys. Res. Lett.*, **38**, 1–6. <http://dx.doi.org/10.1029/2010GL046557>.
- Matsueda, M., and H. Endo, 2011: Verification of medium-range MJO forecasts with TIGGE. *Geophys. Res. Lett.*, **38**, 6–11. <http://dx.doi.org/10.1029/2011GL047480>.
- Matsueda, M., and M. Kyoda, 2016: Wintertime East Asian flow patterns and their predictability on medium-Range timescales. *SOLA*, **12**, 121–126.
- Matsueda, M., T. Nakazawa, 2015: Early warning products for severe weather events derived from operational medium-range ensemble forecasts. *Meteorological Applications*, **22**, 213–222.
- Matsueda, M., and T.N. Palmer, 2018: Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts. *Quart. J. Royal Meteor. Soc.*, **144**, 1012–1027. <https://doi.org/10.1002/qj.3265>.

- Matsueda, M., and H.L. Tanaka, 2008: Can MCGE outperform the ECMWF ensemble? *SOLA*, **4**, 77–80.
- Mori, M., M. Watanabe, H. Shiogama, J. Inoue, and M. Kimoto, 2014: Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. *Nature Geosci.*, **7**, 869–873.
- Melia, N., K. Haines, and E. Hawkins, 2016: Sea ice decline and 21st century trans-Arctic shipping routes, *Geophys. Res. Lett.*, **43**, 9720–9728, doi:10.1002/2016GL069315.
- Murray, R.J., and I. Simmonds, 1991: A numerical scheme for tracking cyclone centers from digital data. Part I: Development and operation of the scheme. *Australian Meteorological Magazine.*, **39**, 155–166.
- Neu, U., M. G. Akperov, N. Bellenbaum, R. Benestad, R. Blender, R. Caballero, A. Cocozza, H.F. Dacre, Y. Feng, K. Fraedrich, J. Grieger, S. Gulev, J. Hanley, T. Hewson, M. Inatsu, K. Keay, S.F. Kew, I. Kindem, G.C. Leckebusch, M.L.R. Liberato, P. Lionello, I.I. Mokhov, G. Pinto, C.C. Raible, M. Reale, I. Rudeva, M. Schuster, I. Simmonds, M. Sinclair, M. Sprenger, N.D. Tilinina, I.F. Trigo, S. Ulbrich, U. Ulbrich, X.L. Wang, and H. Wernli, 2013: IMILAST: A community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bull. American Meteor. Soc.*, **94**(4), 529–547.
- Ogi, M., D.G. Barber, and S. Rysgaard, 2016: The relationship between summer sea ice extent in Hudson Bay and the Arctic Ocean via the atmospheric circulation. *Atmos. Sci. Lett.*, **17**(11), 603–609.
- Ogi, M., K. Yamazaki, and Y. Tachibana, 2004: The summertime annular mode in the Northern Hemisphere and its linkage to the winter mode. *J. Geophys. Res. Atmos.*, **109**(20), D20114, 1–15. <http://doi.org/10.1029/2004JD004514>
- Ono, J., J. Inoue, A. Yamazaki, K. Dethloff, and H. Yamaguchi, 2016: The impact of radiosonde data on forecasting sea-ice distribution along the Northern Sea Route during an extremely developed cyclone, *J. Adv. Model. Earth Syst.*, **8**, 292–303, doi:10.1002/2015MS000552.
- Osborn, T.J. 2011: Winter 2009/2010 temperatures and a record-breaking North Atlantic Oscillation index. *Weather*, **66**, 19–21.
- Overland, J. E., K. Dethloff, J.A. Francis, R.J. Hall, E. Hanna, S.-J. Kim, J.A. Screen, T.G. Shepherd, and T. Vihma, 2016: Nonlinear response of mid-latitude weather to the changing Arctic. *Nature Climate Change*, **6**(11), 992–999.
- Park, Y.Y., R. Buizza, and M. Leutbecher, 2008: TIGGE: Preliminary results on comparing and combining ensembles. *Quart. J. Royal Meteor. Soc.*, **134**(637), 2029–2050.
- Parkinson, C.L., and J.C. Comiso, 2013: On the 2012 record low Arctic sea ice cover: Combined impact of preconditioning and an August storm, *Geophys. Res. Lett.*, **40**, 1356–1361.

- Pezza, A.B., A.R. Harun, and I. Simmonds, 2012: Climate links and recent extremes in antarctic sea ice, high-latitude cyclones, Southern Annular Mode and ENSO, *Clim. Dyn.*, **38**, 57–73.
- Prieto, R., B.D.Mcnoldy, S.R.Fulton, and W.H. Schubert, 2003: A classification of binary Tropical Cyclone – Like vortex interactions. *Mon. Wea. Rev.*, **131**, 2656–2666.
- Rinke, A., M. Maturilli, R.M. Graham, H. Matthes, D. Handorf, L. Cohen, S.R. Hudson, and J.C. Moore, 2017: Extreme cyclone events in the Arctic: Wintertime variability and trends. *Env. Res. Lett.*, **12(9)**. <http://doi.org/10.1088/1748-9326/aa7def>
- Robinson, D.A. 2016: Northern Hemisphere continental snow cover extent, in “State of the Climate in 2015”, (edited by J. Blunden, and D. S. Arndt), *Bull. American Meteor. Soc.*, **95(7)**, S22–S23, <https://doi.org/10.1175/2016BAMSSStateoftheClimate.1>
- Sato, K., J. Inoue, A. Yamazaki, J.-H. Kim, M. Maturilli, K. Dethloff, S.R. Hudson, and M.A. Granskog, 2017: Improved forecasts of winter weather extremes over midlatitudes with extra Arctic observations. *J. Geophys. Res. Oceans*, **122**, 775–787.
- Sato, K., J. Inoue, A. Yamazaki, J.-H. Kim, A. Makshtas, V. Kustov, M. Maturilli, and K. Dethloff, 2018: Impact on predictability of tropical and mid-latitude cyclones by extra Arctic observations. *Nature Scientific Reports*, **8(1)**, 12104.
- Screen, J.A., I. Simmonds, and K. Keay, 2011: Dramatic interannual changes of perennial Arctic sea ice linked to abnormal summer storm activity. *J. Geophys. Res. Atmos.*, **116(15)**, 1–10. <http://doi.org/10.1029/2011JD015847>
- Screen, J.A. 2017a: The missing Northern European winter cooling response to Arctic sea ice loss. *Nature Communications*, **8**, 1–9. <http://doi.org/10.1038/ncomms14603>
- Screen, J.A. 2017b: Simulated atmospheric response to regional and pan-arctic sea ice loss. *J. Clim.*, **30(11)**, 3945–3962.
- Screen, J.A. 2017c: Climate science: Far-flung effects of Arctic warming. *Nature Geosci.*, **10(4)**, D15105, 253–254. <http://doi.org/10.1038/ngeo2924>
- Semmler, T., L. Stulic, T. Jung, N. Tilinina, C. Campos, S. Gulev, and D. Koracin, 2016: Seasonal atmospheric responses to reduced Arctic sea ice in an ensemble of coupled model simulations. *J. Clim.*, **29**, 5893–5913.
- Sepp, M., and J. Jaagus, 2011: Changes in the activity and tracks of Arctic cyclones. *Climatic Change*, **105**, 577–595.
- Serreze, M.C., Barrett, A.P., 2008. The summer cyclone maximum over the central Arctic Ocean. *J. Clim.*, **21**, 1048–1065.
- Serreze, M.C., and J.A. Francis, 2006: The arctic amplification debate. *Climatic Change*, **76**, 241–264.

- Serreze, M.K., A.M. Lynch, M.P. Clark, 2001: The arctic frontal zone as seen in NCEP/NCAR reanalysis. *J. Clim.*, **14**, 1550–1567.
- Serreze, M.C., J. Stroeve, A.P. Barrett, and L.N. Boisvert, 2016: Summer atmospheric circulation anomalies over the Arctic Ocean and their influences on September sea ice extent: A cautionary tale. *J. Geophys. Res. Atmos.*, **121(19)**, 11463–11485.
- Shepherd, B.T.G., 2016: Effects of a warming Arctic. *Science*, **353(6303)**, 989–990.
- Simmonds, I., C. Burke, and K. Keay, 2008: Arctic climate change as manifest in cyclone behavior. *J. Clim.* **21**, 5777–5796.
- Simmonds, I., and K. Keay, 2009: Extraordinary September Arctic sea ice reductions and their relationships with storm behavior over 1979–2008. *Geophys. Res. Lett.*, **36**, L19715, doi:10.1029/2009GL039810.
- Simmonds, I., and I. Rudeva, 2012: The great Arctic cyclone of August 2012. *Geophys. Res. Lett.*, **39**, L23709, doi:10.1029/2012GL054259.
- Simmonds, I., and I. Rudeva, 2014: A comparison of tracking methods for extreme cyclones in the Arctic basin. *Tellus A: Dynamic Meteorology and Oceanography.*, **66**, 25252, doi: 10.3402/tellusa.v66.25252.
- Swinbank, R., M. Kyouda, P. Buchanan, L. Froude, T.M. Hamill, T.D. Hewson, J.H. Keller, M. Matsueda, J. Methven, F. Pappenberger, M. Scheuerer, H.A. Titley, L. Wilson, and M. Yamaguchi, 2016: The TIGGE project and its achievements. *Bull. American Meteor. Soc.*, **97**, 49–67.
- Tanaka, H.L., A. Yamagami, and S. Takahashi, 2012: The structure and behavior of the arctic cyclone analyzed by the JRA-25/JCDAS data. *Polar Science*. **6**, 54–69.
- Tao, W., J. Zhang, and X. Zhang, 2017a: The role of stratosphere vortex downward intrusion in a long-lasting late-summer Arctic storm. *Quart. J. Royal Meteor. Soc.*, **143**, 1953–1966.
- Tao, W., J. Zhang, Y. Fu, and X. Zhang, 2017b: Driving roles of tropospheric and stratospheric thermal anomalies in intensification and persistence of the Arctic superstorm in 2012. *Geophys. Res. Lett.*, **44**, 10017–10025. <http://doi.org/10.1002/2017GL074778>
- Thompson, D.W.J., and J.M. Wallace, 1998: The Arctic oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.*, **25(9)**, 1297–1300. <http://doi.org/10.1029/98GL00950>
- Welch, B.L. 1947. The generalization of ‘Student’s’ problem when several different population variances are involved. *Biometrika*. **34**, 28–35, doi: 10.2307/2332510.

- Wei, M.Z., Z. Toth, and Y. J. Zhu, 2010: Analysis differences and error variance estimates from multi-centre analysis data. *Australian Meteorological and Oceanographic Journal*, **59**, 25–34. <http://doi.org/10.1.1.223.110>
- Yamagami, A., M. Matsueda, H.L. Tanaka, 2017: Extreme Arctic Cyclone in August 2016. *Atmos. Sci. Lett.*, **18**, 307–314, doi: 10.1002/asl.757.
- Yamagami, A., M. Matsueda, H.L. Tanaka, 2018a: Predictability of the 2012 great Arctic cyclone on medium-range timescales. *Polar Sci.*, **15**, 13–23, doi: 10.1016/j.polar.2018.01.002.
- Yamagami, A., M. Matsueda, H.L. Tanaka, 2018b: Medium-range forecast skill for extraordinary Arctic cyclones in summer of 2008–2016. *Geophys. Res. Lett.*, **45**, 4429–4437, doi: 10.1029/2018GL077278.
- Yamaguchi, M., T. Nakazawa, and S. Hoshino, 2012: On the relative benefits of a multi-centre grand ensemble for tropical cyclone track prediction in the western North Pacific. *Quart. J. Royal Meteor. Soc.*, **138(669)**, 2019–2029.
- Yamazaki, A., J. Inoue, K. Dethloff, M. Maturilli, G. König-Langlo, 2015: Impact of radiosonde observations on forecasting summertime Arctic cyclone formation. *J. Geophys. Res. Atmos.*, **120**, doi:10.1002/2014JD022925.
- Yang, X.-Y., X. Yuan, and M. Ting, 2016: Dynamical link between the Barents-Kara Sea ice and the Arctic Oscillation. *J. Clim.*, **29**, 5103–5122. <http://doi.org/10.1175/JCLI-D-15-0669.1>
- Yoshimori, M., A. Abe-Ouchi, M. Watanabe, A. Oka, and T. Ogura, 2014. Robust seasonality of Arctic warming processes in two different versions of MIROC GCM. *J. Clim.*, **27**, 6358–6375.
- Zhang, J., R. Lindsay, A. Schweiger, and M. Steele, 2013a: The impact of an intense summer cyclone on 2012 Arctic sea ice retreat. *Geophys. Res. Lett.*, **40**, 720–726, doi:10.1002/grl.50190.
- Zheng, M., E.K.M. Chang, and B.A. Colle, 2013b: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting.*, **28**, 1133–1156.
- Zhang, X., J.E. Walsh, J. Zhang, U.S. Bhatt, and M. Ikeda, 2004: Climatology and interannual variability of Arctic cyclone activity: 1948–2002. *J. Clim.*, **17**, 2300–2317.