UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

IMPACT OF DATA ASSIMILATION AND MODEL PHYSICS ON THE PREDICTABILITY OF THE 2012 GREAT ARCTIC CYCLONE

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE IN METEOROLOGY

By

ZHIHONG CHEN
Norman, Oklahoma
2022

IMPACT OF DATA ASSIMILATION AND MODEL PHYSICS ON THE PREDICTABILITY OF THE 2012 GREAT ARCTIC CYCLONE

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY THE COMMITTEE CONSISTING OF

Dr. Xuguang Wang, Chair

Dr. Steven M. Cavallo

Dr. Aaron Johnson

© Copyright by ZHIHONG CHEN 2022
All Rights Reserved.

Acknowledgments

I would like to thank my advisors, Dr. Xuguang Wang and Dr. Aaron Johnson, for their guidance, advice and support through out three years of my research for this thesis. I would also like to thank my committee member Dr. Steven Cavallo, for his support and feedback at key stages of my research. I am extremely grateful to my fiancée, Taoyu Long, for her love and support over the past three years. This study was conducted with funding support from the Office of Naval Research. The computational resources for this study are provided by the OU Supercomputing Center for Education & Research.

TABLE OF CONTENTS

Abstract

A case study of the Great Arctic Cyclone 2012 (AC12) is used to understand the role of initial condition errors and model physics errors in the 2-3-day range predictability of high-impact summer Arctic Cyclones (ACs). The control forecasts initialized with analyses assimilating conventional in-situ observation demonstrate improved skills in predicting the peak intensity with less than 3 hPa difference from the verifying reanalysis. However, the forecasted AC12 reaches its peak intensity 18 hours earlier than in the verifying reanalysis and the cyclone track is biased towards the southwest. Using ensemble sensitivity analysis (ESA), the upstream trough, downstream ridge, and the tropopause polar vortex (TPV) to the northeast (NE TPV) of the AC12 are identified to be correlated with the deepening trend and the cyclone track error, but they are not well observed by current in-situ observation networks. Pseudo-observations are constructed from reanalysis and are added to the three features separately to study the impact of the initial condition error in each feature on the predictability of AC12. The cyclone deepening trend error and track error are greatly reduced when the initial condition is better constrained in either the NE TPV or the jet stream wind along the trough and ridge, as the former leads to a southward expansion of the NE TPV and the latter leads to successfully capturing of a shortwave trough at the 2PVU surface above the AC12. Varying the choice of model physics parameterization schemes does not further improve the cyclone track prediction. The cyclone intensity prediction is sensitive to the choice of longwave radiation schemes and planetary boundary layer (PBL) schemes. Varying longwave radiation schemes creates a large ensemble spread (~5 hPa) in the cyclone intensity prediction as the longwave cooling gradient near the tropopause affects the strength of TPVs. The Yonsei University (YSU) PBL scheme further improves the prediction of the deepening trend and sustained intensity of AC12, as it reduces the

PBL model physics error of the convection in the lower-tropopause front areas compared to the Mellor–Yamada–Nakanishi–Niino scheme widely used for Arctic. The choice of shortwave radiation schemes and microphysics schemes is found to have little impact on the predictability of AC12. These results have implications for further improvements in the intensity and track prediction of ACs at shorter time scales to better serve the growing human activity in the Arctic.

1. Introduction

The scientific interest in the predictability of Arctic Cyclones (ACs) is increasing in recent years. New shipping opportunities emerging in the Arctic due to decreased sea ice coverage require skillful AC prediction to avoid safety risks. The complex interactions between ACs and the sea ice on multiple time scales (e.g., Koyama et al. 2017, Luckovich et al. 2021, Clancy et al. 2022) also play a critical role in assessing the availability and risks of the ship routes. Moreover, the dynamical linkages between the Arctic and the midlatitude atmosphere have been revealed in recent studies that improving AC prediction may also lead to the practical predictability improvement in midlatitudes on weather time scales (e.g., Jung et al. 2014, Jung et al. 2015, Sato et al. 2018). The prediction of ACs faces two great challenges which are distinct from tropical and extratropical cyclone prediction. The poor coverage of observation networks in the Arctic presents a challenge to data assimilation (DA) to reduce the initial condition error (Yamazaki et al. 2015, Johnson and Wang 2021). Furthermore, understanding is relatively lacking of how errors in the physical process parameterization (Bromwich et al. 2009, Hines et al. 2011, Hines and Bromwich 2017) in the Arctic contribute to limited predictability of ACs, which presents a challenge to model development and ensemble design for AC prediction.

With the goal of improving the AC predictability, numerous efforts have been made to understand the mechanisms associated with the development and maintenance of summer ACs. Same as in extratropical cyclones, baroclinicity has been shown to play a key role in the genesis and evolution of Summer ACs (Inoue and Hori 2011, Tao et al. 2018). The Eurasian continent interior and the Arctic coastline, featuring strong temperature gradients between land and the ocean, known as the Arctic frontal Zone (AFZ), have been shown to be one of the main regions for AC genesis (Crawford and Serreze 2015, Crawford and Serreze 2016). However, ACs demonstrate a longer lifetime than extratropical cyclones (Tanaka et al. 2012, Aizawa and Tanaka 2016, Yamagami et al. 2016, Tao et al. 2018), implying ACs have unique dynamical properties from extratropical cyclones. Simmonds and Rudeva (2012) suggested a tropopause polar vortex (TPV) vertically stacked with the Great Arctic Cyclone 2012 (AC12) plays a strong role in the evolution of the storm. Aizawa and Tanaka (2016) revealed the vertical structure of long-lasting summer ACs in their mature stage, featuring the deep barotropic cyclonic circulation, downward intrusion of the stratosphere, coupled warm core in the stratosphere and cold core in the troposphere, suggesting the lifetime of ACs is extended by their coupling with TPVs. Tao et al. (2018) found the increased static stability created by vertical coupling with TPV strengthens and sustains AC12. Besides the increased stability, the merging of TPVs has also been shown to extend ACs' lifetime (Yamagami et al. 2016). TPV, as a PV anomaly at the jet stream level, may also trigger Rossby wave initiation events and lead to cyclone genesis (Röthlisberger et al. 2018, Johnson and Wang 2021). The intensification and sustaining mechanisms of TPVs have also been studied for their close interaction with ACs. Cavallo and Hakim (2009) applied the Ertel PV diagnostics to quantify the impact of the diabatic mechanisms on a TPV and showed that the cloud-top radiational cooling dominates the Ertel PV generation near the tropopause. Cavallo and Hakim (2013) demonstrated the water vapor gradient resulting from stratosphere dry air intrusion enhances the cloud-top radiative cooling and PV generation in the vicinity of the TPV. While these studies focus on features that explain AC evolution, there is a need to understand how errors in the initial condition or evolution of these features are contributing to AC forecast errors.

While physical understanding of the development of ACs progresses, previous studies on their predictability are still limited. Yamagami et al. (2018) studied the predictability of 10 high-impact summer ACs at medium-range timescales using ensemble forecasts from five operational prediction centers, showing that the operational ensemble forecasts can predict the cyclone center position error of less than 469km at a lead time of 2.5 days, which is a shorter lead time than tropical cyclones and extratropical cyclones. Yamagami et al. (2019) further examined the predictability of 26 high-impact summer ACs from 1986 to 2016 in Global Ensemble Forecast System, showing only at forecast lead times shorter than 3 days, most ensemble members are able to predict the formation of the cyclone, with <=433.1km track error and ~6.9 hPa center SLP error, suggesting the mid-range predictability of ACs are worse than extratropical cyclones.. The Great Arctic Cyclone 2012 (AC12), with its record-breaking center SLP and sea ice reduction, was studied by Yamagami et al. (2016). They found the predictability of AC12 is around 2-3 days with five operational ensemble forecast products, and successfully predicting the merging of two warm cores at 250 hPa is the common feature of top performance ensemble members. Their study also found that the cyclone tracks are commonly biased towards the southeast in all ECMWF and NCEP ensemble members, suggesting systematic errors are present in the IC and/or model physics at the 3-day lead time in these ensembles. While these studies have mainly emphasized predicting the formation of the ACs in the mid-range time scale >3 days, factors that impact the track and intensity prediction at shorter lead times have not been emphasized in previous studies

Further AC practical predictability improvements can come from either IC or physics improvements. Previous DA studies supported the potential to further improve AC forecasts by assimilating additional observations. Yamazaki et al. (2015) conducted an observing system experiment (OSE) to study the impact of an Arctic radiosonde observation on the predictability of AC12. Although the location of the extra observation is far from the cyclone center at the mature stage, the forecast initialized with the analysis assimilating the extra observation predicts a much stronger cyclone than the forecast initialized by the analysis without the extra observation. Further diagnostics shows that the reproduction of upper tropospheric circulation is improved and the location of the TPV leading to the cyclogenesis is better predicted with the extra observation. Johnson and Wang (2021) conducted an OSE of the July 2018 AC case and found that the denial of three well-placed temperature observations sampling the TPV and the upper-level jet degrades the TPV amplitude and its interaction with the waveguide. The degraded forecast of this interaction leads to a substantial increase of the cyclone track error. Both studies demonstrated that even a few additional observations can improve the prediction of ACs if they are well placed to sample dynamically relevant features. However, due to limited observation facilities in the Arctic, features potentially concerning the AC development, such as troughs, ridges, and TPVs, may not be well observed. To guide future observation network designs, further studies are needed to investigate the impact of the initial condition error on the AC development of each feature of interest in order to guide the decision on which feature should be prioritized for additional observations.

The studies of the impact of model physics on the predictability of ACs, are even more limited. Past studies focus on diagnosing and optimizing model physics in Numerical Weather Prediction (NWP) models to match the Arctic environment in general. Wilson et, al. (2012) examined Polar WRF (PWRF, details in Section 2) performance in forecasts of the atmospheric hydrological cycle in the Arctic, revealing an underprediction of cloud cover, which leads to an excess in incident shortwave radiation and longwave radiation. Hines and Bromwich (2016) improve the short-term cloud and radiation forecast in the Arctic by reducing specified liquid cloud droplet concentration in Morrison 2 microphysics scheme. These results suggest that model physics schemes and parameters performing well in the tropical or extratropical regions may not be the best option for the Arctic. It is also possible that model physics schemes and parameters performing well in forecasts of the Arctic in general may not be the best option for the AC prediction. Based on the previous studies of the mechanisms of the AC development and maintenance, the choice of longwave and shortwave radiation schemes can affect the diabatic cooling and heating near the tropopause, which in turn affects the intensity of TPVs and ACs. The choice of microphysics schemes has the potential to affect the AC structure by varying the cloud structure and latent heat patterns. The choice of Planetary Boundary Layer (PBL) schemes, which model the surface heat and moisture flux and convection differently (Hu et al. 2010), may be of great importance in improving the prediction of the baroclinic phase of the AC evolution. What's more, if initial condition is not well-constrained, the performance of model physics is also affected. For example, bias in moisture near the tropopause can lead to bias in longwave cooling, thus affects the TPV's intensity (Cavallo and Hakim, 2013). Therefore, there is a need to understand the sensitivity of AC forecast performance to different physical parameterization schemes on top of well-constrained initial conditions.

The present study aims to investigate the impact of initial condition error and model physics error on the predictability of the great AC12. This work complements the past studies in several aspects. First, this paper focuses on addressing the AC prediction at a short 2-to-3-day lead time. Second, an ensemble sensitivity analysis (ESA) is used to identify key features and processes affecting the prediction of AC12. By adding pseudo-observations (details in Section 2) into each key feature, the impact of initial condition error on the predictability of AC12 of each feature is isolated and studied. Third, sensitivity experiments are conducted for different physics parameterization schemes initialized with the best analysis created from virous pseudo-observation experiments to investigate whether and how different parameterization schemes can improve the AC12 prediction on top of reduced initial condition error.

The rest of this thesis is organized as follows. Sections 2 includes a description of the DA system and NWP models used for forecasting experiments, an overview of AC12, and experiment designs for the baseline forecast, pseudo-observation experiments, and model physics sensitivity experiments. The results of the experiments are discussed in Section 3. The main conclusions are summarized and suggestions for future work are provided in section 4.

2. Methods

2.1 Case Overview

AC12 is chosen for this study because of two reasons. AC12 is one of the strongest Arctic cyclones on the record since 1979 (Simmonds and Rudeva 2012), and the sea ice coverage in the Arctic broke the lowest record after its impact (Koyama et al. 2017). Understanding the predictability of AC12 will therefore benefit the prediction of extreme AC events and associated sea ice reduction. Secondly, the mechanisms critical to the formation of AC12 and its mid-range predictability have been relatively well studied compared to other AC cases (Yamagami et al. 2018), giving us precious insights when studying its 2-to-3-day predictability.

AC12 formed over Siberia on Aug 2nd according to Simmonds and Rudeva (2012). At 1200UTC 3Aug (fig 1a), it deepened below 995 hPa to the east of a trough and a TPV (‘SW TPV’ in later parts of the paper). To the northeast of AC12 on the east Siberian Sea, there was another AC vertically stacked with a TPV (‘NE TPV’ in the later parts of the paper). AC12 continued to progress to the northeast and reached the coastline with stronger troposphere baroclinicity while the other AC traveled to the west with the TPV above it. Meanwhile, a trough-ridge structure became well defined in the upper troposphere (fig 1b). At 1200UTC 5 Aug (fig 1c), the two ACs merged into one at the edge of the lowered tropopause, and the center SLP reached below 980 hPa, while NE TPV moved to the southwest and SW TPV moved to the east. The TPVs merged on the next day, which has been shown to significantly affect the predictability of AC12 (Yamagami et al. 2018). At 1200UTC 6 Aug (fig 1d), the north-moving AC12 was close to its maximum intensity (~962 hPa in GFS-ANL at 1800UTC 6Aug), and it was superimposed with the upper-level high PV air. AC12 nearly stopped moving for one day, maintaining its strength (fig 1e), before moving to the east and weakening. It dissipated on Aug 15th with a lifespan of 13 days.

2.2 Data Assimilation and Forecast System Configuration

For the DA system, the Gridpoint Statistical Interpolation (GSI)-based EnKF (Whitaker and Hamill 2002, Wang et al. 2013, Wang and Lei 2014) coupled with the Advanced Research version of the Weather Research and Forecasting model (Shamrock et al. 2008) with Polar modifications (PWRF, Hines and Bromwich 2008, Bromwich et al. 2009, Hines et al. 2011) version 4.1 is used. In-situ conventional surface and upper-air observations from the operational GTS data stream are assimilated. Satellite observations are not assimilated because of the presence of unique challenges related the effective satellite data retrieval in the Arctic (e.g., Kwok et al. 2009) that are not the focus of this work. Since conventional observations are sparse in the Arctic compared to mid-latitude and tropical regions, enhancing the understanding of the effect of conventional observations on the predictability of AC12 will potentially benefit future observation infrastructure development in the Arctic.

PWRF modifies WRF physics parameterization in three aspects. It improves the surface energy balance and heat transfer of the Noah Land Surface Model to fit the ice-covered surface in the Arctic (Hines et al. 2011). It reduces the specified liquid cloud droplet concentration in Morrison 2-moment (Morrison et al. 2009) microphysics to match the observed cloud structure in the Arctic (Hines and Bromwich 2017). PWRF also allows the sea ice quantities and associated land masks to update during the simulation, though in this preliminary study the sea ice remains constant as the interactions between ACs and sea ice are not the focus of this work.

Using the above systems, a 40-member cycled ensemble DA and forecast experiment is run from 0600UTC 3 Aug to 0000UTC Aug 8, covering the deepening, maximum intensity, and weakening of AC12 (fig 2a). Twenty of the ensemble members are initialized from NCEP Global Ensemble Forecast System analyses at 0600UTC 3 Aug and the other twenty members are from 6-hour ensemble forecasts of GEFS initialized at 0000UTC 3 Aug. 3-hour DA cycles from 0600UTC 3 Aug to 1200UTC 5 Aug are applied, with Relaxation to Prior Spread posterior inflation (Gaspari and Cohn 1999) set to 0.95, horizontal and vertical localization parameter (Whitaker and Hamill 2012) set to 1000km and 0.35 scale height units for moisture observations, 1.1 scale height units for other observations. This is the same configuration as in Johnson and Wang (2021). A deterministic forecast is initialized from the ensemble mean analysis every 12 hours starting from 1200UTC 3 Aug to 0000UTC 5 Aug, and ends at 0000UTC 8 Aug. The model domain consists of 371*371 grid points centered at the North Pole (Fig 2b) with 27-km spacing and 90 vertical levels are used with the model top set at 10 hPa. Following the Arctic System Reanalysis (Bromwich et al. 2018), the rapid radiative transfer model (RRTMG, Ianoco et al. 2008) is chosen for both longwave and shortwave radiation parameterization, Mellor–Yamada Nakanishi Niino (MYNN, Nakanishi and Niino 2009) Level 2.5 scheme is chosen for PBL parameterization, Kain-Fritsch Cumulus Potential Scheme (Berg et al. 2013) is chosen for cumulus parameterization, the modified Noah Land Surface Model (Hines et al. 2011) is chosen for land surface, and the modified Morrison 2-moment Scheme (Hines and Bromwich 2017) is chosen for microphysics parameterization.

This experiment is referred to as the Control experiment in later sections. The Control forecasts are verified against the Global Forecast System Model Analysis (GFS-ANL), which is considered as the reanalysis. However, the forecast skill of Control is not satisfactory since available rawinsonde observations does not cover the TPVs and ACs mentioned in Section 2.1. The detailed results are shown in Section 3.1.

2.3 Experiment Designs

2.3.1 Pseudo-observation Experiments and Ensemble Sensitivity Analysis (ESA)

The control experiment (Section 3.1) shows that the initial condition in synoptic features related to the evolution of AC12 is not well constrained because of the lack of in-situ conventional observations in the Arctic. Therefore, pseudo-observations are introduced to the features that lack observations. By adding pseudo-observations in different features to make them ‘well observed’, the impact of the initial condition error in each feature can be isolated and studied. This experiment is to address questions including what features need to be prioritized to be observed to improve short-term predictability of AC12, and how the improved initial condition in each feature benefits the prediction of AC12.

One pseudo-observation includes vertical profiles of pressure, geopotential height, virtual temperature, water vapor mixing ratio, and U V wind at one grid point in GFS-ANL, which are pre-processed by PWRF to have the exact same vertical levels as in ensemble mean analyses (an example is shown in the appendix). The six vertical profiles are then formatted to a rawinsonde observation with default observation error variance in GSI. This pseudo rawinsonde is then converted into PREPBUFR (Keyser 2013) format and assimilated into the analysis together with real observations. GFS-ANL is chosen because at various forecast lead times, PWRF forecasts initialized with GFS-ANL outperform PWRF forecasts initialized with the ensemble mean analysis in the Control, in both cyclone intensity and track prediction (not shown), which implies that GFS-ANL constrains the initial condition better than the control analysis. Both model difference between GFS and PWRF and data assimilation difference, including data assimilation algorithms and setting difference and difference in observations assimilated, are contributing to the advantage of GFS-ANL over the Control ensemble mean analysis.

ESA has been widely used to analyze the impact of initial conditions to the forecast (Ancell and Hakim 2007, Torn and Hakim 2008). For an ensemble of size M, the ensemble sensitivity is defined as the linear regression coefficient between a scalar forecast metric J and an analysis state variable x, where J and x are 1 × M ensemble estimates of the metric and the state (Torn and Hakim 2008):

$$\frac{\partial J}{\partial x} = \frac{Cov(J, x)}{Var\,(x)}$$

It reflects how much the forecast metric is likely to change given a unit increase of the state variable in the analysis. An ensemble forecast is run initialized with analyses assimilating real in-situ conventional observations at 1200UTC 4 Aug and ESA (Ancell and Hakim 2007, Torn and Hakim 2008) is used to identify features at the analysis time that are correlated with evaluation metrics, which include the cyclone track error at 1200UTC 6 Aug, and the time that cyclone reaches its peak intensity (referred as ‘phase’ in later sessions). Based on the result of the ESA, which is discussed in detail in Section 3.2, six pseudo-observation experiments are designed and shown in table 1.

2.3.2 Model Physics Sensitivity Experiments

The analysis with best forecasts of AC12 produced by assimilating different combinations of pseudo-observations is used to initialize a set of model physics sensitivity experiments. The experiments are to answer the questions including whether AC12 prediction can be further improved by tuning model physics schemes on top of the improved initial condition, what group of physics schemes create the largest spread in intensity and track forecast of AC12 and whether there are schemes outperforming others in AC prediction.

Four longwave radiation schemes are selected for their potential to affect TPV intensity through cloud-top cooling near the tropopause. RRTMG longwave scheme is the default. RRTM longwave scheme (Mlawer et al. 1997), Goddard longwave scheme (Chou et al. 2001) and CAM longwave scheme (Collins et al. 2004) are selected for the sensitivity experiments as they have been used in related studies for Arctic regions or other regions. Four shortwave radiation schemes are also selected for their difference in modeling the cloud top heating and surface heating. RRTMG shortwave Schemes is the default. Dudhia shortwave scheme (Dudhia 1989), Goddard shortwave scheme (Chou et al. 2001) and CAM shortwave scheme (Collins et al. 2004) are selected as they have been used in related studies for Arctic regions or other regions (e.g., Chandrasekar and Balaji 2012, Wilson 2012 et al). Three PBL schemes are selected for their differences in modeling surface fluxes and convection. MYNN Level 2.5 Schemes is the default. Mellor-Yamada-Janic scheme (MYJ, Janjic 1994) is selected as a representative of local closure schemes, which may not perform well when simulating convection (Hu et al. 2010). Yonsei University scheme (YSU) is selected as it has been proven to outperform MYNN in the previous study (Johnson and Wang 2021). Four microphysics schemes are selected for their differences in modeling cloud structure and latent heating, as the resulting difference in diabatic heating may affect the generation or destruction of the Ertel PV. Morrison 2 is the default. Purdue Lin scheme (Chen and Sun 2002) is selected as a representative of one-dimensional microphysics models. WRF Single-moment 6-class scheme (WSM6, Hong and Lim 2006) and Thompson scheme (Thompson et al. 2008) are selected as they have been used in related studies for Arctic regions (Wilson et al. 2012, Hines and Bromwich 2017). The impact of cumulus schemes and land surface schemes is beyond the scope of this study. The physics sensitivity experiment groups are summarized in Table 2.

3. Results

3.1 The Control Experiment

Figure 3 shows the cyclone center SLP and tracks of the forecasted AC12 in the five deterministic forecasts initialized from the ensemble mean analysis at 1200UTC 3 Aug, 0000 UTC 4 Aug, 1200UTC 4 Aug, and 0000UTC 5 Aug (referred to as ‘0312’, ‘0400’, ‘0412’, ‘0500’). 0312 predicts a minimum SLP around 960 hPa, with less than 1 hPa difference compared to the reanalysis. However, the forecasted AC12 in 0312 reaches its maximum intensity at around 0000UTC 6 Aug, which is 18 hours earlier than the reanalysis. The cyclone is forecasted to deepen rapidly after 0900UTC 4 Aug, whereas in the verifying reanalysis, the cyclone deepens less than 5 hPa in 24 hours from 0900UTC 4 Aug to 0900UTC 5 Aug. For the track prediction, 0312 is biased towards the southwest compared to reanalysis, which is consistent with the cyclone track forecasts of the GEFS and EMCWF ensembles studied in Yamagami et. al (2018). The AC12 in 0400, after assimilating four cycles of surface observations and one cycle of rawinsonde observations, is forecasted with almost the same SLP and track as in 0312. This reflects the limitation of the sparse conventional in-situ observation network, as only few rawinsonde observations are available near the cyclone-related features such as the upper-level trough and the TPVs (locations of rawinsondes are shown in Fig 4). In 0412, the AC12 is forecasted to be weaker than in 0312 and 0400 for 2 hPa. However, just as in the forecasts of previous cycles, it starts to rapidly deepen after initialization from analysis and reaches its maximum intensity at 0000UTC 6 Aug. The track shifts slightly to the northeast, showing improvement compared to 0312 and 0400. 0500 further shift the cyclone track to the northeast, demonstrating significant improvement compared to 0412, the center SLP at analysis time is also improved, but the AC12 still reaches its maximum intensity too early at 0000UTC 6 Aug.

While all four experiments have demonstrated good skills in predicting minimum center SLP prediction (<5 hPa error), two biases are consistently observed. Firstly, for the cyclone intensity prediction, experiments predict a faster cyclone deepening rate and shorter deepening time, implying the cyclone reaches its mature stage earlier than in the verifying reanalysis. Assimilating conventional observations only corrects the cyclone center SLP error at analysis time but cannot improve the deepening trends and the cyclone stage prediction. Secondly, for the track prediction, all groups are biased towards the southwest, though the DA process is fixing the track error cycle by cycle, with the most significant improvement in 0500. These two biases can be contributed by either the initial condition error or the model physics error, or both. Whether and how improved initial condition and model physics can help to reduce the two biases, will be the focus of the pseudo-observation experiments and model physics experiments.

The initial condition error is assumed to be the major contributor to the deepening trend bias and cyclone track bias because the coverage of conventional observations is poorer in the Arctic compared to the midlatitudes and tropics. Since previous studies (Yamazaki et al. 2015, Johnson and Wang 2021) have demonstrated that upper-air observations have a dominant effect on the cyclone evolution for a longer forecast lead time compared to lower levels, the 300 hPa geopotential height background error and analysis increments errors is examined for the four experiments from 0312 to 0500. The analysis at 1200UTC 3 Aug and 0000UTC 4 Aug shows little increment because the cyclone and TPVs are not covered by the available radiosonde observations (not shown). At 1200UTC 4 Aug, the trough axis in the background is negatively tilted (Fig 4a, black thick line) while in reanalysis the trough is neutrally tilted trough, indicated by the dipole on both sides of the trough axis. This implies the low pressure (AC12 in this case) associated with the trough is closer to its mature stage (Schultz et al. 2019) in the background than in the verifying reanalysis, which explains why the forecasted AC12 deepens faster and reaches its maximum intensity earlier. The dipole color shading near the ridge axis (Fig 4a, black dash line) indicates the initial condition error of the location of the ridge, with the background to the west and the reanalysis to the east. This difference in ridge locations is potentially related to the location difference of the 300 hPa cyclone center to the north of the ridge. The cyclone center, which is vertically stacked with the NE TPV (fig 1c), moves further to the west in the background than in the verifying reanalysis. The trough, the ridge, and the cyclone center are not well-observed at 1200UTC Aug 4, although there are a few rawinsonde observations (Fig 4b, blue dots) to the east of the trough, creating a small and localized analysis increment. However, as the trough propagates further east at 0000UTC 5 Aug, there are more available observations, and the tilting angle error of the tough is fixed (not shown), which can contribute to the track forecast improvement shown in Fig 3b. But the initial condition error at the 300 hPa cyclone center and the ridge grows larger as they are not observed, which could be the contributor to the deepening trend bias and unsolved track bias.

Since most upper-level features potentially related to the development of AC12 are not observed, the widely used data denial approach (Johnson and Wang 2021) to study the observations' impact on the forecast error is not a good fit for this case study. Instead, a pseudo-observation approach is proposed with details described in Section 2.3.1. Pseudo-observations is added at 1200UTC 4 Aug for two reasons. Firstly, the initial condition error grows larger at 1200UTC 4 Aug compared to at 1200UTC 3 Aug and 0000UTC 4 Aug. With larger initial condition errors comes greater improvement potential for pseudo-observations. Secondly, the initial condition error at 1200UTC 4 Aug can be separated into each feature including the trough, the ridge, the 300 hPa cyclone, etc., which allows us to study the impact of the initial condition error in each feature on the short-term predictability of AC12.

3.2 Pseudo-observation Experiments

3.2.1 Using ESA to Identify Critical Features

ESA is conducted using the cyclone track error at 48-h forecast lead time (T48, 1200UTC 6 Aug) and the time cyclone reaches its maximum intensity (referred to as 'phase') as forecast metrics following the equation defined in Torn and Hakim (2008). A 40-member ensemble forecast is initialized from ensemble analyses at 1200UTC 4 Aug described in Section 2.2. The linear regression coefficients between the 300 hPa geopotential field, zonal wind field, and meridional wind field and the two metrics (cyclone track error and phase) are calculated, and coefficients with a p-value smaller than 0.05 (i.e., the 95% confidence level) are plotted in Fig5, together with the initial condition error of the ensemble mean (Fig 5g-5i).

Three features are identified by the ESA to be of importance to the cyclone track and phase prediction, the NE TPV, the upstream trough, and the downstream ridge. The positive correlation between the initial geopotential height at 300 hPa and the cyclone track error in the vicinity of the upper-level cyclone center (also the NE TPV center) and the ridge to the northeast of it, suggests a stronger upper-level cyclone, or a stronger TPV, is correlated with less cyclone track error, while the negative correlation with phase suggests a stronger TPV correlating with extended cyclone deepening time. A less tilted trough is correlated with less track error, inferred from the sensitivity to the southeast and southwest of the V-wind maximum (Fig 5c). An east shift of the ridge axis correlates with reduced track error and extended cyclone deepening time, inferred from the sensitivity dipole at both sides of the ridge axis (Fig 5a) and the V-wind maximum (Fig 5c, 5f). The cyclone track sensitivity patterns of the 300 hPa wind field (Fig 5b, 5c) agree with the initial condition error patterns (Fig 5h, 5i) in all the three features identified to be important, suggesting reducing initial condition error in the 300 hPa wind field (Fig 5h, 5i) can plausibly lead to improved cyclone track prediction. This agreement leads to the design of an additional experiment group that only assimilates wind observation along the jet stream, aiming to isolate and study the impact of the initial condition error of the jet wind on the track and deepening trend prediction of AC12. The 500 hPa and 850 hPa sensitivity patterns mostly agree with 300 hPa and are shown in the appendix.

Pseudo-observations are added to the NE TPV and its northeast edge (Fig 5g, green dots), the trough (Fig 5g, cyan dots), and the ridge (Fig 5g, dark blue dots) based on the results of the ESA, covering the area that is not observed by available in-situ rawinsonde observations and has considerable initial condition error in either the geopotential height field or the wind field (Fig 5g-5i). The location of each pseudo-observation is subjectively picked, and the distance between two pseudo-observations is approximately 400 km, close to the distance between real rawinsonde observations along the Eurasian coastline. Three experiments (ROUGH, RIDGE and NETPV) including pseudo-observations in only one feature are designed to isolate and investigate the impact of the initial condition error on AC12 prediction in each feature. Two experiments (T&R and T&R_wind) combining the pseudo-observations in both the trough and the ridge are designed to study the impact of the jet stream on the predictability of AC12. And one additional experiment (ALL) assimilating all the pseudo-observations is designed to study the 'what-if' scenario in which all the synoptic features closely related to the development of the AC12 are well observed.

3.2.2 Pseudo-observation Experiments Results

The cyclone center SLP and track error for all pseudo-observation experiments are shown in Figure 6. At the analysis time, the initial condition error of cyclone center SLP and location is reduced for all experiments. AC12 is forecasted to have extended deepening time with peak intensity reached at 0600UTC 6 Aug compared to 0000UTC 6 Aug in Control with reduced track error. This verifies the ESA result that the trough, the ridge, and the NETPV are features closely correlated with the cyclone deepening trend prediction and track prediction. Among the three experiments with pseudo-observations in only one feature, THOUGH, RIDGE, and NETPV, NETPV (Fig 6c) outperforms the other groups (Fig 6a, 6b) as it predicts a slower deepening rate before 0000UTC 6 Aug and greater track error reduction, which further verifies the strong sensitivity in ESA in the vicinity of the NE TPV. T&R (Fig 6d) outperforms both Trough and Ridge but has a larger track error (around 300km) compared to NETPV (<300km) in most of the forecast lead time. However, if only wind pseudo-observations are assimilated, a weaker, slower-deepening AC12 is forecasted with track error reduced to around 200km, letting T&R_Wind be the best analysis created by assimilating pseudo-observations. The difference between T&R and T&R_wind will be discussed in detail in Section 3.2.3. The ALL experiment, assimilating pseudo-observations from all three features, demonstrates comparable intensity and track prediction skills as NETPV. This implies the NE TPV should be prioritized for additional observations, because once it is well-observed, additional observations in the tough and the ridge will not further benefit the forecast.

While the pseudo-ops impact on forecast track error varied greatly among the experiments, all experiments consistently showed a similar impact on the cyclone forecast deepening rate. All experiments show a maximum intensity around 0600 UTC 6 Aug while the control forecast reaches maximum intensity at 0000 UTC 6 Aug and the verifying reanalysis reaches maximum intensity around 1800 UTC 6 Aug. Diagnostics of the vertical profiles averaged over a 400km diameter circle with cyclone SLP minimum as the center reveals that the delayed superimposition of the AC12 with the TPV is the reason for the extended deepening time. As shown in Fig 7, all the experiments experience delayed stratosphere intrusion as the black thick line deepens to around 500 hPa later than the red thick line (the Control). This agrees with the reanalysis, in which the superimposition with high PV air also happens later than in the Control forecast (Fig 7a). The delayed superimposition leads to delayed tropopause level relative vorticity advection (Fig 7g, using NETPV as an example), resulting in the extended deepening of the surface cyclone. NETPV, T&R, and ALL delay the TPV superimposition more than TROUGH and RIDGE, which agrees with the predicted slower deepening trend shown in Fig 5. The geopotential height difference between experiments and the Control reveals that the magnitude difference of the surface cyclone originates from the tropopause level and propagates into the lower troposphere, reflecting the decisive role of the upper-level dynamics on the evolution of AC12. There is also a geopotential height difference in the 3-hour forecast for all experiments except for the NETPV, which is suspected to be the model imbalance caused by assimilating vertical profiles from another model, as the difference quickly dissipates in a longer forecast lead time.

In addition to the vertical profile difference above the cyclone center, the 2PVU-level potential temperature difference between experiments and the Control at 1200UTC 5 Aug is shown in Fig 8. In the Control forecast, the surface cyclone (the black circle) is already superimposed with the high PV air (black contours). In TROUGH (Fig 8b) and RIDGE (Fig 8c), the surface cyclone is right below the tight gradient of the potential temperature, while in NETPV (Fig 8d), T&R (Fig 8e) and ALL (Fig 8f), the surface cyclone is still to the east of the gradient of the potential temperature. These are in agree with Fig 7 that the superimposition process is further delayed in NETPV, T&R, and ALL compared to TROUGH and RIDGE. The three experiments with the more delayed superimposition are also the three experiments with less track error, implying that the deepening trend and the track prediction of AC12 are correlated, and the upper-level dynamics play the decisive roles in both.

Since it is already demonstrated that the delay of superimposition is the likely cause of the extended deepening, we further verify that the reduced error at the dynamic tropopause leads to reduced cyclone track error. The high PV region, including the NE TPV and SW TPV, is tracked and the RMSE is calculated. For each experiment, a grid point is accounted in the RMSE calculation if first, it is within a certain distance from the surface cyclone center, and second, its potential temperature is lower than 315K at 2PVU surface in either reanalysis or in the experiment. The RMSE of potential temperature is shown in Fig 9d with three TPV tracking examples (Fig 9a-9c) to illustrate the effectiveness of the tracking method – the grid points inside the outmost red contours (315K) are accounted for in the RMSE calculation. At analysis time, all groups, except for NETPV, have a larger initial condition error compared to the Control. Since all the groups except for NETPV contain observation of the jet stream near the trough or the ridge, the spatial correlation of the temperature between the TPVs and the jet stream may not be well captured by the ensemble spread. In NETPV experiment, observations inside the NE TPV reduce the initial condition error without involving the correlation between TPVs and the jet. It implies that assimilating observations inside a TPV better constrains the initial condition of the TPV, compared to assimilating observations outside of it. The RMSE of potential temperature grows larger with a longer forecast lead time for all experiments. TROUGH, RIDGE, and T&R have a 0.5K to 1K reduction of RMSE compared to Control. And at most forecast lead times, NETPV and ALL outperform other experiments, which is consistent with their better forecasting skills of the cyclone track. The linear regression (Fig 10c) between RMSE of potential temperature at 0000UTC 5 Aug (12-hour forecast lead time) with cyclone track error at 1200UTC 6 Aug (48-hour forecast lead time) further illustrates the positive correlation between the error in TPVs and the error of cyclone tracks.

Despite the positive correlation, it is still unclear which feature inside the high PV region, if better forecasted, leads to the cyclone track prediction improvement. Therefore, the ensemble correlations between the potential temperature at the 2PVU surface at T12 and the forecasted cyclone track error at T48 (Fig 11) are examined. The ensemble includes 6 deterministic forecasts, Control, TOUGH, RIDGE, NETPV, T&R, and ALL. While signals are present in various features including the trough and the SW TPV, the positive correlation between the trough and ridge axis is the strongest. The signal suggests that the colder potential temperature in this region is correlated with less track error. In the verifying reanalysis, (Fig 12a), a shortwave trough (red contours in the yellow square) is present at the same region of the strong positive correlation, while the Control (black contours) predicts much higher potential temperature than the reanalysis with the presence of a ridge axis at the region. Therefore, in order to explain the cyclone track prediction improvement, there is a need to examine how pseudo-observation experiments reduce the potential temperature error in the shortwave trough region.

For NETPV (Fig 12d), lower potential temperature is forecasted in the shortwave region at T12, but the shortwave structure is not captured. The NE TPV is forecasted to be less stretched in shape and expand further south compared to the Control, consistent with the reanalysis. The southward expansion of NE TPV is likely to be the cause of the potential temperature decrease in the shortwave region. The same is for RIDGE (Fig 12c), but the magnitude of potential temperature decrease is smaller compared to NETPV. For experiments with pseudo-observations in the upstream trough, the shortwave structure is better captured (the outmost red contours in Fig 12b,e,f). The initial condition error reduction in near the trough is likely to be the cause of the predicted shortwave trough. The NE TPV is also forecasted to shift to the southeast compared to Control. The potential temperature error reduction in TROUGH, T&R and ALL can be contributed to either the better capturing of the shortwave trough or the southeastward shift of the NE TPV, or both. Since NETPV outperforms other groups in track prediction, the initial condition error reduction with pseudo-observation in the NE TPV, which leads to the southward expansion of the NE TPV, is examined in this section. The initial condition error reduction that leads to the successful prediction of the shortwave trough will be studied in the next subsection.

By tracing the forecast difference between NETPV and Control from T12 to T0 (Fig13), it is verified that the vertical structure of the NE TPV is better constrained in analysis with pseudo-observations, which results in the southward expansion of the NETPV, leading to improvement of the cyclone track prediction. At the tropopause, the yellow squared lowered potential temperature in Fig 13d is hard to trace back, while a potential temperature decrease near the southwest edge of the NE TPV at the analysis time (Fig 13a) is still observed. The initial condition improvement of the temperature at 250 hPa is more outstanding compared to the 2PVU potential temperature. The southward expansion of the 250 hPa warmcore is observed at T12 (Fig 13h) and can be traced back to T0 (Fig 13e). The southward displacement of the warm front at the 850 hPa (Fig 13i-13l, yellow squared) is consistent with the southward expansion of the upper-level warmcore, leading to more east-west orientated temperature gradient contours compared to more northeast-southwest orientated in Control. The track of the developing cyclone, following the temperature contours, shifts more to the east in NETPV compared to the north in Control (Fig 12d, the yellow circle), resulting in reduced track error.

3.2.3. T&R_wind v.s. T&R

The T&R_wind experiment demonstrates the best forecast skills of the cyclone track (track error ~ 200km) among all experiments, including T&R. It would appear counterintuitive that T&R_wind outperforms T&R, which assimilates not only wind profiles from reanalysis but also temperature and moisture profiles. A careful investigation of this experiment is therefore conducted in this section to clarify this result. In particular, this discussion focuses on revealing what initial condition error is fixed that results in significant track prediction improvement, and why assimilating only wind from pseudo-observations creates a better analysis than assimilating wind, temperature, and moisture. It should be noted that, assimilating only wind in NE TPV (not shown) does not outperform NETPV, meaning that only temperature observations near the jet stream degrade the analysis.

According to the correlation calculated from the previous subsection (Fig 11), with the least track error, we expect the lowest potential temperature in the shortwave trough region. The T&R_wind forecasts a more established shortwave trough (Fig 14h, yellow dashed line) than T&R (Fig14d, and also than all other groups, Fig11) at T12 with a much lower potential temperature in that area, and the shortwave trough can be traced back to the analysis time (Fig 14e, yellow dashed line) to the east of the SW TPV. As the shortwave propagates faster than longwaves, the cyclone associated with it also moves further to the northeast, resulting in reduced cyclone track error.

To understand why T&R_wind captures the shortwave trough better than T&R, two hypotheses are made. Firstly, assimilating only pseudo wind observations inside the jet stream creates a better initial condition of the jet stream than assimilating both wind, temperature, and moisture, due to imperfect cross-variable correlation in the ensemble backgrounds. Secondly, since physically the jet streak is collocated with a strong potential temperature gradient at 2PVU, a better initial condition of the jet stream results in less initial condition error in the area with a strong potential temperature gradient, including the shortwave to the east of the SW TPV. Further diagnoses are conducted to verify the hypotheses.

Using the tracking method described in the previous section, the RMSE of the jet stream wind speed is calculated. Any grid with windspeed larger than 30 m/s at 2PVU surface (color shading in Fig 9a-9c) in either the reanalysis or in the analysis is counted into the calculation of the RMSE of jet stream windspeed. The result is shown in Fig 15. T&R_wind reduces the initial condition error in the jet wind speed by 0.7 m/s, while T&R increases the initial condition error by 0.3 m/s (Fig 15a). T&R_wind also outperforms T&R in almost all forecast lead times (the RMSE of windspeed grows wild as the track error of the cyclone increases, so it is only plotted to T24). This verifies the first hypothesis that assimilating temperature and moisture pseudo-observations degrades the initial condition of the jet stream compared to only assimilating the wind. It is to be noticed that the moisture observations have minimal impact on the 2PVU surface potential temperature (not shown) and thus will not be mentioned in later paragraphs.

A qualitative diagnosis of the jet stream wind and potential temperature gradient at 2PVU surface is conducted to verify the second hypothesis. At the analysis time, the jet wind speed to the east of the cyclone in T&R_wind (Fig 16i) is stronger than in T&R, along with temperature contours spread over the jet streak, agreeing more with the verifying reanalysis. However, in T&R, the temperature contours are denser at the west side of the jet stream (Fig 16e). This verifies the second hypothesis, that improving jet stream wind will also benefit the temperature gradient analysis. At T3, T6, and T12, compared to T&R, T&R_wind has better jet stream forecasts with stronger jet wind above the AC12, and the stronger jet stream wind is also the reason for the cyclone to move further to the northeast.

Besides increasing initial condition error of the jet stream wind, assimilating temperature observation near the jet stream is also likely to cause the degradation of the potential temperature gradient analysis through an imperfect spatial background error covariance of temperature across the edge of the TPV. At analysis time, assimilating both temperature and wind observations leads to an increased RMSE of potential temperature in the high PV region compared to only assimilating wind (Fig 15b). The increase in the RMSE itself does not necessarily mean the potential temperature gradient is not well captured. However, a detailed look at the potential temperature contours in various forecast lead times in Fig 16 suggests that the TPV centers are better constrained and forecasted in T&R. For example, at T0, the SW TPV center in T&R is constrained to the bottom of the trough (Fig 16e) as in the verifying reanalysis (Fig 16a), while the SW TPV center in T&R_wind is split into two weaker ones and located at further north (Fig 16i). At T12, the SW TPV in T&R and the SW TPV in the reanalysis are almost at the same locations (Fig 16d, 16h), while in T&R_wind it shifts to the north and becomes much weaker than the verifying reanalysis. T&R also forecasts a stronger NE TPV than T&R_wind (Fig 16f-16h, 16j-16l), which is more consistent with the reanalysis. Since T&R predicts the TPVs center locations and magnitudes better than T&R_wind, the location error of the edge of TPVs is more likely to be the major contributor to its higher RMSE.

A detailed investigation of the background error covariance further verifies that assimilating temperature observations near the jet stream can cause an initial condition error increase of both wind speed and potential temperature along the jet, while assimilating wind observations constrains the initial condition better. The covariance between temperature or windspeed at each pseudo-observation location and the temperature or wind speed of the whole domain at the 2PVU surface is examined and one pseudo-observation to the east of the SW TPV stands out (the green dot in Fig 17). Increasing the wind speed at its location is correlated with lowered temperature to the northwest of it, which results in the strong temperature gradient shifting to the east, capturing the shortwave trough. The background covariance of wind speed within the jet stream is also well captured (Fig 17d) as the positive correlation matches the shape of the wind speed error shown in Fig 17f. However, the covariance between the temperature at the observation location and the potential temperature and wind fields is not well captured. It could cause degradation of the analysis if temperature observation is assimilated. The positive correlation to the west and the negative correlation to the southwest of the observation (Fig 17a), does not match the negative potential temperature error spreading around at the southeast edge of the SW TPV (Fig 17e). It can lead to a westward shift of the temperature gradient, which could be the reason why T&R does not capture the shortwave trough. And the correlation between the potential temperature at the observation location and the wind speed field (Fig 17c) does not match the wind speed error pattern (Fig 17f) either, explaining the increased initial condition error of jet wind speed in Fig 16a. Whether the poor spatial correlation of the potential temperature across the edge of the TPV and the poor cross-variable correlation between the potential temperature and jet stream wind speed are just for this case or for more general scenarios needs further study to confirm, but attention is needed when assimilating temperature observation near the jet stream with a strong temperature gradient in future studies.

The above results give insights into what features should be prioritized for additional observations. With the improved initial condition of the jet wind, T&R_wind scores a lower

RMSE of both jet stream wind and potential temperature (Fig 16) than T&R throughout all forecast lead times. With the failure in forecasting the location of magnitude TPV centers, the RMSE reduction of TPV potential temperature in T&R_wind is likely contributed from the improved forecast of the edge of TPVs collocated with jet streaks. This implies the upper-level dynamics are driven more by the jet stream dynamics than TPV center location and magnitudes. The jet stream structure should thus be prioritized for additional observations to improve the short-term predictability of ACs

Section 3.3 Model Physics Sensitivity Experiments

Figure 18 shows the cyclone center SLP prediction and track error of model physics sensitivity experiments. Compared to improving the initial condition (Fig 6, black curve vs. cyan curve), varying model physics schemes have a negligible impact on cyclone track prediction (other curves vs. cyan curve). As for the deepening trend error, in all experiment groups, the forecasted AC12 reaches its maximum intensity at the same time, 0600UTC 6 Aug, except for YSU, which further extends the period of deepening for around 3 hours (Fig 18c). The sustaining of AC12 is also better predicted in YSU with the lowest Center SLP (~970 hPa) valid at 0000UTC 8 Aug. Initial condition error is therefore the main contributor to the forecasting error in the cyclone developing stages, while the model physics errors in the PBL, such as in the vertical mixing of heat, moisture, and momentum associated with turbulent eddies, play a minor role. For the cyclone intensity prediction, varying LW radiation schemes and PBL schemes creates a ~5 hPa spread on the minimum Center SLP of AC12, while for SW and MP schemes there is little spread. RRTMLW and MYJ overpredict the cyclone intensity and GoddardLW underpredicts the cyclone intensity. To summarize, the 2-to-3-day PWRF forecast is more sensitive to the choice of LW and PBL schemes, and less sensitive to SW and MP. The sensitivity of model physics on the related features is further examined to understand how model physics can affect AC prediction.

According to previous studies on the intensification and sustaining of TPV (Cavallo and Hakim, 2009, 2012, 2013), the strong radiative cooling gradient below the tropopause is the main contributor to the PV generation that sustains a TPV. Therefore, we examine the impact of diabatic cooling caused by the longwave radiation on the magnitude of the TPV for different schemes. Fig 19 shows the vertical profile of diabatic cooling due to longwave radiation changing with time-averaged near the AC center. To be noticed, only after 1200UTC 5 Aug the vertical profile captures the vertical structure TPV. However, the cloud top cooling intensity are in consistent before and after 1200UTC 5 Aug, for example, RRTM has stronger cooling compared to RRTMG at both 0000UTC 5 Aug and 1200UTC 6 Aug (Fig 19f). RRTM (Fig 19e) has the strongest cloud-top cooling due to longwave radiation and the strongest diabatic cooling gradient near the dynamic tropopause (the dipole patterns in Fig 19f). Goddard LW is less sensitive to the cloud top (Fig 19c), resulting in a weaker cooling gradient near the tropopause (Fig 19d). CAM has little difference with RRTM (Fig 19b). Fig 20a shows the mean potential temperature of the TPVs identified with the same tracking method of Fig 9 described in Section 3.2.2. The spatially averaged intensity of the TPV in reanalysis (the red curve) remains the same before 1200UTC 5 Aug, and slowly deepens after 1200UTC 5 Aug as it becomes vertically stacked with the surface cyclone. RRTM forecasts the strongest TPV while GLW forecasts the weakest, agreeing with the diabatic cooling gradient shown in Fig 19. With high PV air intruding further into the troposphere in a stronger TPV, the surface cyclone also gets stronger (Tao et al. 2018), which explains why RRTM predicts a much stronger AC12 than RRTMG and CAM, while GoddardLW predicts a weaker surface cyclone. Based only on this result it is not clear which scheme performs best or worst, as there is no observation of the cloud top cooling available. Even for GoddardLW, which unrealistically does not predict a cooling maximum at the cloud top, it predicts the deepening trend of the TPV better than other schemes in which the TPV deepens much faster than in reanalysis.

The same method described in the previous paragraph is applied to SW schemes. The difference in shortwave heating between schemes (Fig 21b,d,f) is of a much smaller magnitude (< 2K/day) than the difference in longwave cooling (~ 5K/day in Fig 18f). As a result, SW schemes create a much smaller spread in cyclone intensity than LW schemes in the 2-3 day forecast range. However, SW schemes do create larger spreads of cyclone intensity at a longer lead time (not shown), suggesting the model physics error in shortwave heating at the cloud top needs a longer time to accumulate and grow before having an impact on the AC.

For PBL schemes, the YSU scheme is found to further reduce the model physics error compared to MYNN and MYJ. From 1200UTC 4Aug to 1200UTC 5 Aug there is PV generation near the surface warm and cold front (horizontal plots not shown) due to latent heat release, PBL processes and friction. YSU predicts weaker PV generation (Fig 22d), reducing the model physics error (Fig 22f) while MYJ predicts stronger PV generation (Fig 22e). For an upper-level positive PV anomaly, upward motion to its downstream is expected (Bluestein 1992). As the high PV air superimposed with the AC12, the upward motion above the surface cyclone advects the high PV air from the surface to around 700 hPa (black thick lines in Fig22f), For MYNN, this results in a positive PV bias in the lower troposphere (blue shading in Fig 22f), YSU reduces this bias (blue shading in Fig 22d) while MYJ enlarges this bias. The choice of PBL also affects the temperature at the lower troposphere, especially in the frontal area. Fig 23 shows the temperature difference at 900 hPa at T36. Both YSU and MYJ predict a warmer cold front and warm front associated with the AC12 compared to MYNN. However, MYJ predicts a colder cyclone center while YSU predicts a warmer center, which agrees more with the verifying reanalysis (Fig 22j). In the relative vorticity perspective, YSU also predicts higher relative vorticity maximized at the surface persisting from T24 to the end of the forecast (Fig 22l), sustaining the intensity of the AC12. The connection between PBL parameterization formulation and the PV generation, temperature, and relative vorticity near the surface fronts is beyond the scope of this study. Future studies, including process-level observations, are needed to further explore the impact of the PBL parameterization on AC predictabilities.

The choice of microphysics schemes affects the cloud structure and precipitation near the center of AC12 (Fig 24). However, there is minimal sensitivity to the intensity and track prediction of AC12 observed for the four schemes in this experiment. The microphysics scheme impact is therefore not investigated further in the present study.

4. Conclusions

4.1 Summary

This study aims to understand the impact of initial condition and model physics errors on the track and deepening trend prediction of the AC12 on the 2-3-day time scale. Using EnKF to assimilate conventional in-situ observations, AC12 simulated by the control experiments reaches its peak intensity 18 hours earlier than in the verifying reanalysis and that the cyclone track is biased towards the southwest. Pseudo-observation experiments and model physics sensitivity experiments were designed to study initial condition error and model physics error contributions to the two primary forecast errors.

ESA using ensemble analyses at 1200UTC 4 Aug showed that reduced initial condition error in three synoptic features, including the upstream trough, the downstream ridge, and the NE TPV, is correlated with reduced cyclone track error and deepening trend error. However, an investigation of analysis increment in the control experiments shows that assimilating conventional observations did not reduce the initial condition error in the three features since those features are not covered by the rawinsonde observation network. To overcome the poor coverage of observations and study the impact of initial condition error in each feature of interest on the predictability of AC12, pseudo-observations were added to the three features separately at 1200UTC 4 Aug. As ESA suggested, the forecasted AC12 with analyses assimilating both real observations and pseudo-observations had reduced track error and reached its peak intensity 6 hours later than in the control experiment, more consistent with the reality. The delayed superimposition with the upper-level high PV air and resulted delayed vorticity advection correspond to the extended cyclone deepening time, suggesting the initial condition error at the upper troposphere contributes to the deepening trend forecast error. The positive correlation between RMSE of TPVs and the cyclone track error further revealed that the upper-level dynamics also govern the cyclone track prediction.

Assimilating pseudo-observation in the NE TPV reduced more initial condition error in the TPV and resulted in the best cyclone track error compared to assimilating pseudo-observation in the trough and the ridge. By tracing back the forecast error of the potential temperature at the 2PVU surface, the shortwave trough to the north of the AC12 at 0000UTC 5 Aug was found to be the most sensitive area of the cyclone track error. Although assimilating pseudo-observations in the NE TPV did not help to capture the shortwave trough, it reduced the forecast error in that area by predicting a southward expansion of the NE TPV. This result agrees with a previous study which found that the accurate prediction of the location and shape of the NE TPV is critical for predicting the AC12 track and magnitude (Yamagami et al, 2016). Tracing back to the analysis time, assimilating additional pseudo-observations in the NE TPV on top of the real conventional observations led to reduced track error in a stronger NE TPV that expands further southward and northeastward. Correspondingly at 850 hPa, the baroclinic zone was shifted further south, leading to the surface cyclone moving further to the east, resulting in improved cyclone track prediction. As the same synoptic pattern, an upstream trough with an SW TPV and a downstream ridge with a NE TPV, has been observed in many intense summer AC cases, such as the Aug 2016 case (Yamagami et al. 2017) and July 2018 case (Johnson and Wang 2021), the contribution of reducing initial condition error in the NE TPV to the AC predictability found in the AC12 case likely characterizes ACs in general, though more case studies are needed to verify this hypothesis.

An additional experiment assimilating pseudo-wind-observations along the jet stream outperformed the analysis with pseudo-observations in the NE TPV. Tracing back to the analysis time, it was the only analysis that captured the upper-level shortwave trough to the northeast of the surface cyclone. It is seemingly counter-intuitive that this analysis better constrains the shortwave trough than the analysis assimilating wind, temperature and moisture at the same locations. However, the RMSE of jet wind at the 2PVU surface revealed that while assimilating wind reduced the initial condition error, assimilating temperature (moisture does not have much impact) increased the error, implying that the cross-variable covariance between temperature and the wind was not well captured along the jet by the ensemble. Assimilating temperature also increased the RMSE of potential temperature near the TPVs, implying the spatial covariance of temperature across the edge of the TPV was not well captured either. This result suggests that cautions are needed when assimilating temperature observations along the jet stream for future studies. Compared to assimilating pseudo-observations in the NE TPV, assimilating jet wind increased the RMSE of TPVs at analysis time but reduced it at longer forecast lead times. Given that the intensity and center location of TPVs were not well captured when assimilating jet wind, it is speculated that a better constraint of the initial condition at the edge of TPVs with strong potential temperature gradients benefits the predictability of the AC more than the strength and the center locations of TPVs, and that the vertical wind profile, which reflects the temperature gradient, should be prioritized for additional observations along the jet stream.

On top of reduced initial condition errors, varying model physics schemes did not lead to substantial improvement of the cyclone track prediction, though some physics configurations did have an impact on the cyclone intensity prediction. Longwave radiation schemes and PBL schemes created the largest spread in peak intensity of the AC12 while shortwave radiation and microphysics schemes had little impact. Different longwave radiation schemes created different cloud-top cooling patterns, which affect the PV generation near the tropopause, leading to intensity differences in TPVs and the AC12, with RRTM Longwave Scheme predicting the strongest TPVs and AC12 and Goddard Longwave Scheme predicting the weakest. The shortwave heating gradients near the tropopause were affected by the choice of the shortwave radiation schemes. However, the shortwave heating difference was of a much smaller magnitude than the longwave cooling difference, implying a longer forecast lead time is needed for the shortwave schemes to create a spread in the cyclone intensity. On the choice of PBL schemes, YSU outperforms MYJ and MYNN as it better simulates the convection near the fronts. YSU is also the only scheme that can further improve the cyclone deepening trend prediction and cyclone sustaining on top of reduced initial condition error. The choice of microphysics schemes had little impact on the strength of the AC12, but it did create spread in cloud structure and precipitation, which could be critical to studies of the interaction between the sea ice and ACs.

4.2 Limitations and Future work

Since this study only considers the AC12 case, the major findings in this study are subjected to further verification with other cases. The validity of the pseudo-observation observation approach also needs further quantification, as it borrows the idea of the observing system simulation experiment (OSSE) (Zeng et al. 2020) but draws ‘observations’ from a different model (GFS in this case) with different dynamics cores and model physics, and signs of model imbalance due to this model difference were found during the assimilation process. Targeting the features identified by this OSSE-like framework in future field campaigns on multiple cases is recommended to further refine the quantitative impacts and their variation across different cases.

On the data assimilation side, there are several opportunities for future studies. Besides conventional observations, satellite observations and sea ice observations have the potential to further reduce initial condition error in the features of interest. Employment of advanced DA algorithms, such as MLGETKF (Wang et al. 2021) and 4DEnVar (Wang and Lei 2014) may also help better constrain the initial condition given a certain set of observations, resulting in improvement of the AC predictability.

On the model physics side, different radiation schemes created spreads in TPV and cyclone intensity, but it is hard to decide whether one scheme outperforms others unless more observations of radiation at the tropopause level are available. Field campaign observations aimed at validating model physics processes are therefore also recommended. A detailed investigation of PBL schemes is needed to reveal how different PBL parameterizations affect the convection near the fronts of ACs. The choice of microphysics schemes should be of concern for sea-ice-related studies of ACs given they create a spread in precipitation. For ensemble prediction employing perturbed physics ensembles, microphysics schemes need to be involved for their potential interaction with radiation schemes and PBL schemes.

References

Aizawa, T., H. L. Tanaka, and M. Satoh, 2014: Rapid development of arctic cyclone in June 2008 simulated by the cloud resolving global model NICAM. Meteor. Atmos. Phys., 126, 105–117, https://doi.org/10.1007/s00703-013-0272-6.

——, ——, 2016. Axisymmetric structure of the long lasting summer Arctic cyclones. Pol. Sci. 10, 192–198.

Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. Mon. Wea. Rev., 135, 4117–4134, https://doi.org/ 10.1175/2007MWR1904.1.

Bauer, P., L. Magnusson, J.N. Thépaut and T.M. Hamill, 2016 Aspects of ECMWF model performance in polar areas. Quarterly Journal of the Royal Meteorological Society, 142, 583– 596.

Berg, L.K., W.I. Gustafson, E.I. Kassianov, E.I., and L. Deng, 2013: Evaluation of a modified scheme for shallow convection: Implementation of CuP and case studies. Mon. Wea. Rev., 141, 134-147. doi:10.1175/MWR-D-12-000136.1

Bluestein, H. B., 1992. Synoptic-dynamic meteorology in midlatitudes: Volume 1, principles of kinematics and dynamics.

Bromwich, D. H., K. M. Hines, and L.-S. Bai, 2009: Development and testing of Polar Weather Research and Forecasting model: 2. Arctic Ocean. J. Geophys. Res., 114, D08122, https://doi.org/10.1029/2008JD010300.

——, A. Wilson, L. Bai, Z. Liu, M. Barlage, C. Shih, S. Maldonado, K. Hines, S.-H. Wang, J. Woollen, B. Kuo, H. Lin, T. Wee, M. Serreze, and J. Walsh, 2018: The Arctic System

Reanalysis Version 2. Bull. Amer. Meteor. Soc., 99, 805-828, doi: 10.1175/BAMS-D-16-0215.1. Full Text (PDF)

Cavallo, S. M., and G. J. Hakim, 2009: Potential vorticity diagnosis of a tropopause polar cyclone. Mon. Wea. Rev., 137, 1358–1371, https://doi.org/10.1175/2008MWR2670.1.

——, and ——, 2013: Physical mechanisms of Tropopause Polar Vortex intensity change. J. Atmos. Sci., 70, 3359–3373, https://doi.org/10.1175/JAS-D-13-088.1.

Chandrasekar, R., C. Balaji, 2012: Sensitivity of tropical cyclone Jal simulations to physics parameterizations. J Earth Syst Sci 121, 923–946. https://doi.org/10.1007/s12040-012-0212-8

Chen, S.-H. and W.-Y. Sun, 2002: A one-dimensional time dependent cloud model. J. Meteor. Soc. Japan., 80(1), 99–118. doi:10.2151/jmsj.80.99

Chou, M. D., M. J. Suarez, X. Z. Liang, and M. M. H. Yan, 2001: A thermal infrared radiation parameterization for atmospheric studies. NASA Tech. Memo., 104606, 19, 68 pp.

Clancy, R., C. Bitz, E. Blanchard-Wrigglesworth, M. McGraw, and S. Cavallo, 2022: A cyclone-centered perspective on the drivers of asymmetric patterns in the atmosphere and sea ice during Arctic cyclones. J. Climate, 35, 73–89, https://doi.org/10.1175/JCLI-D-21-0093.1.

Collins, W. D., et al., 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Tech. Note NCAR/TN–464+STR. 214 pp.

Crawford, A.D., Serreze, M.C., 2016. Does the summer Arctic frontal zone influence Arctic Ocean cyclone activity? J. Clim. 29, 4977–4993.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two–dimensional model. J. Atmos. Sci., 46, 3077–3107. doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2

Hines, K. M., and D. H. Bromwich, 2008: Development and testing of polar Weather Research and Forecasting (WRF) Model. Part I: Greenland ice sheet meteorology. Mon. Wea. Rev., 136, 1971-1989, https://doi.org/10.1175/2007MWR2112.1.

——, and ——, 2017: Simulation of late summer Arctic clouds during ASCOS with Polar WRF. Mon. Wea. Rev., 145, 521–541, https://doi.org/10.1175/MWR-D-16-0079.1.

——, ——, L. Bai, M. Barlage, and A. G. Slater, 2011: Development and testing of polar WRF. Part III: Arctic land. J. Climate, 24, 26–48, https://doi.org/10.1175/2010JCLI3460.1.

Hong, S.-Y., and J.–O. J. Lim, 2006: The WRF single–moment 6–class microphysics scheme (WSM6). J. Korean Meteor. Soc., 42, 129–151.

Hu, X. M., J. W. Nielsen-Gammon, and F. Zhang (2010), Evaluation of three planetary boundary layer schemes in the WRF model, J. Appl. Meteorol. Climatol., 49(9), 1831– 1844, doi:10.1175/2010JAMC2432.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long–lived greenhouse gases: Calculations with the AER radiative transfer models. J. Geophys. Res., 113, D13103. doi:10.1029/2008JD009944

Inoue, J., Hori, M.E., 2011. Arctic cyclogenesis at the marginal ice zone: a contributory mechanism for the temperature amplification? Geophys. Res. Lett. 38, 1–6. https://doi.org/10.1029/2011GL047696. L12502.

Janjic, Zavisa I., 1994: The Step–Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon. Wea. Rev., 122, 927–945. doi:10.1175/1520-0493(1994)122%3c0927:TSMECM%3e2.0.CO;2

Johnson, A., and X. Wang, 2021: Observation impact study of an Arctic cyclone associated with a TPV-induced Rossby wave initiation event. Mon. Wea. Rev., 149, 1577-1591, https://doi.org/10.1175/MWR-D-20-0285.1.

Jung, T., M. A. Kasper, T. Semmler, and S. Serrar, 2014: Arctic influence on subseasonal midlatitude prediction. Geophys. Res. Lett., 41, 3676–3680, https://doi.org/10.1002/2014GL059961.

——, and Coauthors, 2015: Polar lower-latitude linkages and their role in weather and climate prediction. Bull. Amer. Meteor. Soc., 96, ES197–ES200, https://doi.org/10.1175/BAMS-D-15-00121.1.

Kain, John S., 2004: The Kain–Fritsch convective parameterization: An update. J. Appl. Meteor., 43, 170–181. doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Keyser, D, 2013: PREPBUFR processing at NCEP. National Centers for Environmental Prediction Tech. Rep.[Available online at http://www. emc. ncep. noaa. gov/mmb/data_processing/prepbufr. doc/document. htm.].

Koyama, T., J. Stroeve, J. Cassano and A. Crawford, 2017: Sea ice loss and Arctic cyclone activity from 1979 to 2014. Journal of Climate, 30(12), 4735-4754.

Kwok, R., G. F. Cunningham, M. Wensnahan, I. Rigor, H. J. Zwally, and D. Yi, 2009: Thinning and volume loss of the arctic ocean sea ice cover: 2003-2008, J. Geophys. Res., 114, C0700

Lukovich, J. V., J.C. Stroeve, A. Crawford, L. Hamilton, M. Tsamados, H. Heorton, and F. Massonnet, 2021: Summer Extreme Cyclone Impacts on Arctic Sea Ice, J. Climate, 34, 4817–4834, https://doi.org/10.1175/JCLI-D-19-0925.1

Melia, N., K. Haines, and E. Hawkins, 2016: Sea ice decline and 21st century trans-Arctic shipping routes. Geophysical Research Letters, 43, 9720– 9728. https://doi.org/10.1002/2016GL069315.

Mlawer, Eli. J., S.J. Taubman, P.D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated–k model for the longwave. J. Geophys. Res., 102, 16663–16682. doi:10.1029/97JD00237

Morrison, H., G. Thompson, V. Tatarskii, 2009: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One– and Two–Moment Schemes. Mon. Wea. Rev., 137, 991–1007. doi:10.1175/2008MWR2556.1

Nakanishi, M., and H. Niino, 2009: Development of an improved turbulence closure model for the atmospheric boundary layer. J. Meteor. Soc. Japan, 87, 895–912. doi:10.2151/jmsj.87.895

Niu, G–Y., Z–L. Yang, K.E. Mitchell, F. Chen, M.B. Ek, M. Barlage, A. Kumar, K. Manning, D. Niyogi, E. Rosero, M. Tewari, Y. Xia, 2011: The community Noah land surface model with multiparameterization options (Noah–MP): 1. Model description and evaluation with local–scale measurements. J. Geophys. Res., 116, D12109. doi:10.1029/2010JD015139

Röthlisberger, M., O. Martius, and H. Wernli, 2018: Northern Hemisphere Rossby wave initiation events on the extratropical jet—A climatological analysis. J. Climate, 31, 743–760, https://doi.org/10.1175/JCLI-D-17-0346.1.

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. Sci. Rep., 8, 12104,

Schultz, D. M., L.F. Bosart, B.A. Colle, H.C. Davies, C. Dearden, D. Keyser, O. Martius, P.J. Roebber, W.J. Steenburgh, H. Volkert, and A.C. Winters, 2019: Extratropical Cyclones: A Century of Research on Meteorology's Centerpiece, Meteorol. Monogr., 59, 16.1–16.56.

Simmonds, I., and I. Rudeva, 2012: The great Arctic cyclone of August 2012. Geophys. Res. Lett., 39, L230709, https://doi.org/10.1029/2012GL054259.

Shamrock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-4751STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Tanaka, H. L., A. Yamagami, and S. Takahashi, 2012: The structure and behavior of the arctic cyclone in summer analyzed by the JRA-25/JCDAS data. Polar Sci., 6, 55–69, https://doi.org/10.1016/j.polar.2012.03.001.

Tao, W., J. Zhang, and X. Zhang, 2017: The role of stratosphere vortex downward intrusion in a long-lasting late-summer Arctic storm, Q. J. Roy. Meteor. Soc., 143, 1953–1966, https://doi.org/10.1002/qj.3055

Thompson, G., P. Field, R.M. Rasmussen, W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. Mon. Wea. Rev., 136, 5095–5115. doi:10.1175/2008MWR2387.1

Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. Mon. Wea. Rev., 136, 663–677, https://doi.org/10.1175/2007MWR2132.1.

Wang, X., H. G. Chipilski*, C. H. Bishop, E. Satterfield, N. Baker, J. Whittaker, 2020: A Multiscale Local Gain Form Ensemble Transform Kalman Filter (MLGETKF). Mon. Wea. Rev., 149, 605-622, https://doi.org/10.1175/MWR-D-20-0290.1.

——, and T. Lei, 2014: GSI-based four-dimensional ensemble-variational (4DEnsVar) data assimilation: Formulation and single-resolution experiments with real data for NCEP Global Forecast System. Mon. Wea. Rev., 142, 3303–3325, https://doi.org/10.1175/MWR-D-13-00303.1.

——, D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVarbased ensemble-variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. Mon. Wea. Rev., 141, 4098–4117, https://doi.org/10.1175/MWR-D-12- 00141.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. Mon. Wea. Rev., 130,

——, and ——, 2012: Evaluating methods to account for system errors in ensemble data assimilation. Mon. Wea. Rev., 140, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1.1913–1924, https://doi.org/10.1175/15200493(2002)130,1913:EDAWPO.2.0.CO;2

Wilson, A. B., D. H. Bromwich, and K. M. Hines, 2012: Evaluation of polar WRF forecasts on the Arctic system reanalysis domain: 2. Atmospheric hydrologic cycle. J. Geophys. Res., 117, D04107, https://doi.org/10.1029/2011JD016765.

Yamagami, A., M. Matsueda, and H. L. Tanaka, 2017: Extreme arctic cyclone in August 2016. Atmos. Sci. Lett., 18, 307–314, https://doi.org/10.1002/asl.757.

——, ——, and ——. 2018, Predictability of the 2012 great Arctic cyclone on medium-range timescales. Polar Science, 15, 13– 23. https://doi.org/10.1016/j.polar.2018.01.002.

——, ——, and ——,2019, Skill of medium-range reforecast for summertime extraordinary Arctic cyclones in 1986–2016. Polar Science, 20, 107– 116.

——, ——, and ——, 2018: Medium-range forecast skill for extraordinary Arctic cyclones in summer of 2008-2016. Geophys. Res. Lett., 45, 4429–4437, https://doi.org/10.1029/2018GL077278.

Yamazaki, A., J. Inoue, K. Dethloff, M. Maturilli, and G. König-Langlo, 2015: Impact of radiosonde observations on forecasting summertime Arctic cyclone formation. J. Geophys. Res. Atmos., 120, 3249–3273, https://doi.org/10.1002/2014JD022925.

Zeng, X., and Coauthors, 2020: Use of observing system simulation experiments in the United States. Bull. Amer. Meteor. Soc., 101, E1427–E1438, https://doi.org/10.1175/BAMS-D-19-0155.1.

Zhang, X., Walsh, J. E., Zhang, J., Bhatt, U. S., and Ikeda, M.: Climatology and interannual variability of Arctic cyclone activity: 1948–2002, J. Climate, 17, 2300–2317, https://doi.org/10.1175/1520-0442(2004)017<2300:CAIVOA>2.0.CO;2

Tables

Table 1. Experiment groups in with pseudo-observations

Name	Pseudo-observation locations
Control	No pseudo-observation.
TROUGH	Pseudo-observations along the jet stream near the trough to the west of AC12
RIDGE	Pseudo-observations along the jet stream near the ridge to the northeast of AC12
NETPV	Pseudo-observations on the north edge of NE TPV and in the NE TPV.
T&R	Containing pseudo-observations in TROUGH and RIDGE.
T&R_wind	Containing pseudo-observations in TROUGH and RIDGE, without temperature and moisture information.
ALL	Containing pseudo-observations in TROUGH, RIDGE, and NE TPV.

Table 2. Physics sensitivity experiment groups, the default schemes, and sensitivity experiment names.

Groups Name	Default	Experiment names
Longwave Radiation (LW)	RRTMG LW	RRTM, Goddard LW, CAM LW
Shortwave Radiation (SW)	RRTMG SW	Dudhia, Goddard SW, CAM SW
PBL	MYNN	MYJ, YSU
Microphysics (MP)	Morrison2	LIN, WSM6, Thompson

Figures

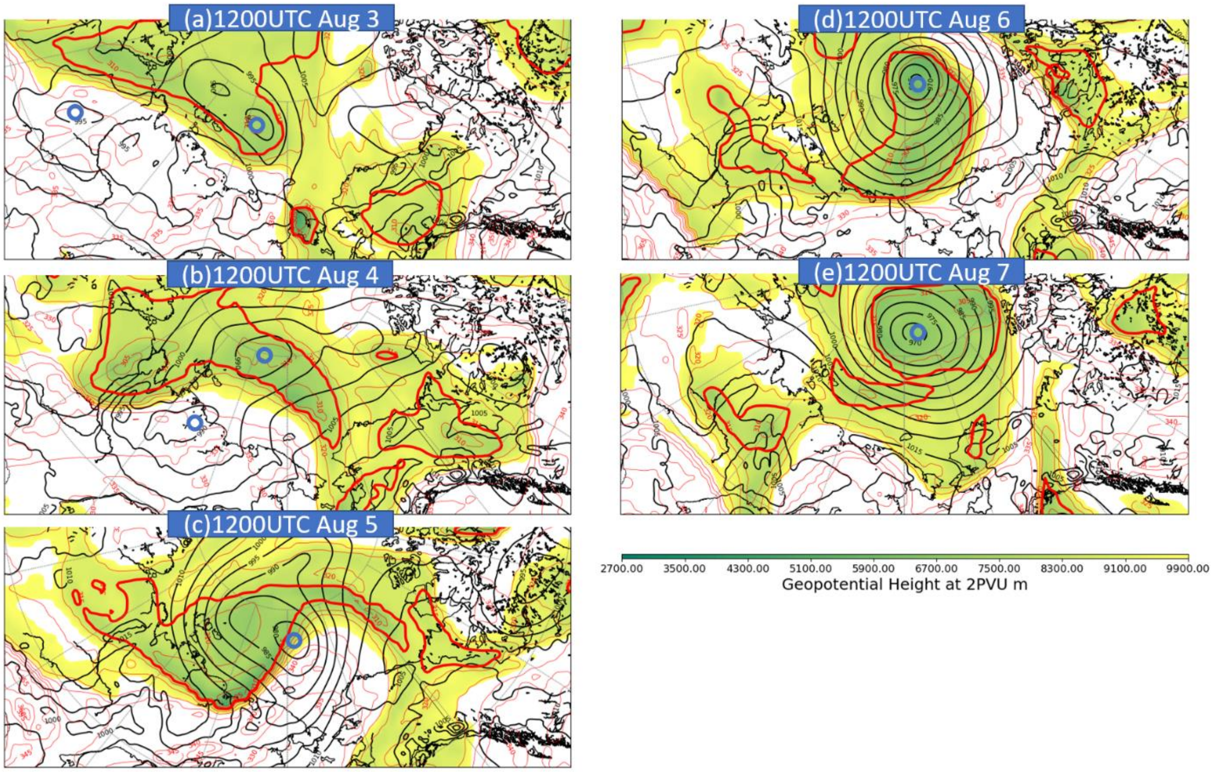


Fig1. Potential temperature (K, red contours), geopotential height (m, color filling) and Sea Level Pressure (SLP, hPa, black contours) of AC12 from a)1200UTC Aug3 to e)1200UTC Aug7 from GFS-ANL. Blue cycles show the location of surface cyclones. The thick red contours show the area with potential temperature at 2 PVU lower than 315K.

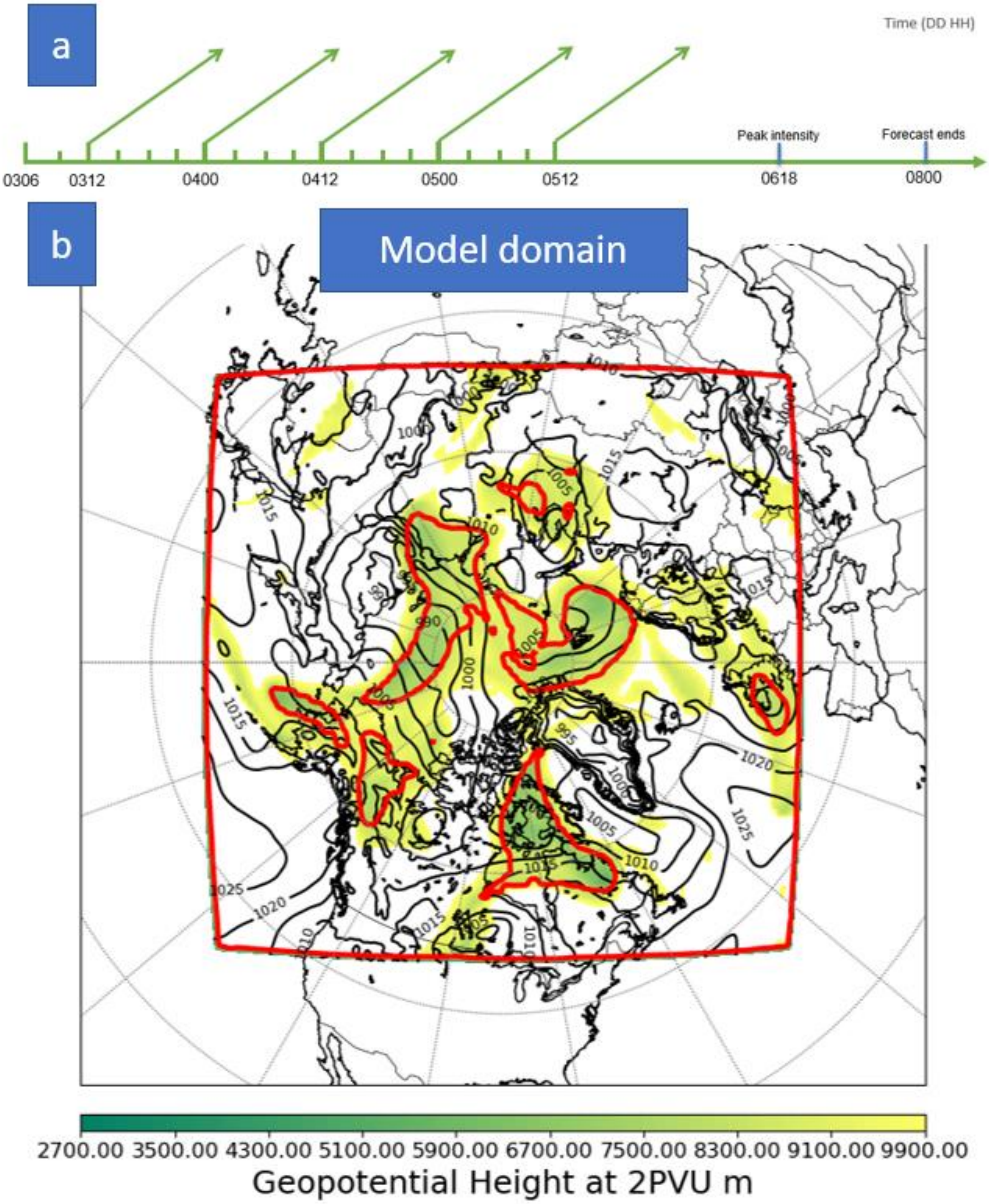


Fig 2. The experiment design of the control experiment. a) Initialization time and end time of the control forecasts. b) Same as in figure 1, at 1200UTC 4 Aug. The outmost red contour shows the boundary of the model domain.

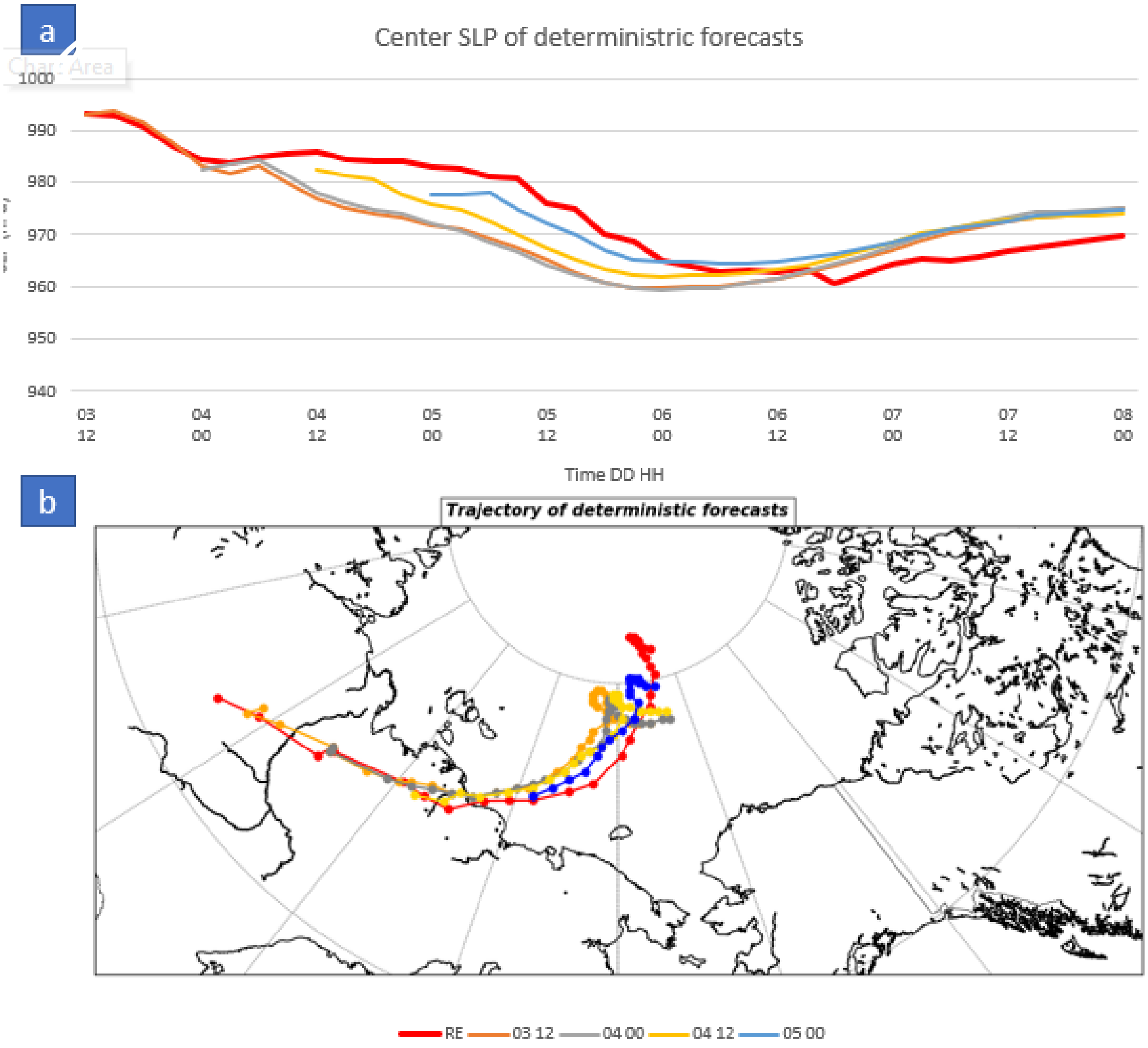


Figure 3. Cyclone center SLP(a) and tracks(b) of five deterministic forecasts initialized from the ensemble mean analysis assimilating conventional surface and upper-air observations.

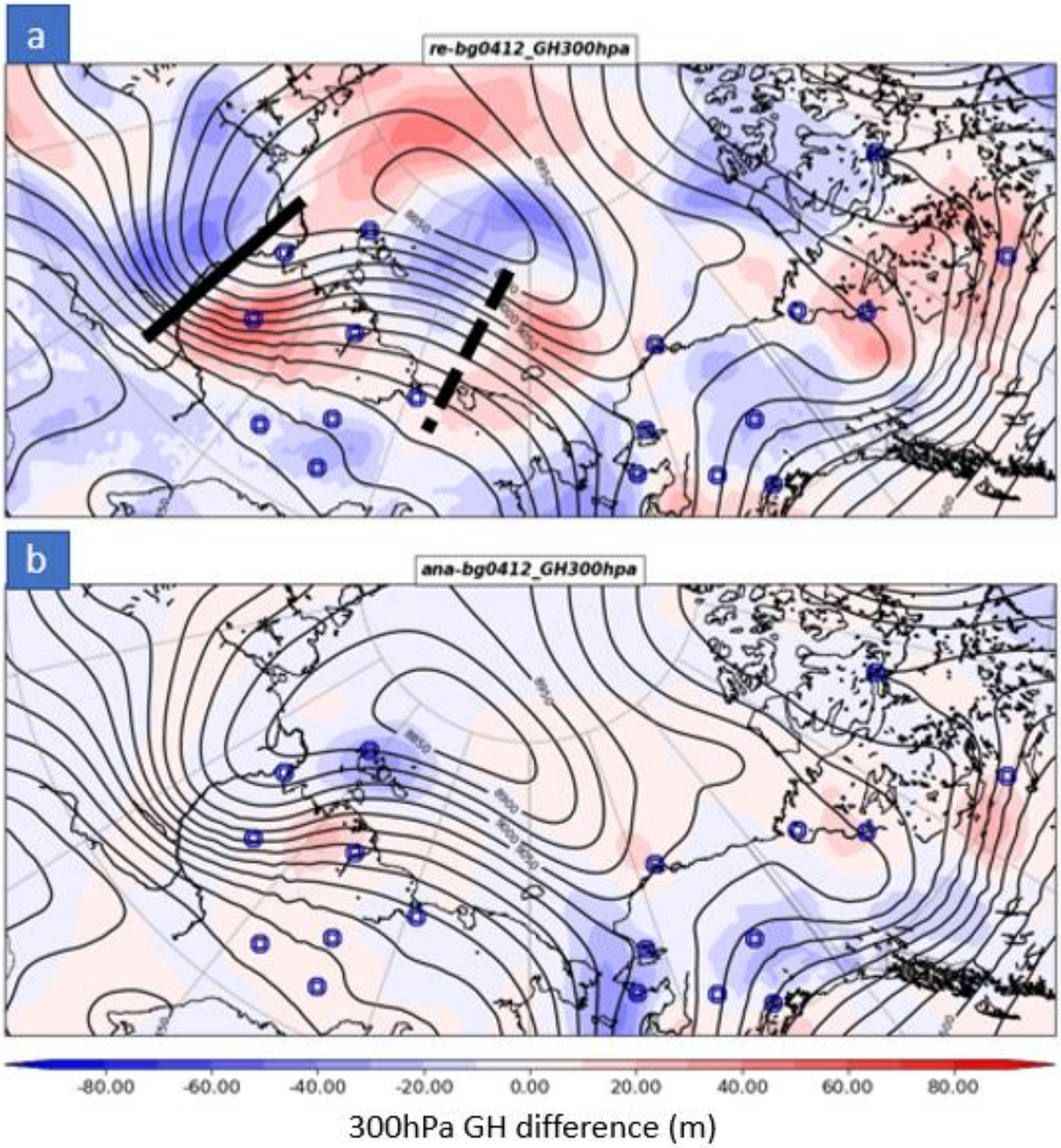


Figure 4. Background error (a) and analysis increment (b) at 1200UTC 4 Aug of 300hPa geopotential height (m, color shading). Black contours are geopotential height contours (m) of the background. The black thick lines are the location of the trough axis in the background, and the black dash lines are the location of the ridge axis in the background (d). Blue dots are locations of rawinsonde observations.

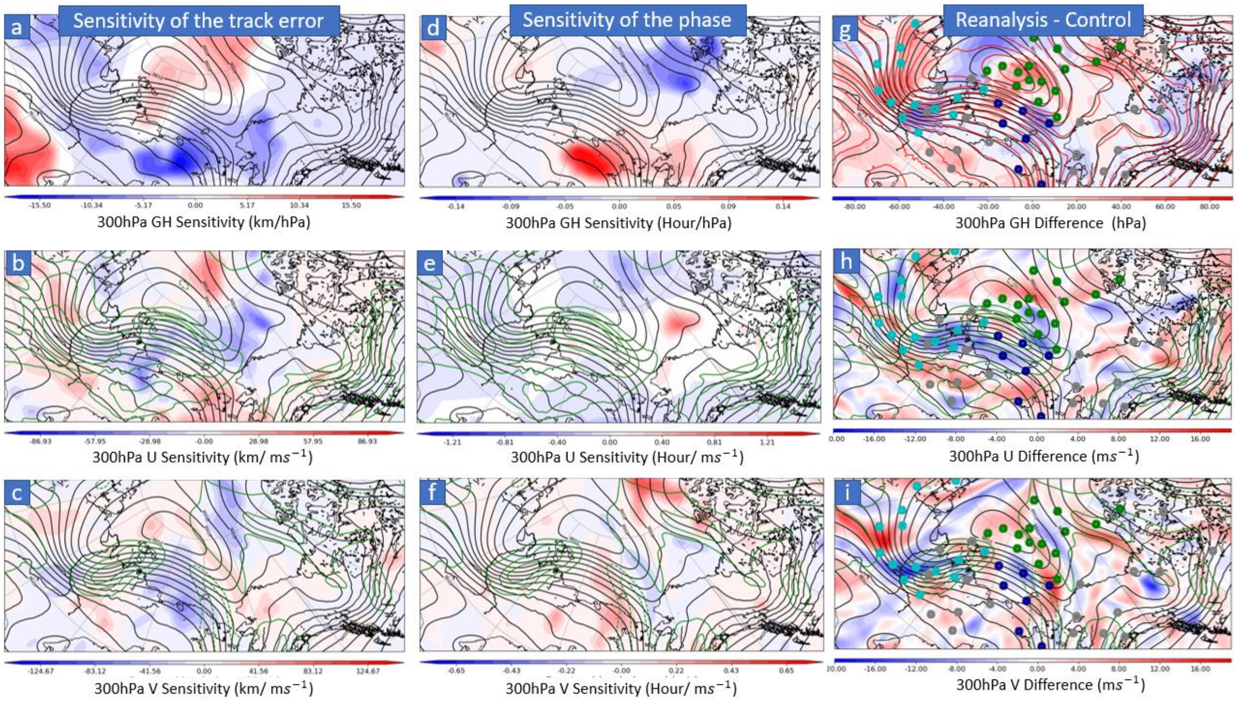


Fig 5. The ensemble sensitivity analysis (a-f) and initial condition error (g-i). Color shading are the sensitivity with p-value <0.05 (a-f) and the difference between the Control analysis and the valid reanalysis of 300 hPa geopotential height (g), U wind speed (h) and V wind speed (i). The black contours show 300 hPa geopotential height (m). The green contours show 300 hPa U or V wind speed (m/s). Grey dots in g) to i) are locations of in-situ rawinsonde observations. Cyan dots are locations of pseudo-observations of TROUGH, dark blue dots are of RIDGE, and green dots are of NETPV.

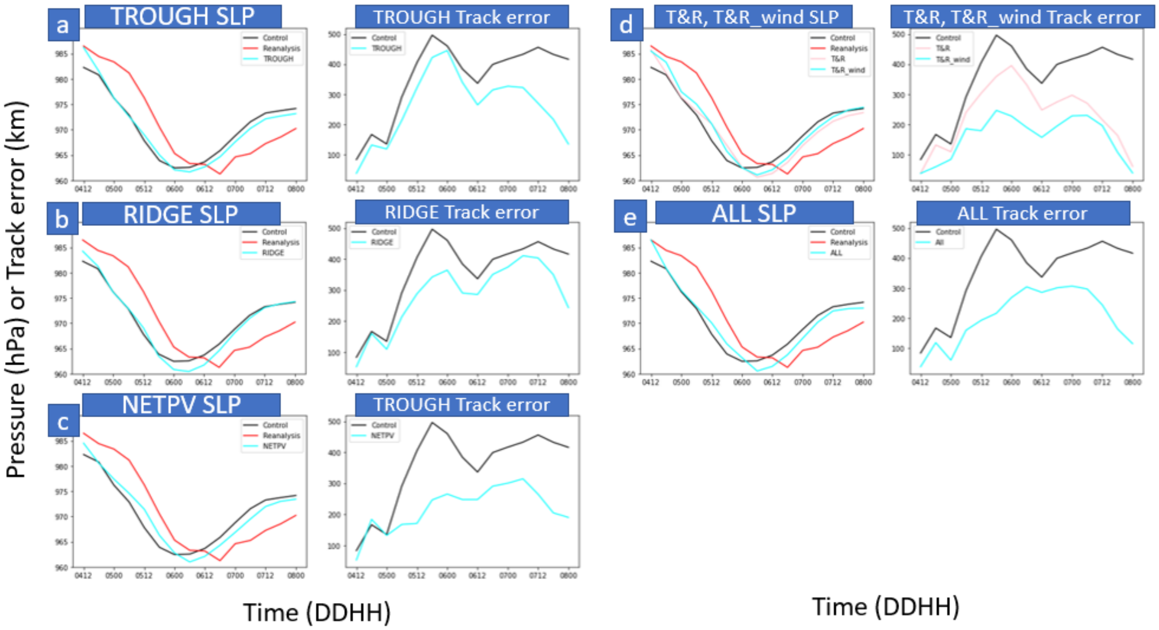


Fig 6. Cyclone center SLP forecasts and the forecasted track error of pseudo-observation experiments.

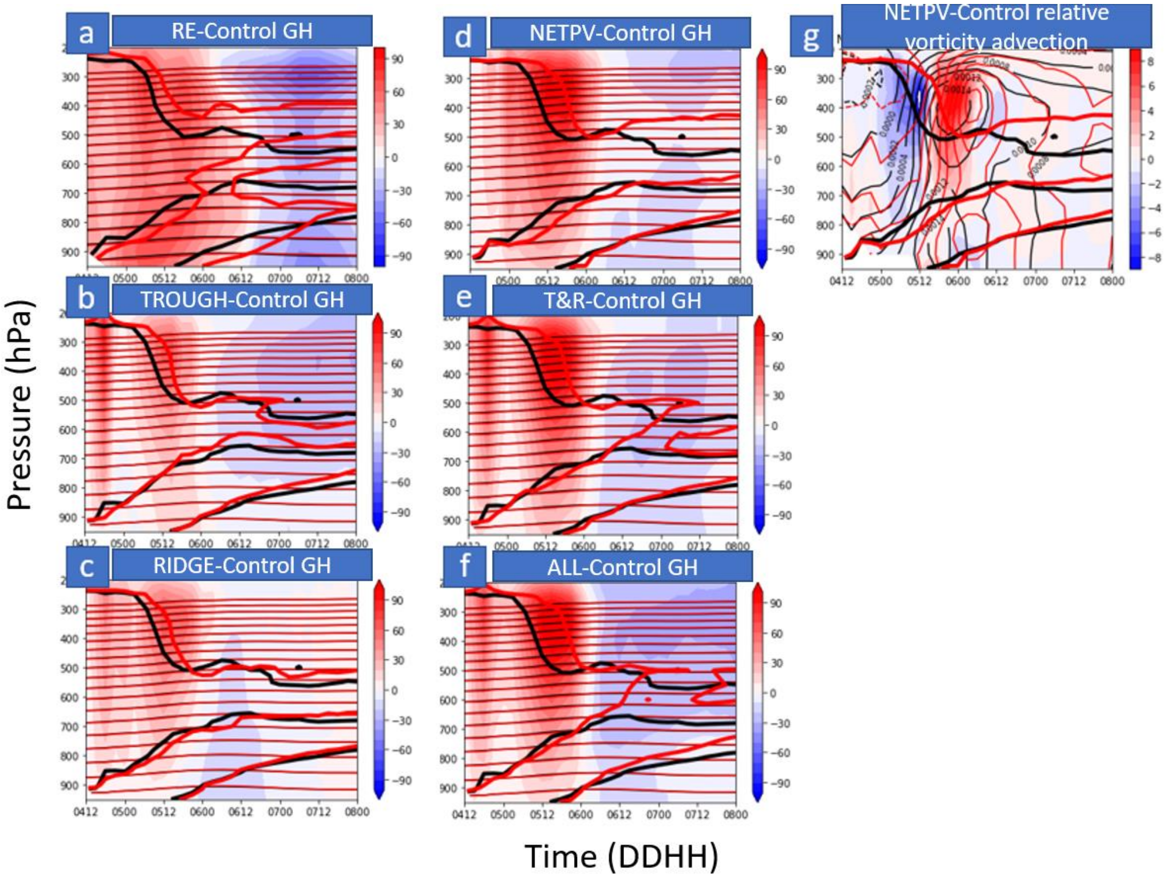


Figure 7. (a) The mean vertical profile with in 200km radius of cyclone center of the geopotential height difference (hPa, color shading) between the reanalysis (red thin contours) and Control (black thin contours). (b)-(f) The geopotential height difference between the pseudo-observation experiments (red thin contours) and Control (black thin contours). (g) The relative vorticity advection difference ($10^{-9}\ s^{-2}$,color shading) between NETPV and Control (red contours: relative vorticity in NETPV, black contours: relative vorticity in Control). Thick lines are the 2PVU surface

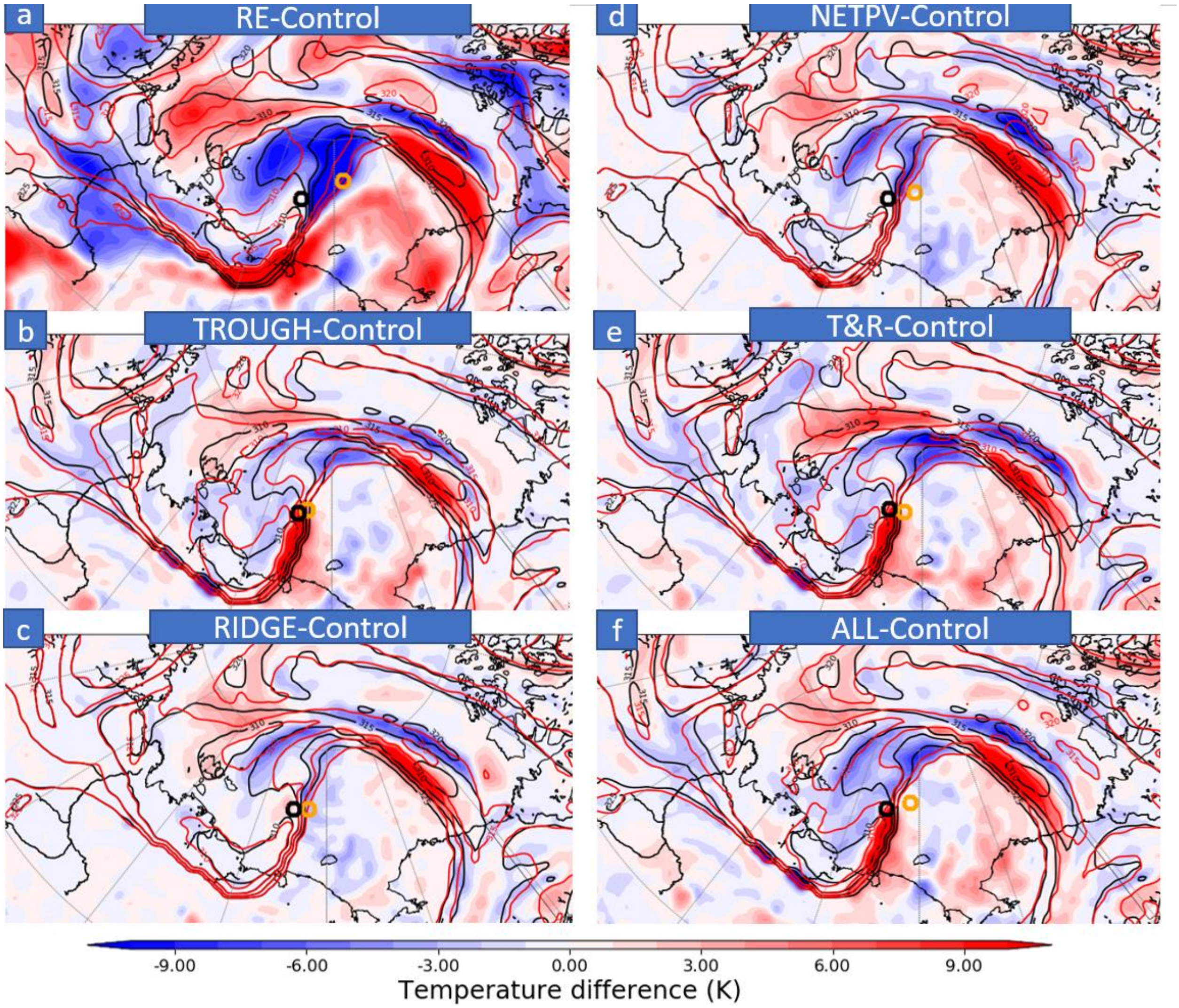


Figure 8. The potential temperature difference on the 2PVU surface between the reanalysis (K, red contours) and Control (black contours) (a) and between the pseudo-observation experiments (red contours) and the Control (black contours) at 1200UTC 5 Aug (b-d). Orange cycles are the cyclone center location of AC12 in the reanalysis (a) and pseudo-observation experiments (b-f), black cycles are the cyclone center location of the Control forecast.

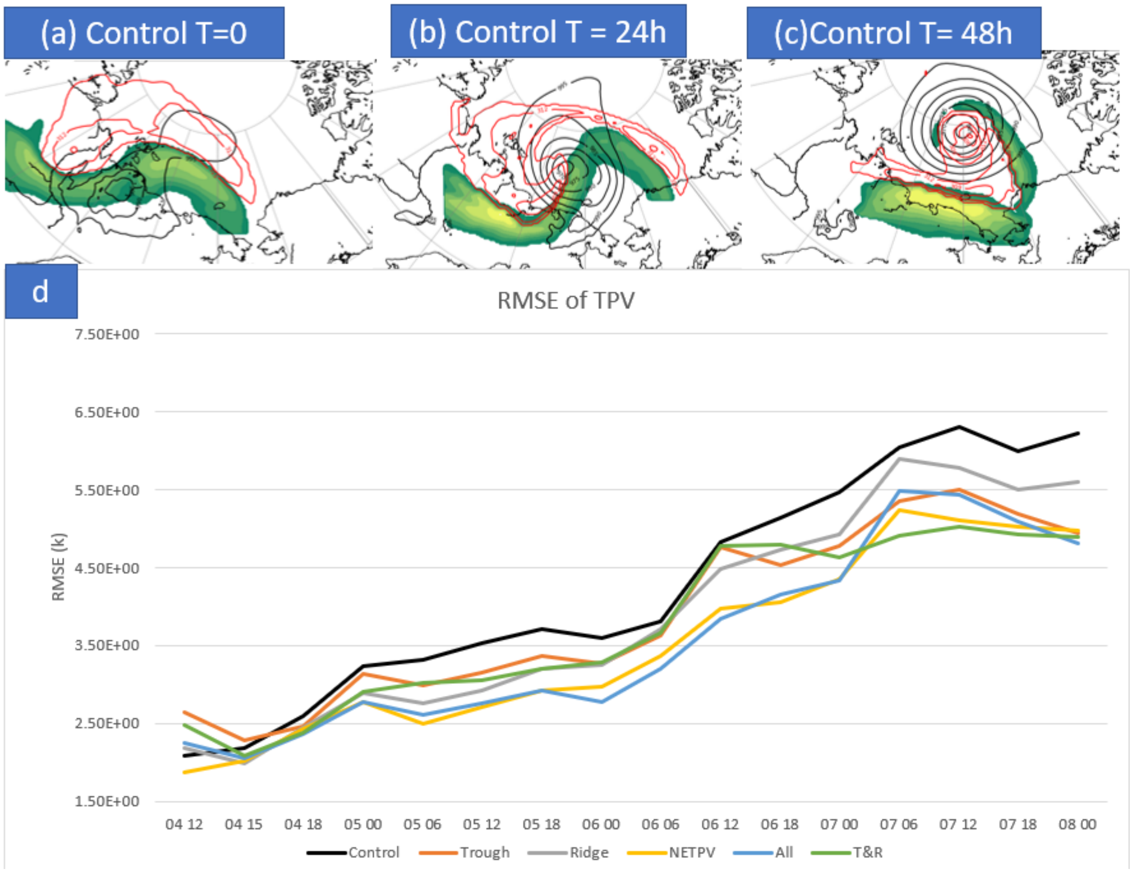


Figure 9. The RMSE of TPVs (d) and three examples of the TPV tracking method (a-c). The grids circled by the outmost red contours (315 K contours of the potential temperature at the 2PVU surface) are taken into account when calculating the RMSE of TPVs (a-c), the grids covered by the color shading (windspeed larger than 30 m/s at the 2PVU surface) are taken into account when calculating the RMSE of jet stream windspeed (a-c).

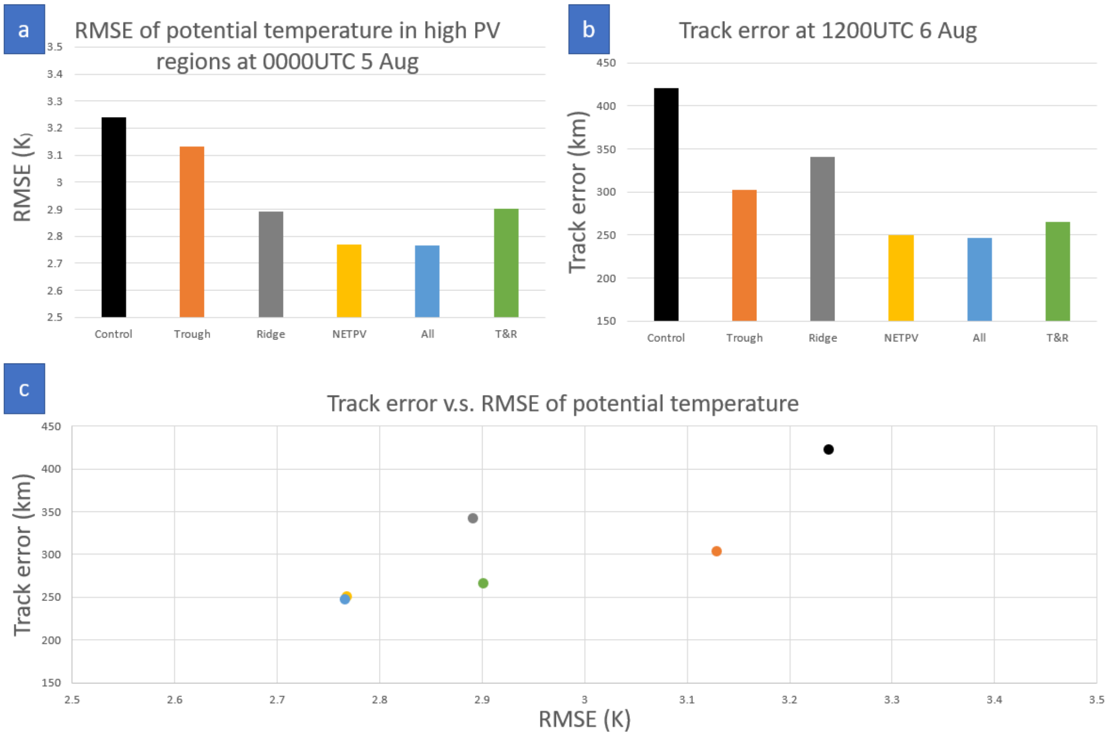


Figure 10. (a)The RMSE of potential temperature in the high PV region (b) The cyclone track error at T48 (c) The scatter plot between the two.

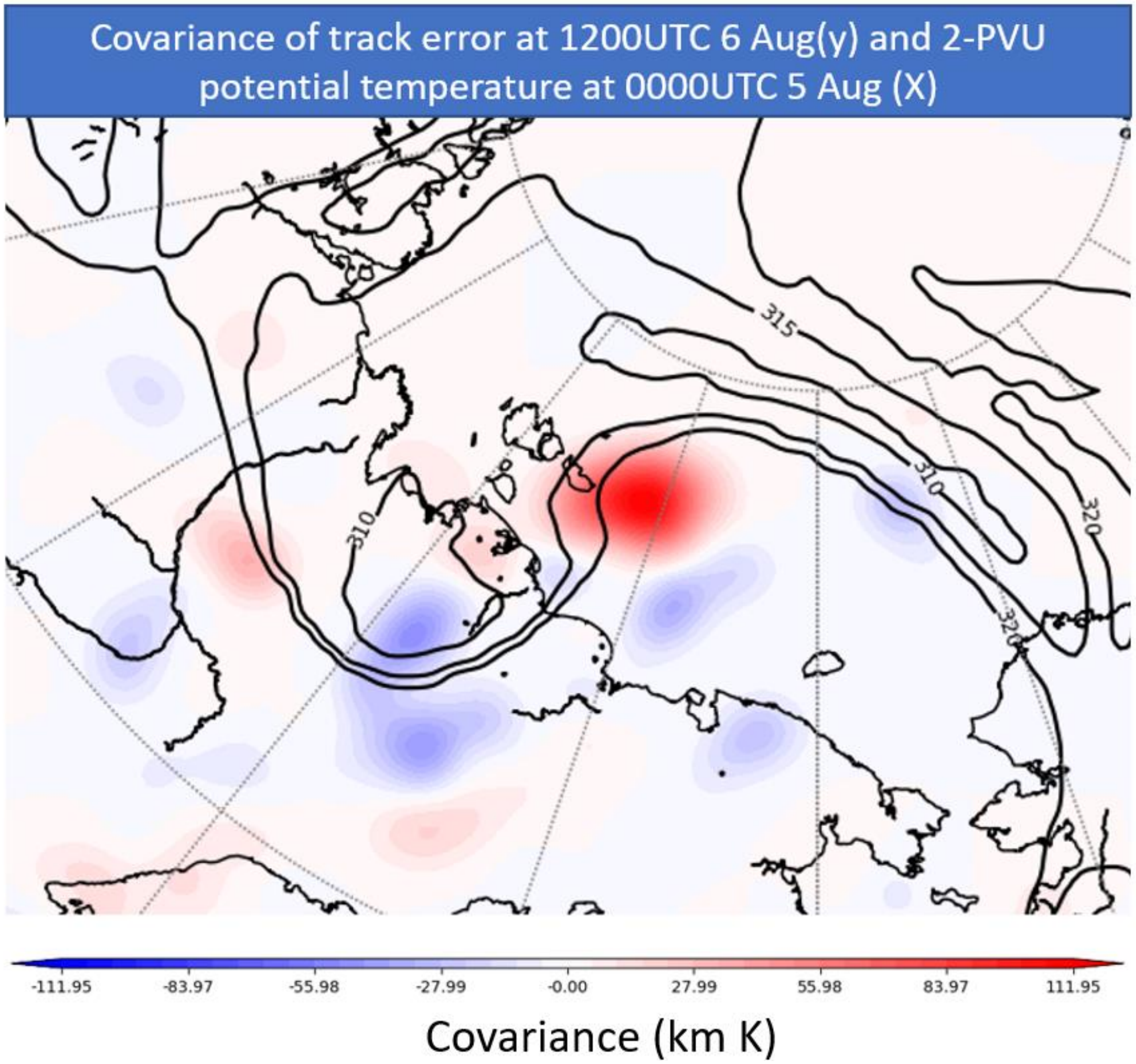


Figure 11. Covariance between the potential temperature field at T12 and cyclone track error at T24.

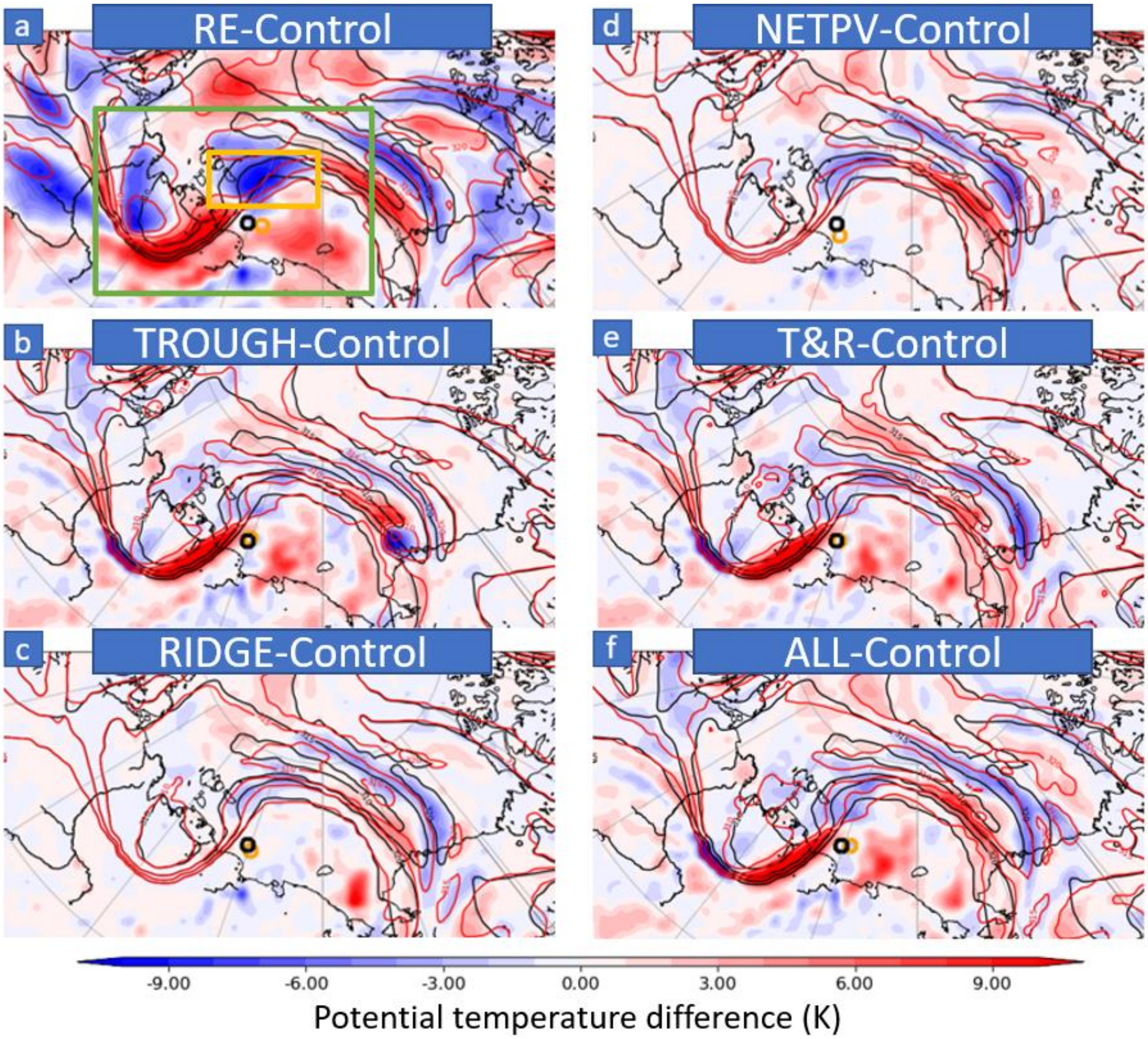


Figure 12. Same as Fig. 8 but at 0000UTC 5 Aug

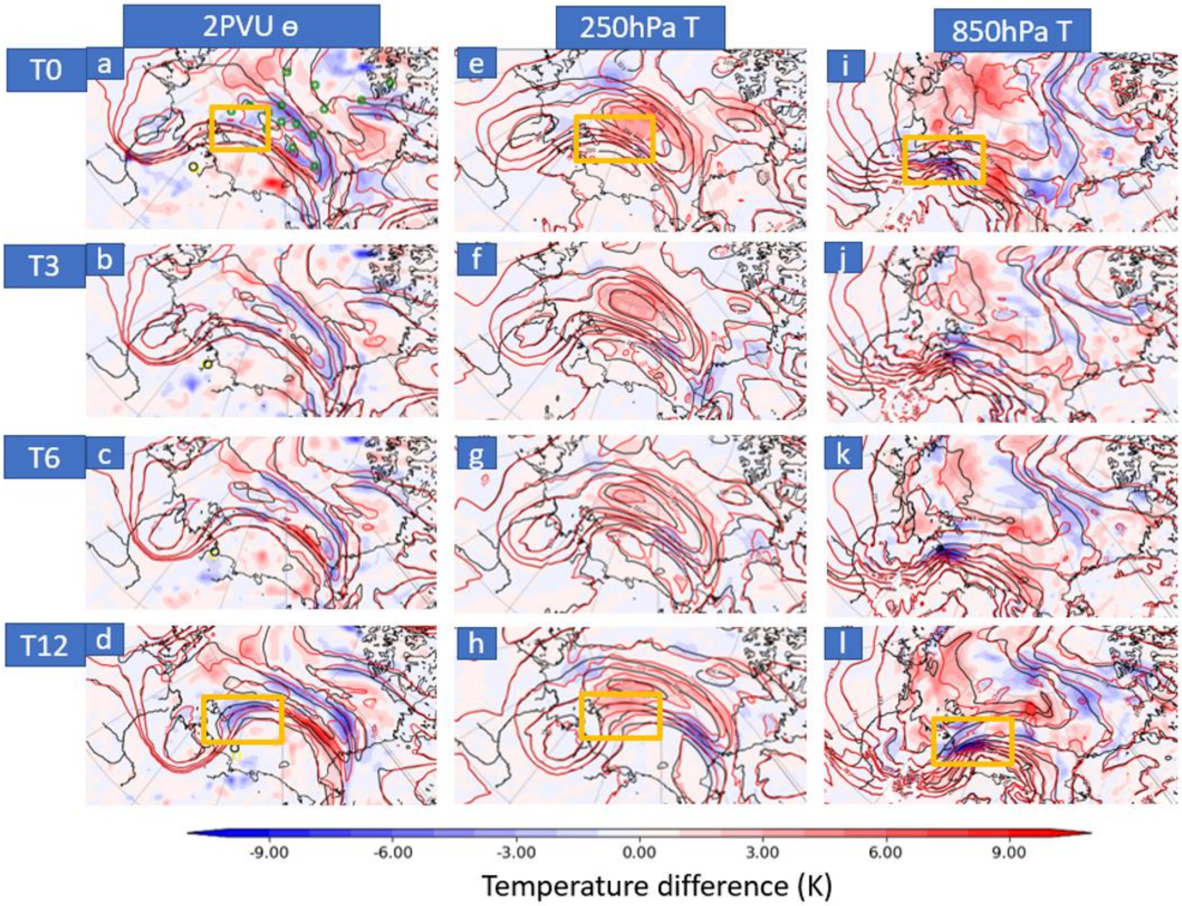


Figure 13. The potential temperature difference at 2PVU (a-d), temperature difference at 250hPa (e-h) and at 850hPa (i-l) between NETPV (K, red contours) and Control (black contours)

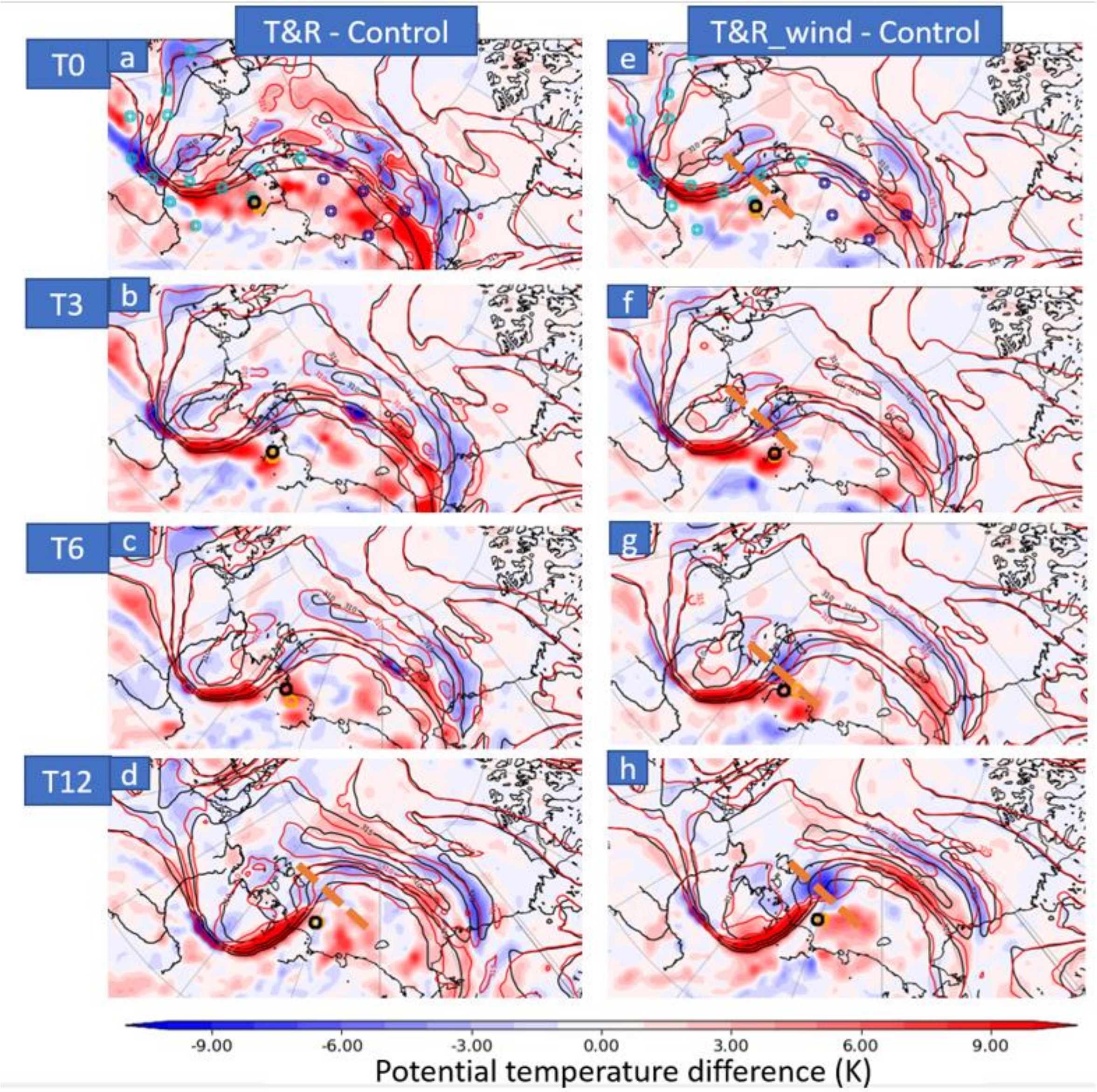


Figure 14. The potential temperature difference at 2PVU from T0 (1200UTC 4 Aug) to T12 (0000UTC 5 Aug) between T&R and Control (a-d), and between T&R_wind (e-h). Orange cycles shows the location of cyclone center in pseudo-observation experiments and black cycles show the locations in Control experiments. The orange dash line shows the location of the shortwave trough.

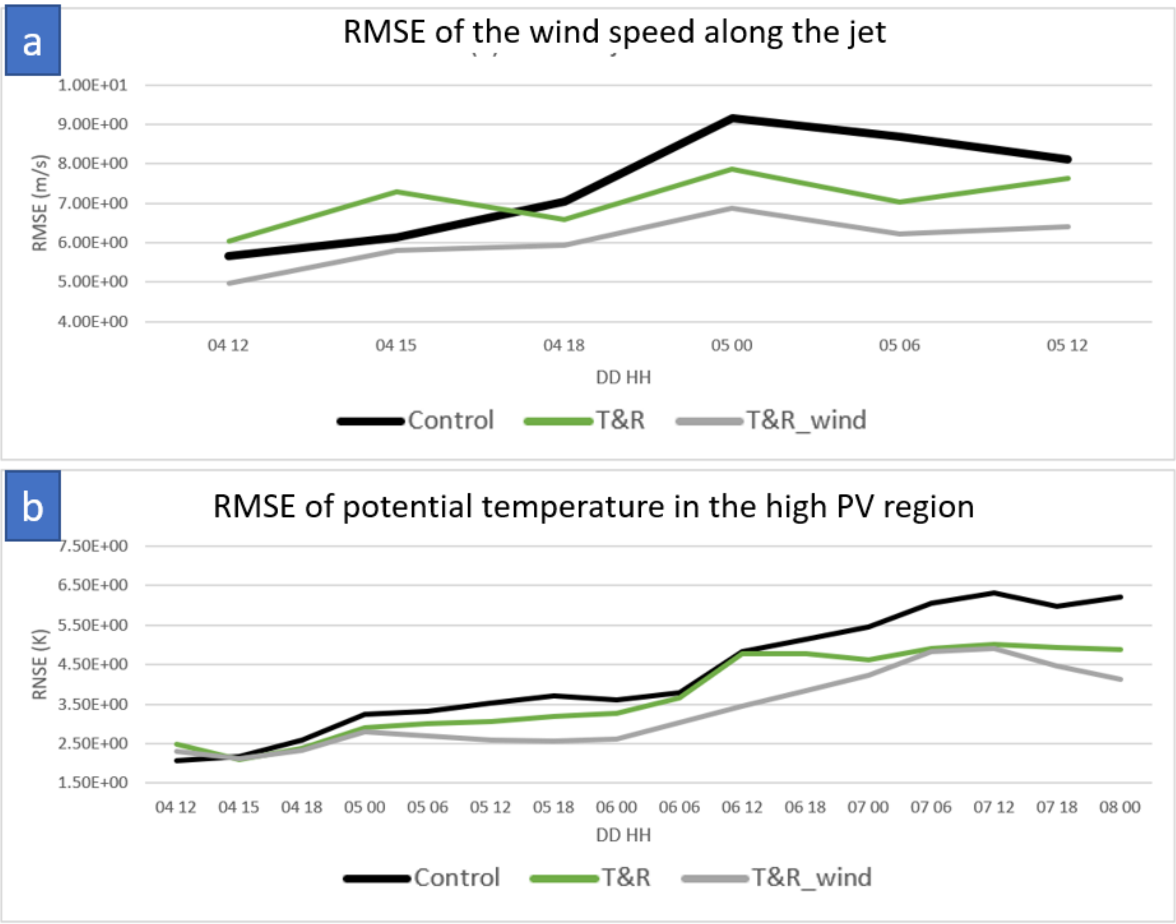


Figure 15. RMSE of jet wind speed (a) and potential temperature in the high PV region (b).

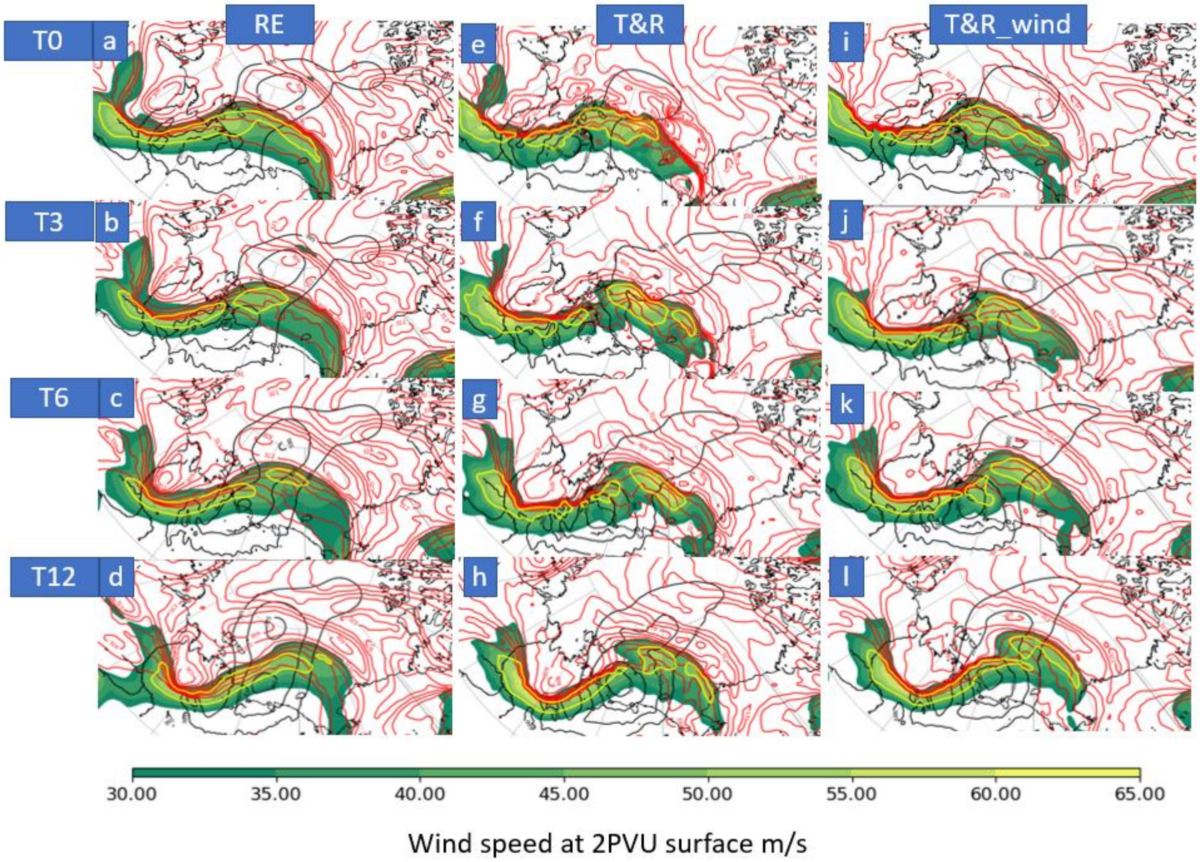
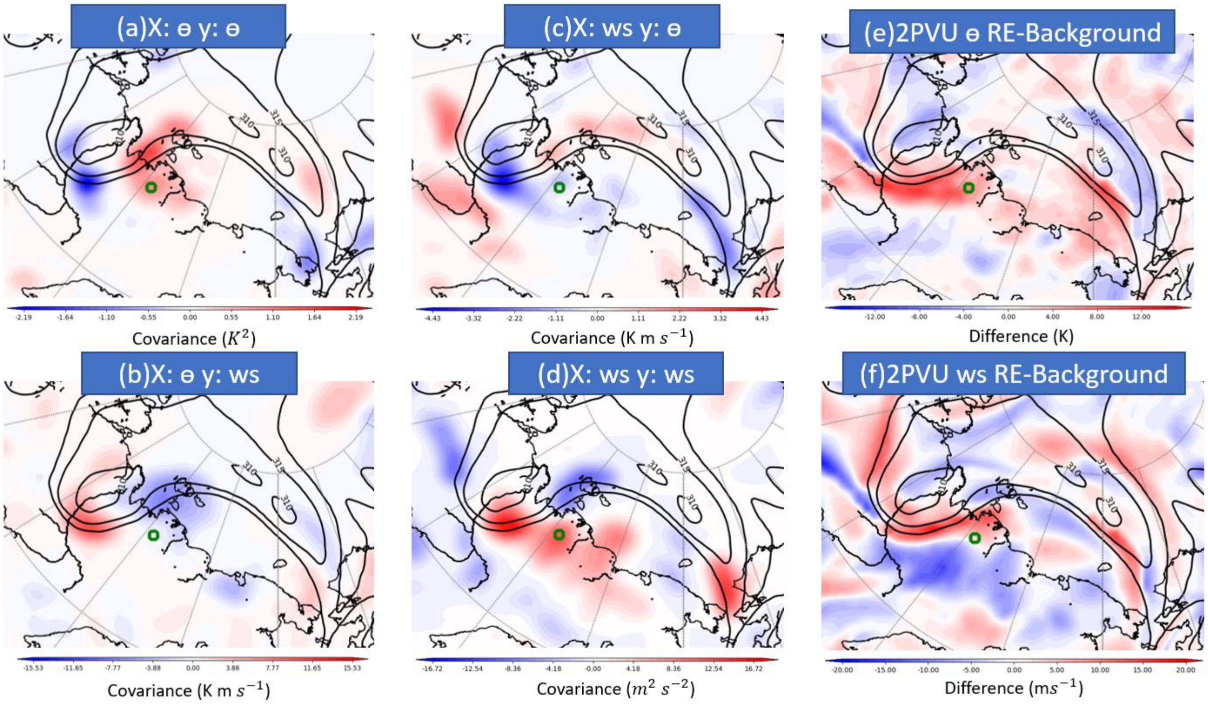


Figure 16. Potential temperature (K, red contours), wind speed (color shading) and SLP (hPa, black contours) of reanalysis (a-d), T&R (e-h) and T&R_wind (i-j) from T0 to T12.



(a)X: ɵ y: ɵ

(c)X: ws y: ɵ

(e)2PVU ɵ RE-Background

(b)X: ɵ y: ws

(d)X: ws y: ws

(f)2PVU ws RE-Background

Fig 17. Covariance plots of potential temperature (a,c), windspeed (b,d) at the observation location (yellow dots) and potential temperature (a,b), windspeed (c,d) field at 2PVU surface. The difference between reanalysis and control forecast at 2PVU surface of potential temperature (e) and wind speed (f).

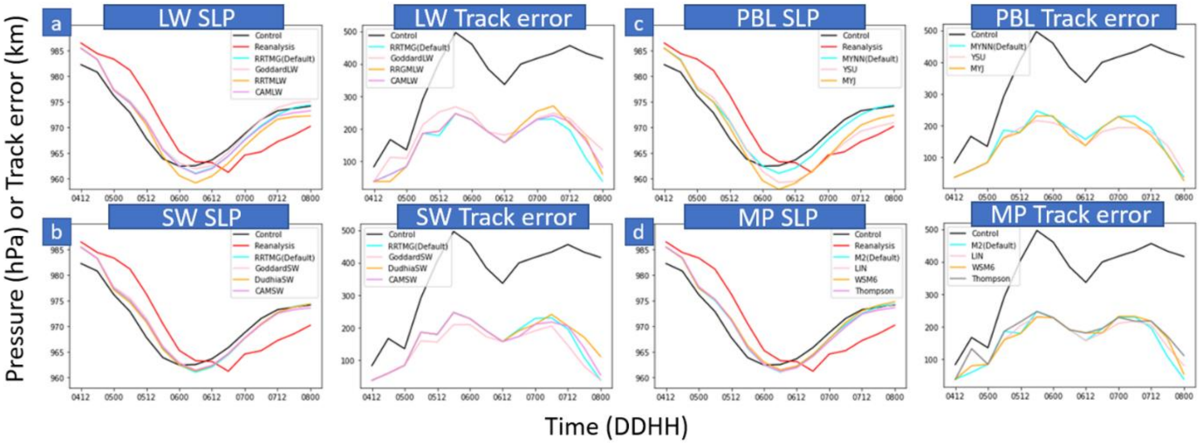


Fig 18. Cyclone center SLP and track error of (a) longwave experiments, (b) shortwave experiments, (c) PBL experiments, and (d) microphysics experiments.

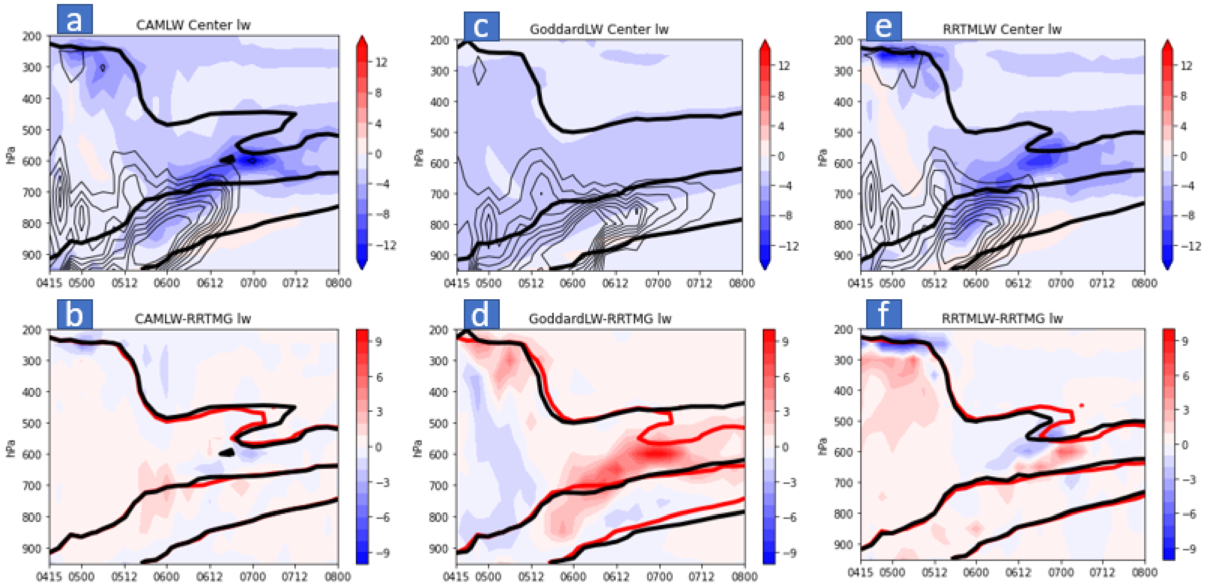


Fig 19. Comparison of LW cooling between different LW schemes. (a)(c)(e) show vertical profile of longwave cooling (K/day, blue color shading) averaged near the cyclone center changing with time. The black thick contours are the 2PVU surface, and the black thin contours are cloud mixing ratio (QICE+QCLOUD). (b)(d)(f) show the difference in longwave cooling (K/day) between each experiment to the default RRTMG LW scheme, the red thick contours are the 2PVU surface of RRTMG LW.

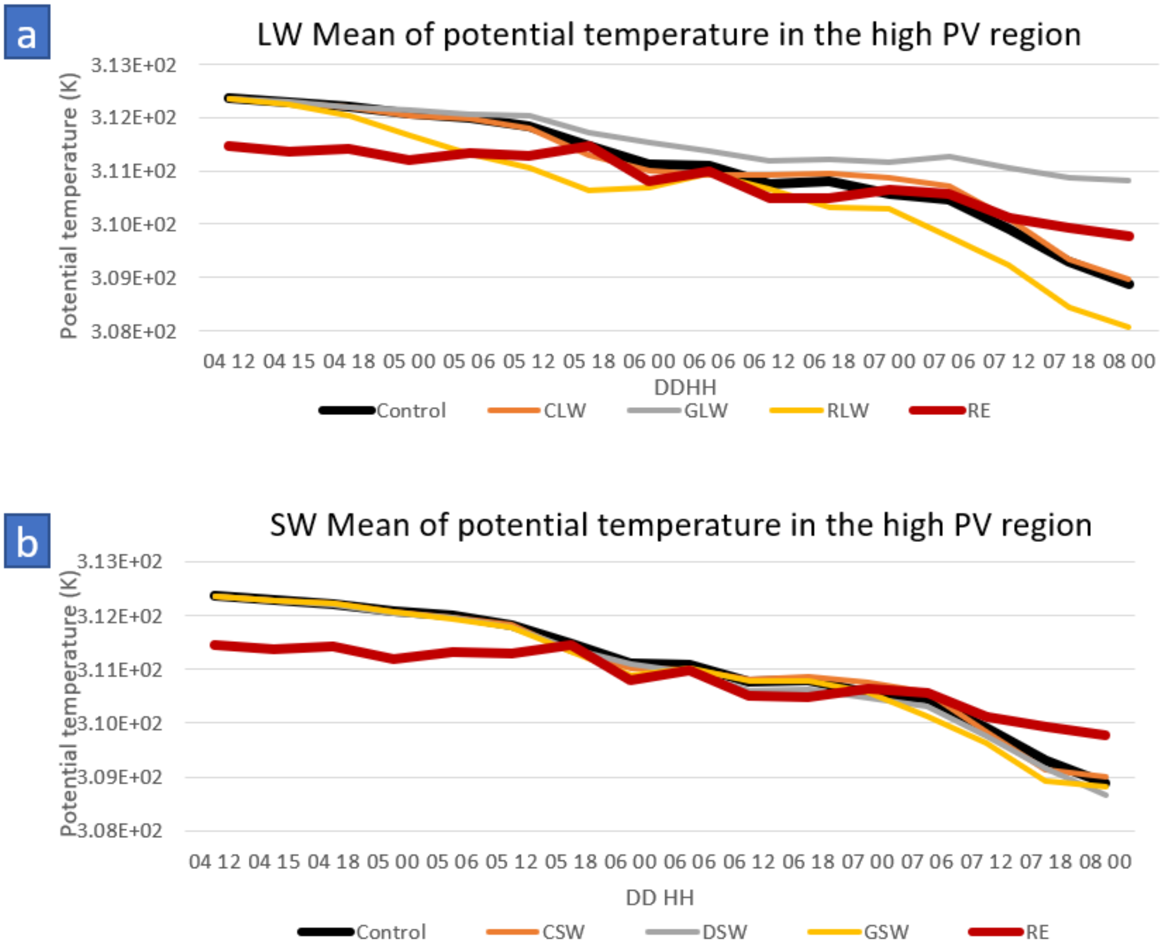


Fig 20. TPV intensity of LW and SW experiments

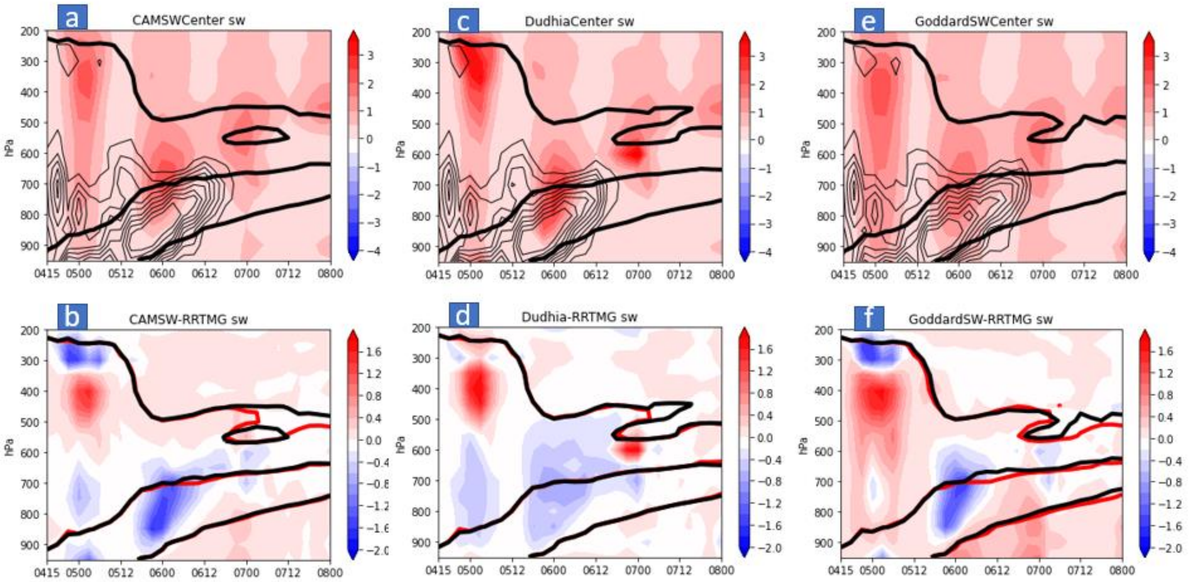


Fig 21 Same as Fig. 20 but for shortwave heating.

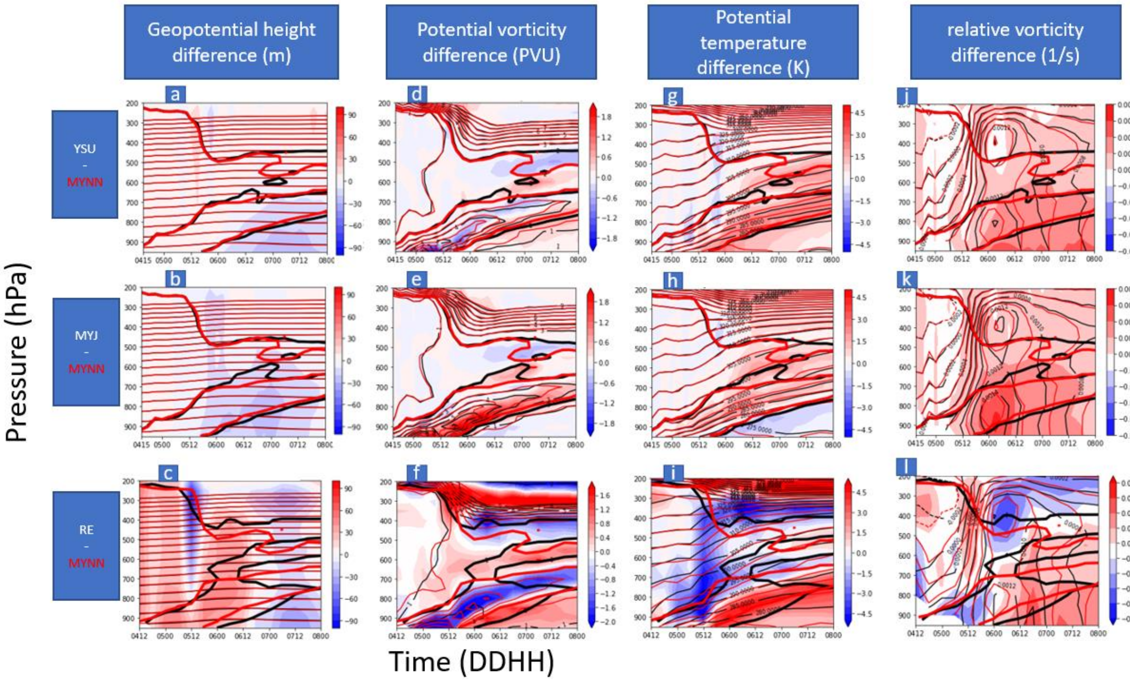


Fig 22. Difference in GH (m, a-c), PV (PVU, d-f), potential temperature (K, g-h), and relative vorticity (s^{-1}, j-l) between YSU (red contours) and MYNN(black contours) (a,d,g,j), MYJ (red contours) and MYNN (black contours)(b,e,h,k), the reanalysis (red contours) and MYNN (black contours) (c,f,I,l). The thick contours are the 2PVU surface.

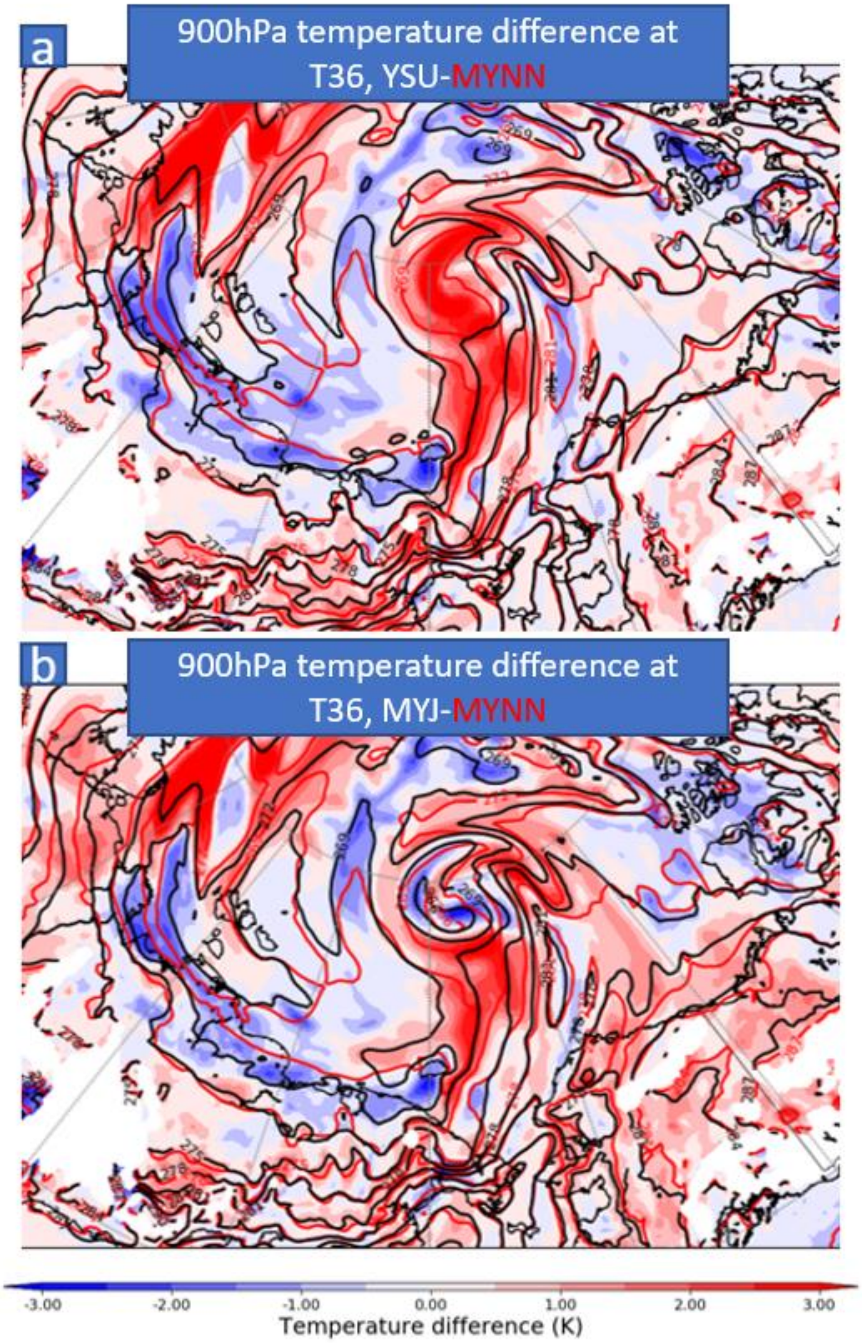


Fig 23. 900hPa Temperature difference between (a)YSU(K, red contours) and MYNN(black contours), (b) MYNN (red contours) and MYNN(black contours).

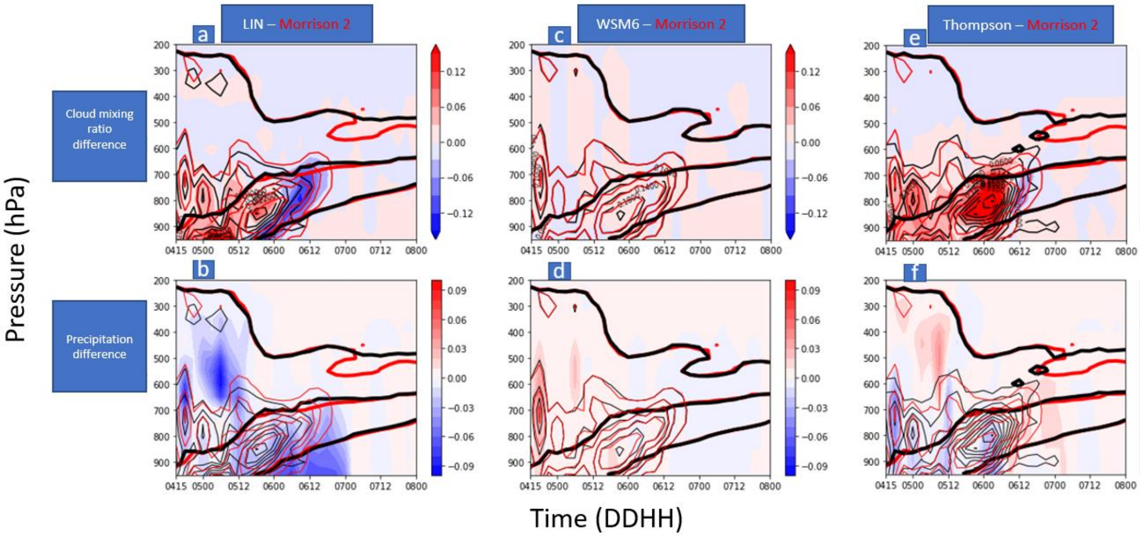
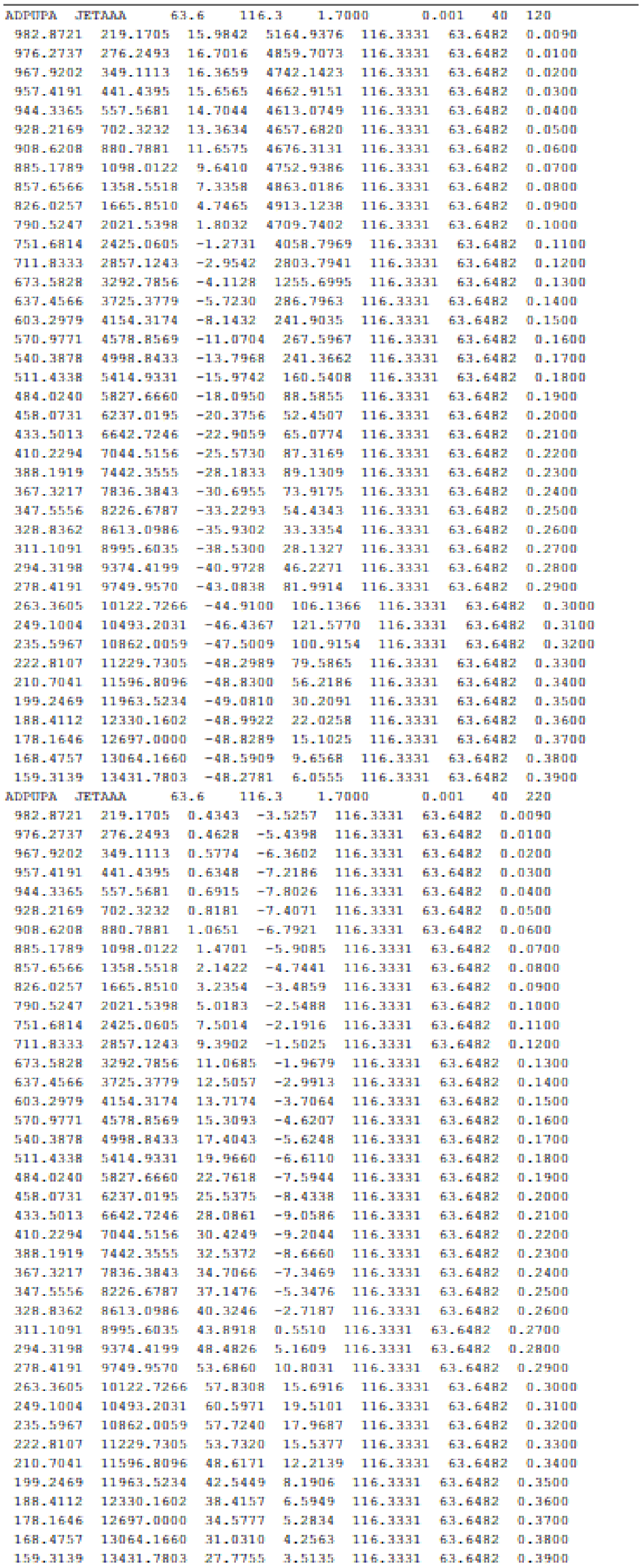


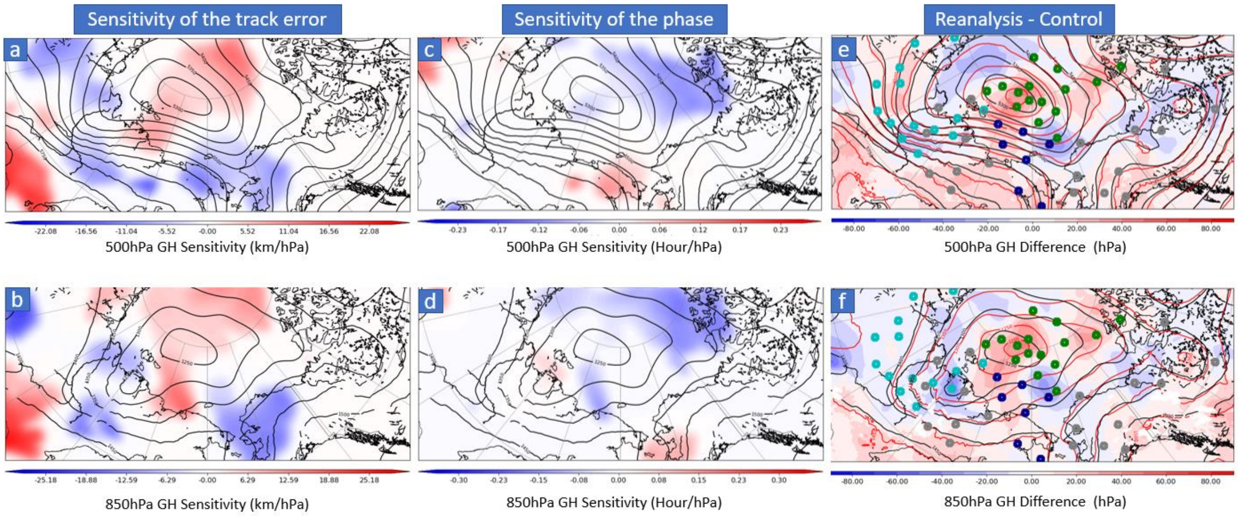
Fig 24. Cloud fraction (a,c,e) and precipitation (b,d,f) difference (g/kg) between (a,b) LIN (red) and Morrison 2-moment(black), (c,d) WSM6 (red) and Morrison 2-moment (black), (e,f) Thompson (red) and Morrison 2-moment (black). Thin contours are cloud mixing ratio and thick contours are 2PVU surface.

Appendix



```
ADPUPA  JETAAA      63.6     116.3     1.7000        0.001    40   120
 982.8721  219.1705  15.9842  5164.9376  116.3331  63.6482  0.0090
 976.2737  276.2493  16.7016  4859.7073  116.3331  63.6482  0.0100
 967.9202  349.1113  16.3659  4742.1423  116.3331  63.6482  0.0200
 957.4191  441.4395  15.6565  4662.9151  116.3331  63.6482  0.0300
 944.3365  557.5681  14.7044  4613.0749  116.3331  63.6482  0.0400
 928.2169  702.3232  13.3634  4657.6820  116.3331  63.6482  0.0500
 908.6208  880.7881  11.6575  4676.3131  116.3331  63.6482  0.0600
 885.1789  1098.0122  9.6410  4752.9386  116.3331  63.6482  0.0700
 857.6566  1358.5518  7.3358  4863.0186  116.3331  63.6482  0.0800
 826.0257  1665.8510  4.7465  4913.1238  116.3331  63.6482  0.0900
 790.5247  2021.5398  1.8032  4709.7402  116.3331  63.6482  0.1000
 751.6814  2425.0605  -1.2731  4058.7969  116.3331  63.6482  0.1100
 711.8333  2857.1243  -2.9542  2803.7941  116.3331  63.6482  0.1200
 673.5828  3292.7856  -4.1128  1255.6995  116.3331  63.6482  0.1300
 637.4566  3725.3779  -5.7230  286.7963  116.3331  63.6482  0.1400
 603.2979  4154.3174  -8.1432  241.9035  116.3331  63.6482  0.1500
 570.9771  4578.8569  -11.0704  267.5967  116.3331  63.6482  0.1600
 540.3878  4998.8433  -13.7968  241.3662  116.3331  63.6482  0.1700
 511.4338  5414.9331  -15.9742  160.5408  116.3331  63.6482  0.1800
 484.0240  5827.6660  -18.0950  88.5855  116.3331  63.6482  0.1900
 458.0731  6237.0195  -20.3756  52.4507  116.3331  63.6482  0.2000
 433.5013  6642.7246  -22.9059  65.0774  116.3331  63.6482  0.2100
 410.2294  7044.5156  -25.5730  87.3169  116.3331  63.6482  0.2200
 388.1919  7442.3555  -28.1833  89.1309  116.3331  63.6482  0.2300
 367.3217  7836.3843  -30.6955  73.9175  116.3331  63.6482  0.2400
 347.5556  8226.6787  -33.2293  54.4343  116.3331  63.6482  0.2500
 328.8362  8613.0986  -35.9302  33.3354  116.3331  63.6482  0.2600
 311.1091  8995.6035  -38.5300  28.1327  116.3331  63.6482  0.2700
 294.3198  9374.4199  -40.9728  46.2271  116.3331  63.6482  0.2800
 278.4191  9749.9570  -43.0838  81.9914  116.3331  63.6482  0.2900
 263.3605  10122.7266  -44.9100  106.1366  116.3331  63.6482  0.3000
 249.1004  10493.2031  -46.4367  121.5770  116.3331  63.6482  0.3100
 235.5967  10862.0059  -47.5009  100.9154  116.3331  63.6482  0.3200
 222.8107  11229.7305  -48.2989  79.5865  116.3331  63.6482  0.3300
 210.7041  11596.8096  -48.8300  56.2186  116.3331  63.6482  0.3400
 199.2469  11963.5234  -49.0810  30.2091  116.3331  63.6482  0.3500
 188.4112  12330.1602  -48.9922  22.0258  116.3331  63.6482  0.3600
 178.1646  12697.0000  -48.8289  15.1025  116.3331  63.6482  0.3700
 168.4757  13064.1660  -48.5909  9.6568  116.3331  63.6482  0.3800
 159.3139  13431.7803  -48.2781  6.0555  116.3331  63.6482  0.3900
ADPUPA  JETAAA      63.6     116.3     1.7000        0.001    40   220
 982.8721  219.1705  0.4343  -3.5257  116.3331  63.6482  0.0090
 976.2737  276.2493  0.4628  -5.4398  116.3331  63.6482  0.0100
 967.9202  349.1113  0.5774  -6.3602  116.3331  63.6482  0.0200
 957.4191  441.4395  0.6348  -7.2186  116.3331  63.6482  0.0300
 944.3365  557.5681  0.6915  -7.8026  116.3331  63.6482  0.0400
 928.2169  702.3232  0.8181  -7.4071  116.3331  63.6482  0.0500
 908.6208  880.7881  1.0651  -6.7921  116.3331  63.6482  0.0600
 885.1789  1098.0122  1.4701  -5.9085  116.3331  63.6482  0.0700
 857.6566  1358.5518  2.1422  -4.7441  116.3331  63.6482  0.0800
 826.0257  1665.8510  3.2354  -3.4859  116.3331  63.6482  0.0900
 790.5247  2021.5398  5.0183  -2.5488  116.3331  63.6482  0.1000
 751.6814  2425.0605  7.5014  -2.1916  116.3331  63.6482  0.1100
 711.8333  2857.1243  9.3902  -1.5025  116.3331  63.6482  0.1200
 673.5828  3292.7856  11.0685  -1.9679  116.3331  63.6482  0.1300
 637.4566  3725.3779  12.5057  -2.9913  116.3331  63.6482  0.1400
 603.2979  4154.3174  13.7174  -3.7064  116.3331  63.6482  0.1500
 570.9771  4578.8569  15.3093  -4.6207  116.3331  63.6482  0.1600
 540.3878  4998.8433  17.4043  -5.6248  116.3331  63.6482  0.1700
 511.4338  5414.9331  19.9660  -6.6110  116.3331  63.6482  0.1800
 484.0240  5827.6660  22.7618  -7.5944  116.3331  63.6482  0.1900
 458.0731  6237.0195  25.5375  -8.4338  116.3331  63.6482  0.2000
 433.5013  6642.7246  28.0861  -9.0586  116.3331  63.6482  0.2100
 410.2294  7044.5156  30.4249  -9.2044  116.3331  63.6482  0.2200
 388.1919  7442.3555  32.5372  -8.6660  116.3331  63.6482  0.2300
 367.3217  7836.3843  34.7066  -7.3469  116.3331  63.6482  0.2400
 347.5556  8226.6787  37.1476  -5.3476  116.3331  63.6482  0.2500
 328.8362  8613.0986  40.3246  -2.7187  116.3331  63.6482  0.2600
 311.1091  8995.6035  43.8918  0.5510  116.3331  63.6482  0.2700
 294.3198  9374.4199  48.4826  5.1609  116.3331  63.6482  0.2800
 278.4191  9749.9570  53.6860  10.8031  116.3331  63.6482  0.2900
 263.3605  10122.7266  57.8308  15.6916  116.3331  63.6482  0.3000
 249.1004  10493.2031  60.5971  19.5101  116.3331  63.6482  0.3100
 235.5967  10862.0059  57.7240  17.9687  116.3331  63.6482  0.3200
 222.8107  11229.7305  53.7320  15.5377  116.3331  63.6482  0.3300
 210.7041  11596.8096  48.6171  12.2139  116.3331  63.6482  0.3400
 199.2469  11963.5234  42.5449  8.1906  116.3331  63.6482  0.3500
 188.4112  12330.1602  38.4157  6.5949  116.3331  63.6482  0.3600
 178.1646  12697.0000  34.5777  5.2834  116.3331  63.6482  0.3700
 168.4757  13064.1660  31.0310  4.2563  116.3331  63.6482  0.3800
 159.3139  13431.7803  27.7755  3.5135  116.3331  63.6482  0.3900
```

The above figure shows an example of a pseudo-rawinsonde observations. The first row include observation type (ADPUPA, upper air radiosonde observations), name (JETAAA), latitude (63.6) and longitude (116.3), altitude (1.7 m), time after the cycle time (0.001 h) , vertical levels (40) and report type (120: temperature and moisture sounding, 220: wind sounding). For the 120 observations, the first 4 columns are pressure level (hPa), height level (m), virtual temperature (℃) and specific humidity (mg/kg), and the last 3 columns are drifting info, including longitude, latitude and time. For the 220 observations, the first 4 columns are pressure level (hPa), height level (m), zonal wind (m/s) and meridional wind (m/s), and the last 3 columns are still drifting info. This text file is then formatted into a PREPBUFR file.



Additional ensemble sensitivity analysis results with 500hPa GH and 850hPa GH as the state variable. The above figure shows the ensemble sensitivity analysis (a-f) and initial condition error (g-i) of the two state variables. The signals at the two levels below 300hPa share consistent features with 300hPa, suggesting the trough, ridge, and NETPV with its northeast edge are critical to the AC12's evolution.