

**Development of Multi-Center Grand Ensemble Prediction
and its Application to High Impact Weather**

July 2008

Mio MATSUEDA

**Development of Multi-Center Grand Ensemble Prediction
and its Application to High Impact Weather**

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

Mio MATSUEDA

Abstract

Recently, the ensemble forecast has become a major component of operational global weather prediction systems. TIGGE (THORPEX Interactive Grand Global Ensemble) is a part of WWRP (World Weather Research Programme) to accelerate the improvements in the accuracy of one-day to two-week high-impact weather forecasts. The TIGGE has enabled us to get operational medium-range ensemble forecast data near real time.

In this study, first, the overall intercomparisons of five operational medium-range ensembles: CMC, ECMWF, JMA, NCEP, and UKMO, were performed. The forecast skill was evaluated using Root Mean Square Error (RMSE) for 500 hPa geopotential height over the Northern Hemisphere (NH, 20°N–90°N) from December 2006 to November 2007. It was found in the control run and ensemble mean forecast that the ECMWF has the best forecast skill for almost all seasons and almost all forecast ranges. In particular, the ECMWF is far superior to the other centers in the early forecast range (day 0–3). The spread of the ECMWF showed good agreement with the RMSE of the ECMWF for almost all forecast ranges. On the other hand, it was found that the performance of the CMC, JMA, NCEP, and UKMO ensemble mean is almost comparable with each other. Their skills can be considered as the second-best.

Second, Multi-Center Grand Ensembles (MCGEs) were constructed using five medium-range ensemble forecasts: CMC, ECMWF, JMA, NCEP, and UKMO. The forecast performance of the MCGEs relative to the ECMWF ensemble was investigated using the seasonal RMSE and Ranked Probability Score (RPS) for 500 hPa height over the Northern Hemisphere from December 2006 to November 2007. It was found in the deterministic and probabilistic verifications that the MCGEs can outperform the ECMWF ensemble at least in the medium forecast range (day 6–9) for all seasons. The forecast time when the MCGEs first outperform the ECMWF ensemble is somewhat different depending on the season. During the northern summer, the advantage of the MCGEs appears as early as at +4 day forecast time. The improvements in the RMSE and RPS are several percentage points in the medium forecast range. These are almost comparable with the rate of improvement in a single-center ensemble forecast during the latest few

years. Although the ensemble spread of the MCGE was larger than the RMSE in the early forecast range, ensemble spread of the MCGE showed good agreement with the RMSE since the medium-range forecast range.

Third, an analysis of atmospheric blocking was performed using ensemble forecast data, ensemble-based simple sensitivity analysis, and multi-analysis ensemble forecasts. The ensemble forecast initialized at 12 UTC on 10th December 2005 was a very interesting case. All NCEP members were not able to predict the location of the blocking occurred on 15th December 2005 correctly, whereas almost all JMA members were able to predict it correctly. According to the multi-analysis ensemble forecasts and ensemble-based sensitivity analysis, it was found that the collective mis-prediction mainly resulted from the NCEP control analysis over the central North Pacific at 12 UTC on 10th December although the decreases of the imperfection of the model formulation were recognized. In the sensitivity area, there was a cut-off cyclone. The difference between the JMA and NCEP control analyses around the cyclone was relatively larger than the other areas. Due to the lack of the effective initial perturbation in the NCEP members, the large uncertainty around the cyclone was not able to be reduced. This led to the collective mis-prediction. In fact, the multi-analysis ensemble forecasts from the NCEP analyses with the regionally amplified initial perturbation showed the improvement in the RMSE over the blocking region without the degradation of the forecast skill over the Northern Hemisphere. These results indicate that the sensitivity area was essential for the prediction of the blocking. The global amplification of the initial perturbations led to the decrease of the RMSE over the blocking region and the increase of the RMSE over the Northern Hemisphere. These results suggested that the excessive amplification of the initial perturbation over non-sensitivity area is undesirable and that the regional amplification technique can lead to better forecast without the degradation of the forecast over the other area. The result in this study shows that such a case dependent estimates may really have value as compared to climatologically based rescaling that is used widely.

Keywords: THORPEX, TIGGE, medium-range ensemble forecast, Multi-Center Grand Ensemble, high-impact weather, ensemble-based sensitivity analysis, multi-analysis ensemble

Contents

Abstract	i
List of Tables	vi
List of Figures	vii
List of Acronyms	xii
1 Introduction	1
1.1 Ensemble forecast	1
1.2 THORPEX and TIGGE	9
1.3 Purpose of this study	11
2 Multi-Center Grand Ensemble using five operational medium-range ensemble forecasts	21
2.1 Operational medium-range ensemble prediction systems	21
2.1.1 Canadian Meteorological Center (CMC)	21
2.1.2 European Center for Medium-Range Weather Forecasts (ECMWF)	22
2.1.3 Japan Meteorological Agency (JMA)	24
2.1.4 National Centers for Environmental Prediction (NCEP)	24
2.1.5 United Kingdom Meteorological Office (UKMO)	26
2.2 Data	27

2.2.1	Ensemble forecast data	27
2.2.2	ECMWF 40 Year Re-Analysis (ERA40)	27
2.3	Methods	28
2.3.1	Statistics on ensemble forecast	28
2.3.2	Verification scores for deterministic forecasts	29
2.3.3	Verification score for probabilistic forecasts	31
2.3.4	Comparison of scores	32
2.4	Comparison of CMC, ECMWF, JMA, NCEP, and UKMO ensemble forecasts	34
2.4.1	Daily and seasonal RMSEs	34
2.4.2	Relationship between RMSE and ensemble spread	36
2.5	Construction of Multi-Center Grand Ensemble	40
2.6	Verifications of forecast skill of Multi-Center Grand Ensemble	41
2.6.1	Deterministic verifications	41
2.6.2	Probabilistic verifications	43
3	Analyses of extreme event using ensemble forecast data	63
3.1	Data and method	63
3.1.1	Ensemble forecast data	63
3.1.2	Ensemble-based sensitivity analysis	63
3.1.3	Multi-analysis ensemble forecasts	65
3.2	Target blocking	65
3.3	Comparison of control runs	66
3.4	Multi-analysis ensemble forecasts with NCEP analyses	67

3.5	Ensemble-based sensitivity analysis	68
3.6	Multi-analysis ensemble forecasts with amplified initial perturbations	69
4	Discussion	86
4.1	Construction method of Multi-Center Grand Ensemble	86
4.2	Identification of causes of forecast error and initial perturbation method	88
5	Conclusion	90
	Acknowledgements	94
	REFERENCES	96

List of Tables

1.1	Operational medium-range ensemble prediction systems as of December 2007.	13
1.1	Operational medium-range ensemble prediction systems as of December 2007 (continued).	14
1.2	Outline of data accumulated in the TIGGE archives as of June 2008.	15
2.1	Ensemble means created in this study.	45
2.2	Configuration of MCGEs created in this study.	46
3.1	Operational medium-range ensemble prediction system at CMC, JMA, and NCEP as of December 2005.	71
3.2	120-hr RMSEs of NCEP-EPS and JMA-GSM runs with NCEP analyses for 500 hPa height over the blocking region.	72
3.3	120-hr RMSEs of NCEP-EPS and JMA-GSM runs with NCEP analyses for 500 hPa height over the Northern Hemisphere.	72

List of Figures

1.1	Anomaly correlation coefficients of 3-, 5-, 7-, and 10-day ECMWF 500 hPa height forecasts for the extratropical Northern and Southern Hemispheres, plotted in the form of annual running means of monthly-mean scores for the period of 1981–2006 (Simmons et al. 2006).	16
1.2	Longest available record of skill of JMA ensemble mean forecast from DJF 2002 to SON 2007. The skill is evaluated by the seasonal RMSE for 500 hPa height over the Northern Hemisphere.	17
1.3	Spaghetti diagram of 500 hPa height (5550 m) for ensemble members of BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, NCEP, and UKMO, initialized on 20th February 2008, valid 12 UTC on 25th February 2008. . .	18
1.4	Time series of 120-hr RMSEs of nine single-center ensemble mean forecasts: BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, NCEP, and UKMO, for 500 hPa height over the Northern Hemisphere from October 2006 to March 2007.	19
1.5	Skills of nine single-center control and ensemble mean forecasts: BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, NCEP, and UKMO, for 500 hPa height over the Northern Hemisphere from December 2007 to February 2008.	20
2.1	Geometric relationship between the AC and RMSE in the phase space. . .	30

2.2	Time series of 120-hr RMSEs of five single-center control and ensemble mean forecasts; CMC, ECMWF, JMA, NCEP, and UKMO, for 500 hPa height over the Northern Hemisphere from December 2006 to November 2007.	47
2.3	Skills of five single-center control and ensemble mean forecasts; CMC, ECMWF, JMA, NCEP, and UKMO, for 500 hPa height over the Northern Hemisphere from December 2006 to February 2007 and from March 2007 to May 2007.	48
2.4	Skills of five single-center control and ensemble mean forecasts; CMC, ECMWF, JMA, NCEP, and UKMO, for 500 hPa height over the Northern Hemisphere from June 2007 to August 2007 and from September 2007 to November 2007.	49
2.5	Ensemble mean and spread of CMC, ECMWF, NCEP, UKMO, and JMA for 500 hPa height, initialized on 17th November 2007, valid at 12 UTC on 22nd November 2007.	50
2.6	Time series of 120-hr ensemble spreads of CMC, ECMWF, NCEP, UKMO, and JMA for 500 hPa height over the Northern Hemisphere from December 2006 to November 2007.	51
2.7	Ensemble spreads of CMC, ECMWF, NCEP, UKMO, and JMA for 500 hPa height over the Northern Hemisphere from December 2006 to February 2007 and from March 2007 to May 2007.	52
2.8	Ensemble spreads of CMC, ECMWF, NCEP, UKMO, and JMA for 500 hPa height over the Northern Hemisphere from June 2007 to August 2007 and from September 2007 to November 2007.	53
2.9	Scatter diagrams of the RMSE versus the ensemble spread of five single-center ensembles for 500 hPa height over the Northern Hemisphere at 24-hr lead time.	54
2.10	Scatter diagrams of the RMSE versus the ensemble spread of five single-center ensembles for 500 hPa height over the Northern Hemisphere at 72-hr lead time.	55

2.11	Scatter diagrams of the RMSE versus the ensemble spread of five single-center ensembles for 500 hPa height over the Northern Hemisphere at 120-hr lead time.	56
2.12	Scatter diagrams of the RMSE versus the ensemble spread of five single-center ensembles for 500 hPa height over the Northern Hemisphere at 168-hr lead time.	57
2.13	Scatter diagrams of the RMSE versus the ensemble spread of five single-center ensembles for 500 hPa height over the Northern Hemisphere at 216-hr lead time.	58
2.14	Relative improvement (%) in seasonal RMSE of MCGEs against ECMWF51 for 500 hPa height over the Northern Hemisphere from December 2006 to February 2007, from March 2007 to May 2007, from June 2007 to August 2007, and from September 2007 to November 2007.	59
2.15	Percent improvement in daily 168-hr RMSE of MCGEs against ECMWF ensemble for 500 hPa height over the Northern Hemisphere from December 2006 to February 2007, from March 2007 to May 2007, from June 2007 to August 2007, and from September 2007 to November 2007.	60
2.16	Scatter diagram of the RMSE versus the ensemble spread of MCGE51 for 500 hPa height over the Northern Hemisphere at 24-, 72-, 120-, 168-, and 216-hr lead times from December 2006 to November 2007.	61
2.17	Relative improvement (%) in seasonal RPS of MCGEs against ECMWF ensemble for 500 hPa height over the Northern Hemisphere from December 2006 to February 2007, from March 2007 to May 2007, from June 2007 to August 2007, and from September 2007 to November 2007.	62
3.1	Conceptual diagram of phase space on disturbance.	73

3.2	Spaghetti diagrams of 500 hPa height (5500 m) for ensemble members of CMC, JMA, and NCEP, initialized on 10th December 2005, valid 12 UTC on 15th December 2005.	74
3.3	Time evolution of 500 hPa height and its forecast error for JMA analysis, control runs of JMA and NCEP, and JMA-GSM run with NCEP control analysis. The initial time is 12 UTC on 10th December 2005.	75
3.4	Spaghetti diagrams of 500 hPa height (5500 m) for JMA-GSM runs with NCEP control and perturbed analyses from 12 UTC on 10th December 2005, valid 12 UTC on 15th December 2005	76
3.5	120-hr forecast of 500 hPa height and its forecast error for NCEP original control and perturbed runs from 12 UTC on 10th December 2005.	77
3.6	Sensitivity area obtained from JMA ensemble forecast initialized at 12 UTC on 10th December 2005.	78
3.7	Sensitivity area for components, u' , v' , T' , and p'_s at each pressure level obtained from JMA ensemble forecast initialized at 12 UTC on 10th December 2005.	79
3.8	Initial difference between JMA and NCEP analyses at 12 UTC on 10th December 2005 measured by dry total energy.	80
3.9	Time evolution of sea level pressure for JMA and NCEP control runs.	81
3.10	Sea level pressure and its analysis error based on AFES-LETKF Experimental Re-Analysis (ALERA) at 12 UTC on 10th December 2005.	82
3.11	Dry total energy for NCEP initial perturbations at 12 UTC on 10th December 2005.	83
3.12	120-hr forecast of 500 hPa height and its forecast error for JMA-GSM runs with NCEP analyses with the regionally amplified perturbations, initialized at 12 UTC on 10th December 2005.	84

3.13 120-hr ensemble mean forecasts of 500 hPa height and its forecast error for NCEP original ensemble forecast and JMA-GSM runs with NCEP analyses with the original, the globally amplified, and the regionally amplified perturbations, initialized at 12 UTC on 10th December 2005. 85

List of Acronyms

AC	Anomaly Correlation
ALERA	AFES-LETKF Experimental Re-Analysis
BOM	Bureau of Meteorology
BS	Brier Score
BSS	Brier Skill Score
BV	Bred Vector
CMA	China Meteorological Administration
CMC	Canadian Meteorological Center
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos
DA	Data Assimilation
DEMETER	DEvelopment of a Multi-model Ensemble system for seasonal to inTERannual prediction
DJF	December-January-February
DOTSTAR	Dropwindsonde Observation for Typhoon Surveillance near the TAIwan Region
ECMWF	European Center for Medium-Range Weather Forecasts
EAKF	Ensemble Adjustment Kalman Filter
EnKF	Ensemble Kalman Filter
EOF	Empirical Orthogonal Function
EPS	Ensemble Prediction System
ERA40	ECMWF 40 Year Re-Analysis
ET	Ensemble Transform
ETKF	Ensemble Transform Kalman Filter

ETR	Ensemble Transform with Rescaling
FNMOCC	Fleet Numerical Meteorology and Oceanography Center
GEM	Global Environmental Multi-scale
GSI	Gridpoint Statistical Interpolation
GSM	Global Spectral Model
GIFS	Global Interactive Forecast System
JMA	Japan Meteorological Agency
JJA	June-July-August
KMA	Korean Meteorological Administration
LAF	Lagged Average Forecasting
LEKF	Local Ensemble Kalman Filter
LETKF	Local Ensemble Transform Kalman Filter
MAM	March-April-May
MCGE	Multi-Center Grand Ensemble
MCS	Mesoscale Convective System
MOGREPS	Met Office Global and Regional Ensemble Prediction System
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NH	Northern Hemisphere
NWP	Numerical Weather Prediction
OI	Optimal Interpolation
PO	Perturbed Observation
PV	Potential Vorticity
RMSE	Root Mean Square Error
RP	Random Parameters
PROVOST	PRediction Of climate Variations On Seasonal to interannual Time-scales
RPS	Ranked Probability Score

SCV	Stochastic Convective Vorticity
SEF	Spectral Finite-Element
SON	September-October-November
SSI	Spectral Statistical Interpolation
SV	Singular Vector
THORPEX	THE Observing system Research and Predictability EXperiment
TIGGE	THORPEX Interactive Grand Global Ensemble
VAREPS	Variable Resolution Ensemble Prediction System
WMO	World Meteorological Organization
WWRP	World Weather Research Programme

Chapter 1

Introduction

1.1 Ensemble forecast

Recently, the ensemble forecast has become a major component of operational global weather prediction systems, and has drawn more attention in various timescales, such as short-, medium-, and long-ranges for both operational and research purposes. The ensemble simulation also adopted in the global warming research in order to estimate the uncertainty of future climate.

In the ensemble forecasts, several model forecasts are performed by introducing perturbations in the initial conditions or in the models themselves. Ensemble forecast is employed to achieve two main goals: the first one is to provide that an ensemble average forecast beyond the first few days is more accurate than individual forecast, because the components of the forecast that are most uncertain tend to be averaged out. The second and more important goal is to provide forecasters with an estimation of the reliability of the forecast, which, because of changes in atmospheric predictability, varies from day to day and from region to region (Kalnay 2003).

As of October 2006, about 12 and 5 meteorological operational and research centers all over the world operate their own global and regional ensemble prediction system, respectively (WMO 2007). These ensemble systems differ in the resolution of the analysis and forecast modes, in the initial perturbation method, in the schemes used to parameterize physical processes, in the ensemble size, and in the forecast length. The long-range

ensemble forecast is also operated in about 10 operational centers under the different systems.

In December 1992, the European Center for Medium-Range Weather Forecasts (ECMWF; Molteni et al. 1996) and the National Centers for Environmental Prediction (NCEP; Toth and Kalnay 1997) employed the ensemble techniques for operational medium-range forecast for the first time followed by other Numerical Weather Prediction (NWP) centers: the Japan Meteorological Agency (JMA; JMA 2007), the Canadian Meteorological Center (CMC; Houtekamer et al. 1996), the Bureau of Meteorology (BOM; Bourke et al. 2004), and so on. The details of each medium-range Ensemble Prediction System (EPS) as of December 2007 are summarized in Table 1.1. Each NWP center adopts different EPS. This fact depends on the computer resources, the development cost, the aim of the EPS in each NWP center, and so on. For example, NCEP performs medium-range ensemble forecast 4 times per day to cover the lack of ensemble size per run due to computer resource limitations. Also, JMA has changed the initial perturbation method from the Bred Vector (BV) method to the Singular Vector (SV) method on 21st November 2007 in order to gain the more useful information from the own EPS. Each EPS is sometimes changed along with the changes of their computer resources. Many researches on the basic performance of the EPS have been performed with a view to the operational use. For example, the impact of model resolution on the performance of an EPS was investigated (Buizza et al. 2003, 1998; Mullen et al. 2002; Pellerin et al. 2003; Szunyogh and Toth 2002) mainly along with the update of the computer resources. Also, the impact of ensemble size was investigated (Buizza and Palmer 1998; Buizza et al. 1998; Mitchell and Houtekamer 2002; Mullen et al. 2002; Richardson 2001b). Based on these researches, own computer resources, and own development cost, each NWP center constructs a better, but not the best, EPS at least for oneself.

Intercomparisons of these EPSs were done by Buizza et al. (2005) for the CMC, ECMWF, and NCEP over the Northern Hemisphere (NH) from May to July 2002, Bourke et al. (2004) for the BOM and ECMWF over the Southern Hemisphere from 2nd April to 31st August 2001, Matsueda et al. (2006) for the CMC, JMA, and NCEP over the NH in September 2005, Froude et al. (2007) for the ECMWF and NCEP against the

extratropical storm tracks from 6th January to 5th April 2005, and Matsueda et al. (2007) for the CMC, JMA, and NCEP over the NH from August 2005 to February 2006.

The forecast errors in the short- and medium-range forecasts are mainly affected by the chaotic nature of the atmosphere, the uncertainties of the initial conditions, and the imperfection of the numerical model.

The initial perturbation method is one of the important factors in determining the EPS performance. The basic concept of the initial perturbation method is that uncertainties included in the initial data are reduced by introducing the initial perturbations. The major initial perturbation methods are the Lagged Average Forecasting (LAF) method, the total energy norm based SV method, the BV method, the Perturbed Observation (PO) method, the Ensemble Kalman Filter (EnKF) method, and the Ensemble Transform Kalman Filter (ETKF) method, the Ensemble Transform (ET) method, and the ET with Rescaling (ETR) method. The LAF, SV, BV, and PO methods are recognized as the first generation methods. The EnKF, ETKF, ET, and ETR methods are recognized as the second generation methods.

The LAF method is the simplest initial perturbation method (Hoffman and Kalnay 1983). This method has been basically used in the long-range forecasts and climate researches. This method uses the forecasts initialized at different times. In the medium-range ensemble forecast, the LAF method is not used alone. The LAF method is not used alone but used with the following initial perturbation methods, which are based on dynamical and mathematical concepts. This method is effective in case that the ensemble size by the following initial perturbation methods is not enough because of the limitation of the computer resources. This method is adopted at the BOM, CMC, Centro de Previsão de Tempo e Estudos Climáticos (CPTEC; Ferreira 2007), ECMWF, Korean Meteorological Administration (KMA; Park 2007), NCEP, and United Kingdom Meteorological Office (UKMO; Bowler et al. 2008) as of December 2007.

The SV method is first introduced at the ECMWF (Buzza and Palmer 1995; Molteni et al. 1996). The SV method assumes that each possible pattern with unit total energy in the three dimensional and multivariate space of a numerical model is equally likely to the

error pattern in a numerical analysis. Sampling out of the assumed uniform distribution is done through the SV method that results in perturbations that produce the maximum linear growth over a pre-specified optimization period (typically, 1 to 2 days). Growth and perturbation patterns beyond a transitional period are determined by the state dependent local Lyapunov characteristics of the system (supporting sustainable growth), while those during the transitional period are strongly influenced by the choice of the norm used in the definition of SVs. The SV method uses an adjoint (Errico 1997) and linear tangent of the forecast model to determine the growth directions, and the perturbations are designed in this subspace. This method is adopted at the BOM, China Meteorological Administration (CMA; WMO 2007), ECMWF, and JMA as of December 2007.

Toth and Kalnay (1993, 1997) introduced the BV method at the NCEP. The BV method is based on the argument that fast-growing perturbations develop naturally in a data assimilation cycle and will continue to grow as short- and medium-range forecast error. The BV method generates perturbations in directions where past forecast errors have grown rapidly. To simulate the error breeding process in an analysis cycle, a perturbation field in a breeding cycle is dynamically cycled through the use of two non-linear forecast integrations, where the difference between the two forecasts are periodically rescaled and then repositioned onto consecutive analysis fields (Toth and Kalnay 1993). Although the SV method needs an adjoint code (Ericco 1997), the BV method generates initial perturbations without it. So, it is easy to introduce the BV into own EPS. This method is adopted at the Fleet Numerical Meteorology and Oceanography Center (FNMOC; WMO 2007) and KMA as of December 2007.

The PO method (Houtekamer et al. 1996), used at the CMC, relied on an Optimal Interpolation (OI) analysis scheme. This was due to the fact that the OI scheme was much cheaper to run. In addition, the POs may introduce undesirable noise into the forecast (Whitaker and Hamill 2002). As most observations are perturbed by a simulated observational error, followed by multiple numerical analyses, perturbations with the PO method represent random samples from the suboptimal analysis PDF, containing both fast growing and neutral/decaying perturbation patterns with realistic amplitude. Since the PO method samples analysis error, therefore it is closer to a second generation method

than other first generation methods.

The PO and BV methods are related in that they aim at providing a random sample of analysis errors. The resulting perturbations generated by the BV method are similar to those by the PO method except that perturbations represent only the growing components and ignore the neutral and decaying components of the analysis error. The BV method is also related to the SV method in that both methods aim at sampling the fastest-growing forecast errors. Legras and Vautard (1996) show that the BV and SV methods are related through notions of "backward" and "forward" Lyapunov vectors. The difference between these two methods is that while the BV method attempts to provide a random sample of growing analysis errors, the SV method gives a selective sample of perturbations that can produce the fastest linear growth in the future.

The EnKF method is currently drawing attention as a data assimilation technique and an ensemble forecast technique for the next generation because of its simple conceptual formulation and relative ease of implementation, e.g., it requires no derivation of a tangent linear operator or adjoint equations, and no integrations backward in time. The EnKF produces the perturbed members of ensemble forecast with consideration of the flow-dependant analysis error. The EnKF was first introduced to a quasi-geostrophic system by Evensen (1994). Recent reviews and overviews of the EnKF are given by Evensen (2007), Evensen (2003), and Hamill (2003), which provide detailed information on the formulation, interpretation and implementation of the EnKF, and now serve a reference document for the basic methodology. However, the EnKF has later been further developed and examined in a large number of published papers to improve the accuracy and reduce the computational cost, such as Ensemble Adjustment Kalman Filter (EAKF) method (Anderson 2001), Ensemble Transform Kalman Filter (ETKF) method (Bishop et al. 2001), Local Ensemble Kalman Filter (LEKF) method (Ott et al. 2004), four-dimensional Ensemble Kalman Filter (4D EnKF) method (Hunt et al. 2004), and Local Ensemble Transform Kalman Filter (LETKF) method (Hunt et al. 2007). In the operational medium-range ensemble forecast, the EnKF method has been adopted as an initial perturbation method at the CMC since 12th January 2005 (Candille et al. 2007; Pellerin et al. 2005) and at the UKMO since March 2006 (Bowler et al. 2008). Also, the JMA

is developing this method with a view to an operational use (Miyoshi and Yamane 2007; Miyoshi et al. 2007; Miyoshi and Sato 2007; Miyoshi and Aranami 2006). The development costs can be reduced by introducing the EnKF methods into own EPS because of unnecessary of the adjoint code and so on.

The ETKF method is based on the application of a Kalman Filter. The application of the ETKF to ensemble forecast can be found in Wang and Bishop (2003) and Wang et al. (2004). Although the ETKF formulation is derived from the EnKF theory which is used for data assimilation, Wang and Bishop (2003) showed how it could be used to generate ensemble perturbations in an idealized observation framework without having to perform data assimilation. The ETKF was further extended to an operational environment with the NCEP operational model and real-time observations by Wei et al. (2006a). In the ETKF framework, the perturbations are dynamically cycled with orthogonalization in the normalized observational space. The ensemble variance is constrained by the distribution and error variance of observations. However, there are still some challenging issues in the ETKF based ensemble with real observations, such as perturbation inflation. Flow dependent inflation factors are hard to construct due to the fact that the number and positions of observations change rapidly from one cycle to the next. Since the ensemble mean from the ETKF has yet to be improved to the level of the analysis from a mature variational Data Assimilation (DA) like NCEP SSI (Spectral Statistical Interpolation; Parrish and Derber 1992), the perturbations generated by the ETKF are added to the analysis field produced by an operational DA system. In addition, the ETKF is much more expensive than breeding in an operational environment with real-time observations.

The ET/ETR method is one of the new initial perturbation methods as in the EnKF method. The ET technique was first proposed by Bishop and Toth (1999) in target observation studies. The research and experiments on using the ET and ET with Rescaling (ETR) for ensemble forecasts first started at the NCEP before 2004, and the initial results were presented in the THORPEX Symposium in 2004 (Wei et al. 2005a). Since then, more experimental results with ensemble using the ET and ETR have been presented and documented in Wei et al. (2005b, 2006b,c). Both ET and ETR are generalizations of the BV method. The ET and ETR methods create globally orthogonal and no paired initial

perturbations. In the ET and ETR methods, the initial perturbations are restrained by the best available analysis variance from the operational DA system and centered around the analysis field generated by the same DA system. In this way, the ensemble system will be consistent with the DA. The perturbations are also flow dependent and orthogonal with respect to the inverse of analysis error covariance. This will overcome some drawbacks in the current operational system resulting from paired perturbations (Wei et al. 2006a). Another advantage is that the ET/ETR technique is considerably cheaper than ETKF if the analysis variance information is available. The ETR method is used at the NCEP, as of December 2007.

There are many quantitative comparisons of these initial perturbation methods using a low-order Lorenz Model (Lorenz 1963, 1996) and/or a low resolution atmospheric circulation model (Houtekamer and Derome 1995; Anderson 1997; Szunyogh et al. 1997; Hamill et al. 2000; Bowler 2006; Descamps and Talagrand 2007). Houtekamer and Derome (1995), using a quasi-geostrophic model, found that the SV, BV, and PO methods produce comparable ensemble forecasts. Bowler (2006), using the low-order Lorenz model, concluded that the BV and SV methods are less efficient than the EnKF method. Descamps and Talagrand (2007), using a low-order Lorenz model and a three-level quasi-geostrophic atmospheric model, found that the EnKF and ETKF methods produce much better ensemble forecast than the SV and BV methods, and that the EnKF method has better performance than the ETKF method.

The BV, ETKF, ET, and ETR methods belong to the same class of methods based on the concept of breeding, involving the dynamical cycling of ensemble perturbations. This is based on the observation that modern NWP analysis methods strongly rely on short range forecasts (Toth and Kalnay 1993). This is supported by Errico et al. (2007) who found that: analysis error characteristics (e.g., statistics) are similar (to first approximation) to those of 6-hr forecast error. Wei et al. (2008), using the NCEP operational analysis/forecast and observation systems, found that the ETR method performed best in the probabilistic scores and in terms of the forecast error explained by the perturbations. A common feature of the second generation techniques is that the initial perturbations are more consistent with the DA system. A good DA system will provide accurate estimates

of the initial analysis error variance for the EPS, while a good, reliable EPS will produce accurate flow dependent background covariance for the DA system.

Furthermore, a multi-analysis ensemble proposed by Richardson (2001a) is also a simple perturbation method. This technique uses the available operational analyses obtained from other NWP centers. The analysis uncertainties caused by the DA system in single NWP center can be reduced by introducing the analyses of the different NWP centers. Evans et al. (2000) and Richardson (2001a) also created the multiple analyses ensemble, which is generated by adding the ECMWF initial perturbations to the UKMO control analysis.

While the initial perturbation methods can reduce the uncertainties of the initial condition under the assumption that the imperfection of the numerical model is small enough (small imperfection of numerical model is not always a relevant assumption), most famous ensemble techniques to reduce the imperfection of numerical model is a multi-model ensemble. This technique is adopted in not only the operational short- and medium-range ensembles but also seasonal- to inter-annual-range ensembles and global warming predictions. This technique combines the independent forecast data obtained by different models in order to represent not only the uncertainties of the initial condition but also the imperfection of the model formulation. Krishnamurti et al. (1999) first introduced the multi-model superensemble to the weather and climate forecasts, and the further researches have been done (Krishnamurti et al. 2000a, 2000b; Krishnamurti et al. 2001). The aim of the multi-model superensemble is to construct a deterministic forecast, which is more skillful than each NWP center's forecast, by collecting some NWP center's single or several (not ensemble) forecasts and then using a multiple regression technique. On the other hand, the poor man's ensemble introduced by Ziehmann (2000) and Buizza et al. (2003) is regarded as collections of some NWP center's single or several (not ensemble) deterministic forecasts as 'one' ensemble forecast.

Also, Evans et al. (2000) and Mylne et al. (2002) quasi-operationally performed a multi-model multi-analysis medium-range ensemble, which combines the multi-analysis ensemble and multi-model ensemble. The use of the multi-model multi-analysis ensemble

is a standard practice in seasonal-range forecasting where the benefits were demonstrated in the PROVOST (PRediction Of climate Variations On Seasonal to interannual Time-scales; Graham et al. 2000) and the European DEMETER (DEvelopment of a Multi-model Ensemble system for seasonal to inTERannual prediction; Palmer et al. 2004) projects. The multi-model multi-analysis techniques are also being used increasingly for short-range ensemble forecasting (Stensrud et al. 1999).

In the operational medium-range ensemble forecast, the maintenance of multi-model is more expensive than that of single model. Only CMC adopted this technique and used two kinds of model in the operational medium-range ensemble forecast until 9th July 2007.

There are other ways to represent model uncertainty, e.g., stochastic physics (Buizza et al. 1999; Palmer 2001) for the operational medium-range ensemble forecast in the ECMEW, CMC, and UKMO, multi-parameterization approach (Houtekamer 1996) for the operational medium-range ensemble forecast in the CMC, and perturbed-parameter approach (Murphy et al. 2004; Stainforth et al. 2005) for climate prediction.

1.2 THORPEX and TIGGE

The NWP technique has progressed rapidly along with the development of the computer science. As shown in Fig. 1.1, a five-day weather forecast today is as reliable as a three-day weather forecast 25 years ago, which is a major scientific advance (Simmons et al. 2006). The skill of the ensemble forecast has also been improved continuously (Fig. 1.2). The World Meteorological Organization (WMO) began THE Observing system Research and Predictability EXperiment (THORPEX; WMO 2005) project in 2005. The THORPEX is a 10-year international research and development programme to accelerate improvements in the accuracy of one-day to two-week high-impact weather forecasts for the benefit of society, the economy and the environment. The THORPEX establishes an organizational framework that addresses weather research and forecast problems whose solutions will be accelerated through international collaboration among academic institutions, operational forecast centers and users of forecast products. The THORPEX will contribute to

the development of a future global interactive forecasting system, which would generate numerical probabilistic products, available to all nations – developed or developing. At the heart of the THORPEX is the research needed for the design and demonstration of a global interactive forecasting system that allows information to flow interactively between forecast users, numerical forecast models, DA systems and observations. Such a system can also be adapted to allow the observing system, observations, assimilation and the model to be configured to maximize forecast skills for specific societal and economic uses.

The THORPEX Interactive Grand Global Ensemble (TIGGE) is a key component of the THORPEX which in turn is a major component of the World Weather Research Programme (WWRP) of the WMO. The objectives of the TIGGE are: (i) enhancing collaboration on ensemble prediction, internationally and between operational centers and universities; (ii) developing new methods to combine ensembles from different sources and to correct for systematic errors (biases, spread over-/under-estimation); (iii) achieving a deeper understanding of the contribution of observation, initial and model uncertainties to forecast error; (iv) exploring the feasibility and the benefit of interactive ensemble systems responding dynamically to changing uncertainty; (v) enabling evolution towards an operational system, the "Global Interactive Forecast System (GIFS)". As of June 2008, ten operational NWP centers: BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, Meteo-France, NCEP, and UKMO, producing daily global ensemble forecasts to 1-2 weeks ahead delivered in near-real-time a selection of forecast data to the TIGGE data archives at the CMA, ECMWF, and National Center for Atmospheric Research (NCAR). This data stored in the TIGGE archives is accumulating at a daily rate of approximately 300GB from these providers. The outline of data accumulated in the TIGGE archives is shown in Table 1.2. The ensemble size accumulated within one day is 498 until 60-hr forecast time, and 487 after 66-hr forecast time (Fig. 1.3). Although intercomparison of the skill of the deterministic forecast has been performed before the TIGGE project started, overall intercomparison of the skill of the ensemble forecast has not been performed. The TIGGE archives, starting from October 2006, enable us to compare the skill of ensemble forecasts (Figs. 1.4 and 1.5). It is well known that the ECMWF has the best performance in the deterministic forecast. It is found that the ECMWF also has the best performance in the

ensemble forecast. In particular, the ECMWF is far superior to the other centers in the early forecast range. This might result from the high quality of the numerical model and the DA techniques in the ECMWF NWP system.

Prior to the TIGGE project, Matsueda et al. (2006, 2007) constructed the Multi-Center Grand Ensemble (MCGE) forecast, consisting of three operational medium-range ensemble forecasts (maximum ensemble size of 86) by the JMA, NCEP, and CMC, on a quasi-operational basis. Although many researches indicated the advantages of multi-model and poor man's ensemble against a deterministic forecast, and the increase in the ensemble size was expected to lead to the more skillful probabilistic forecast, it was difficult to increase the ensemble size by various limitations (e.g., a NWP center performs only several deterministic forecasts, and the data transfer of operational ensemble forecast data is difficult due to its huge data size and the slow speeds of internet connection). Matsueda et al. (2006) have revealed that MCGE predictions are more skillful than single-center ensemble predictions with equal weights among ensemble members and no bias correction using monthly deterministic and probabilistic scores, such as Anomaly Correlation (AC), Root Mean Square Error (RMSE), and Brier Skill Score (BSS) for 500 hPa geopotential height (Z500) and 850 hPa temperature over the Northern Hemisphere (NH, 20°N–90°N) in September 2005. Matsueda et al. (2007) also verified the daily performance of the MCGE for ensemble mean of the Z500 over the NH from August 2005 to February 2006. They found that the improvement in the daily RMSE of the MCGE relative to a single-center ensemble was reduced up to about 20% whether the atmospheric field was more predictable or not. Their works are a pioneer of the intercomparison of operational medium-range ensemble forecast and the construction of a new type of ensemble.

1.3 Purpose of this study

In this study, the overall intercomparisons of five operational medium-range ensemble forecasts: CMC, ECMWF, JMA, NCEP, and UKMO, were performed. A new MCGE was constructed using these operational medium-range ensembles from December 2006 to November 2007, following the method presented by Matsueda et al. (2006, 2007). The

forecast performances of the MCGE were investigated, and the possibility of the construction of the MCGE that outperforms the ECMWF ensemble, having the best performance in the world, was searched. Also, the analysis of an extreme event, atmospheric blocking, was performed using ensemble forecast data, ensemble-based simple sensitivity analysis, and multi-analysis ensemble forecasts.

The outline of the paper is as follows. Chapter 2 introduces the five operational medium-range ensemble prediction systems, and presents the data and analysis procedure used in this study. The overall intercomparisons of the five operational medium-range ensemble forecasts, and the construction and validation of MCGE are performed. In Chapter 3, the analysis of an extreme event, atmospheric blocking, is performed using ensemble forecast data, ensemble-based simple sensitivity analysis, and multi-analysis ensemble forecasts. Based on the results obtained from Chapter 2-3, the forecast performances of the single-center ensemble and MCGE, and the analysis of extreme event using the ensemble forecast are discussed in Chapter 4. Conclusions of this study are summarized in Chapter 5.

Table 1.1: Operational medium-range ensemble prediction systems as of December 2007.

	BOM ¹ Australia	CMA ² China	CMC ³ Canada	CPTEC ⁴ Brazil	ECMWF ⁵ Europe
Model Uncertainty	NO	NO	Stoch. Phys. + Multi-Parameterization	NO	Stoch. Phys.
Initial Perturbation	SVs+LAF	SVs	EnKF+LAF	EOF-based+LAF	SVs+LAF
Int. Prtb. Model Resolution	TL42L19	T21L19	0.9degL28	T126L28	TL42L62
Forecast Model Resolution	TL119L19	T106L19	0.9degL28	T126L28	TL399L62 (0-7d) TL255L62 (8-15d)
Initial UTC	00, 12	00, 12	00, 12	00, 12	00, 12
Forecast Length (days)	10	10	16	15	15
Member/run	33	15	21	15	51
Member/day	66	30	42	30	102
Reference	Bourke et al. (2004)	WMO (2007)	Candille et al. (2007)	Ferreira (2007)	Buizza et al. (2007)

1. BOM: Bureau of Meteorology (<http://www.bom.gov.au/index.shtml>)
2. CMA: China Meteorological Administration (<http://www.cma.gov.cn/english/>)
3. CMC: Canadian Meteorological Center (http://www.weatheroffice.gc.ca/ensemble/index_e.html)
4. CPTEC: Centro de Previsão de Tempo e Estudos Climáticos (<http://www.cptec.inpe.br/>)
5. ECMWF: European Centre for Medium-Range Weather Forecasts (<http://www.ecmwf.int/>)

Table 1.1: (continued).

	FNMOC ⁶ United States	JMA ⁷ Japan	KMA ⁸ Korea	NCEP ⁹ United States	UKMO ¹⁰ United Kingdom
Model Uncertainty	NO	NO	NO	NO	Stoch. Phys.
Initial Perturbation	BVs	SVs	BVs+LAF	ETR+LAF	ETKF+LAF
Int. Prtb. Model Resolution	T119L24	TL63L40	?	T126L28	?
Forecast Model Resolution	T159L30	TL319L60	T213L40	T126L28	N144 (0.833° × 1.25°) L38
Initial UTC	00	12	00, 12	00, 06, 12, 18	00, 12
Forecast Length (days)	10	9	10.5	16	15
Member/run	16	51	17	21	24
Member/day	16	51	34	84	48
Reference	WMO (2007)	JMA (2007)	Park (2007)	Wei et al. (2008)	Bowler et al. (2008)

6. FNMOC: Fleet Numerical Meteorology and Oceanography Center (<https://www.fnmoc.navy.mil/efs/efs.html>)

7. JMA: Japan Meteorological Agency (<http://www.jma.go.jp/jma/indexe.html>)

8. KMA: Korea Meteorological Administration (<http://www.kma.go.kr/intro.html>)

9. NCEP: National Centers for Environmental Prediction (<http://www.emc.ncep.noaa.gov/gmb/ens/>)

10. UKMO: United Kingdom Meteorological Office (<http://www.metoffice.gov.uk/index.html>)

Table 1.2: Outline of data accumulated in the TIGGE archives as of June 2008.

Portal Site	CMA (http://wisportal.cma.gov.cn/tigge/index.jsp) ECMWF (http://tigge.ecmwf.int/) NCAR (http://dss.ucar.edu/pub/tigge/)
Data Provider	BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, Meteo-France, NCEP, UKMO
Data Format	GRIB2
Surface level	psea, ps, prcp, t2m, u10m, v10m, etc.
Pressure level	hgt, tmp, u, v, q (1000, 925, 850, 700, 500, 300, 250, 200 hPa)
Potential Temperature Level	PV (320 K isentropic level)
Potential Vorticity Level	Potential Temperature, u, v (2PV)
Horizontal Resolution	original resolution of each NWP center
Data Period year.mon.day	2006.10.01– (ECMWF, JMA, UKMO) 2007.03.05– (NCEP) 2007.05.15– (CMA) 2007.09.03– (BOM) 2007.10.03– (CMC) 2007.10.25– (Meteo-France) 2007.12.28– (KMA) 2008.02.01– (CPTEC)

Improvement of operational ECMWF forecasts

Anomaly correlation of 500hPa height forecasts

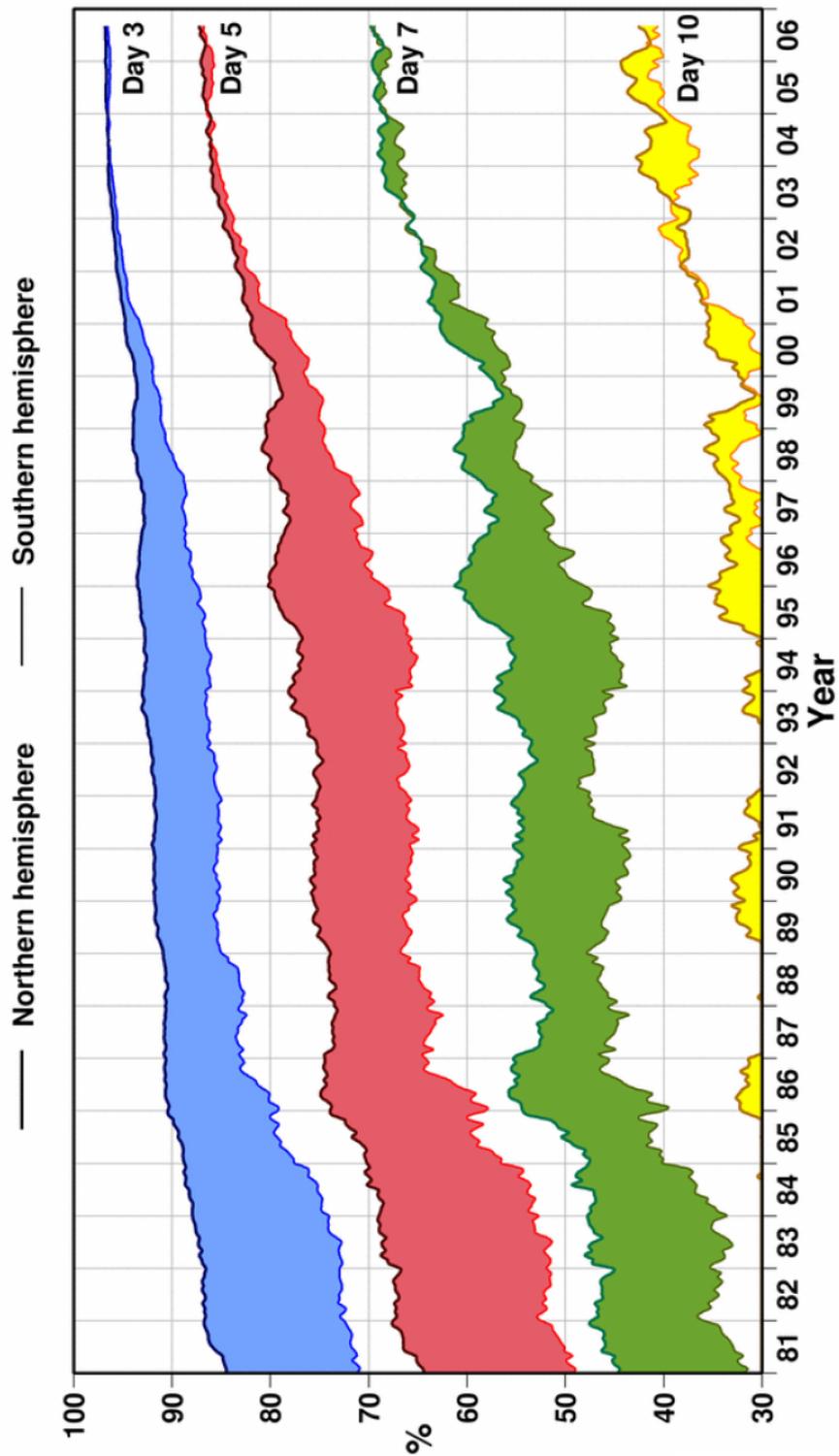


Fig. 1.1: Anomaly correlation coefficients of 3-, 5-, 7-, and 10-day ECMWF 500 hPa height forecasts for the extratropical Northern and Southern hemispheres, plotted in the form of annual running means of monthly-mean scores for the period of 1981–2006. The shading shows the differences in scores between the two hemispheres at the forecast ranges indicated (Simmons et al. 2006).

JMA Ensemble Mean Forecast NH Z500 RMSE (2002DJF-2007SON)

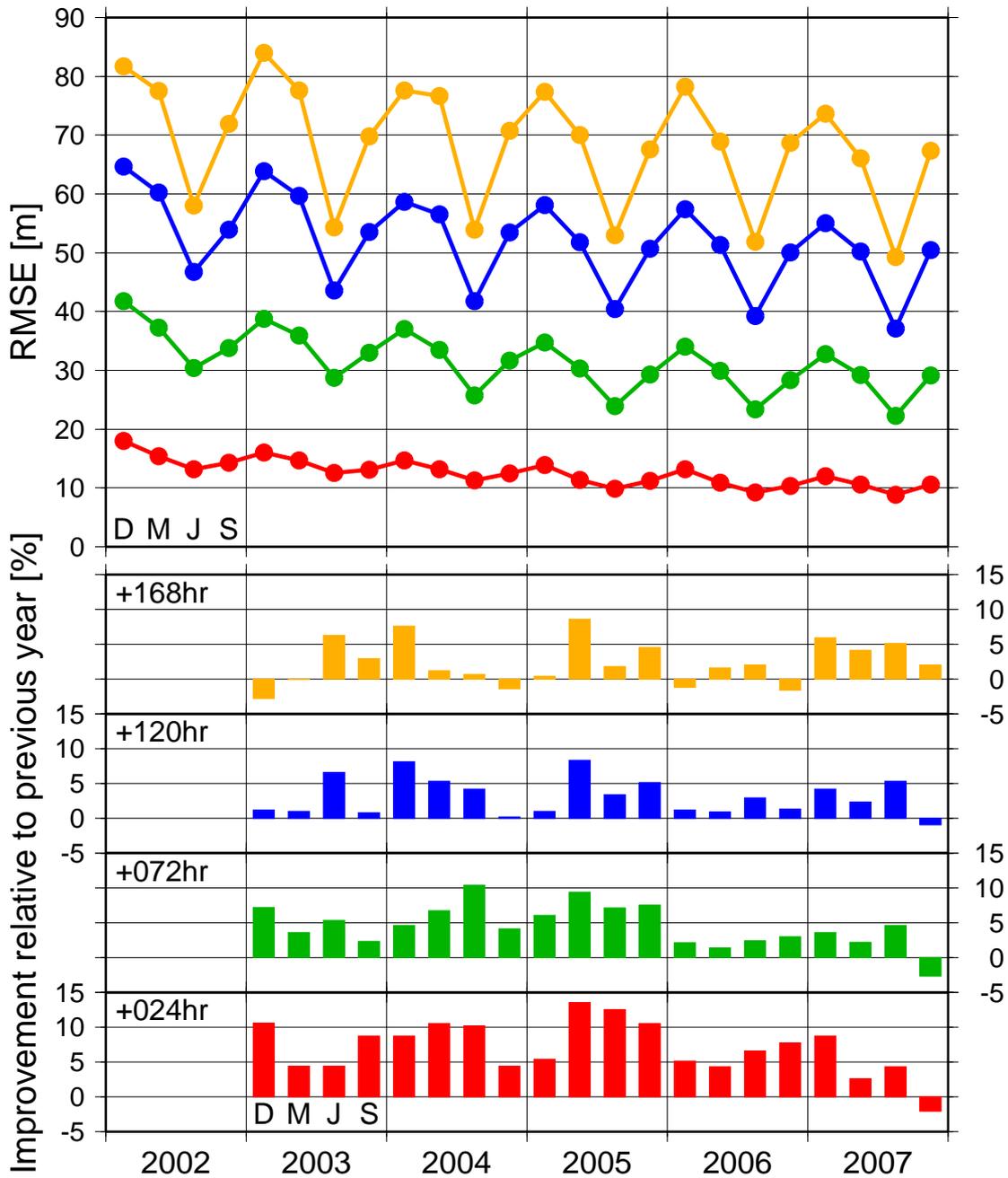


Fig. 1.2: Longest available record of skill of JMA ensemble mean forecast from DJF 2002 to SON 2007. The skill is evaluated by the seasonal RMSE for 500 hPa height over the Northern Hemisphere (20°N–90°N). The D, M, J, and S shown in the figure indicate DJF, MAM, JJA, and SON, respectively. The red, green, blue, and yellow lines are for 24-, 72-, 120-, and 168-hr forecast times, respectively. The red, green, blue, and yellow bars indicate improvement in the seasonal RMSE relative to that of the previous year for 24-, 72-, 120-, and 168-hr forecast times, respectively.

Medium-Range Ensemble Forecasts

Z500 Spaghetti Diagram (5550m)

Initial Time: 20080220

Valid Time: 20080225 12UTC

BOM CMA CMC CPTEC ECMWF JMA KMA NCEP UKMO

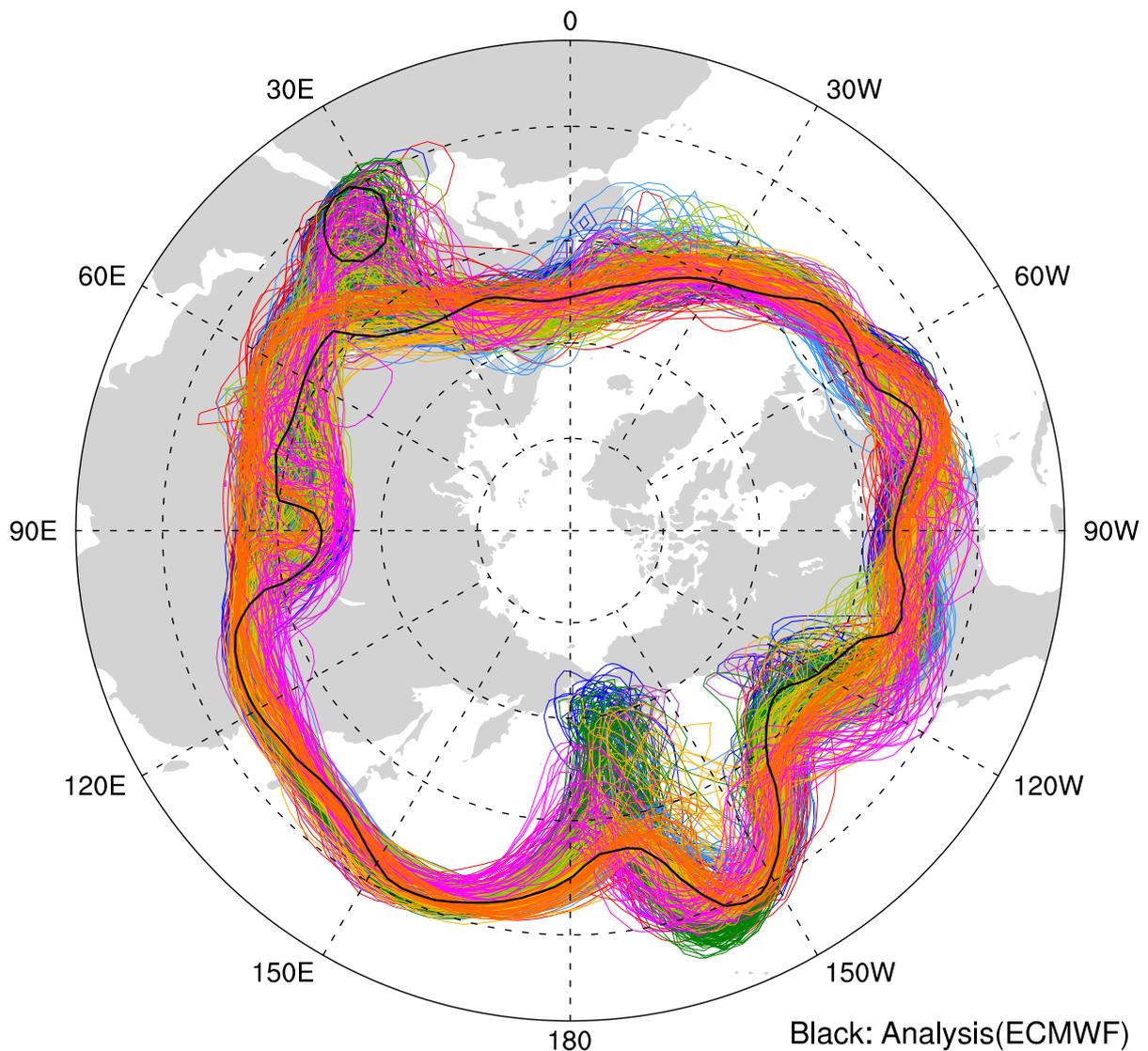


Fig. 1.3: Spaghetti diagram of 500 hPa height (5550 m) for ensemble members of BOM (yellow green), CMA (aqua), CMC (yellow), CPTEC (orange), ECMWF (blue), JMA (red), KMA (pink), NCEP (green), and UKMO (purple), initialized on 20th February 2008, valid 12 UTC on 25th February 2008. The black line is for ECMWF analysis at the valid time. The ensemble size is 487.

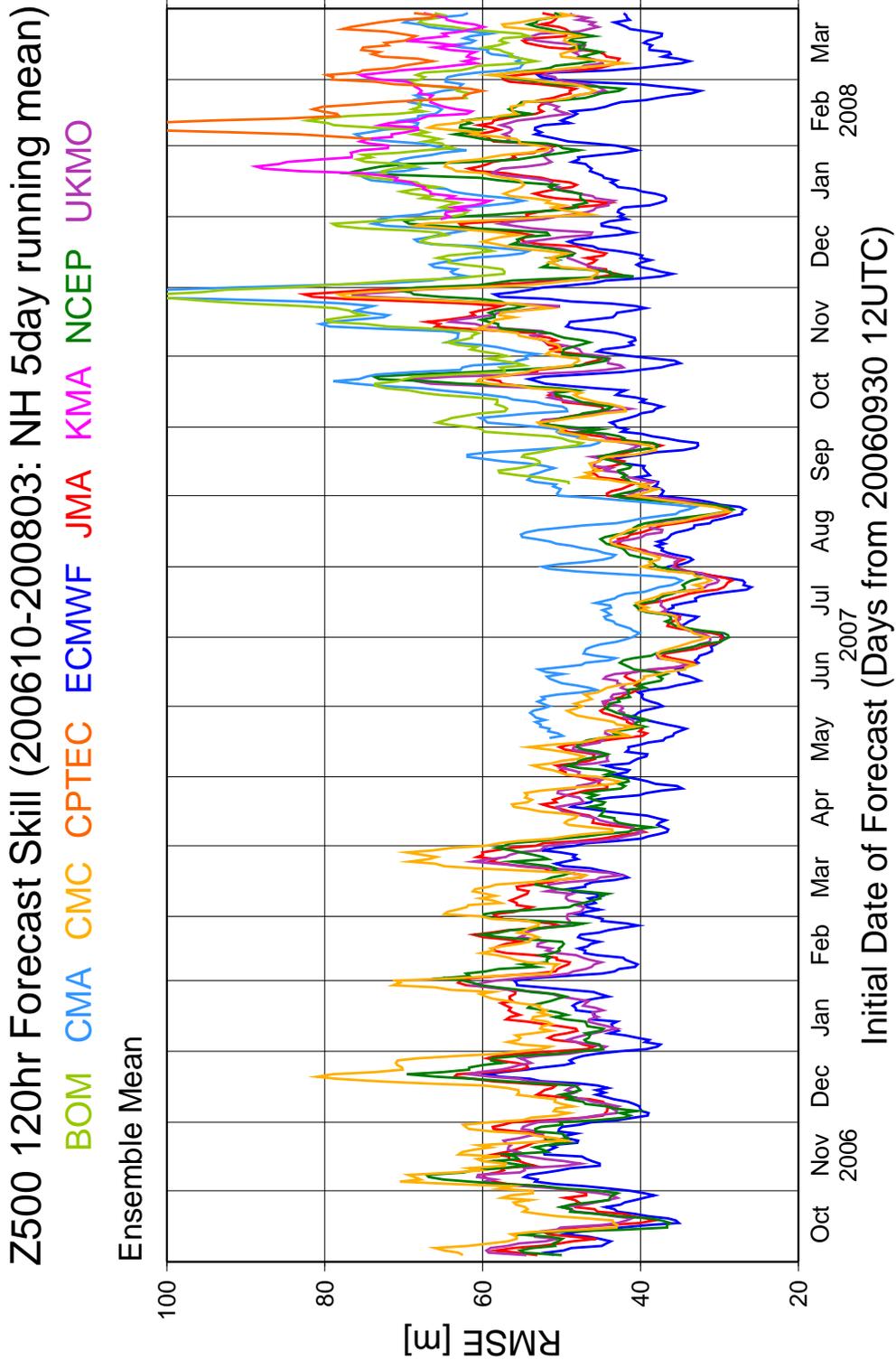


Fig. 1.4: Time series of 120-hr RMSEs of nine single-center ensemble mean forecasts: BOM (yellow green), CMA (aqua), CMC (yellow), CPTEC (orange), ECMWF (blue), JMA (red), KMA (pink), NCEP (green), and UKMO (purple), for 500 hPa height over the Northern Hemisphere (20°N–90°N) from October 2006 to March 2007. 5-days running mean is applied for each solid line.

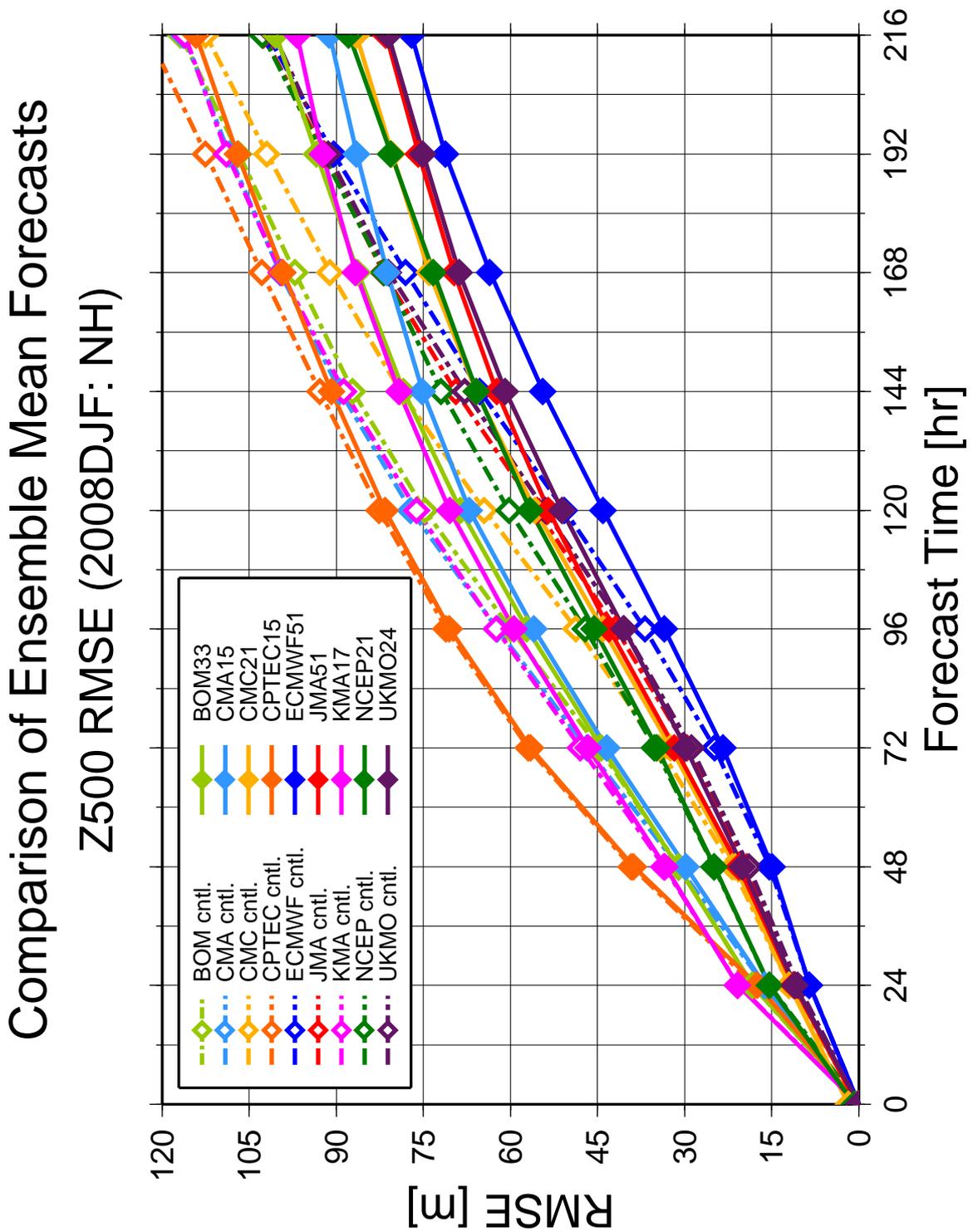


Fig. 1.5: Skills of nine single-center control and ensemble mean forecasts for 500 hPa height over the Northern Hemisphere (20°N–90°N) from December 2007 to February 2008. The dashed-dotted line with open diamond and the solid line with closed diamond are for control and ensemble mean forecasts, respectively. The yellow green, aqua, yellow, orange, blue, red, pink, green, and purple lines are for the BOM, CMA, CMC, CPTEC, ECMWF, JMA, KMA, NCEP, and UKMO, respectively.

Chapter 2

Multi-Center Grand Ensemble using five operational medium-range ensemble forecasts

2.1 Operational medium-range ensemble prediction systems

2.1.1 Canadian Meteorological Center (CMC)

On 24th January 1996, the Canadian Meteorological Center (CMC) started running a medium-range ensemble prediction system with the perturbed observation (PO) method and 8 ensemble members (Houtekamer et al. 1996; Lefaiivre et al. 1997). This set using 8 different versions of the Spectral Finite-Element (SEF; Ritchie 1991; Ritchie and Beaudoin 1994) model was extended to sixteen members on 24th August 1999 by adding 8 different versions of the Global Environmental Multi-scale (GEM; Côté et al. 1998a, b) model. The models differ in their physical parameterizations and their dynamical cores. Eight different configurations of each model are described in Pellerin et al. (2003). The horizontal resolution was increased in June 2001: the SEF models went from TL95 to TL149 with an equivalent increase from 1.875 degrees to 1.2 degrees for the grid point GEM model (Pellerin et al. 2003). On 12th January 2005, the independent assimilation cycles driven by 8 versions of the SEF model and the OI technique were replaced by a 96 member ensemble with each member using the same version of the GEM model and the EnKF

for the operational EPS. The EnKF introduced at the CMC (Houtekamer and Mitchell 2005) is a 4D-DA method that uses a Monte-Carlo ensemble of short-range forecasts to estimate the flow-dependent covariances of the guess fields. There is thus a reservoir of initial conditions which are all suitable to initiate a medium-range EPS. Sixteen of these analyses were arbitrarily chosen to initiate and run the SEF and GEM models up to 10 days. In 10th July 2007, 4 more members were added to produce a 20-member ensemble. All perturbed forecasts and the control forecast are made with the GEM model up to 16 days twice a day. The horizontal resolution of the GEM model is increased to 0.9 degrees. The 20 models have different physical parameterizations, DA cycles and sets of perturbed observations. The detail of the different physics parameterizations is described in the CMC Web site (http://www.weatheroffice.gc.ca/ensemble/verifs/model_e.html).

2.1.2 European Center for Medium-Range Weather Forecasts (ECMWF)

The European Center for Medium-Range Weather Forecasts (ECMWF) is a first NWP center that started to perform a medium-range ensemble forecast in December 1992 together with the NCEP. When first implemented in 1992, the ECMWF-EPS was based on 33 forecasts produced with a T63L19 resolution version of the ECMWF model (Molteni et al. 1996). The initial uncertainties were simulated by starting 32 members from perturbed initial conditions defined by T21L31 perturbations with the fastest growth during the first 36 hours of the forecast range (the singular vectors of the tangent forward model version; Buizza and Palmer 1995). Between December 1992 and December 2005, the EPS has been upgraded several times, benefiting both from any change of the ECMWF DA and forecasting system, and from modifications of the EPS configuration specifically designed to improve the simulation of initial and model uncertainties. It is worth identifying the few of them that affected the EPS overall configuration: In 1996 the system was upgraded to a 51-member TL159L31 system (Buizza et al. 1998), with T42L31 singular vectors. In March 1998, initial uncertainties due to perturbations that had grown during 48 hours previous to starting time (evolved singular vectors; Barkneijer et al. 1999) were included. In October 1998, a scheme to simulate model uncertainties due to random model error

in the parameterized physical processes was introduced (Buizza et al. 1999). As a result of these changes, the upgraded 51-member system had a better level of spread, a more skilful ensemble mean, a higher chance of including the verification analysis inside the forecast distribution and more accurate probabilistic predictions. In October 1999, following the increase of the number of vertical levels in the DA and high resolution deterministic model from 31 to 60, the number of vertical levels in the EPS was increased from 31 to 40. In November 2000, following the increase of resolution of the ECMWF DA and high resolution deterministic forecast from TL319L60 to TL511L60, the EPS resolution was upgraded from TL159L40 to TL255L40 (Buizza et al. 2003), with T42L40 singular vectors. Since April 2003, the ECMWF-EPS has been running twice a day, with 00 and 12 UTC initial times. In 2003, work started to investigate the possibility of further increasing the resolution of the ensemble prediction and to extend the forecast length of ensemble system. During this work, different ensemble configurations were tested to identify the best candidates to replace the operational ensemble configurations. At many NWP centers which perform the medium-range ensemble forecast, the horizontal resolution used in the ensemble forecast is constant with the forecast range. Following the example of the NCEP (Szunyogh and Toth 2002), the ECMWF adopted a configuration that had a variable resolution, with a higher spectral truncation in the early forecast range and a lower spectral truncation in the later forecast range. They call this configuration VAREPS (Variable Resolution Ensemble Prediction System; Buizza et al. 2007). It is hoped that introducing the VAREPS leads to skilful predictions of small-scale, severe-weather events in the early forecast range, and provision of accurate large-scale forecast guidance in the extended forecast range (say beyond forecast day 7) under the same computing resource as traditional EPS using the constant horizontal resolution. On 1st February 2006, the horizontal resolution of the 10-days operational EPS was increased from TL255L40 to TL399L62. On 12th September 2006, the 10-days operational EPS was change to the 15-days VAREPS with a TL399L62 resolution up to forecast day 10 and a TL255L62 resolution between forecast days 10 and 15.

2.1.3 Japan Meteorological Agency (JMA)

In March 2001, the Japan Meteorological Agency (JMA) started running a medium-range ensemble prediction system with the BV method and 25 ensemble members once daily. A low-resolution version of the JMA Global Spectral Model (GSM), T106L40, was integrated from perturbed initial conditions. Thus, the dynamical framework and physical processes are identical with those of the operational GSM except for the horizontal resolution. In March 2006, the JMA-EPS changed the horizontal resolution into TL159L40 with a semi-Lagrangian advection scheme, and the ensemble size into 51. The process of the BV method is as follows. First, short-range forecast errors are added to a control analysis as an initial perturbation. Second, both perturbed analysis and control analysis are integrated up to 12 hours. Then, the difference of the two fields at 12 hours is scaled down to normalize. The normalized perturbation is added again to the analysis at that time. The second and third processes, known as a breeding cycle, are integrated every 12 hours. In the EPS, 50 perturbed initial fields are generated from 25 independent breeding cycles by adding each perturbation positively and negatively. The amplitudes of the initial perturbation, which are confined in the Northern Hemisphere and tropics, are adjusted so that the variance of 500 hPa height over the extra-tropical Northern Hemisphere is equal to 14.5% of the climatological variance, and also reflect analysis errors according to a geographical distribution. The normalized perturbations are also orthogonalized to each other before adding to the analysis. In addition, the JMA-EPS was changed on 21st November 2007. The horizontal resolution, vertical levels, and initial perturbation methods, were changed into TL319, 60 layers, and SV method with T63L40 initial perturbation, respectively. It is noted that the verification period in this study, from December 2006 to November 2007, contains these 10 days in which the JMA-EPS uses the SV method.

2.1.4 National Centers for Environmental Prediction (NCEP)

On 7th December 1992 the National Centers for Environmental Prediction (NCEP) began an operational medium-ensemble forecast using the BV initial perturbation method

(Tracton and Kalnay 1993; Toth and Kalnay 1993). This initial ensemble configuration consisted of high- and low-resolution control forecasts and one pair of perturbed forecast at 00 UTC, and high-resolution control forecast at 12 UTC. The high-resolution control forecasts were run at T126 resolution with 28 vertical levels out to day-6 and day-3 lead times at 00 and 12 UTC, respectively, after which the fields got truncated to T62 resolution and the runs were extended to day-12 lead time due to the limitation of the computer resources. Also, the NCEP used the low-resolution (T62) analysis cycle in the following way: a low-resolution control forecast (as before), and a BV perturbation was added to and subtracted from the analysis to provide initial conditions for two perturbed forecasts. The BV perturbation was based on the rescaled difference between a member of the ensemble and control forecast at 1-day lead time. This difference was then used as the initial perturbation for the next-day ensemble member. With this self-breeding method, efficient perturbations were generated without any extra computational cost beyond running the forecasts themselves. Following the installation at the NCEP of a new Cray C90 supercomputer, the ensemble forecast system was upgraded on 30th March 1994. The NCEP replaced the one BV pair of forecast at 00 UTC by five BV pairs of forecasts at 00 UTC and two BV pairs at 12 UTC, and all the forecasts are extended to 16 days, without the change of model resolution. These 7 BV pairs were based on 7 independent breeding cycles. After that, many small implementations have been done as follows for about ten years: On 15th June 1998, the high-resolution control forecasts were run at T170L42 resolution out to day-7 and day-3 lead times at 00 and 12 UTC, respectively, after which the resolution remained at T62L28. On 11th May 2000, the perturbed ensemble size at 12 UTC was increased from 4 to 10. On 27th June 2000, the horizontal resolution of all perturbed members at 00 and 12 UTC and the low-resolution control run at 12h UTC was increased from T62L28 to T126L42 out to day-3 lead times. On April 2003, the high-resolution control runs up to 3.5 days at 00 and 12 UTC were run at T254L64. On 9th March 2004, a major implementation was done. The ensembles have been available four times daily at a T126L28 resolution up to 7.5 days (a T62L28 resolution for the remaining 8.5 days). On 16th August 2005, the breeding cycle was changed from 24 hr cycle to 6 hr cycle. The horizontal resolution of perturbed member after 7.5 days was changed from T62L28 to T126L28. The high-resolution control forecasts at 00, 06, 12,

and 18 UTC were run at T382L64 up to 7.5 days and at T190L64 after 7.5 days. On 30th May 2006, a new initial perturbation method using the ETR technique was introduced into the NCEP-EPS (Wei et al. 2008). Also, the perturbed ensemble size was increased from 10 to 14, and the low-resolution control runs at T126L28 were added for 06, 12, and 18 UTC. On 27th March 2007, the perturbed ensemble size was increased from 14 to 20, which is most recent implementation.

2.1.5 United Kingdom Meteorological Office (UKMO)

The medium-range ensemble forecast at the United Kingdom Meteorological Office (UKMO) is based on the Met Office Global and Regional Ensemble Prediction System (MOGREPS; Bowler et al. 2008). This system consists of global and regional ensembles, with the global ensemble providing the boundary conditions and initial condition perturbations for the regional ensemble. The UKMO global ensemble forecast was started in June 2005 with 24 members twice daily. Perturbations to the initial conditions are generated using an ETKF method (Bowler 2006). The global Unified Model is run at a horizontal resolution of 0.83° latitude by 1.25° longitude (approximately 90 km in the mid-latitudes), with 38 levels. In the MOGREPS, two stochastic physics schemes are included to represent the effects of structural and sub-grid-scale model uncertainties: the Random Parameters (RP) scheme and the Stochastic Convective Vorticity (SCV) scheme. The RP scheme aims to account for the uncertainty associated with empirical parameters in many physical processes (convection, boundary layer exchange, etc) and to simulate the non-deterministic processes not explicitly accounted for by the different parameterizations. The idea behind RP to treat a selected group of parameters as stochastic variables, is similar to the approach followed by Lin and Neelin (2000) and Bright and Mullen (2002). A total of 8 parameters from 4 different physical parameterizations are included. The physical parameterizations affected by the scheme are: large-scale precipitation, convection, boundary layer, and gravity wave drag. The main aim of the SCV scheme (Gray and Shutts 2002) is to represent a Potential Vorticity (PV) anomaly dipole similar to the one typically associated with a Mesoscale Convective System (MCS), because the PV signatures of MCSs are not well represented in the Unified Model. The SCV scheme is used only by the global ensemble.

2.2 Data

2.2.1 Ensemble forecast data

In this section, five operational medium-range ensemble forecast data: CMC, ECMWF, JMA, NCEP, and UKMO, are used. These data are obtained from the TIGGE portals constructed at the CMA, ECMWF, and NCAR. Analysis period is from December 2006 to November 2007. The configurations of these EPSs in the verification period is same as that in Table 1.1, except for the changes of the ensemble size in the CMC and NCEP, of the numerical model in the CMC and JMA, and of the initial perturbation method in the JMA. Although the horizontal resolutions of the original ensemble forecast data transferred from each NWP center to the TIGGE portal are different from each other, it is possible to interpolate each original horizontal resolution to the common horizontal resolution before we download these ensemble forecast data. In this study, 2.5° latitude \times 2.5° longitude grid is adopted.

2.2.2 ECMWF 40 Year Re-Analysis (ERA40)

In the deterministic and probabilistic verifications, climate data and climatological standard deviation are needed. These were calculated from the ECMWF 40 year reanalysis data (ERA40; Uppala et al. 2005). First, the climatology and climatological standard deviation are calculated for each calendar day. After that, a 60-day low-pass Lanczos filter (Duchon 1979) is applied for the climatology and climatological standard deviation. Although we generally need to use the climatology and climatological standard deviation of its own model to verify the forecast skill, the climatology and climatological standard deviation calculated from the ERA40 are used in this study for simplicity.

2.3 Methods

2.3.1 Statistics on ensemble forecast

The most compact expressions of the information contained in an ensemble forecast are ensemble mean and ensemble spread.

Ensemble mean

The ensemble mean is obtained by averaging all ensemble forecasts:

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n \quad (2.1)$$

where x_n is the ensemble member value for member n , N is the number of ensemble member. This has the effect of filtering out features of the forecast that are less predictable. These features might differ in position, intensity and even presence among the members. The averaging retains those features that show agreement among the members of the ensemble. This is also, but to a lesser extent, the case with the central cluster in the tubing method (Atger 1999).

The averaging technique works best some days into the forecasts when the evolution of the perturbations are dominantly non-linear. During the initial phase, when the evolution of the perturbations has a strong linear element, the ensemble average is almost identical to the control because of the “ mirrored ” perturbations (added to and subtracted from the control run).

Ensemble spread

If 50 ensemble members are quite different from each other, it is obvious that many of them are wrong. If there is a good agreement among the members, there are more reasons to be confident about the forecast and that most of them are close to the truth.

The ensemble spread measures the differences between the members in the ensemble forecast. The ensemble spread is the rms-difference between the ensemble members and

ensemble mean defined as:

$$spread = \sqrt{\frac{1}{D \cdot N} \sum_{d=1}^D \sum_{n=1}^N (x_n^d - \overline{x^d})^2} \quad (2.2)$$

where N is the number of ensemble members, D is the number of grid-points in the spatio-temporal (for seasonal score) or temporal domains (for daily score), namely which indicates all grid points over the Northern Hemisphere (NH, 20°N–90°N) in the verification period. x_n^d is the ensemble member value for member n at the grid-point d , and $\overline{x^d}$ is the ensemble mean at the same grid-point. A small (large) spread indicates low (high) forecast uncertainty. However, a small (large) spread does not necessarily indicate high (low) skill, although it could give an indication of high (low) predictability.

2.3.2 Verification scores for deterministic forecasts

In deterministic verifications, Root Mean Square Error (RMSE) is used to evaluate the forecast skills of the control run and ensemble mean forecast.

Root Mean Square Error (RMSE)

The RMSE is defined by the following equation:

$$RMSE = \sqrt{\frac{1}{D} \sum_{d=1}^D (x_f^d - x_a^d)^2} \quad (2.3)$$

where x_f^d and x_a^d indicate the forecast and analysis values at the grid-point d , respectively. In this study, D is the number of grid points in the spatio-temporal (for seasonal score) or temporal domains (for daily score), namely which indicates all grid points over the NH in the verification period. Each ensemble mean and control forecasts are verified against its own analysis. The control run of each ensemble at initial time is regarded as the analysis data. The RMSE indicates a forecast error, and the RMSE score of zero (0.0) demonstrates a perfect skill. The RMSE is expected to be comparable with the ensemble spread at the same verification time.

In general, Anomaly Correlation (AC) represented by the following equation is also

used in the deterministic verification.

$$\text{AC} = \frac{\sum_{d=1}^D (x_f^d - x_c^d)(x_a^d - x_c^d)}{\sqrt{\sum_{d=1}^D (x_f^d - x_c^d)^2} \sqrt{\sum_{d=1}^D (x_a^d - x_c^d)^2}}, \quad (2.4)$$

where x_f^d , x_a^d and x_c^d indicate the forecast, analysis and climatology values at the grid-point d , respectively. D is the number of total grid-points in the temporal domain. The AC indicates a patterns correlation between forecast and analysis anomalies, so the AC decreases with time. The AC is 1.0 for the perfect forecast. Based on experience with the anomaly correlation, a score above 0.6 suggests that the forecast is sufficiently good, while a score below 0.6 signifies the forecast is not useful. In general, the time when the AC first reaches 0.6 is called the limitation of predictability. Calculating the AC requires not only forecast and analysis data but also climatology. Easily expected from geometric relationship between the AC and RMSE in phase space shown in Fig. 2.1, the AC is sensitive to the choice of the climatological reference, whereas the RMSE is not influenced by climatology. Although the model climatology is different from each other, it is not easy to obtain all model climatologies of the NWP centers. Based on these facts, the deterministic verification is performed using the RMSE.

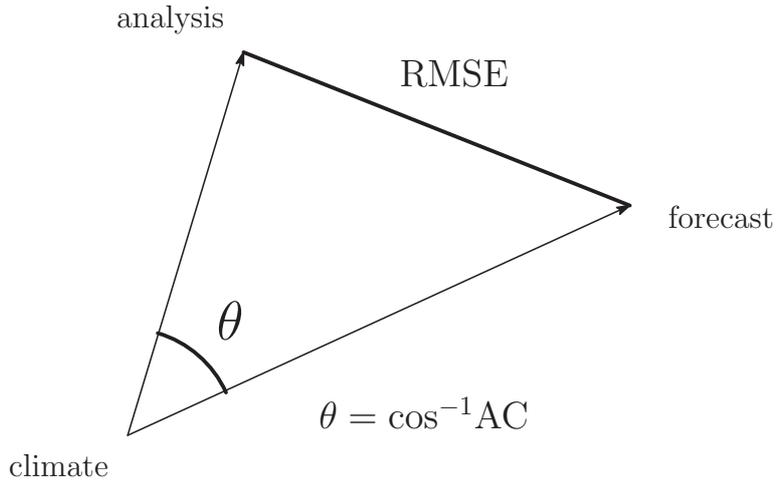


Fig. 2.1: Geometric relationship between the AC and RMSE in the phase space.

2.3.3 Verification score for probabilistic forecasts

In the ensemble forecast, many predictions are performed. One can obtain the occurrence probabilities of the weather events by counting the ensemble members which are included in arbitrary ranked categories. In probabilistic verifications, Ranked Probability Score (RPS; Epstein 1969; Murphy 1971) was used to evaluate the skill of ensemble probabilistic forecast.

Ranked Probability Score (RPS)

One of the most commonly used measure in the probabilistic verification is the RPS (Wilks 2006). The RPS is essentially a generalization of the Brier Score (BS; Brier 1950) to the multi-category situation. That is, RPS is a squared-error score with respect to the observation 1 if the forecast event occurs, and 0 if the event does not occur. However, in order for the score to be sensitive to distance, the squared errors are computed with respect to the *cumulative* probabilities in the forecast and observation vectors.

In this study, the RPS is calculated based on 10 climatologically equally likely categories ($J = 10$). The climatological anomaly of each ensemble member, normalized by a climatological standard deviation, is classified into 10 categories: < -2.0 , $[-2.0, -1.5)$, $[-1.5, -1.0)$, \dots , $[1.0, 1.5)$, $[1.5, 2.0)$, ≥ 2.0 . The predicted and observed probabilities included in the j -th ($j = 1, 2, \dots, J$) category are represented as p_i and o_i . The cumulative predictions and observations, denoted P_m and O_m , are defined as functions of the components of the prediction vector and observation vector, respectively, according to

$$P_m = \sum_{j=1}^m p_j, \quad m = 1, 2, \dots, J, \quad (2.5)$$

and

$$O_m = \sum_{j=1}^m o_j, \quad m = 1, 2, \dots, J. \quad (2.6)$$

Note that since P_m and O_m are both cumulative functions of probability components that must add to one, the final sums P_J and O_J are always both equal to one by definition. The RPS is the sum of squared differences between the components of the cumulative

prediction and observation vectors in Eqs. 2.5 and 2.6, given by

$$\text{RPS} = \frac{1}{J-1} \sum_{m=1}^J (P_m - O_m)^2, \quad (2.7)$$

or, in terms of the predicted and observed vector components p_j and o_j ,

$$\text{RPS} = \frac{1}{J-1} \sum_{m=1}^J \left[\left(\sum_{j=1}^m p_j \right) - \left(\sum_{j=1}^m o_j \right) \right]^2. \quad (2.8)$$

A perfect forecast would assign all the probability to the single p_i corresponding to the event that subsequently occurs, so that the prediction and observation vectors would be the same. In this case, the RPS is equal to zero. Forecasts that are less than perfect receive scores that are positive numbers, so the RPS has a negative orientation. Note also that the final ($m = J$) term in Eqs. 2.7 and 2.8 is always zero, because the accumulations in Eqs. 2.5 and 2.6 ensure that $P_J = O_J = 1$. Therefore, the worst possible score is 1. For two forecast categories ($J = 2$), the RPS is same as the BS. Note that since the last term, for $m = J$, is always zero, in practice it needs not actually to be computed.

In this study, the RPS is calculated for each grid point over the NH in the verification period, and summed over the spatio-temporal domain for seasonal score.

2.3.4 Comparison of scores

In order to compare the skill of forecast with that of a reference forecast, a skill score is usually calculated.

Skill score

The skill score indicates the improvement rate of a forecast relative to a reference forecast. The detail is as follows. For any verification diagnostic, X , the skill of a forecast relative to a reference forecast is given by

$$\text{SS} = \frac{X_r - X_f}{X_r - X_p} \quad (2.9)$$

where X_f is the score of X for the forecast, X_r for the reference forecast and X_p for a perfect deterministic or probabilistic forecast. A skill score has a maximum value of unity (or

100%) for a perfect forecast ($X_f = X_p$) and a value of zero for performance equal to that of the reference ($X_f = X_r$). The SS has no lower limit, with negative values representing poorer skill than the reference. Normally the reference forecast used is a standard baseline such as persistence or climatology. In order to evaluate the performance of the MCGE relative to the ECMWF ensemble, the ECMWF ensemble is used as a reference in this study. Also, the X_p is equal to zero for the RMSE and RPS. The skill score for the RMSE and RPS is described simply as follows:

$$\text{SS} = \frac{X_r - X_f}{X_r} = 1 - \frac{X_f}{X_r} \quad (2.10)$$

2.4 Comparison of CMC, ECMWF, JMA, NCEP, and UKMO ensemble forecasts

2.4.1 Daily and seasonal RMSEs

First, the comparisons among the deterministic skills of five control runs: CMCcntl, ECMWFcntl, JMAcntl, NCEPcntl, and UKMOcntl, and among that of five ensemble mean forecasts: CMC17 (CMC21), ECMWF51, JMA51, NCEP11 (NCEP21), and UKMO24, are conducted using the RMSE for 500 hPa geopotential height (Z500) over the NH from December 2006 to November 2007. It is noted that the ensemble sizes of the CMC and NCEP have changed within the verification period. The single-center ensemble means created in this study are listed in Table 2.1.

Figure 2.2a indicates the time series of 120-hr RMSEs of five single-center control runs for the Z500 over the NH from December 2006 to November 2007. Each forecast was verified against its own analysis (defined as the control run at the initial time) on a regular 2.5 degree grid. It is noted that the 5 day running mean are conducted against the time series of the RMSEs. It is found that each forecast skill varies slowly throughout the verification period irrespective to seasons. The fluctuation of the RMSE is large during the winter season, and small during the summer season. Paying attention to daily scores, it is also found that the RMSEs vary considerably depending on atmospheric flow of the day, as seen in Kimoto et al. (1992), who have investigated the daily scores with single deterministic forecasts from three operational centers: ECMWF, JMA, and NCEP. Although there are some differences among the magnitude of the RMSEs, the RMSEs tend to show similar variations along the atmospheric flow. The ECMWF control run seems to be most skilful than the other control runs, except for the summer season. This might result from the high quality of the numerical model and the DA techniques in the ECMWF NWP system. Also, the CMC control run seems to have worst skill than the other control runs, especially during the winter and spring seasons. There seem to be no apparent differences among the forecast skills of the JMA, NCEP, and UKMO throughout the verification period.

Figure 2.2b indicates the time series of 120-hr RMSEs of five single-center ensemble means for the Z500 over the NH from December 2006 to November 2007. Each ensemble mean forecast was verified against its own analysis (defined as the control run at the initial time) on a regular 2.5 degree grid. It is noted that the ordinate for ensemble means (Fig. 2.2b) is not that for the control runs (Fig. 2.2a). As in the control runs, the RMSEs vary depending on atmospheric flow of the day. Each ensemble mean has better forecast skill than each control run. In terms of the ensemble mean, the ECMWF tends to be most skilful than any other ensemble means, especially during the autumn season. Although the CMC control run has apparent worst skill especially during the winter and spring seasons, the forecast skill of the CMC ensemble mean is almost comparable to that of the JMA, NCEP, and UKMO during the whole verification period. Although the CMC adopted the multi-model (multi-parameterization) ensemble until 9th July 2007 and the EnKF method, a part of this improvements in the RMSE of the CMC might have resulted from the use of these advanced techniques.

Seasonal mean scores are shown in Fig. 2.3 and 2.4. For all ensembles in four seasons, the RMSE of the ensemble mean forecast is lower than the RMSE of the corresponding control forecast. During the winter season, it is found that the ECMWF has the best control and ensemble mean forecasts at all the lead times. In particular, the ECMWF is far superior to the other centers in the early forecast range. The high performance of numerical model and DA system in the ECMWF NWP system is recognized. Also, the CMC has the worst forecast skill in terms of the control run and ensemble mean forecast. In terms of the control run, the UKMO has second-best forecast skill until 168-hr lead time. The NCEP and JMA has worse forecast skill than the ECMWF in the early forecast range. The JMA has second-worst forecast skill at all the lead times, whereas the NCEP has comparable with or better forecast skill than the UKMO after 168-hr lead time. On the other hand, in terms of the ensemble mean, the UKMO has the better forecast skill than the JMA and NCEP at all the lead times. The NCEP is slightly more skilful than the JMA at the medium-lead time, while the JMA is slightly more skilful than the NCEP at the short- and long-lead times

During the spring season, the NWP center which has the best or worst performance

is the same as that during the winter season. However, the JMA, NCEP, and UKMO have similar forecast skills in terms of the ensemble mean, while the control run of the NCEP is more skilful than that of the JMA and UKMO after 120-hr lead time.

During the summer season, the ECMWF has the best forecast skill in terms of the control run and ensemble mean, except for the control run after 168-hr lead time . It is interesting that the CMC ensemble mean is almost comparable to the JMA, NCEP, and UKMO ensemble means although the CMC control run has the worst forecast skill. The improvement of the CMC ensemble mean might be due to upgrade of the CMC-EPS on 9th July 2007. Also, the JMA ensemble mean has the second-best forecast skill after 120-hr lead time, whereas the NCEP ensemble mean has the worst forecast skill.

During the autumn season, the ECMWF has the best performance, except for the control run after 168-hr lead time. Compared with the other seasons, the ECMWF is far superior to the other centers in the early forecast range. The skill of the ECMWF at 120-hr lead time is comparable with that of the remains at 96-hr lead time. Although the control runs of the CMC, JMA, NCEP, and UKMO have different forecast skill, these ensemble mean forecasts have similar forecast skill.

2.4.2 Relationship between RMSE and ensemble spread

Figure 2.5 shows the Z500 ensemble mean and spread for the CMC, ECMWF, NCEP, UKMO, and JMA ensembles initialized at 17th November 2007 valid at 12 UTC on 22nd November 2007. A atmospheric blocking occurred over the west coast of North America on 22nd November 2007. A jet stream widely meandered over north Atlantic-Europe region. In terms of the hemispheric characteristics, it is easily found that the magnitudes of ensemble spread are different from each other. As shown later, the JMA ensemble has the largest spread, whereas the NCEP ensemble has the smallest spread. The spatial pattern of each spread is also different from each other. For example, the JMA ensemble showed high uncertainty on a trough over the north Pacific, whereas the ECMWF ensemble at 12 UTC did not show notable uncertainty on a trough. This indicates that the JMA members predicted a mutually different trough, unlike the ECMWF members. In other words,

this indicates a possibility that JMA ensemble captures an atmospheric phenomenon not predicted in the ECMWF ensemble. In contrast, all ensembles have large uncertainties over the Scandinavian Peninsula and the south of the Greenland, although the spatial extent of large uncertainties is somewhat different. Many of ensembles also have large spread over the south of Alaska. These areas correspond to the upstream of a ridge.

Figure 2.6 illustrates the time series of 120-hr spreads of five single-center ensembles for the Z500 over the NH from December 2006 to November 2007. As in the RMSE, the ensemble spreads vary depending on atmospheric flow of the day. The magnitude of the spread is large (small) during the winter (summer) season. Also, there are well-defined differences among five spreads. The JMA has the largest spread in the five single-center ensembles, while the NCEP has the smallest spread in the five single-center ensembles. The spreads of the CMC, ECMWF, and UKMO are comparable with each other. These differences might result from the different policy to the EPS. In fact, the JMA-EPS has a policy that the ensemble spread is comparable with the RMSE in latter half of the forecast period. So, the JMA-EPS has larger initial perturbations than the other NWP centers.

Seasonal means of the ensemble spreads are shown in Fig. 2.7 and 2.8. The following features are seen regardless of the season. The features at 120-hr lead time, as seen in Fig. 2.6, are shown at the other lead time, except for the short-lead time. The NCEP has the smallest ensemble spread after 48-hr lead time, while the JMA has the largest ensemble spread at all the lead times. The ensemble spreads of the CMC, ECMWF, and UKMO are comparable with each other, although the UKMO tends to have smaller ensemble spread at the latter half of the lead time except for the winter season.

As described in section 2.3, the ensemble spread is expected to be comparable with the RMSE at the same verification time. A good relationship between the RMSE and ensemble spread is required for a good EPS. From this viewpoint, the relationships between the RMSE and ensemble spread are investigated using daily data.

Figures 2.9–2.13 illustrate a scatter diagram of the RMSE versus the ensemble spread for each EPS at 24-, 72-, 120-, 168-, and 216-hr lead times. Easily expected from Figs.

2.2b and 2.6, the characteristics are different from each other. Also, even if the certain center is paid attention, each characteristic at different lead time is different from each other.

Figure 2.9 shows a scatter diagram at 24-hr lead time. The red, blue, green, and yellow circles are for DJF, MAM, JJA, and SON, respectively. The spreads of the JMA and UKMO are larger than the RMSEs of them. The scatter diagrams of the CMC and NCEP "stand", that is, the magnitudes of the spreads of the CMC and NCEP are almost same regardless of the season and magnitudes of the RMSEs. The spread of the CMC tends to be larger (smaller) than the RMSE of CMC during the JJA (DJF) and SON (MAM). The scatter diagram of the NCEP indicates a good relationship between the spread and RMSE during JJA and SON, while the spread of the NCEP tends to be smaller than the RMSE of the NCEP during DJF and MAM. In terms of the ECMWF, the spread tend to be a little larger than the RMSE. Furthermore, the scatter diagram of the ECMWF does not extend widely like the other centers. At 24-hr lead time, the ECMWF seems to indicate a good relationship between the spread and RMSE.

Figure 2.10 shows a scatter diagram at 72-hr lead time. The CMC indicates a good relationship between the spread and RMSE during JJA and SON, while the spread of the CMC tends to be smaller than the RMSE of the CMC during DJF and MAM. The spread of the ECMWF tends to be slightly larger than the RMSE of the ECMWF regardless of the season, as in 24-hr lead time. The spread of the JMA is larger than the RMSE of JMA, while the spread of the NCEP is smaller than the RMSE of the NCEP. In terms of the UKMO, the spread is comparable to the RMSE. At 72-hr lead time, the UKMO indicates a good relationship between the spread and RMSE regardless of the season.

Figure 2.11 shows a scatter diagram at 120-hr lead time. The scatter diagrams of the CMC and NCEP indicate the same feature as that at 72-hr lead time. The spread of the JMA is larger than the RMSE of the JMA, as in 24- and 72-hr lead times. The spread of the UKMO is slightly smaller than the RMSE of the UKMO. For ECMWF, the scatter diagram indicates a good relationship between the spread and RMSE. At 120-hr lead time, the ECMWF seems to indicate a good relationship between the spread and

RMSE.

Figure 2.12 shows a scatter diagram at 168-hr lead time. The scatter diagrams for all EPS indicate the same features as that at 120-hr lead time. The ECMWF seems to indicate a good relationship between the spread and RMSE. All EPS, however, has an interesting common feature. The magnitude of the spread peaks out when the magnitude of the RMSE is particularly large (e.g., over 80 m). This might indicate that there is a limit in the width of the atmospheric phenomena which the single-center ensemble can capture. This feature also appears for the CMC, JMA, NCEP, and UKMO at 120-hr lead time and for the CMC and NCEP at 72-hr lead time.

Figure 2.13 shows a scatter diagram at 216-hr lead time. As in the 168-hr lead time, the NCEP and UKMO have the small spreads, compared to the RMSE. The spread of the CMC is also smaller than the RMSE of the CMC. The ECMWF and JMA indicate a good relationship between the spread and RMSE. However, the spread against the large RMSE tends to be small for all EPS, as in 120- and 168-hr lead times.

In terms of the relationship between the ensemble spread and RMSE, it is found that the ECMWF-EPS has an advantage over the other EPSs for the whole forecast range. Also, the JMA shows a good relationship between the spread and RMSE in the medium forecast range. The UKMO has the large and small spreads in the early and medium forecast ranges, respectively. The CMC and NCEP tend to have the small spread in all the forecast ranges.

2.5 Construction of Multi-Center Grand Ensemble

As introduced in section 1, there are various types of the multi-model ensemble. It had been difficult to construct the multi-model ensemble consisting of the operational medium-range ensemble forecasts because of the huge data size and the internet connection speeds. However, with the recent increase of the internet connection speeds and the beginning of the TIGGE project, one can get operational medium-range ensemble forecast data compared with the past. Prior to the TIGGE project, Matsueda et al. (2006, 2007) quasi-operationally collected three medium-range ensemble forecast data: CMC, JMA, and NCEP, and constructed the new multi-model ensembles. This new type of multi-model ensemble is called Multi-Center Grand Ensemble (MCGE). They investigated the forecast skills of the MCGE, and indicated that the MCGE outperforms the single-center ensembles in terms of the deterministic and probabilistic verifications. It has been more than one year since the TIGGE database started. One can easily get various operational medium-range ensemble forecast data for long period. It is well known that the ECMWF has the best forecast skill in the deterministic and ensemble forecasts, as shown in the above subsection. With recent advancement of NWP technique, however, it is difficult for each NWP center to develop the more skilful EPS compared with the past. It has been recognized that new ensemble techniques: multi-model ensemble, multi-analysis ensemble and so on, are very important to develop the more skilful EPS. This can be easily imagined from the fact that the ECMWF, which has the best forecast skill, actively promotes the TIGGE project which has huge ensemble data. In Matsueda et al. (2006, 2007), the MCGE was constructed using the CMC, JMA, and NCEP medium-range ensembles, which have a comparable forecast skill. In this study, the MCGE is constructed using the best and second-best ensembles: CMC, ECMWF, JMA, NCEP, and UKMO. One of the main themes in this study is "Can MCGE outperform the ECMWF ensemble using the operational medium-range ensemble forecast data obtained in the TIGGE database?" The MCGEs are constructed by combining the five medium-range ensemble forecasts: CMC, ECMWF, JMA, NCEP, and UKMO with equal weights and no bias correction. The MCGEs created in this study are listed in Table 2.2.

2.6 Verifications of forecast skill of Multi-Center Grand Ensemble

2.6.1 Deterministic verifications

Figure 2.14 illustrates the RMSE skill score of the MCGEs relative to the ECMWF ensemble for the Z500 over the NH. During DJF 2007, it is found that the MCGEs can outperform the ECMWF ensemble at least in the medium forecast range (day 6–9). Although the ECMWF ensemble outperforms the MCGE51 and MCGE154 up to 144-hr lead time, the MCGE51 and MCGE154 outperform the ECMWF ensemble after 144-hr lead time. The forecast time when the MCGE279 starts to outperform the ECMWF ensemble is 156 hr. The seasonal RMSE is reduced by up to 2% by constructing the MCGEs. Although the improvement in the RMSE is a few percents in the medium forecast range, it is interesting that the MCGE51 outperforms the ECMWF ensemble by replacing many members in the ECMWF ensemble with the members in the other centers which are not more skillful than the ECMWF members. The disadvantage of the MCGEs in the early forecast range might be due to the model biases. The improvement rates of the MCGE51 and MCGE154 relative to the ECMWF ensemble are comparable to each other, except after 192-hr lead time. Although the introduction of the LAF method in the MCGE279 tends to lead to degradation of the forecast skill especially in the early forecast range, the MCGE279 has a comparable skill with the MCGE51 and MCGE154 in the medium forecast range. This indicates that the demerit of the LAF method disappears in the long-range ensemble forecast, as shown in the past study (Hoffman and Kalnay 1983).

Similar results are obtained for MAM 2007. The MCGEs at least outperform the ECMWF ensemble after 156-hr lead time. The seasonal RMSE is reduced by up to 2% by constructing the MCGEs. The MCGE154 is slightly more skillful than MCGE51 for all forecast steps. Although the MCGE279 has the worst score in early forecast range, the MCGE279 has a comparable skill with the MCGE51 and MCGE154 in the medium forecast range.

A little different results are obtained for JJA 2007. The MCGE51 and MCGE154 at

least outperform the ECMWF ensemble after 96-hr lead time, whereas the MCGE279 outperforms it after 120-hr lead time. The seasonal mean improvement rate of the MCGE51 and MCGE154 relative to the ECMWF ensemble is roughly 2% with a slight decrease after 120-hr lead time. Also, the MCGE154 has a slightly better skill than the MCGE51 for the whole forecast range. Although the MCGE279 starts to outperform the ECMWF more late than the MCGE51 and MCGE154, the seasonal mean improvement rate of the MCGE279 relative to the ECMWF after 192-hr lead time is comparable with that of the MCGE51 and MCGE154.

During SON 2007, the effect of the MCGE seems to be small for the whole forecast range compared to the other seasons. The forecast time when the MCGEs start to outperform the ECMWF ensemble is roughly 180 hr. Although the seasonal mean improvement rate of the MCGEs relative to the ECMWF ensemble after 192-hr lead time is comparable to that of the other seasons, the MCGEs up to 180-hr lead time have lower improvement rate (of course, it indicates a deterioration relative to the ECMWF ensemble) than the other seasons. As in Fig. 2.4, the ECMWF ensemble during SON 2007 had much better forecast skill than the other single-center ensembles in the early forecast range. This may lead to the notable disadvantage of the MCGE relative to the ECMWF ensemble in the early forecast range.

Figure 2.15 shows daily RMSE improvement of the MCGE51 relative to the ECMWF ensemble for 168-hr lead time. As is not shown in the other lead times, it is found that the MCGE51 can reduce the forecast error up to approximately 20% whether the atmospheric field is more predictable or not in the medium forecast range. Also, the forecast skill tends to be improved when the ECMWF ensemble has especially large forecast error, whereas the forecast skill leads to a relatively large deterioration when the ECMWF ensemble has small forecast error. These results indicate that the MCGE can avoid the worst forecast skill instead of abandonment of the best score.

Figure 2.16 illustrates a scatter diagram of the RMSE versus the ensemble spread for the MCGE51 at 24-, 72-, 120-, 168-, and 216-hr lead times. In the early forecast range of 24- and 72-hr lead times, the ensemble spread is larger than the RMSE. This result

is easily expected from the fact that the MCGE consists of five single-center ensembles. Also, a part of large spread might be due to large model biases in each ensemble data. However, the ensemble spread is comparable to the RMSE with lead time. Compared with the relationship for the ECMWF in Figs. 2.9(b)-2.13(b), the "standing" relationship in the ECMWF ensemble at 216-hr lead time, appearing around the maximum spread, seems to disappear in the MCGE due to effect of multi-model.

2.6.2 Probabilistic verifications

Figure 2.17 illustrates the RPS skill score of the MCGEs relative to the ECMWF ensemble for the Z500 over the NH. During DJF 2007, it is found that the MCGEs can outperform the ECMWF ensemble in at least medium forecast range, as in the deterministic verification. The MCGE154 outperforms the ECMWF ensemble after 96-hr lead time, whereas the others outperform the ECMWF after 120-hr lead time. The MCGE154 has the best forecast skill in all the MCGEs for the whole forecast range. It is interesting that the forecast time when the MCGEs start to outperform the ECMWF ensemble in the probabilistic forecast is earlier than that in the deterministic forecast. Also, the MCGEs reduced the seasonal RPS by up to 3%. The daily RPS can be reduced by up to 15% (not shown). The relatively large degradation of the RPS up to 72-hr lead time might be due to the model biases. Although the MCGE279 has the worst score in the early forecast range, the MCGE279 has a comparable skill with the MCGE154 in the medium forecast range. The MCGE51 has the worst score in all the MCGEs after 168-hr lead time.

During MAM 2007, the MCGEs have a comparable skill with the ECMWF ensemble in 96-hr to 120-hr lead time, and have better skill than the ECMWF ensemble after 120-hr lead time. The construction of the MCGE leads to the reduction of the seasonal RPS by 4%. The daily RPS can be reduced by up to 20% (not shown). Also, the MCGE154 has the best forecast skill in all the MCGEs for the whole forecast range. Unlike DJF 2007, the MCGE279 does not have the worst score even in the early forecast range. The MCGE51 has the worst score in all the MCGEs after 120-hr forecast time, as in DJF 2007. It is noted that the further degradations of the RPS appear at 72-hr lead time. The model biases seem to strongly affect the forecast skill up to 72-hr lead time.

During JJA 2007, the effect of the MCGE appears even in the early forecast range. However, the MCGEs become once less skilful than the ECMWF ensemble at 72-hr forecast range. This feature seems to be due to the strong model biases up to 72-hr lead time. After 120-hr lead time, the seasonal RPS is consistently reduced by up to approximately 4%. The worst score for the MCGE51 also appears after 120-hr lead time. The MCGE279 is comparable to the MCGE154 in 144-hr to 168-hr lead time, and is more skilful than that after 168-hr lead time.

During SON 2007, the MCGEs can at last outperform the ECMWF ensemble in the medium forecast range. In particular, the MCGE51 outperforms the ECMWF ensemble at only 216-hr lead time. The MCGE154, which has the best skill in all the MCGEs, has a almost comparable skill with the ECMWF in 120-hr to 156-hr lead time. After that, the MCGE154 starts to outperform the ECMWF ensemble. The MCGE279, which includes the LAF method, is comparable with the MCGE154 after 192-hr lead time. The MCGE154 and MCGE279 reduce the seasonal RPS by up to 2%, whereas the MCGE51 reduces it by up to 1%. The further degradations of the RPS at 72-hr lead time also appear. As in the deterministic verification, much better forecast skill of the ECMWF ensemble than the other ensembles may lead to further disadvantage of the MCGE relative to the ECMWF ensemble in the early to medium forecast ranges.

Table 2.1: Ensemble means created in this study. Left column is abbreviated ensemble name. The numbers listed in the table are ensemble size of each single-center ensemble. The numbers in parentheses indicate initial UTC.

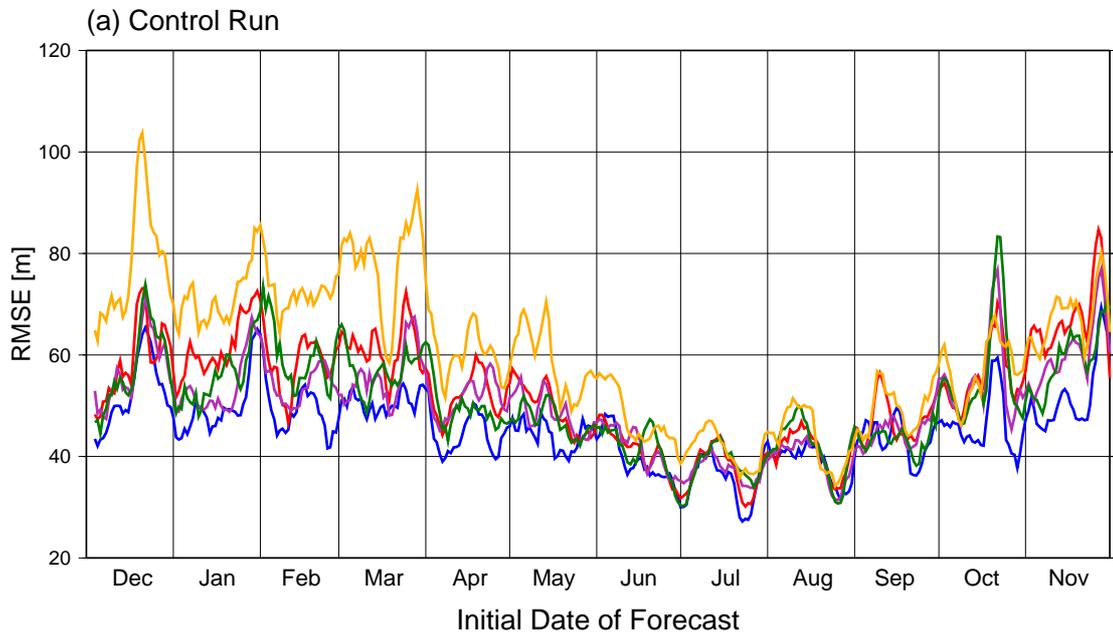
Name	Period year.mon	CMC mem. (UTC)	ECMWF mem. (UTC)	JMA mem. (UTC)	NCEP mem. (UTC)	UKMO mem. (UTC)
CMC17	2006.12–2007.08	17 (12)	–	–	–	–
CMC21	2007.09–2007.11	21 (12)	–	–	–	–
ECMWF51	2006.12–2007.11	–	51 (12)	–	–	–
JMA51	2006.12–2007.11	–	–	51 (12)	–	–
NCEP11	2006.12–2007.02	–	–	–	11 (12)	–
NCEP21	2007.03–2007.11	–	–	–	21 (12)	–
UKMO24	2006.12–2007.11	–	–	–	–	24 (12)

Table 2.2: Configuration of MCGEs created in this study. Left column is an abbreviated ensemble name. The numbers listed in the table are ensemble size of each single-center ensemble included in each MCGE. The numbers in parentheses indicate initial UTC.

Name	CMC (UTC)	ECMWF (UTC)	JMA (UTC)	NCEP (UTC)	UKMO (UTC)
MCGE51	11 (12)	10 (12)	10 (12)	10 (12)	10 (12)
MCGE154	17 (12)	51 (12)	51 (12)	11 (12)	24 (12)
MCGE279	34 (00, 12)	102 (00, 12)	51 (12)	44 (18, 00, 06, 12)	48 (00, 12)

Z500 120hr Forecast Skill (200612-200711: NH 5day running mean)

CMC ECMWF JMA NCEP UKMO



Z500 120hr Forecast Skill (200612-200711: NH 5day running mean)

CMC ECMWF JMA NCEP UKMO

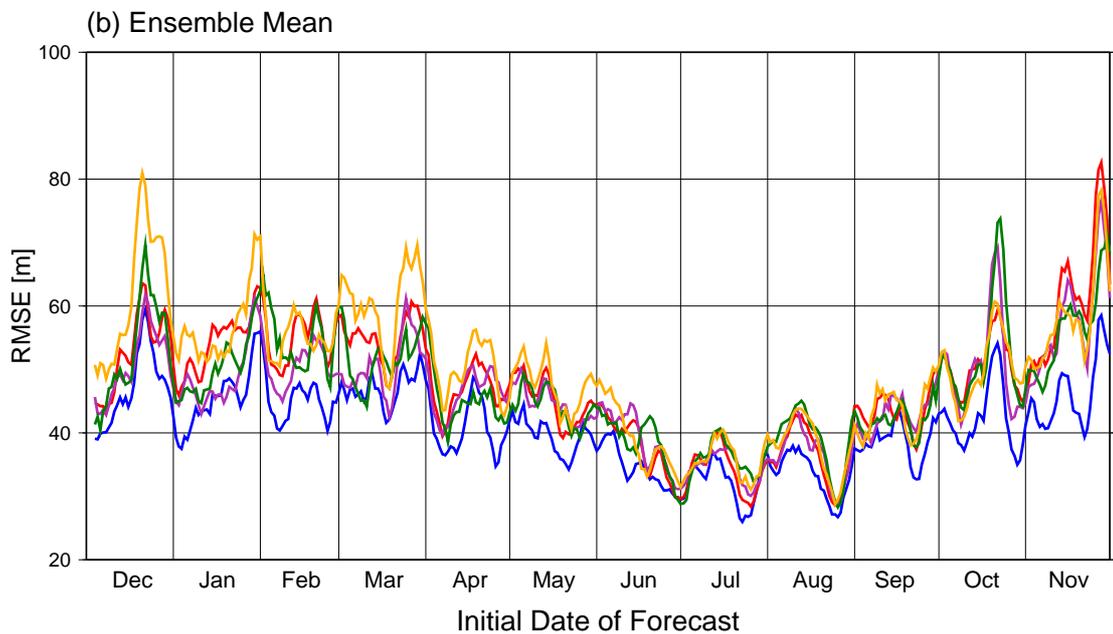
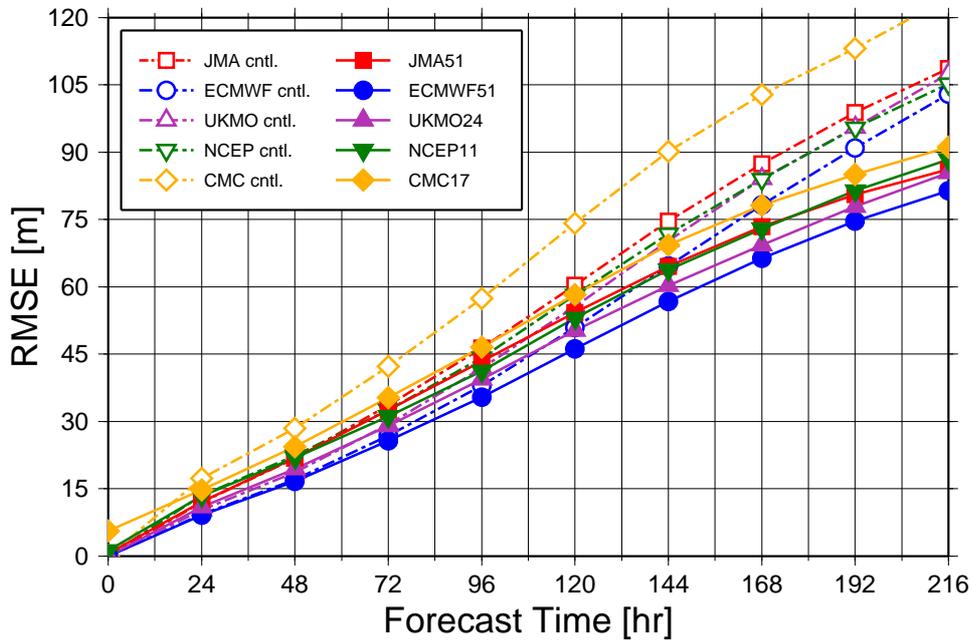


Fig. 2.2: Time series of 120-hr RMSEs of (a) five single-center control runs and (b) five single-center ensemble mean forecasts for 500 hPa height over the Northern Hemisphere (20°N – 90°N) from December 2006 to November 2007. Yellow, blue, red, green, and purple solid lines are for CMC, ECMWF, JMA, NCEP, and UKMO, respectively. 5-days running mean is applied for each solid line.

Comparison of Ensemble Mean Forecasts Z500 RMSE (2007DJF: NH)



Comparison of Ensemble Mean Forecasts Z500 RMSE (2007MAM: NH)

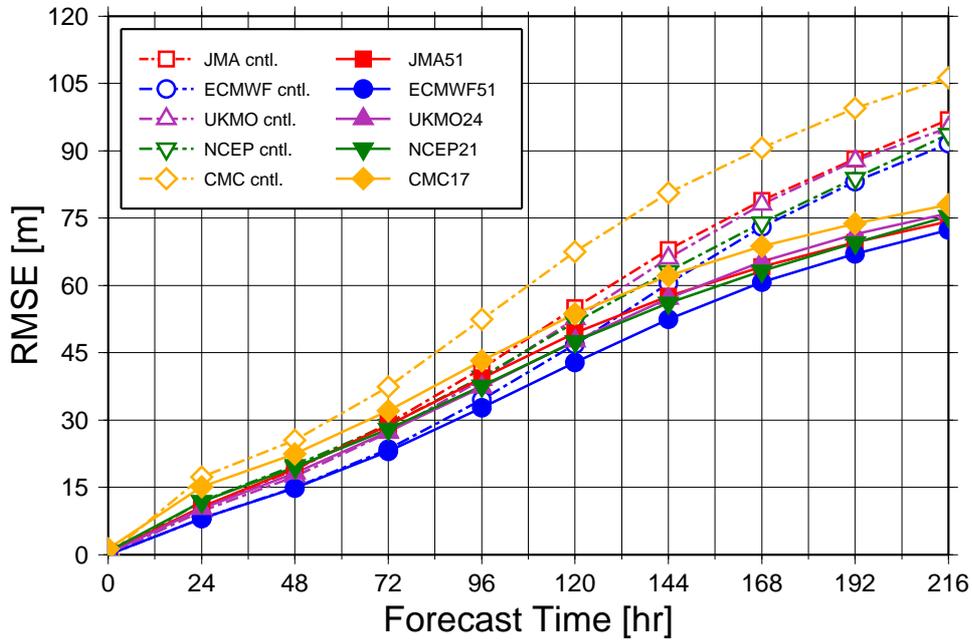
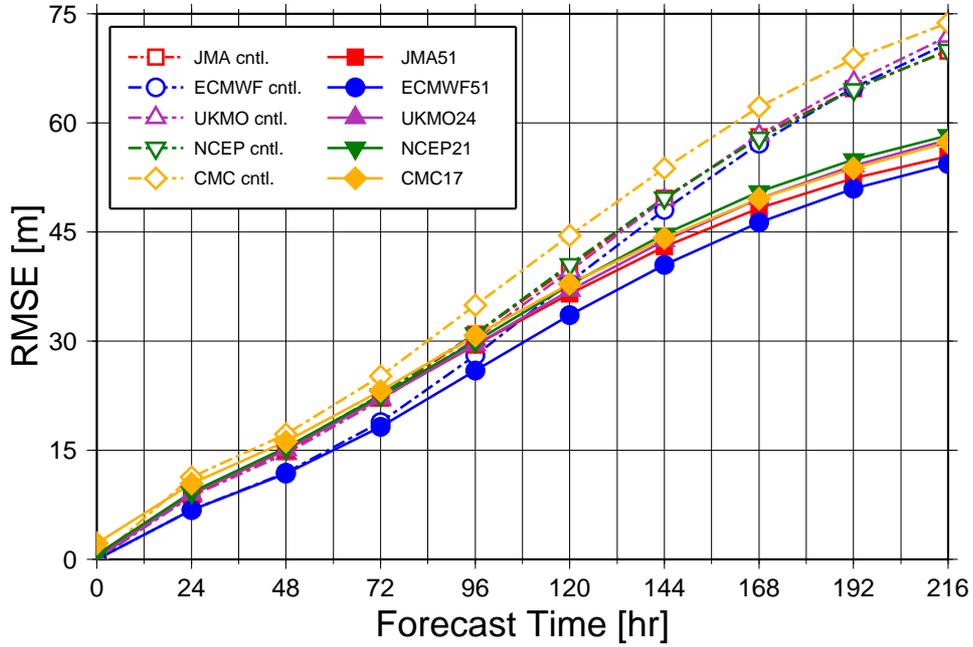


Fig. 2.3: Skills of five single-center control and ensemble mean forecasts for 500 hPa height over the Northern Hemisphere (20°N–90°N) from December 2006 to February 2007 (upper) and from March 2007 to May 2007 (lower). Yellow, blue, red, green, and purple solid lines are for CMC, ECMWF, JMA, NCEP, and UKMO, respectively.

Comparison of Ensemble Mean Forecasts Z500 RMSE (2007JJA: NH)



Comparison of Ensemble Mean Forecasts Z500 RMSE (2007SON: NH)

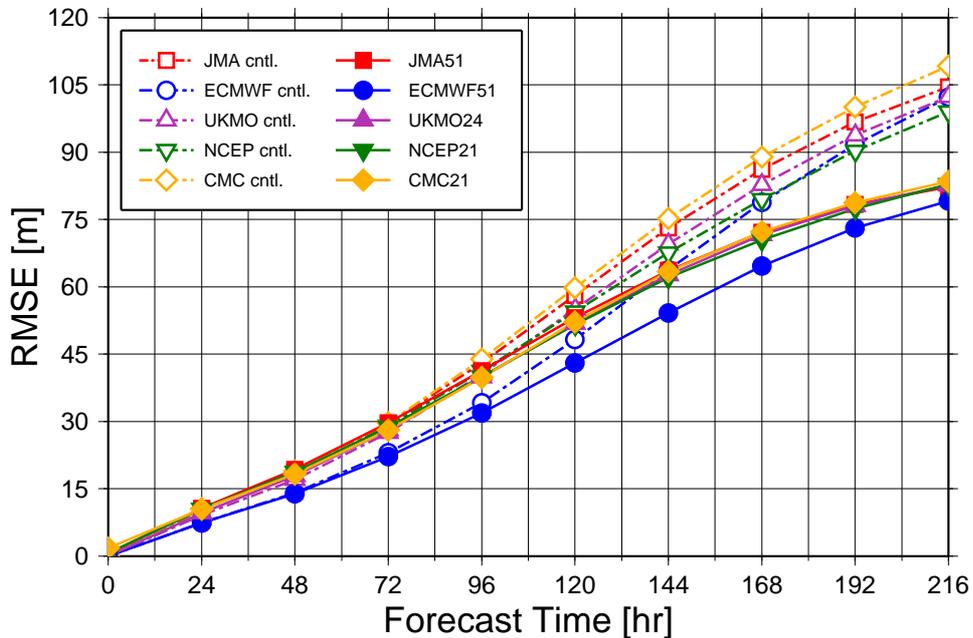


Fig. 2.4: Same as Fig. 2.3, but for (upper) from June 2007 to August 2007 and (lower) from September 2007 to November 2007.

Z500 Ensemble Mean and Spread

Initial Time: 20071117

Valid Time: 20071122 12UTC

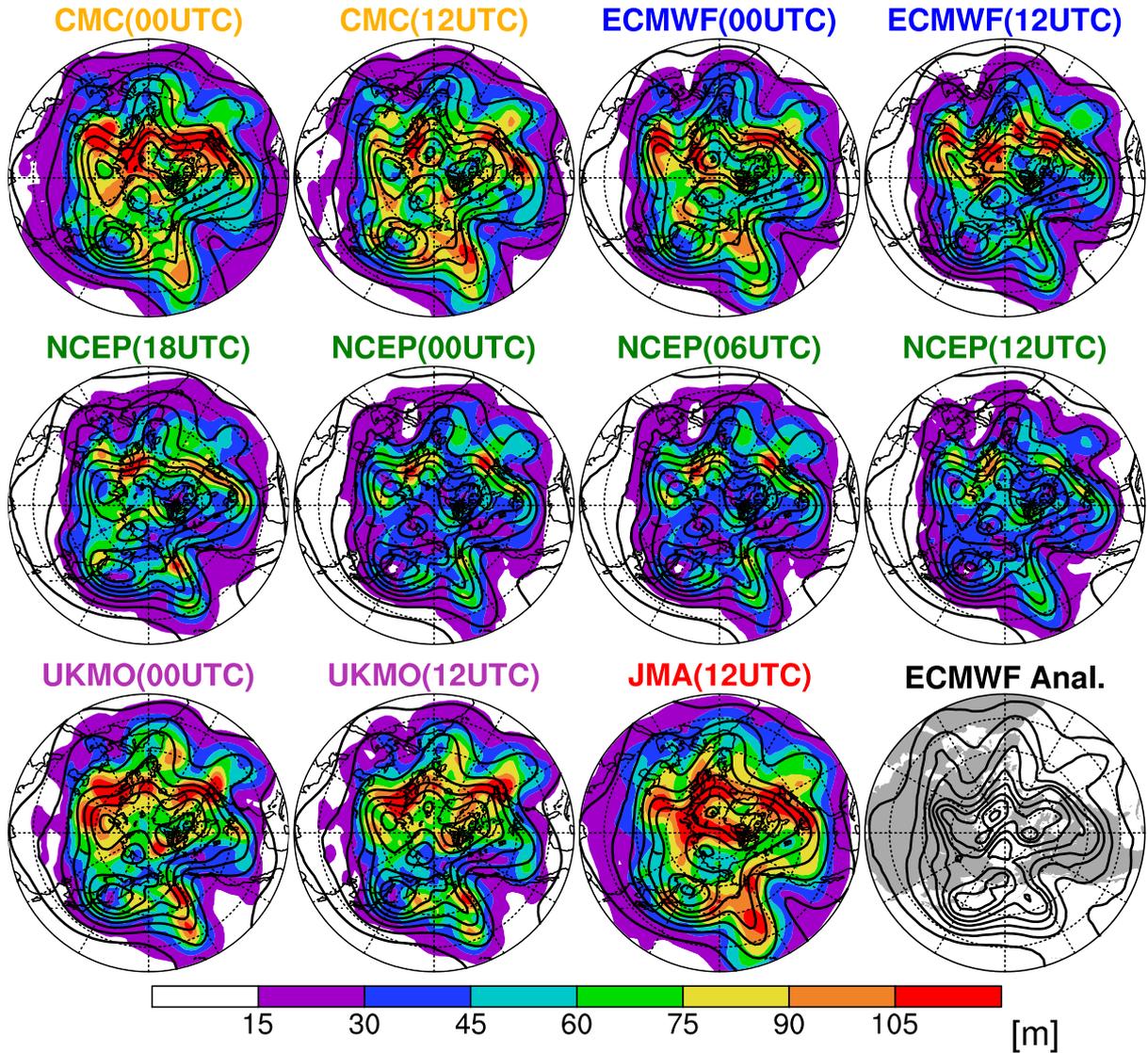


Fig. 2.5: Ensemble mean (full lines, with a contour interval of 120 m) and spread (shading) of CMC, ECMWF, NCEP, UKMO, and JMA for 500 hPa height, initialized on 17th November 2007, valid at 12 UTC on 22nd November 2007. The last panel shows the ECMWF analysis.

Z500 120hr Ensemble Spread (200612-200711: NH 5day running mean)

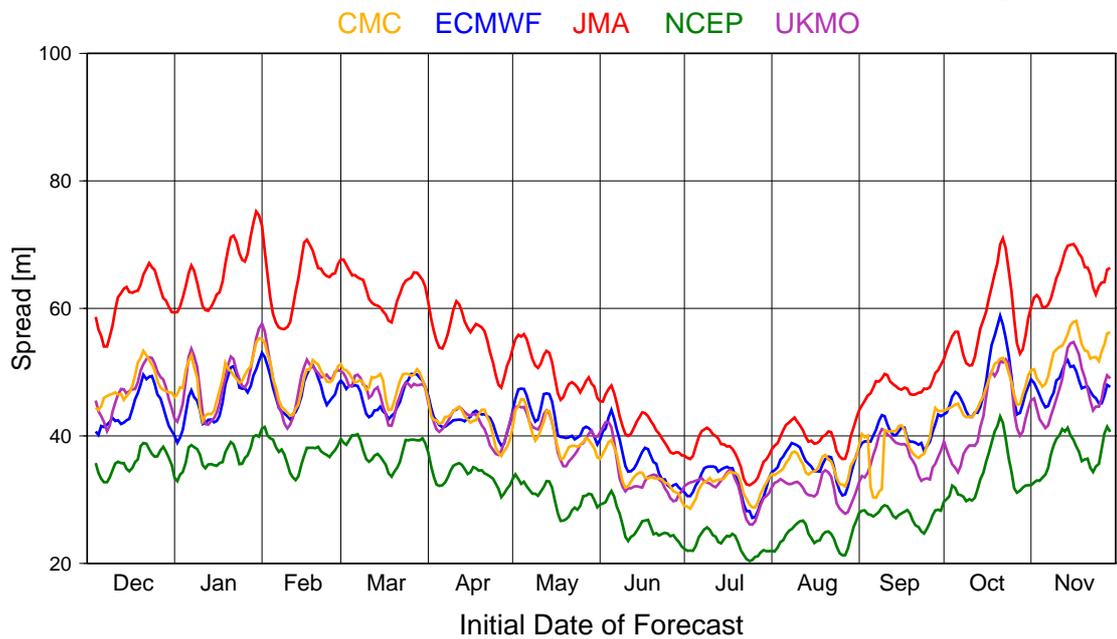
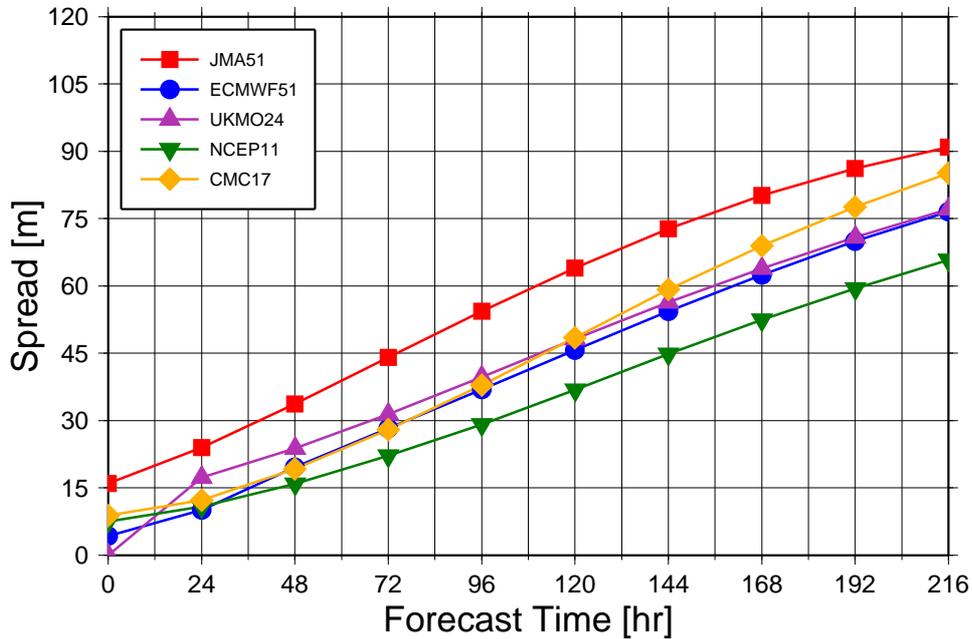


Fig. 2.6: Time series of 120-hr ensemble spreads of five single-center ensemble forecasts; CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), for 500 hPa height over the Northern Hemisphere (20°N–90°N) from December 2006 to November 2007. 5-days running mean is applied for each solid line.

Comparison of Ensemble Forecasts Z500 Ensemble Spread (2007DJF: NH)



Comparison of Ensemble Forecasts Z500 Ensemble Spread (2007MAM: NH)

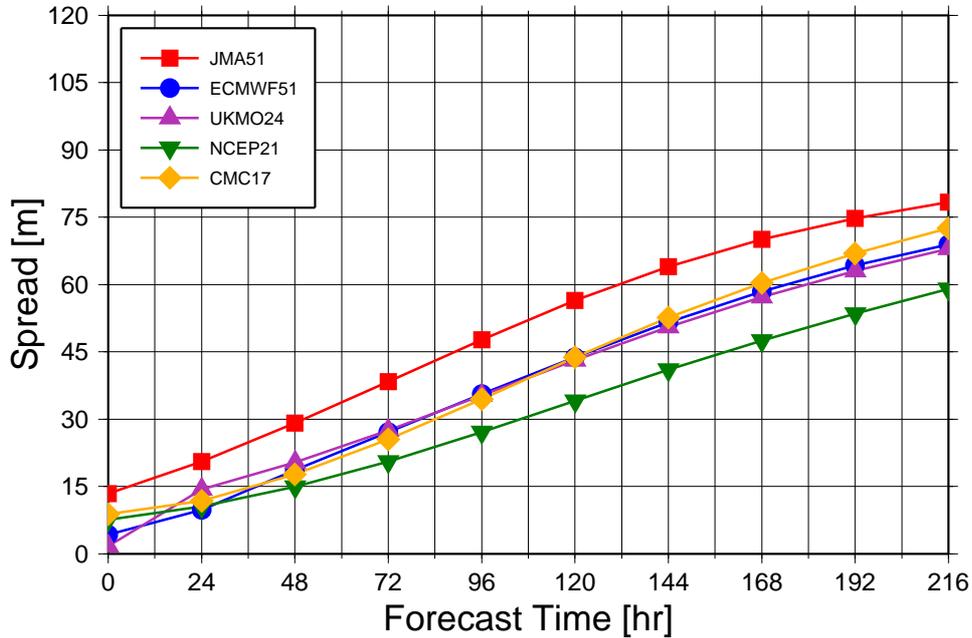
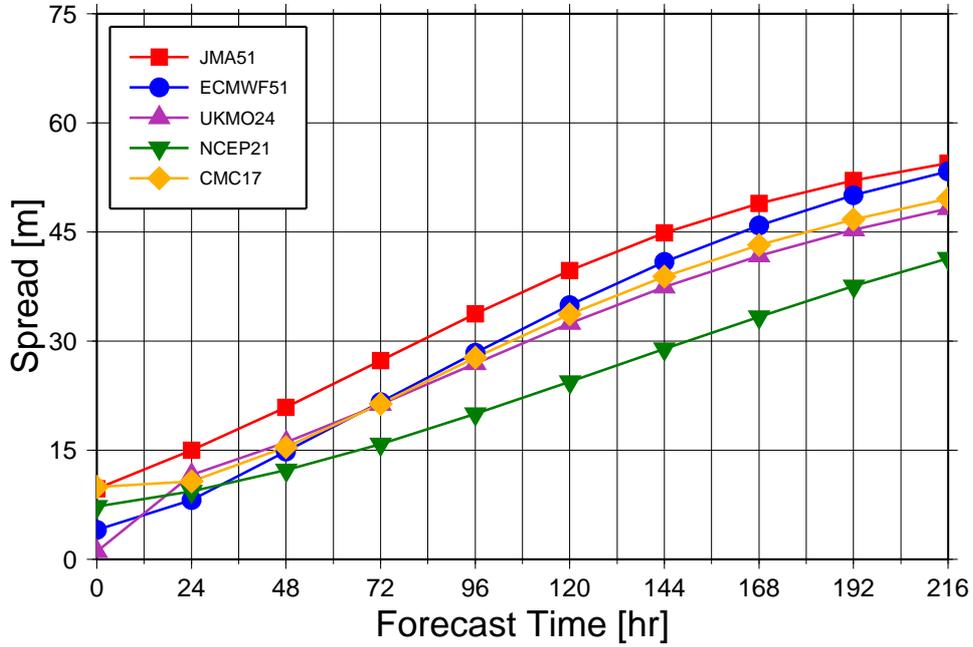


Fig. 2.7: Ensemble spreads of single-center ensemble forecasts; CMC (yellow), ECMWF (blue), JMA (red), NCEP (green), and UKMO (purple), for 500 hPa height over the Northern Hemisphere (20°N–90°N) from December 2006 to February 2007 (upper) and from March 2007 to May 2007 (lower).

Comparison of Ensemble Forecasts Z500 Ensemble Spread (2007JJA: NH)



Comparison of Ensemble Forecasts Z500 Ensemble Spread (2007SON: NH)

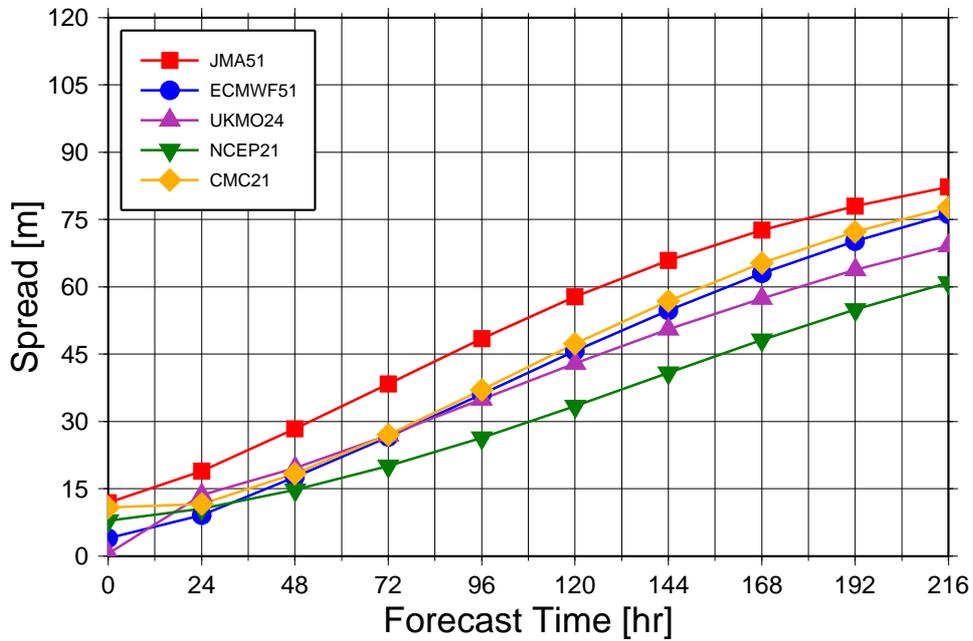


Fig. 2.8: Same as Fig. 2.7, but for (upper) from June 2007 to August 2007 and (lower) from September 2007 to November 2007.

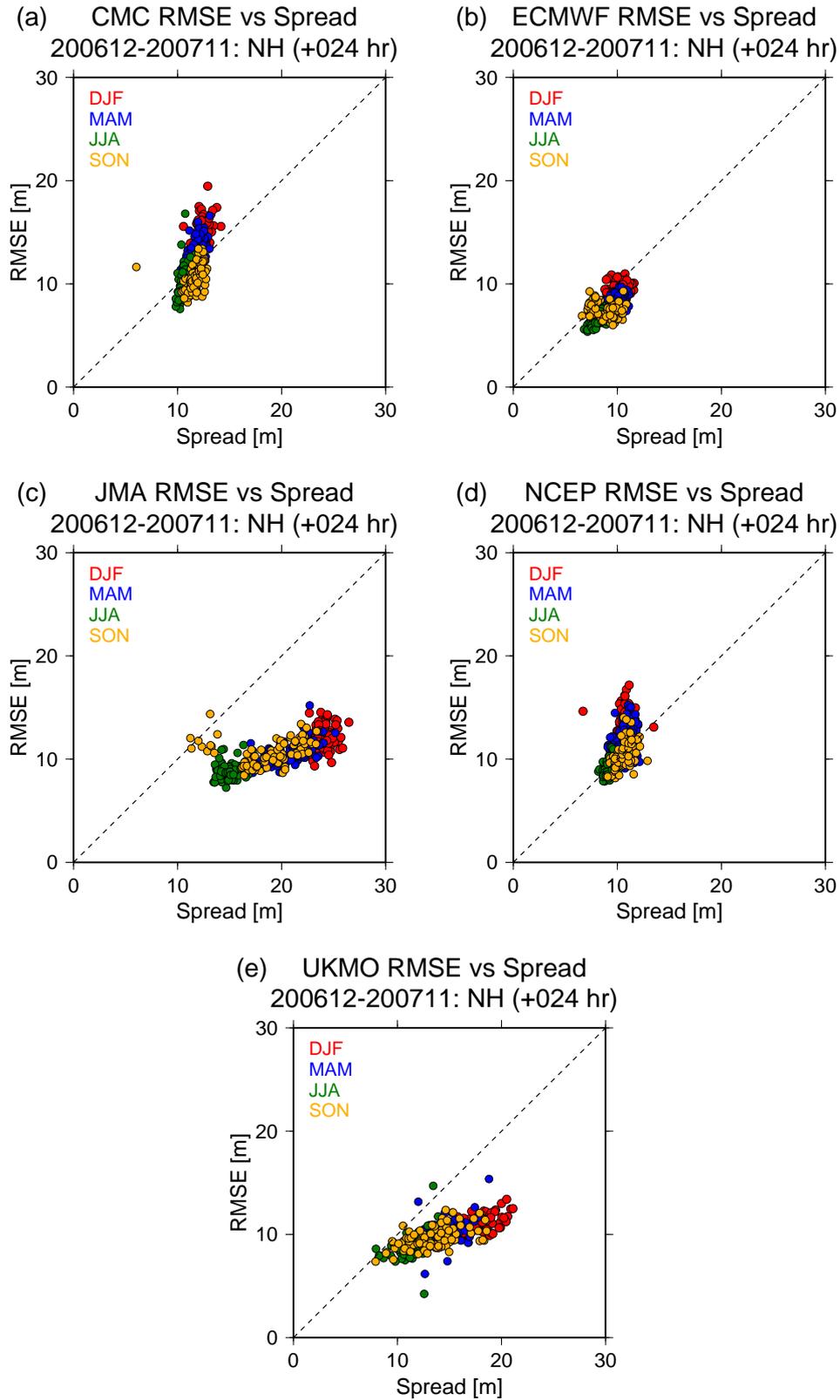


Fig. 2.9: Scatter diagrams of the RMSE versus the ensemble spread of (a) CMC, (b) ECMWF, (c) JMA, (d) NCEP, and (e) UKMO for 500 hPa height over the Northern Hemisphere (20°N–90°N) at 24-hr lead time from December 2006 to November 2007.

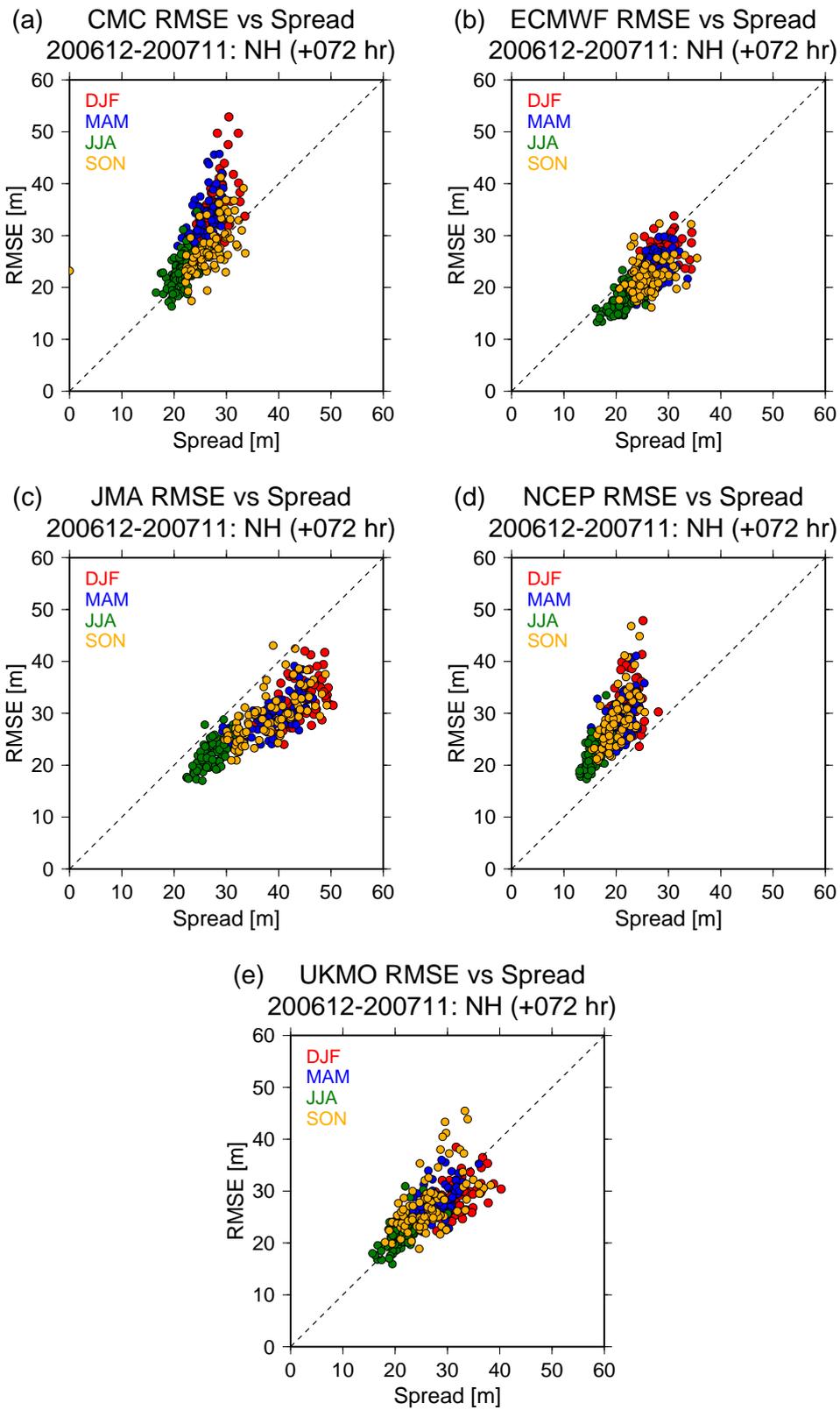


Fig. 2.10: Same as Fig. 2.9, but for at 72-hr lead time.

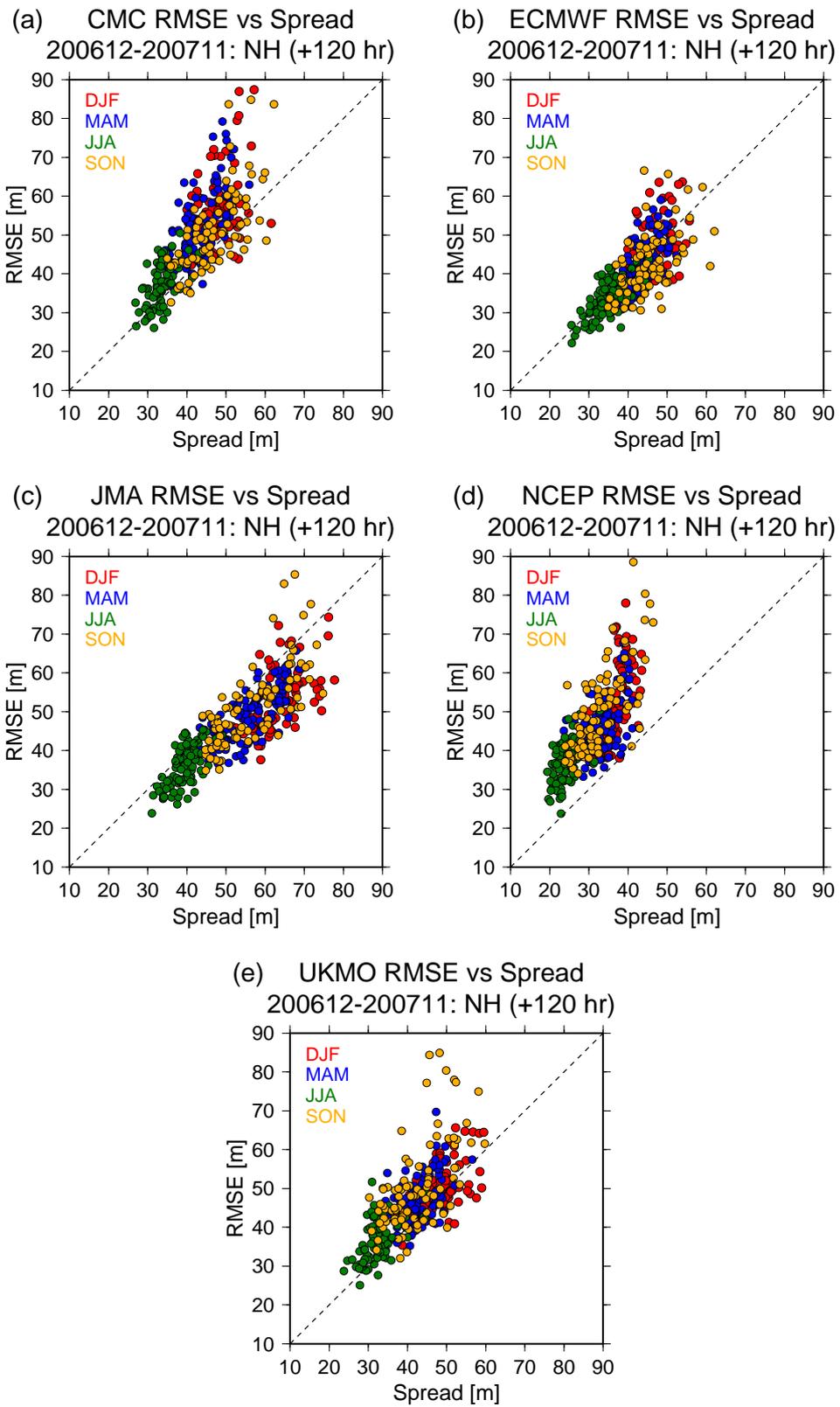


Fig. 2.11: Same as Fig. 2.9, but for at 120-hr lead time.

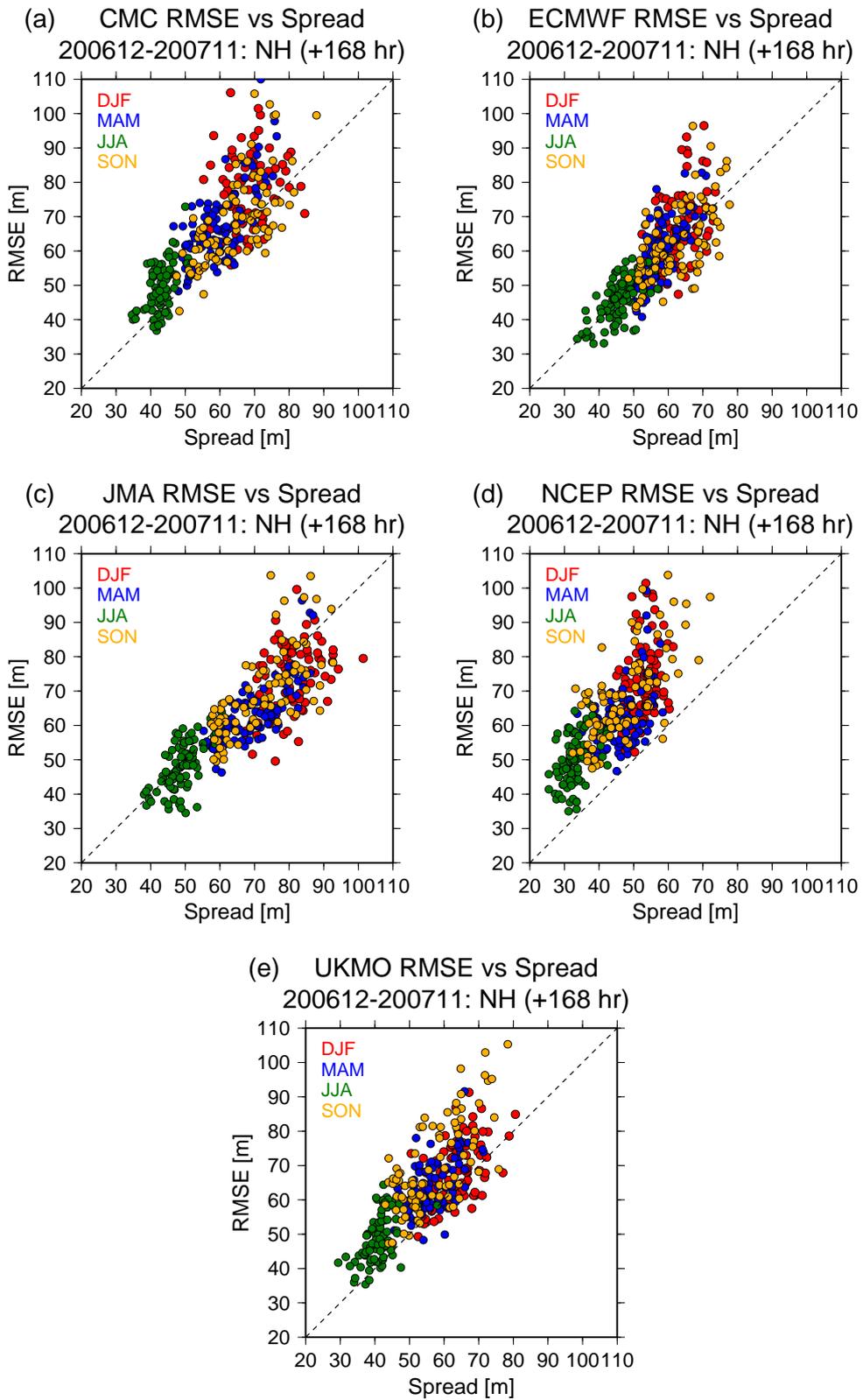


Fig. 2.12: Same as Fig. 2.9, but for at 168-hr lead time.

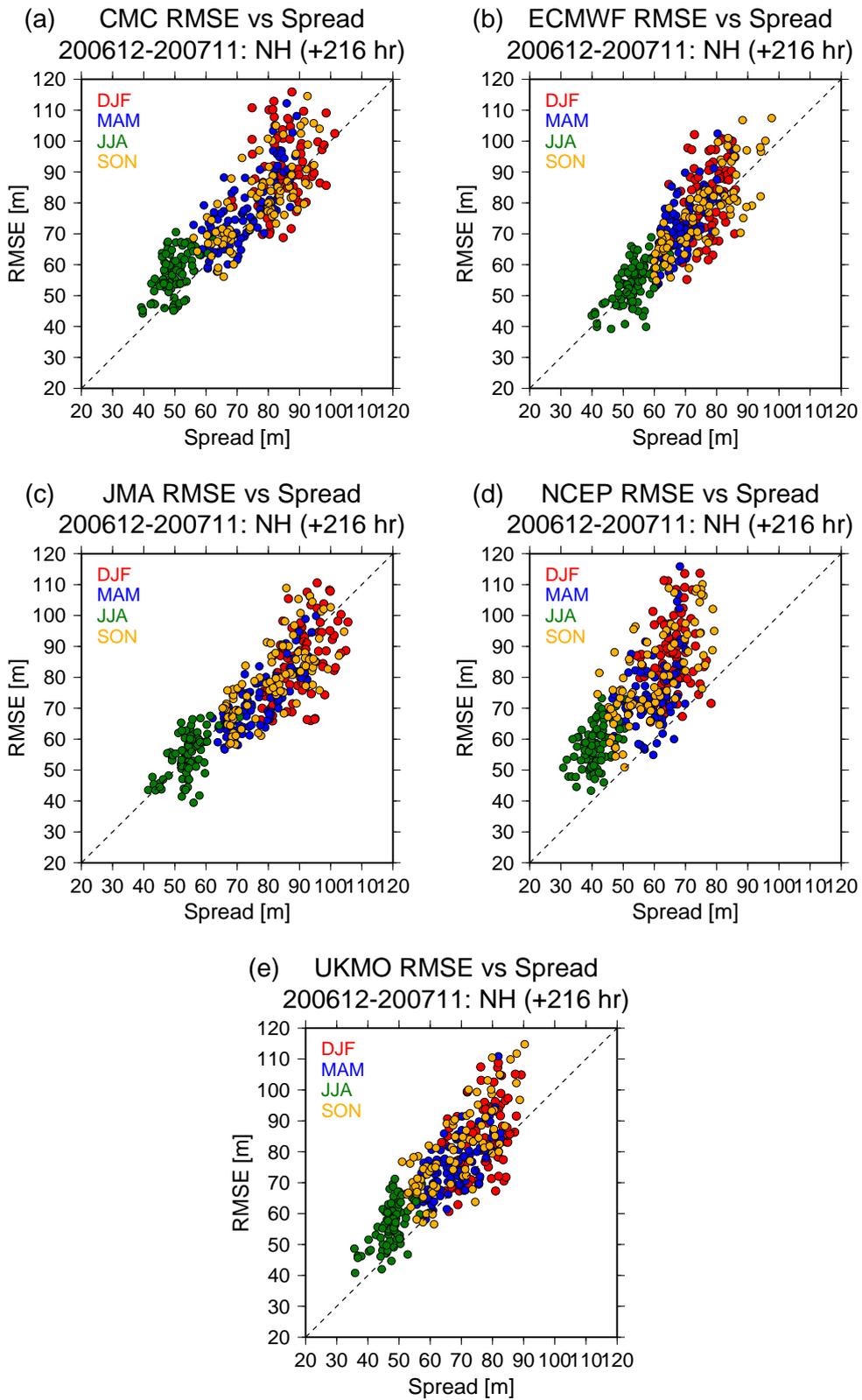


Fig. 2.13: Same as Fig. 2.9, but for at 216-hr lead time.

Comparison of Ensemble Mean Forecasts Z500 at NH grids

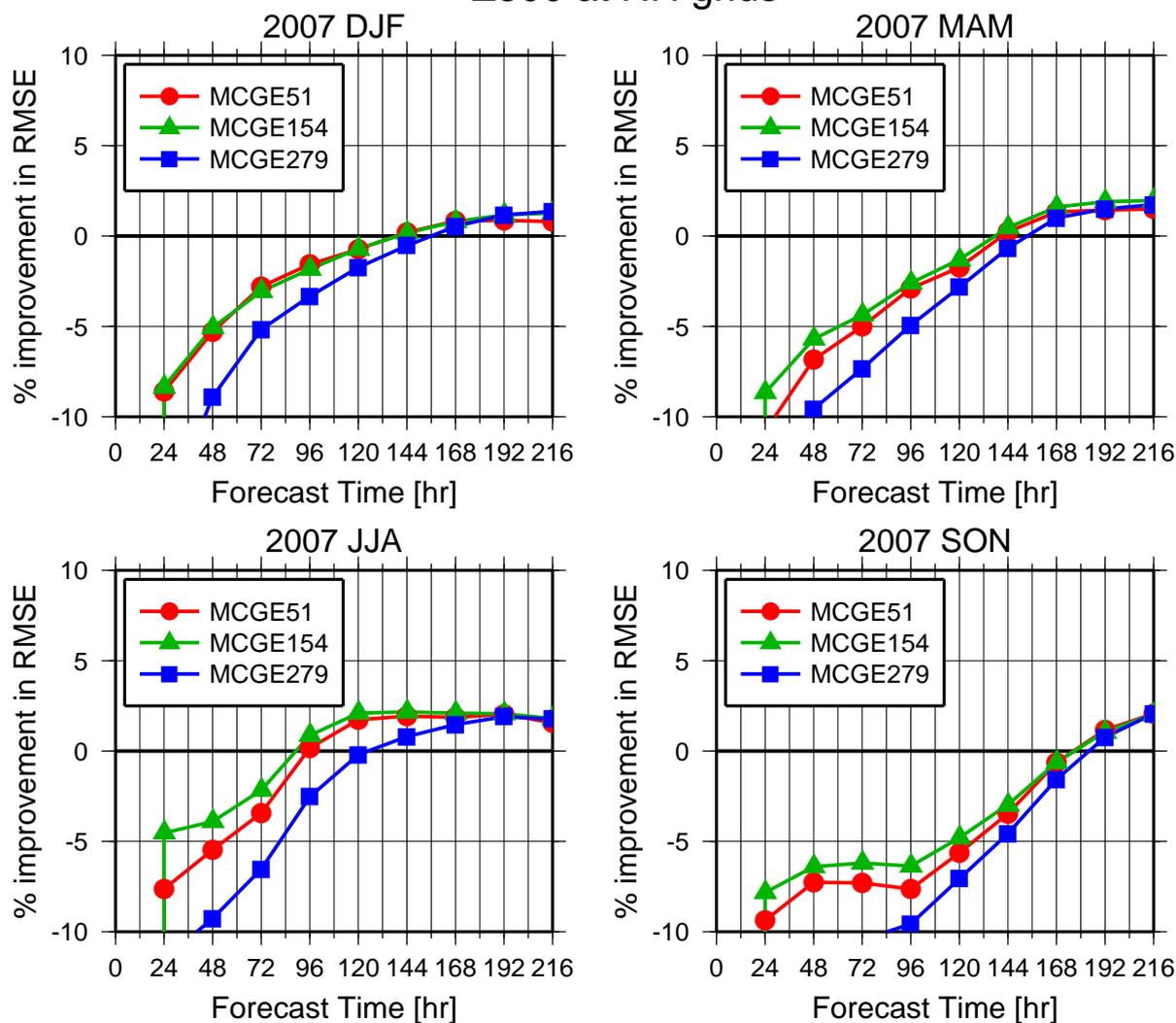


Fig. 2.14: Relative improvement (%) in seasonal RMSE of MCGEs against ECMWF51 for 500 hPa height over the Northern Hemisphere (20°N – 90°N) from December 2006 to February 2007 (a), from March 2007 to May 2007 (b), from June 2007 to August 2007 (c), and from September 2007 to November 2007 (d). Red, green, and blue solid lines are for MCGE51, MCGE154, and MCGE279, respectively.

168hr RMSE for Z500 at NH grids
ECMWF51 vs MCGE51

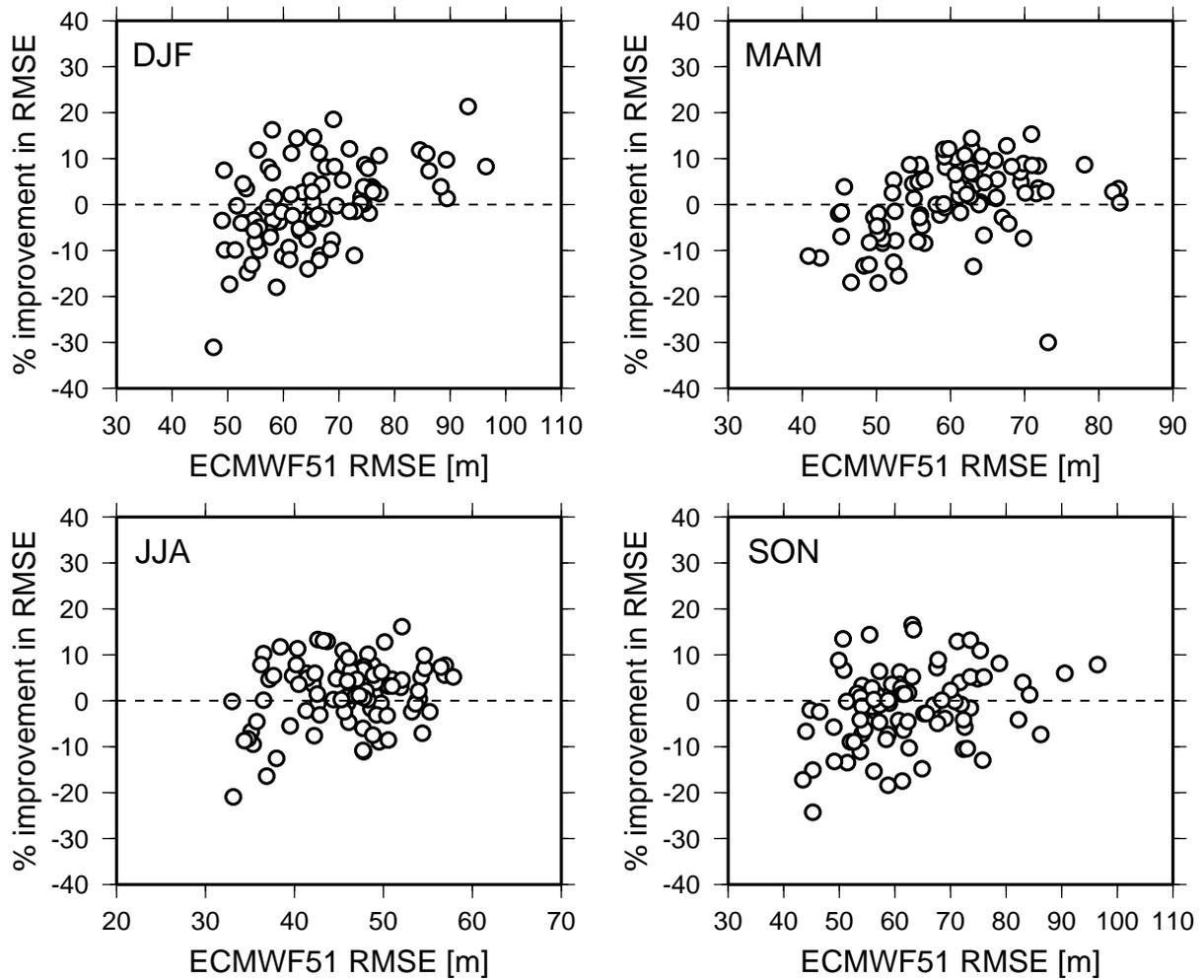


Fig. 2.15: Percent improvement in daily 168-hr RMSE of MCGEs against ECMWF ensemble for 500 hPa height over the Northern Hemisphere (20°N – 90°N) from December 2006 to February 2007 (a), from March 2007 to May 2007 (b), from June 2007 to August 2007 (c), and from September 2007 to November 2007 (d).

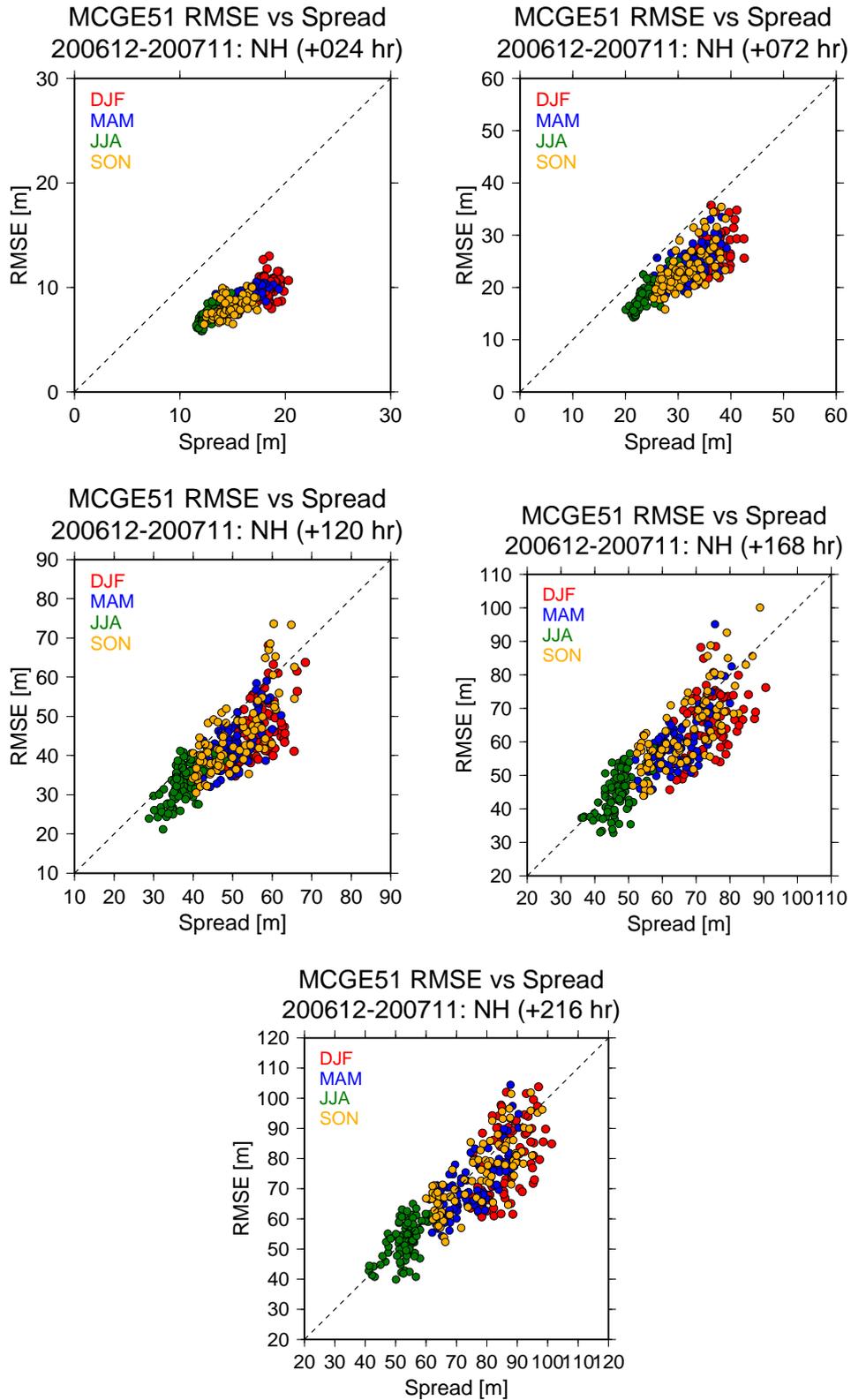


Fig. 2.16: Scatter diagram of the RMSE versus the ensemble spread of MCGE51 for 500 hPa height over the Northern Hemisphere (20°N–90°N) at 24-, 72-, 120-, 168-, and 216-hr lead times from December 2006 to November 2007.

Comparison of Ensemble Probabilistic Forecasts Z500 at NH grids

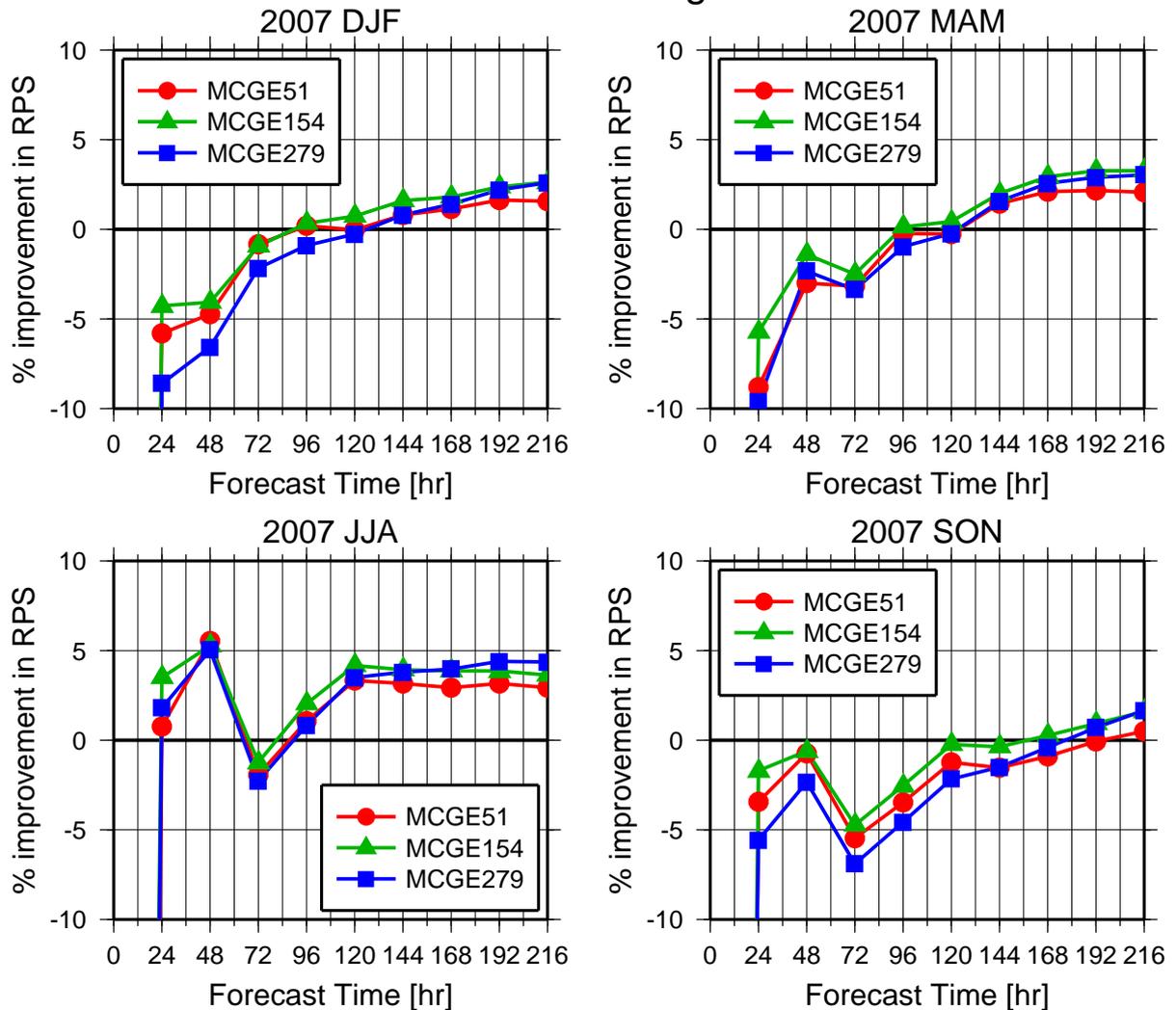


Fig. 2.17: Relative improvement (%) in seasonal RPS of MCGEs against ECMWF ensemble for 500 hPa height over the Northern Hemisphere (20°N–90°N) from December 2006 to February 2007 (a), from March 2007 to May 2007 (b), from June 2007 to August 2007 (c), and from September 2007 to November 2007 (d). Red, green, and blue solid lines are for MCGE51, MCGE154, and MCGE279, respectively.

Chapter 3

Analyses of extreme event using ensemble forecast data

In this chapter, an ensemble-based simple sensitivity analysis and ensemble simulations are performed against an atmospheric blocking shown in Matsueda et al. (2007) using the operational medium-range ensemble forecast data.

3.1 Data and method

3.1.1 Ensemble forecast data

In this chapter, three operational medium-range ensemble forecast data: CMC, JMA, and NCEP, are used. The details of these EPS as of December 2005 are summarized in Table 3.1. The horizontal resolution of the forecast model is comparable to each other. Both JMA and NCEP used the BV method as the initial perturbation, and the CMC used the EnKF method. Compared with Table 1.1, one can find the remarkable progresses of the operational ensemble forecast along with the progress of computer science. The ensemble size of each center doubled in about two years.

3.1.2 Ensemble-based sensitivity analysis

Enomoto et al. (2007) proposed a SV-like simple sensitivity analysis using ensemble forecast data. This method does not need the numerical prediction model and adjoint

code, and needs only ensemble forecast data that has already been calculated. Using this technique, the sensitivity area in the prediction of atmospheric phenomenon can be identified. The detail is as follows. A goal of this technique is to find the initial perturbation which grows up fastest, that is, \mathbf{y} in Fig. 3.1. The linear time evolution of \mathbf{y} is represented as follows:

$$\mathbf{z} = \mathbf{M}\mathbf{y}. \quad (3.1)$$

Also, the linear time evolution for each ensemble member is assumed

$$\mathbf{z}_i = \mathbf{M}\mathbf{y}_i, \quad i = 1, 2, \dots, n, \quad (3.2)$$

where \mathbf{z}_i is the forecast departure from the control run at a target lead time for i th member, \mathbf{y}_i is the initial perturbation for i th member. Consider the initial perturbation \mathbf{y} by a linear combination of the original initial perturbations:

$$\mathbf{y} = p_1\mathbf{y}_1 + p_2\mathbf{y}_2 + \dots + p_n\mathbf{y}_n. \quad (3.3)$$

Using matrix notations:

$$\mathbf{Y} = (\mathbf{y}_1 \ \mathbf{y}_2 \ \dots \ \mathbf{y}_n), \quad \mathbf{Z} = (\mathbf{z}_1 \ \mathbf{z}_2 \ \dots \ \mathbf{z}_n), \quad \mathbf{p} = (p_1 \ p_2 \ \dots \ p_n)^\top, \quad (3.4)$$

equations (3.2) and (3.3) may be written as

$$\mathbf{Z} = \mathbf{M}\mathbf{Y}, \quad (3.1')$$

$$\mathbf{y} = \mathbf{Y}\mathbf{p}, \quad (3.2')$$

and then equation (3.1) may be written as

$$\mathbf{z} = \mathbf{M}\mathbf{y} = \mathbf{M}\mathbf{Y}\mathbf{p} = \mathbf{Z}\mathbf{p}. \quad (3.5)$$

The constrained maximization problems can be solved using the Lagrange's method of undetermined multipliers. Lagrangian function and its variations are represented as follows:

$$\begin{aligned} \mathbf{F}(\mathbf{p}, \lambda) &= \langle \mathbf{z}, \mathbf{z} \rangle + \lambda(1 - \langle \mathbf{y}, \mathbf{y} \rangle) \\ &= \langle \mathbf{Z}\mathbf{p}, \mathbf{Z}\mathbf{p} \rangle + \lambda(1 - \langle \mathbf{Y}\mathbf{p}, \mathbf{Y}\mathbf{p} \rangle), \end{aligned} \quad (3.6)$$

$$\delta\mathbf{F}(\mathbf{p}, \lambda) = 2 \langle \delta\mathbf{p}, \mathbf{Z}^\top\mathbf{Z}\mathbf{p} - \lambda\mathbf{Y}^\top\mathbf{Y}\mathbf{p} \rangle - \delta\lambda(1 - \langle \mathbf{Y}\mathbf{p}, \mathbf{Y}\mathbf{p} \rangle). \quad (3.7)$$

As a result, it is only necessary to solve the eigenvalue problem:

$$(\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Z}^\top \mathbf{Z} \mathbf{p} = \lambda \mathbf{p} \quad (3.8)$$

The size of matrix $(\mathbf{Y}^\top \mathbf{Y})^{-1} \mathbf{Z}^\top \mathbf{Z}$ is the ensemble size. This problem can be solved easily. If each initial perturbation is mutually orthogonal and has the same norm, $(\mathbf{Y}^\top \mathbf{Y})^{-1}$ becomes a scalar matrix. It is only necessary to perform the singular value decomposition of the matrix \mathbf{Z} . It is noted that one can use only a half of original initial perturbations if the NWP center uses the positive-negative perturbation pairs.

3.1.3 Multi-analysis ensemble forecasts

Based on the ensemble-based simple sensitivity analysis, multi-analysis ensemble forecasts were performed using the JMA Global Spectral Model (JMA-GSM; JMA 2007). The GSM used in this study is same as the operational GSM used in the current JMA-EPS. The horizontal resolution of JMA-GSM is TL159L40. The JMA-GSM used in this study is a semi-Lagrangian model, whereas the operational JMA-GSM as of December 2005 was not.

3.2 Target blocking

The target blocking shown in Matsueda et al. (2007) occurred over the Rocky Mountain at 12 UTC on 15th December 2005. The mature time of this blocking was 18th December 2005. This blocking did not persist for a long time, and decayed within several days. Ensemble forecasts initialized at 12 UTC on 10th December 2005, was very interesting. Figure 3.2 illustrates the spaghetti diagrams of the Z500 for the CMC, JMA, and NCEP ensemble forecasts and JMA analysis, at 0-hr to 120-hr lead times. The initial times are 12 UTC on 10th December 2005 for the JMA and NCEP, and 00 UTC on 10th December 2005 for the CMC. Until 48-hr lead time, it is found that each ensemble member captured the analysis well. However, at 72-hr lead time, the NCEP members started to mis-predict the blocking. At 96-hr lead time, the NCEP members predicted the ridge of the blocking more upstream than the analysis. At 120-hr lead time, all NCEP members predicted

the wrong location of the blocking (Fig. 3.5), whereas the JMA members and most of the CMC members predicted the right location of the blocking. Interestingly, most of the CMC members initialized at 00 UTC on 11th December 2005 and a half of the JMA members initialized at 12 UTC on 11th December 2005 predicted the wrong location of the blocking (not shown), as in the NCEP members initialized at 12 UTC on 10th December 2005. In order to identify the cause of the NCEP's collective mis-prediction initialized at 12 TUC on 10th December 2005, the simple sensitivity analysis and multi-analysis ensemble forecasts were performed.

3.3 Comparison of control runs

Based on the fact that almost all JMA members predicted the right location of the blocking and all NCEP members predicted the wrong location of the blocking, the time evolution of each control run is first focused on. Figure 3.3 illustrates the time evolution of (a) the JMA analysis, (b) the JMA control run, (c) the NCEP control run, and (d) the JMA-GSM run with the NCEP control analysis up to 120-hr lead time. It must be noted that the RMSE shown in each panel was calculated over the blocking region (170°E – 260°E , 20°N – 80°N). It is found that the JMA control run (b) predicted the location of the blocking correctly (of course, the JMA control run was not the perfect forecast). The time evolution of the NCEP control run (c) is almost similar to that of the JMA control run (b) until 48-hr lead time. After 48-hr lead time, however, the time evolution of the NCEP control run is quite different from that of the JMA control run. At 72-hr lead time, positive and negative forecast errors, located over the north Pacific at 48-hr lead time, developed further. These forecast errors were not shown in the JMA control run. It is found that the origin of these forecast errors occurred in the NCEP control run at 24-hr lead time. The negative forecast error at 72-hr lead time corresponded to a cyclone (Fig. 3.9). Also, there was a small negative forecast error over the Rocky Mountain. In the JMA control run, this forecast error appeared not over the Rocky Mountain but over the south of Alaska. Although the location of these negative forecast errors was different from each other, the origin of these negative forecast error seems to be located over the south of Alaska at 48-hr lead time. At

96-hr lead time, the blocking predicted by the NCEP control run started to develop more upstream than that of the analysis. The negative forecast error over the Rocky Mountain at 72-hr lead time rapidly developed with its southward travel, and seems to block the eastward shift of the blocking ridge. On the other hand, the negative forecast error over the south of Alaska in the JMA control run seems to remain there due to the ridge of developing blocking. The negative error in the NCEP control run further developed until 120-hr lead time, and generated the blocking with remarkable meandering. Although the JMA control run had large positive bias around the blocking region, the location of the blocking was predicted correctly. The 120-hr RMSE of the NCEP control run was about 1.7 times that of the JMA control run.

3.4 Multi-analysis ensemble forecasts with NCEP analyses

Second, the multi-analysis ensemble forecasts were conducted with the JMA-GSM (TL159 L40) using the NCEP control and perturbed analyses. If the NCEP members on the JMA-GSM cannot predict the right location of blocking accurately, it can be concluded that the main cause of the NCEP's collective mis-prediction was due to the initial condition of the NCEP.

The characteristics shown in the NCEP control run were also shown in the JMA-GSM run with the NCEP control analysis (Fig. 3.3d). Despite the change of the numerical model, the NCEP control analysis on the JMA-GSM (d) led to a wrong prediction of the location of the blocking. It can be concluded that one of the causes of the collective mis-prediction of the blocking is due to the NCEP control analysis at the initial time. It is, however, interesting that the 120-hr RMSE of the JMA-GSM run with the NCEP control analysis is smaller than that of the NCEP original control run. This might indicate the decreases of the imperfection of the model formulation by introduction of other numerical models.

For the JMA-GSM runs with the NCEP perturbed analyses, it is found that all of them were not able to predict the right location of the blocking (Fig. 3.4), as in the NCEP

original perturbed runs. However, the JMA-GSM runs with the NCEP analyses predicted the location of the blocking somewhat accurately than the NCEP original runs, but more inaccurately than the JMA original runs. In fact, as shown in Table 3.2, the JMA-GSM runs with the NCEP analyses, except for 02m, 03m, and 05p (m and p indicate ensemble member in which the initial perturbation is subtracted from and added to the control run, respectively), had better forecast skill over the blocking region than the NCEP original perturbed runs. This result also might indicate the decreases of the imperfection of the model formulation.

3.5 Ensemble-based sensitivity analysis

In the previous section, it was found that the collective mis-prediction resulted from the initial value of the NCEP members. In order to detect the sensitivity area against the blocking, the ensemble-based sensitivity analysis was performed. In this study, the dry total energy norm (Talagland 1981; Ehrendorfer and Errico 1995) was used:

$$\text{TE} = \frac{1}{2} \iint_A u'^2 + v'^2 + \frac{c_p}{T_r} T'^2 + RT_r \left(\frac{p'_s}{p_r} \right)^2 dA dp, \quad (3.9)$$

where u' , v' , T' , and p'_s are perturbed components of zonal and meridional velocity, temperature, and surface pressure, respectively, c_p the specific heat at constant pressure, R the gas constant for dry air, T_r ($=300$ K) and p_r ($=800$ hPa) are the reference temperature and pressure. The target area is set to 190°E – 250°E , 30°N – 75°N , and 1000 – 200 hPa. The target time is set to 12 UTC on 15th December 2005, that is, 120-hr lead time.

Figure 3.6 illustrates the sensitivity area measured by vertically integrated dry total energy norm obtained from the JMA ensemble data. There are well-defined signals over the central north Pacific. These signals exist at each pressure level for each component (Fig. 3.7). This region, 150°E – 190°E , 30°N – 50°N , was defined as the sensitivity area.

In this sensitivity area, it is found that the difference between the JMA and NCEP control analyses at 12 UTC on 10th December 2005 measured by the dry total energy norm is relatively larger than the other areas (Fig. 3.8). In other words, the sensitive area had large uncertainty. The difference seems to be due to a cut-off cyclone (Fig. 3.9).

The difference between the JMA and NCEP control analyses at the center of the cut-off cyclone (180°E , 40°N) was about 4 hPa at 12UTC on 10th December. The difference reached about 8 hPa at 48-hr lead time. After 72-hr lead time, the cut-off cyclone traveled quite different direction from each other. The cut-off cyclone simulated by the NCEP control run traveled toward the southeast, whereas that simulated by the JMA control run traveled toward the northeast. The southeast travel of the cyclone predicted by the NCEP control run can be detected in the Z500 field (Fig. 3.3). It might be considered that the synoptic field around the cyclone over the central north Pacific at 12 UTC on 10th December affected the blocking formation. In fact, the AFES-LETKF Experimental Re-Analysis (ALERA; Miyoshi et al. 2007) shows large analysis error, that is, large uncertainty around the cyclone (Fig. 3.10).

3.6 Multi-analysis ensemble forecasts with amplified initial perturbations

It was suggested in the previous section that the NCEP control analysis had a weaker cyclone than the JMA control analysis, and the cyclone affected the blocking formation. However, even if the control analysis has large initial uncertainty, there is a possibility that the initial perturbations in the ensemble forecast reduced it. Figure 3.11 illustrates the dry total energy for the NCEP initial perturbations at 12 UTC on 10th December 2005. It is found that the initial perturbations: 02 and 03, did not have well-defined signals around the cyclone. Also, even if there are any signals around the cyclone (perturbations 01, 04 and 05), the amplitude of the initial perturbation seems to be small compared with the analysis difference shown in Fig. 3.8. Base on the fact that there is large uncertainty in the sensitivity area, these results might suggest that the NCEP ensemble did not have effective initial perturbations to predict the blocking formation more accurately. As described in the previous chapter, the NCEP has smaller initial perturbations than the other EPSs. There is a possibility that amplification of the initial perturbation leads to the improvement of the forecast skill at least in this case.

Based on the ensemble-based sensitivity analysis, the multi-analysis ensemble fore-

casts were performed. First, the amplitude of the NCEP initial perturbations was globally increased by a factor of 1.5. The fourth column in Table 3.2 shows the 120-hr RMSE for the Z500 over the blocking region (170°E–260°E, 20°N–80°N). Compared with the third column in Table 3.2, it is found that for most of runs the global amplification led to the improvement in the RMSE over the blocking region. However, the global amplification of initial perturbations led to the degradation of the forecast over the Northern Hemisphere (Table 3.3). Compared with the third and fourth columns in Table 3.3, it is found that the global amplification led to worse skill, on the hemispherical scale, than the JMA-GSMs with the original amplitude, except for 04p and 05p.

Based on these results, additional multi-analysis ensemble forecasts with regionally amplified initial perturbations were performed. The amplitude of the NCEP initial perturbation was increased by a factor of 1.5 only over the sensitivity area. If the regional amplification led to the improvement of the forecast over the blocking region, it can be concluded that the sensitivity area is a key component of the prediction of the blocking. For many members, the regional amplification reduced forecast error over the blocking region without the degradation of the forecast skill over the Northern Hemisphere (see the fifth columns in Tables 3.2 and 3.3). This indicates that the amplification of the perturbations over the sensitivity area was essential for the improvement of the prediction of the blocking. The perturbed members: 01p, 02p, 04m, and 05m, with the regionally amplified perturbations have the lowest RMSE over the blocking region. It is found the predicted location of the blocking in these members was closer to the analysis than that in the NCEP original ones (Fig. 3.12). These members did not have well-defined negative forecast error shown in the NCEP original EPS. This seems to enable the blocking ridge to shift somewhat eastward. In terms of the ensemble mean, the improvement of the forecast by the regional amplification is also obvious (Fig. 3.13). These results indicate that the sensitivity area was a key component of the prediction of the blocking. They also indicate that the excessive amplification of the initial perturbation over non-sensitivity area is undesirable and that the regional amplification technique can lead to better forecast without the degradation of the forecast over the other area.

Table 3.1: Operational medium-range ensemble prediction system at CMC, JMA, and NCEP as of December 2005.

	CMC Canada	JMA Japan	NCEP US
Model Uncertainty	Multi-Model Stoch. Phys.	NO	NO
Initial Perturbation	EnKF	BVs	BVs+LAF
Forecast Model Resolution	TL149L23-41 1.2degL28	T106L40	T126L28
Initial UTC	00	12	00, 06, 12, 18
Forecast Length (interval)	240hr (12hr)	216hr (12hr)	0-180hr (6hr) 180-384 (12hr)
Member/run	17	25	11
Member/day	17	25	44

Table 3.2: 120-hr RMSEs of NCEP-EPS and JMA-GSM runs with NCEP analyses for 500 hPa height over the blocking region (170°E–260°E, 20°N–80°N).

20051210 12UTC+120hr	NCEP original EPS	JMA-GSM runs with NCEP analyses		
		Amp: 1.0	Amp: 1.5	Amp: 1.5area
00	139	122	-	-
01p	143	95.6	84.1	81.9
01m	123	115	136	128
02p	131	71.7	71.3	61.9
02m	103	111	101	110
03p	143	127	141	134
03m	100	140	151	138
04p	148	102	89.5	93.7
04m	88.3	73.9	68	56.5
05p	116	125	98.4	101
05m	128	68.9	63.7	63.1
Ensemble Mean	117	91.1	78.0	79.6

JMA ensemble Mean: 61.6 m

Table 3.3: 120-hr RMSEs of NCEP-EPS and JMA-GSM runs with NCEP analyses for 500 hPa height over the Northern Hemisphere (20°N–90°N).

20051210 12UTC+120hr	NCEP original EPS	JMA-GSM runs with NCEP analyses		
		Amp: 1.0	Amp: 1.5	Amp: 1.5area
00	96.9	87.1	-	-
01p	104	79.0	82.6	75.8
01m	90.8	104	122	109
02p	102	69.0	88.3	66.5
02m	82.7	97.0	104	97.4
03p	110	105	117	108
03m	79.1	87.5	95.5	86.8
04p	99.3	79.9	78.3	78.6
04m	81.2	67.0	72.9	63.1
05p	96.3	87.6	76.4	80.2
05m	98.5	78.4	85.7	76.2
Ensemble Mean	85.8	72.5	70.7	69.1

JMA ensemble Mean: 57.1 m

Sensitivity Analysis

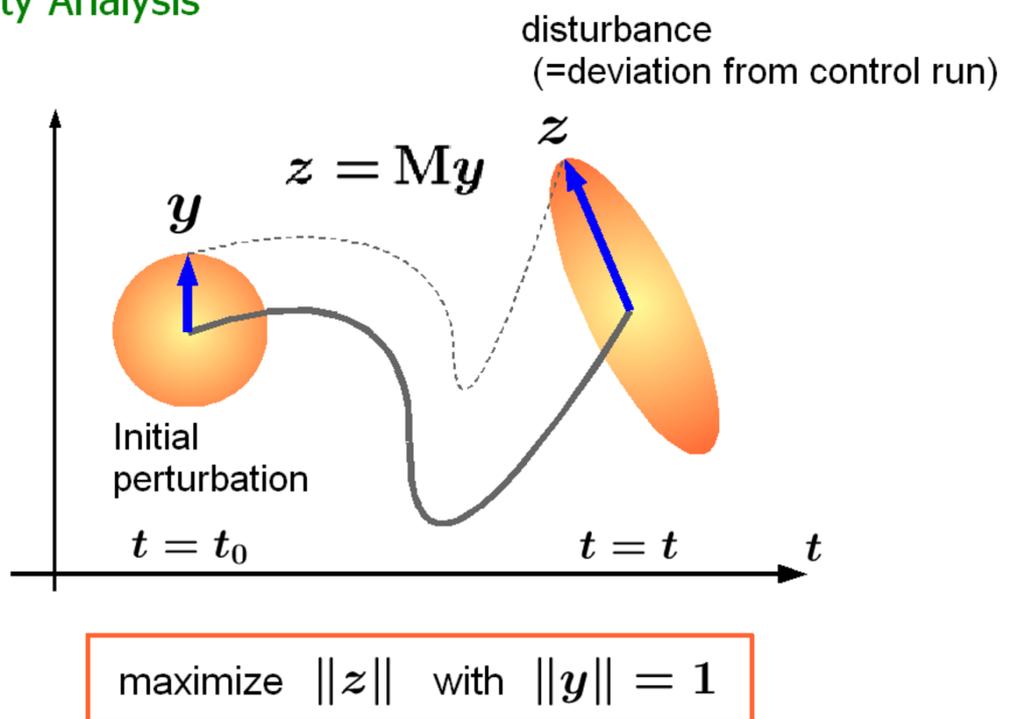


Fig. 3.1: Conceptual diagram of phase space on disturbance.

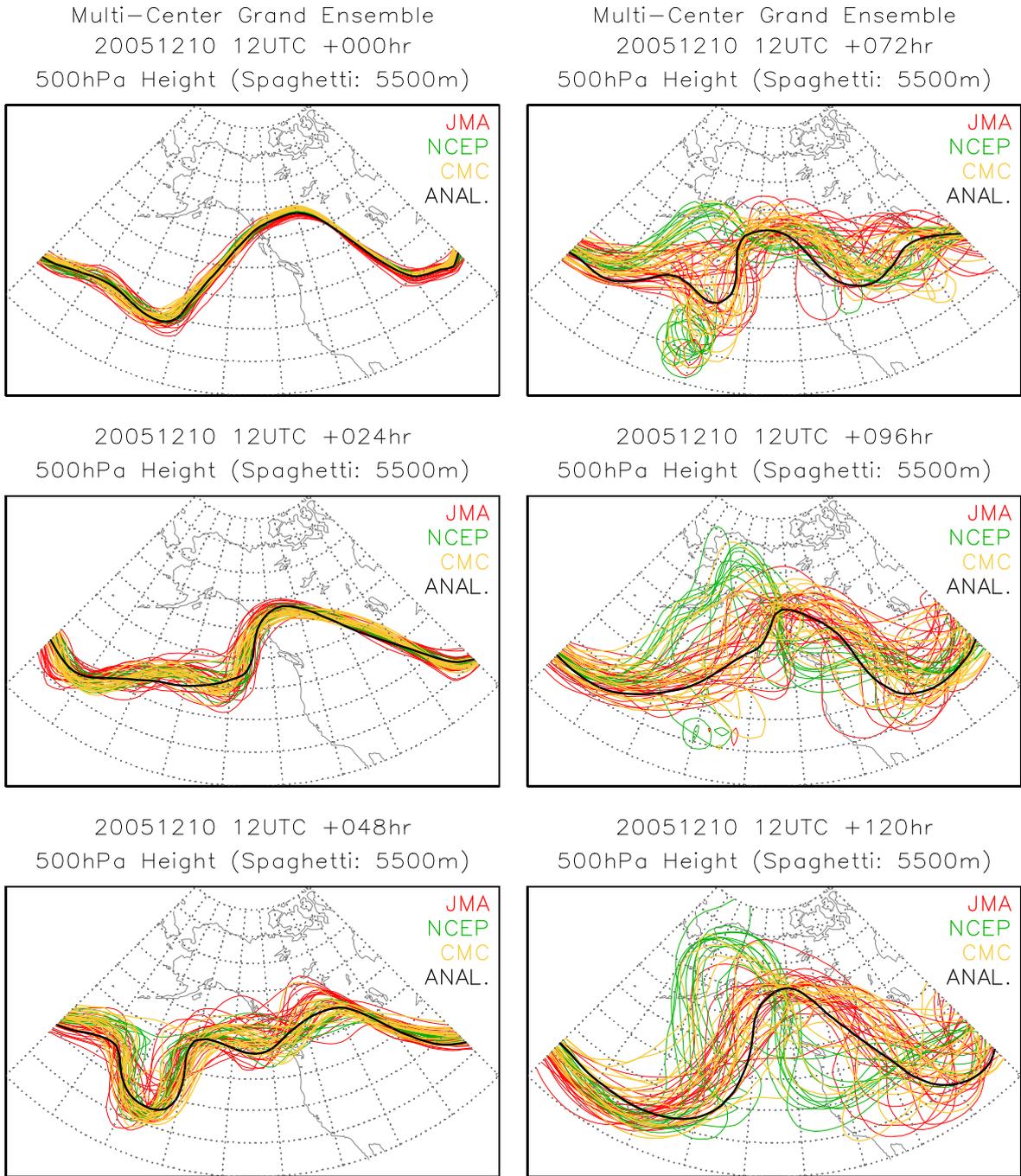
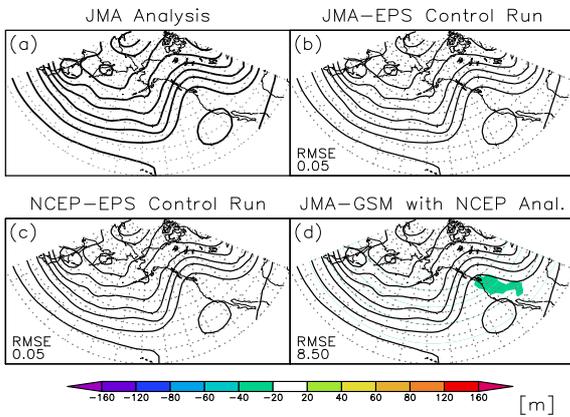
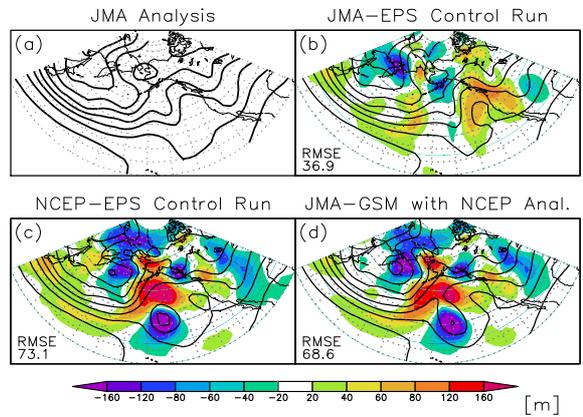


Fig. 3.2: Spaghetti diagrams of 500 hPa height (5500 m) for ensemble members of CMC (yellow), JMA (red), and NCEP (green), initialized at 00 UTC (CMC) or 12 UTC (JMA and NCEP) on 10th December 2005, valid 12 UTC on 15th December 2005. Thin solid line is for each ensemble member forecast and thick solid line for JMA analysis at the valid time.

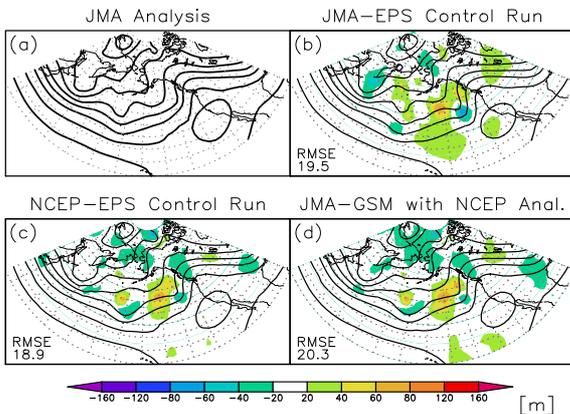
Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121012UTC



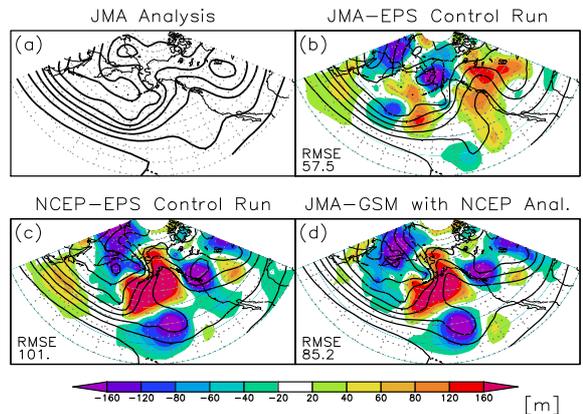
Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121312UTC



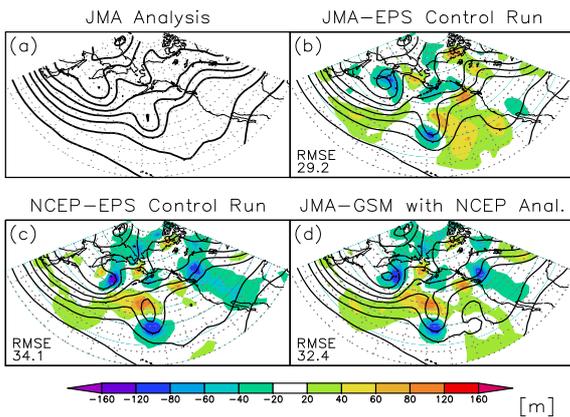
Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121112UTC



Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121412UTC



Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121212UTC



Single-Center vs Multi-Analysis (Z500)
 Init: 2005121012UTC Valid: 2005121512UTC

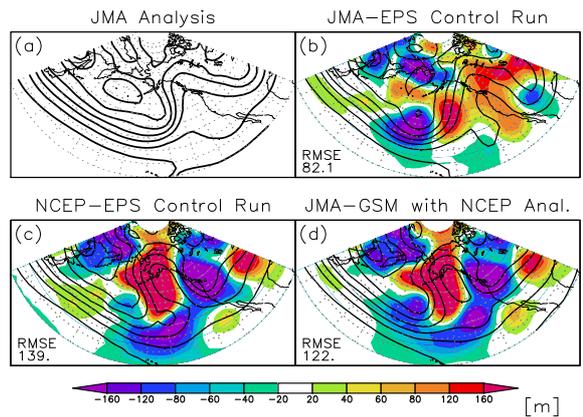


Fig. 3.3: Time evolution of 500 hPa height (contour) and its forecast error (shaded) for (a) JMA analysis, the control runs of (b) JMA and (c) NCEP, and (d) JMA-GSM run with NCEP control analysis. The initial time is 12 UTC on 10th December 2005.

JMA Model Run from NCEP Analysis
500hPa Height (Spaghetti: 5500m)
20051210 12UTC +120hr

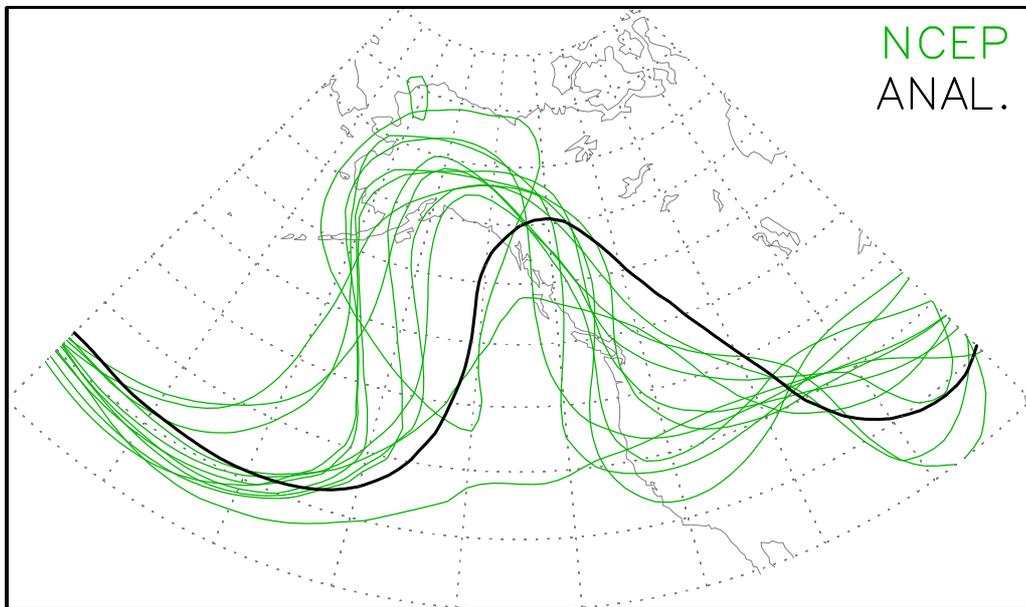


Fig. 3.4: Spaghetti diagrams of 500 hPa height (5500 m) for JMA-GSM runs with NCEP control and perturbed analyses from 12 UTC on 10th December 2005, valid 12 UTC on 15th December 2005. Thin solid line is for each ensemble member forecast and thick solid line for the NCEP analysis at the valid time.

NCEP Medium-Range Ensemble (Z500)
20051210 12UTC +120hr

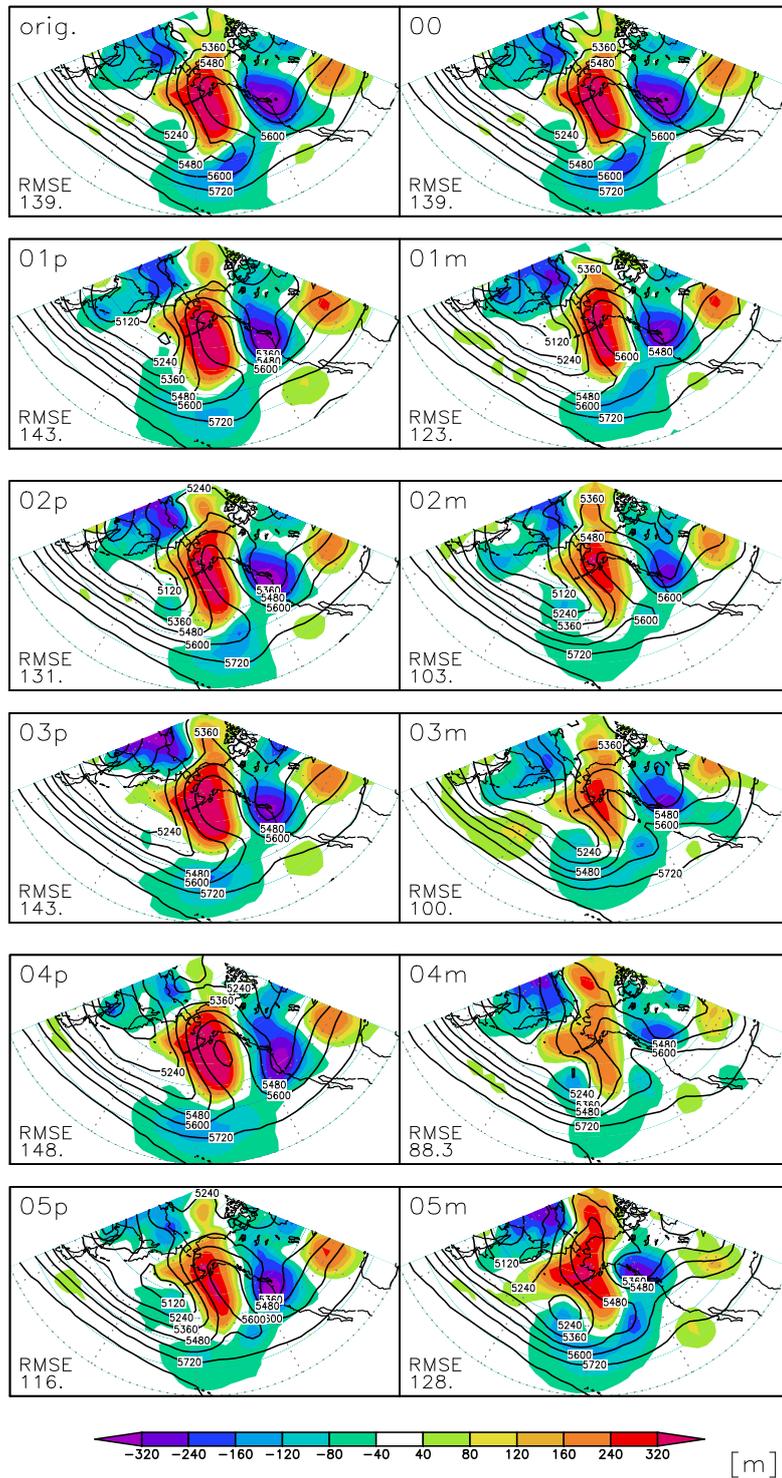


Fig. 3.5: 120-hr forecast of 500 hPa height (contour) and its forecast error (shaded) for NCEP original control and perturbed runs from 12 UTC on 10th December 2005. RMSE is calculated over the blocking region (170°E–260°E, 20°N–80°N).

Ensemble based Sensitive Analysis (JMA)

Valid: 2005121012UTC +120hr (SV1: 43.6%)

TARGET AREA: 190–250E,30–75N,1000–200hPa(Total Dry Energy)

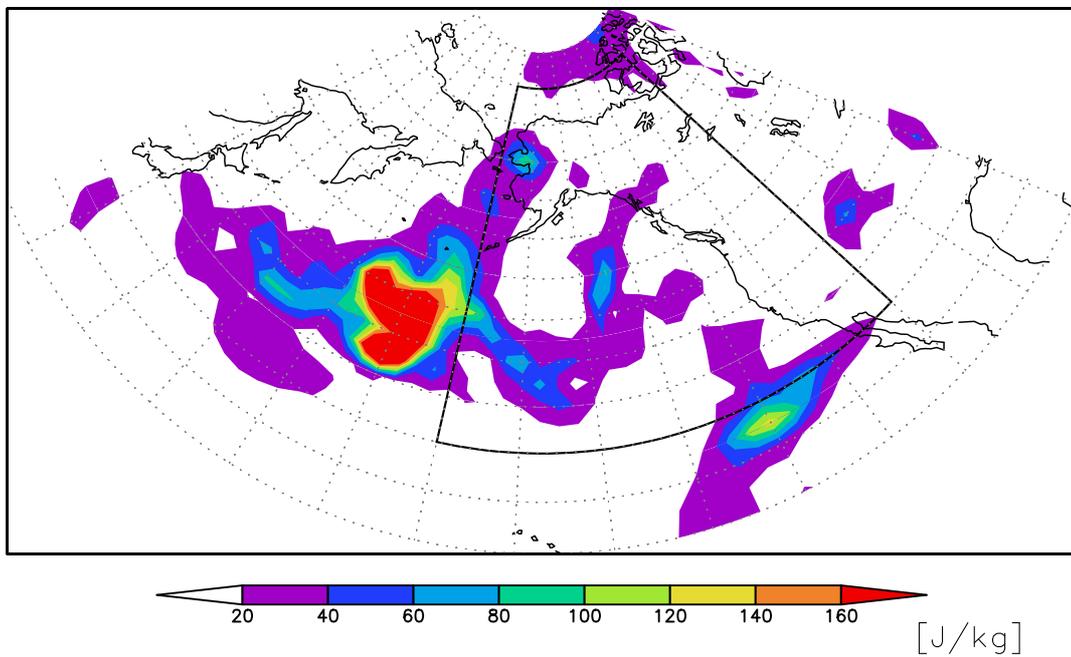


Fig. 3.6: Sensitivity area obtained from JMA ensemble forecast initialized at 12 UTC on 10th December 2005. The target time is 12 UTC on 15th December 2005, that is, 120-hr lead time. The target area is surrounded by black solid line.

Ensemble based Sensitive Analysis (JMA)

Valid: 2005121012UTC +120hr (SV1: 43.6%)

TARGET AREA: 190–250E,30–75N,1000–200hPa

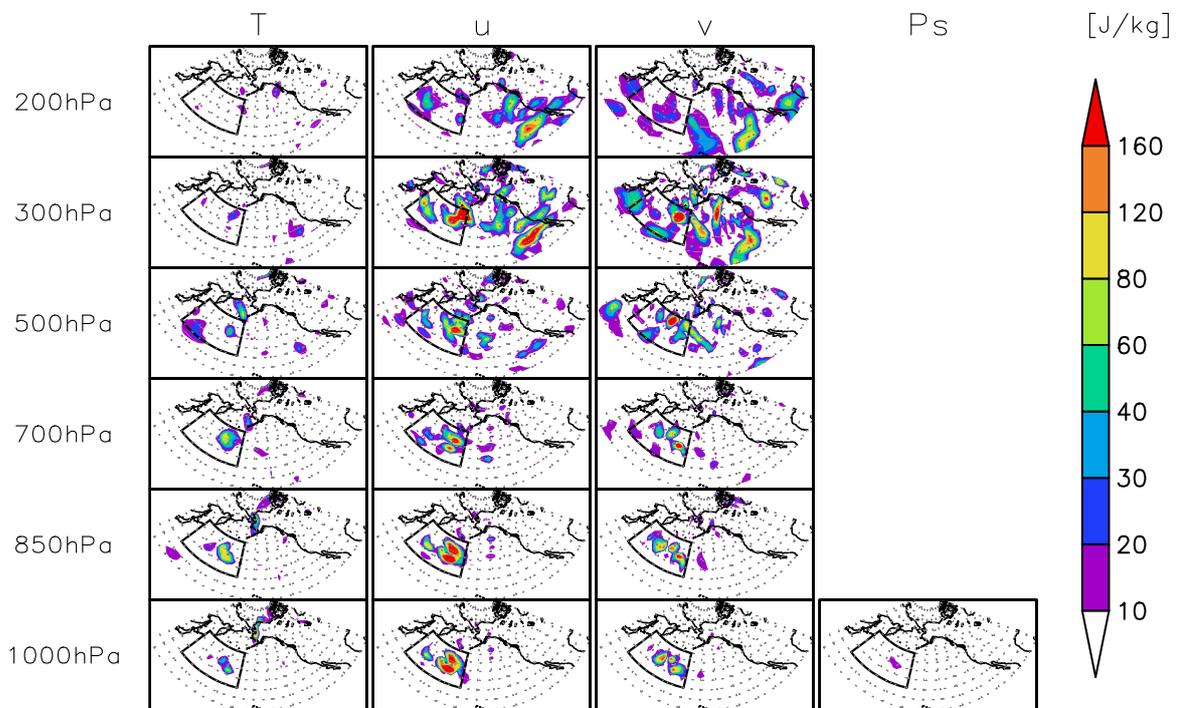


Fig. 3.7: Same as Fig. 3.6, but for components, u' , v' , T' , and p'_s at each pressure level. The sensitivity area is surrounded by black solid line.

Dry Total Energy of Initial Difference (1000–200hPa)
Valid: 2005121012UTC

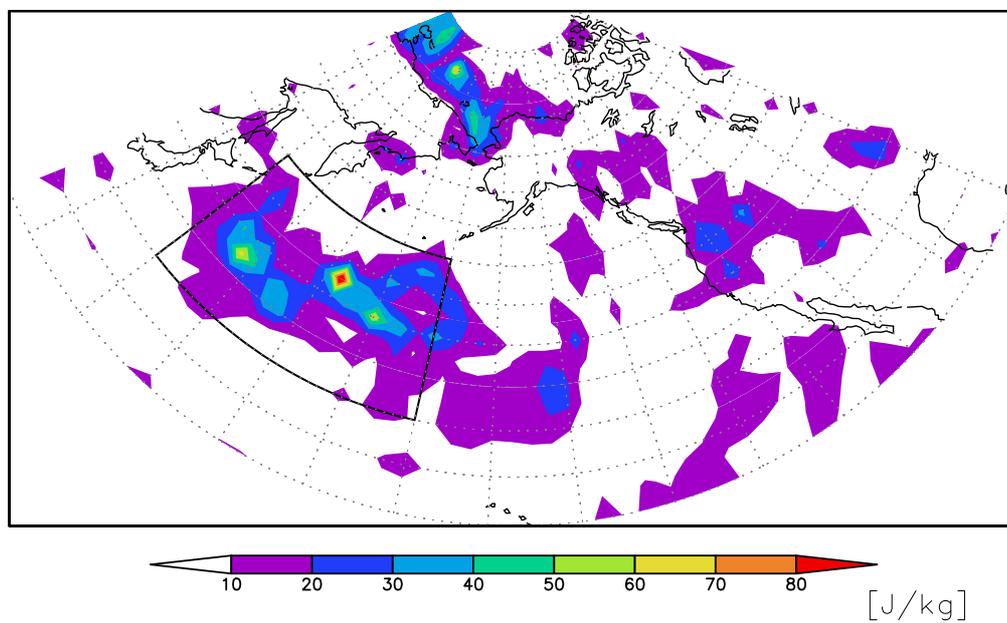


Fig. 3.8: Initial difference between JMA and NCEP analyses at 12 UTC on 10th December 2005 measured by dry total energy. The sensitivity area is surrounded by black solid line.

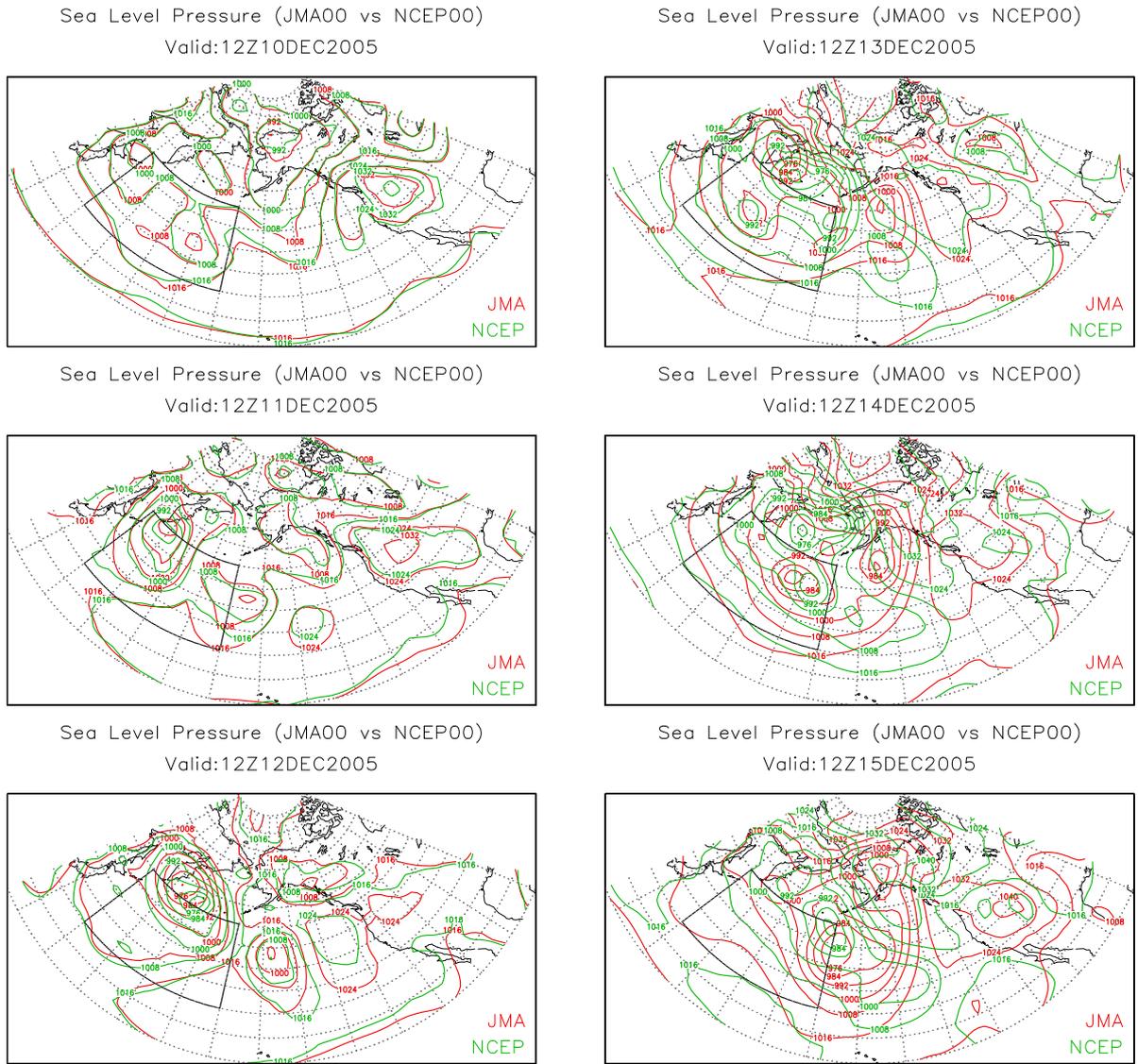


Fig. 3.9: Time evolution of sea level pressure for JMA (red) and NCEP (green) control runs. The sensitivity area is surrounded by black solid line.

AFES-LETKF Experimental Reanalysis (SLP)

Valid: 2005121012UTC

Sensitivity Area: 150-190E,30-50N

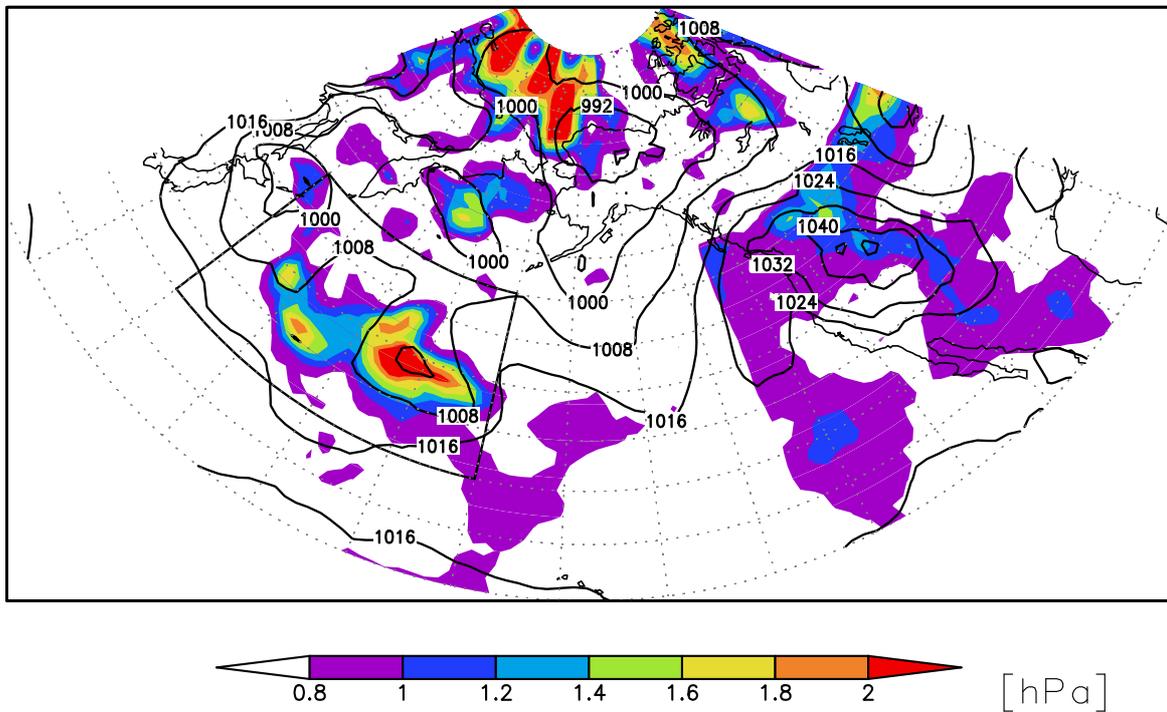


Fig. 3.10: Sea level pressure (contour) and its analysis error (shaded) based on AFES-LETKF Experimental Re-Analysis (ALERA) at 12 UTC on 10th December 2005.

Total Dry Energy (1000–200hPa)
NCEP Initial Perturbations (2005121012UTC)

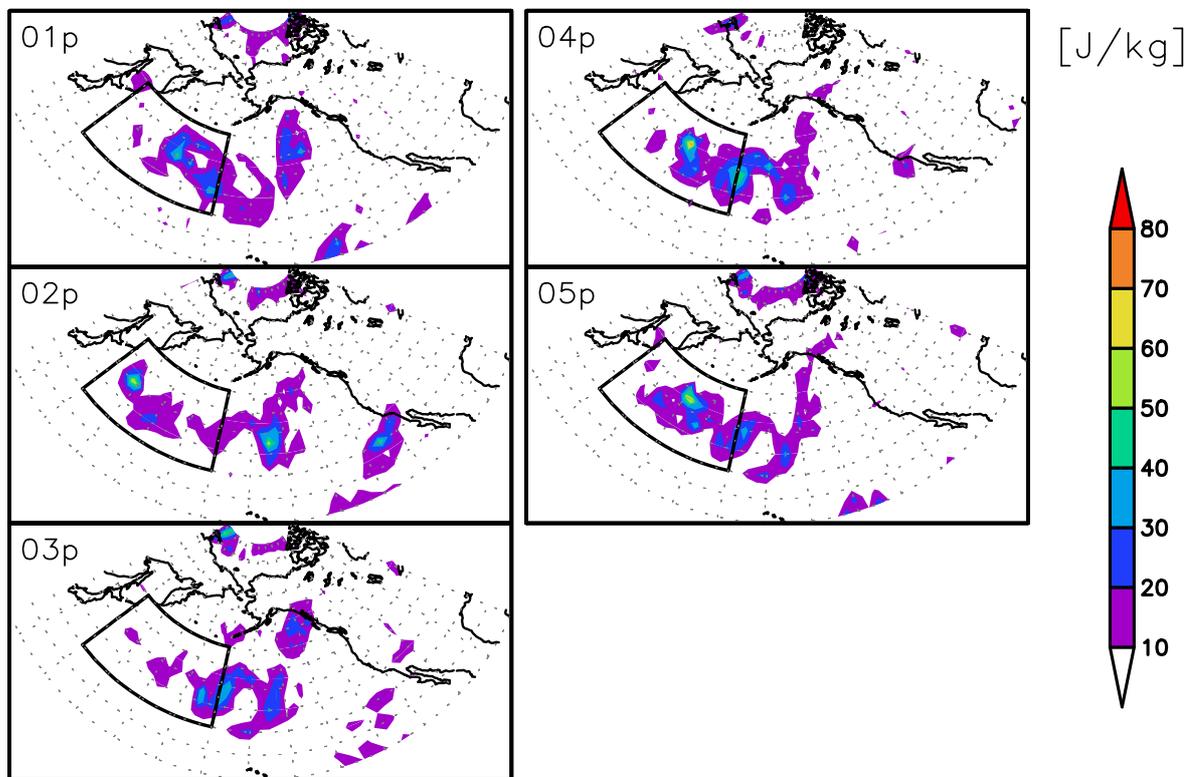


Fig. 3.11: Dry total energy for NCEP initial perturbations at 12 UTC on 10th December 2005. The sensitivity area is surrounded by black solid line.

NCEP Multi-Analysis Ensemble (Z500)
 20051210 12UTC +120hr (AMP:1.5area)

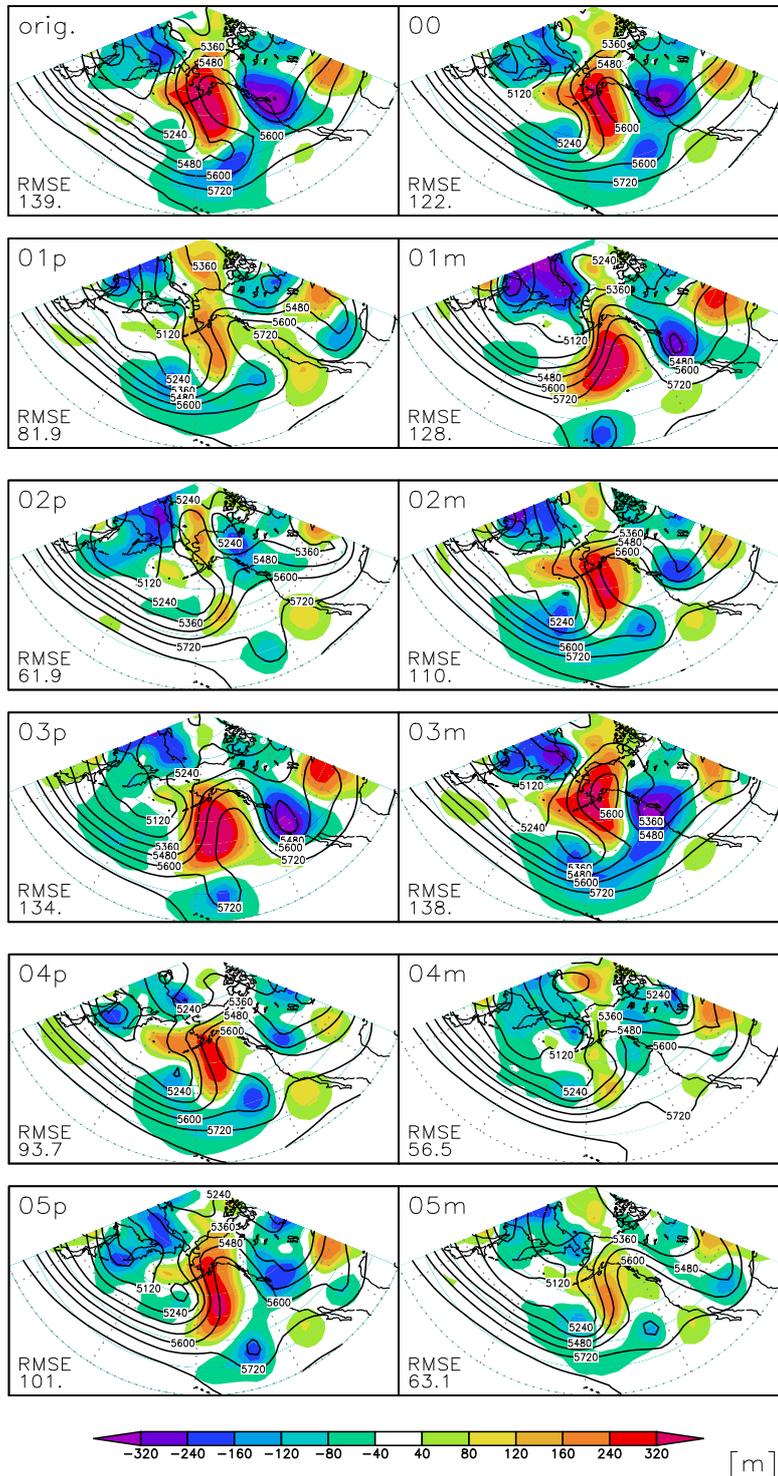


Fig. 3.12: 120-hr forecast of 500 hPa height (contour) and its forecast error (shaded) for JMA-GSM runs with NCEP analyses with the regionally amplified perturbations, initialized at 12 UTC on 10th December 2005. The initial perturbations only in the sensitivity area were amplified by a factor of 1.5.

NCEP Multi-Analysis Ensemble Mean (Z500)
20051210 12UTC +120hr

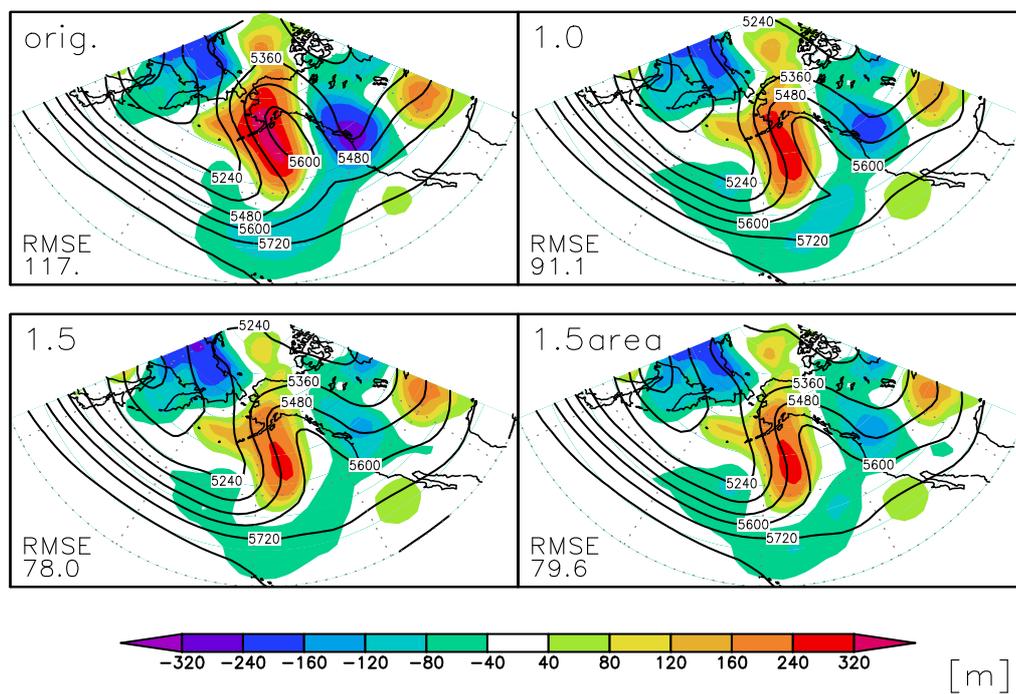


Fig. 3.13: 120-hr ensemble mean forecasts of 500 hPa height (contour) and its forecast error (shaded) for (a) NCEP original ensemble forecast, (b) JMA-GSM runs with NCEP analyses, (c) same as (b) but for globally amplified perturbations by a factor of 1.5, (d) same as (b) but for regionally amplified perturbations by a factor of 1.5, initialized at 12 UTC on 10th December 2005. RMSE is calculated over the blocking region (170°E–260°E, 20°N–80°N).

Chapter 4

Discussion

4.1 Construction method of Multi-Center Grand Ensemble

In this study, the MCGEs were constructed for the Z500 with equal weights among ensemble members and no bias correction. More than 1 year has passed since the TIGGE data archive started. The data accumulation by all NWP centers has just started. Although it was difficult to obtain the operational ensemble forecast data quasi-operationally due to the huge data size before the TIGGE project, one can easily obtain the operational ensemble forecast data quasi-operationally from the TIGGE data archives. As shown in Chapter 2, it has been possible to compare the single-center ensembles all over the world, and the forecast skills and forecast characteristics have been revealed. In order to construct better MCGEs, the appropriate weights among the members and bias correction might be needed. Especially in the MCGEs for the surface components, such as temperature at 2 m, U and V winds at 10 m, and precipitation, the weights among members and bias correction might be very important. In fact, the basic research on the combine methods was performed using the low-order Lorenz 1963 model (Johnson 2006). However, the data period is not enough to estimate the weights and bias correction appropriately. This might be easily expected from the following example. For example, the low-frequency variability, such as the Arctic Oscillation and atmospheric blocking, dominated in the previous winter, and the low-frequency variability does not dominate in this winter. In this case, it is obvious that it is not appropriate to apply the weights and bias estimated

from the data in the previous winter to the ensemble data in this winter. Through the TIGGE project, we have to investigate which atmospheric fields are more predicted and biased or not for each single-center ensemble. Also, to avoid the deficiency of enough ensemble data obtained from TIGGE portal, introduction of the Kalman filter into the bias correction technique might be possible. The Kalman filter technique repeatedly learns the information of the model bias obtained until immediately before, and updates it. So, the model bias derived from the Kalman filter might be more appropriate for the bias correction than the model bias based on the past same season.

Also, the ECMWF analysis was used in the verifications of the ECMWF ensemble and MCGEs. Each ensemble member of each NWP center was integrated on its own numerical model. Although the model climate of each numerical model is similar to each other, there are some differences between them. In other words, each model climate has some biases not only against the climate of its own model but also against the climate of the ECMWF model. In this respect, the comparison of the skills of the ECMWF ensemble and MCGEs in this study is advantageous for the ECMWF ensemble. The bias correction of the CMC, JMA, NCEP, and UKMO against the ECMWF analysis might lead to further improvement in the skill of the MCGEs. The development of methods of the weights and bias correction in the MCGE is one of the objectives of the TIGGE project. As shown in the Fig. 1.2, due to the rapid developments of the NWP technologies, it tends to be difficult for each single-center ensemble to improve the skill of the ensemble forecast in the last few years. This is easily expected from the fact that the ECMWF, having the best forecast skill in the world, actively promotes the TIGGE project. In spite of the simplest MCGEs with equal weights and no bias correction, it is remarkable that the MCGE can outperform the ECMWF ensemble at least in the medium forecast range. Also, it is valuable that the improvement rate of the MCGE against the ECMWF ensemble is almost comparable with that in the single-center ensemble forecast during the latest few years.

4.2 Identification of causes of forecast error and initial perturbation method

In Chapter 3, the collective mis-prediction of the blocking in the NCEP ensemble was shown. In order to identify the cause of the collective mis-prediction, the ensemble data, which is enough to conduct multi-analysis ensemble forecasts, was provided by Dr. Toth and Dr. Woubs. The total data size provided is 500GB. Recent internet speed enables us to transfer the huge ensemble forecast data. Such collaborations among the NWP centers and universities is one of the objectives of the TIGGE project.

It is found that the collective mis-prediction mainly resulted from the NCEP control analysis over the central north Pacific at 12 UTC on 10th December. The initial condition seems to be related to a cut-off cyclone. The difference of the DA system between the JMA and NCEP seems to appear remarkably due to the cut-off cyclone. Without the collaboration between the JMA and NCEP, it is difficult to identify the causes of the analysis differences. Many causes of the difference might be considered. For example, if a part of the observation and satellite data used in the JMA DA system were not used in the NCEP DA system, this might lead to the difference of the initial data around the cut-off cyclone. Possibly, by using a particular data, not used in the NCEP DA system but used in the JMA DA system, in the NCEP DA system, the initial difference related to the cut-off cyclone might be improved. In fact, the JMA DA system used 5 aircraft, 10 buoy, and 2 ship data around the cyclone over the sensitivity area when the global early analysis was made (Mr. Iriguchi and Dr. Miyoshi, personal communication). If these data were not used in the NCEP DA, these data can affect the control analysis of the JMA. In the interactive forecast system aimed by the THORPEX, after the identification of the sensitivity area against a target area using the SV method, an aircraft is driven to observe meteorological components over the sensitivity area, and then the DA is performed again with these new data, and reforecasts are performed. Yamaguchi et al. (2008) showed the improvement in the predicted typhoon track by the interactive forecast system for the 2004 DOTSTAR (Dropwindsonde Observation for Typhoon Surveillance near the TAIwan Region; Wu et al. 2007) cases. If additional JMA-GSM experiment using the JMA control

analysis made without the aircraft, buoy, and ship data shows same result as the NCEP original control run, it can be concluded that these data was essential for the analysis difference between the JMA and NCEP.

On the other hand, even if the control analysis has large uncertainty, there is a possibility that the initial perturbations of the ensemble forecast reduce it. However, the NCEP-EPS has smaller initial perturbation than the other EPSs in general. The initial perturbations seem to not reduce the initial uncertainty of the control analysis. As a result, all perturbed members predicted the wrong location of the blocking, as in the control run. Although the JMA-GSM runs with the NCEP analyses led to a little improvement in terms of the location of the blocking, the RMSE over the blocking region decreased for many members. This indicates the advantage of the multi-model and multi-analysis ensembles. In the near future, the multi-analysis ensemble forecast might be performed operationally. Also, additional ensemble forecasts with regionally amplified initial perturbations over the sensitivity area lead to the further improvement of the location of the blocking without the degradation of the forecast over the Northern Hemisphere. The NCEP plans to derive case dependent estimates from operational Gridpoint Statistical Interpolation (GSI) analysis, and use that when rescaling the initial ET perturbation (Dr. Toth, personal communication). The result in this study shows that such an approach may really have value as compared to climatologically based rescaling that is used widely.

Chapter 5

Conclusion

The NWP technique has progressed rapidly along with the development of the computer science. A five-day weather forecast today is as reliable as a three-day weather forecast 25 years ago which is a major scientific advance. Recently, the ensemble forecast has become a major component of the operational global weather prediction systems, and has drawn more attention in various timescales, such as short-, medium-, and long-ranges for both operational and research purposes.

The WMO began THE Observing system Research and Predictability experiment (THORPEX) project in 2005 in order to accelerate improvements in the accuracy of one-day to two-week high-impact weather forecasts for the benefit of society, the economy, and the environment. The THORPEX establishes an organizational framework that addresses weather research and forecast problems whose solutions will be accelerated through international collaboration among academic institutions, operational forecast centers, and users of forecast products. At the heart of the THORPEX is the research needed for the design and demonstration of a global interactive forecasting system that allows information to flow interactively between the forecast users, numerical forecast models, DA systems, and observations.

The THORPEX Interactive Grand Global Ensemble (TIGGE), which is a key component of the THORPEX, has enabled us to get operational medium-range ensemble forecast data quasi-operationally, to compare the medium-range ensemble forecasts, to construct the new ensemble forecast, and to analyze extreme events.

In this study, first, the overall intercomparisons of five operational medium-range ensembles: CMC, ECMWF, JMA, NCEP, and UKMO, were performed. In the deterministic verifications of each control run and each ensemble mean forecast, the daily and seasonal RMSE for 500 hPa geopotential height (Z500) over the Northern Hemisphere (NH, 20°N–90°N) from December 2006 to November 2007 were used.

In terms of the control runs, although the daily RMSEs are different from each other, the daily RMSEs tend to show similar variations along the atmospheric flow throughout the verification period. The ECMWF control run had the lowest RMSE until 168-hr lead time for all season. In particular, the ECMWF is far superior to the other centers in the early forecast range (day 0–3). The ECMWF control run was most skillful for all lead time in winter, but is comparable with the other control runs after 168-hr lead time in other seasons. The CMC control run had the largest RMSE for all season and all lead time. The JMA control run tended to have second-worst forecast skill for all season and all lead time. The NCEP control run tended to have better RMSE than the UKMO, and have sometimes comparable RMSE with the ECMWF in the medium forecast range.

In terms of the ensemble mean forecast, the ECMWF ensemble was most skillful for all season and all lead time. Especially in the autumn season, the ECMWF is far superior to the other centers. The second-best centers depended on the seasons. There seem to be no apparent differences among the forecast skills of the JMA, NCEP, and UKMO throughout the season. Although the CMC had largest RMSE in DJF 2007 and MAM 2007, the CMC is comparable with the JMA, NCEP, and UKMO after the system upgrade of the CMC-EPS on 9th July 2007. Based on these results, the CMC, JMA, NCEP, and UKMO can be considered as the second-best center.

The ensemble spreads of the single-center ensemble were also investigated. The JMA and NCEP ensembles have the largest and lowest spread for all seasons, respectively. The spread of the ECMWF showed good agreement with the RMSE of the ECMWF for almost all forecast range. In the medium forecast range, however, the spread against the large RMSE tended to be too small for all EPS.

Second, MCGEs were constructed using five medium-range ensemble forecasts: CMC,

ECMWF, JMA, NCEP, and UKMO. The forecast performance of the MCGEs relative to the ECMWF ensemble, having the best forecast skill in the world, was investigated using seasonal Root Mean Square Error (RMSE) and Ranked Probability Score (RPS) for the Z500 over the NH from December 2006 to November 2007.

As a result, it was found in the deterministic and probabilistic verifications that the MCGEs can outperform the ECMWF ensemble at least in the medium forecast range (day 6–9) for all seasons. The forecast time when the MCGEs first outperform the ECMWF ensemble is somewhat different depending on the season. During the summer season, the advantage of the MCGEs appears as early as at +4 day forecast time. The improvements in the RMSE and RPS are several percentage points in the medium forecast range. These are almost comparable with the rate of improvement in a single-center ensemble forecast during the latest few years.

In the early forecast range, the ensemble spread of the MCGE was larger than the RMSE as expected easily. The ensemble spread of the MCGE showed good agreement with the RMSE in the medium-range forecast range.

Third, the analysis of an extreme event, atmospheric blocking, was performed using ensemble forecast data, ensemble-based simple sensitivity analysis, and multi-analysis ensemble forecasts. The ensemble forecasts initialized at 12 UTC on 10th December 2005 were the very interesting case. All NCEP members were not able to predict the location of the blocking occurred on 15th December 2005 correctly, whereas almost all JMA members were able to predict it correctly. In order to identify the cause of the collective mis-prediction of the blocking, the multi-analysis ensemble forecast was performed using the NCEP analyses with the original initial perturbations. Although the decreases of the imperfection of the model formulation were recognized, it was found that the collective mis-prediction mainly resulted from the NCEP control analysis at 12 UTC on 10th December 2005. Next, the ensemble-based sensitivity analysis was performed in order to detect the sensitivity area against the blocking. As a result, the sensitivity area was detected over the central north Pacific. It was found that this was related to a cut-off cyclone. In the sensitivity area, the difference between the JMA and NCEP control anal-

yses measured by the dry total energy norm was relatively larger than the other areas. In other words, the sensitive area had large uncertainty. The NCEP ensemble, however, did not have effective initial perturbations to predict the blocking formation more accurately. Based on this fact, the multi-analysis ensemble forecasts were performed using the NCEP analyses with globally and regionally amplified initial perturbations. The amplitude of the NCEP initial perturbations was increased by a factor of 1.5 over the global area or only over the sensitivity area. Although the global amplification of the initial perturbation led to decrease of the RMSE over the blocking region, that also led to the degradation of the forecast skill over the NH. On the other hand, the regional amplification of the initial perturbation led to decrease of the RMSE over the blocking without the degradation of the forecast skill over the NH. These results indicate that the sensitivity area was essential for the prediction of the blocking. Also, they indicate that the excessive amplification of the initial perturbation over non-sensitivity area is undesirable and that the regional amplification technique can lead to better forecast without the degradation of the forecast over the other area. The result in this study shows that such a case dependent estimates may really have value as compared to climatologically based rescaling that is used widely.

Acknowledgements

I would like to show my appreciation to all members who have helped the progress of this study. Without the supports from members all over the world, this study could not have been done.

First, I would like to express special thanks to Prof. Hiroshi L. Tanaka, Center for Computational Sciences, University of Tsukuba, for his valuable comments and encouragements. The discussion on extreme events and ensemble forecast with him was very helpful in improving the presentation of this study. I also thanks to Profs. Fujio Kimura, Yousei Hayashi, Kenichi Ueno, Hiroaki Ueda, Hiroyuki Kusaka, and Mr. Keiichi Kondo, Graduate School of Life and Environment Sciences, University of Tsukuba, whose comments were valuable for this study.

Secondly, a part of this study was done under collaboration with Numerical Prediction Division of Japan Meteorological Agency (JMA). I acknowledge Mr. Masayuki Kyouda, Dr. Tadashi Tsuyuki, Mr. Yoshiaki Takeuchi, Mr. Ryouta Sakai, Mr. Takeshi Iriguchi, Mr. Mr. Hitoshi Yonehara, Dr. Takemasa Miyoshi, and Mr. Munehiko Yamaguchi in the JMA, and Dr. Takeshi Enomoto in the Earth Simulator Center. The stay in the operational numerical weather center for two years is priceless experience for me. In particular, I would like to express my appreciation to Mr. Masayuki Kyouda, Climate Prediction Division, JMA, for his kind supports, valuable discussion and encouragements. He taught me not only the basis of the ensemble forecast but also the application of the ensemble forecast. The authour is captivating by the ensemble forecast. Also, I appreciate to Dr. Tadashi Tsuyuki, Dr. Takeshi Enomoto, Dr. Takemasa Miyoshi, and Mr. Munehiko Yamaguchi for their valuable comments and discussion. Thanks are due to Mr.

Yoshiaki Takeuchi for providing me with the latest information on the TIGGE Working Group, to Mr. Hitoshi Yonehara for exchanging the latest information on the TIGGE archive, to Mr. Ryouta Sakai for providing me with the historical skill of the JMA-EPS, and to Mr. Takeshi Iriguchi for providing me with the list of the observation data. Also, the ensemble experiment with JMA-GSM was performed on the JMA supercomputer.

Thirdly, I would like to appreciate to overseas scientists, Dr. Zoltan Toth, Dr. Richard Wobus, Dr. Mozheng Wei in the NCEP, Prof. Eugenia Kalnay in the University of Maryland, Dr. Kenneth R. Mylne, Dr. Christine Johnson in the UKMO, and members of GIFS-TIGGE Working Group. In particular, I would like to express my appreciation to Dr. Zoltan Toth, Dr. Richard Wobus, and Dr. Mozheng Wei for providing me with huge ensemble forecast data and for valuable comments and discussion, and encouragements. A part of this study was made possible largely through the ensemble data. I would like to acknowledge here the generosity of them. Also, I appreciate to Prof. Eugenia Kalnay for reading draft of this study and encouraging, members of GIFS-TIGGE Working Group for valuable comments. Thanks are due to Dr. Kenneth R. Mylne and Dr. Christine Johnson for their valuable comments and discussion.

Finally, I would like to show special thanks to my parents. They have kept supporting me mentally and financially.

The figures in this study were made using GrADS (Grid Analysis and Display System) provided by the COLA (Center for Ocean Land Atmosphere Studies, Maryland) and GMT (Generic Mapping Tools) provided by the University of Hawaii.

References

- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903.
- Anderson, J. L., 1997: Impact of dynamical constraints on the selection of initial conditions for ensemble predictions: Low-order perfect model results. *Mon. Wea. Rev.*, **125**, 2969–2983.
- Atger F, 1999: The Skill of Ensemble Prediction Systems. *Mon. Wea. Rev.*, **127**, 1941–1953.
- Barkmeijer J., R. Buizza, and T. N. Palmer, 1999: 3D-Var Hessian singular vectors and their potential use in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2333–2351.
- Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive Sampling with Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.*, **129**, 420–436.
- Bishop, C. H. and Z. Toth, 1999: Ensemble Transformation and Adaptive Observations. *J. Atmos. Sci.*, **56** 1748–1765.
- Bourke, W., R. Buizza and M. Naughton, 2004: Performance of the ECMWF and the BoM Ensemble Prediction Systems in the Southern Hemisphere. *Mon. Wea. Rev.*, **132**, 2338–2357.
- Bowler, N. E., 2006: Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus*, **58A**,

538–548.

- Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 703–722.
- Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1–3.
- Bright, D. R. and Mullen S. L. 2002: Short-range ensemble forecasts of precipitation during the southwest monsoon, *Wea. Forecasting*, **17**, 1080–1100.
- Buizza, R., D. S. Richardson, and T. N. Palmer, 2003: Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man’s ensembles. *Quart. J. Roy. Meteor. Soc.*, **129**, 1269–1288.
- Buizza, R., and J. -R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Quart. J. Roy. Meteor. Soc.*, **133**, 681–695.
- Buizza, R., M. Miller, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908.
- Buizza, R., P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems. *Mon. Wea. Rev.*, **133**, 1076–1097.
- Buizza, R., and T. N. Palmer, 1995: The singular–vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456.
- Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Mon. Wea. Rev.*, **126**, 2503–2518.
- Buizza, R., T. Petroligis, T. N. Palmer, J. Barkmeijer, M. Hamrud, A. Hollingsworth, A. Simmons, and N. Wedi, 1998: Impact of model resolution and ensemble size on the

- performance of an ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **124**, 1935–1960.
- Candille, G., C. Côté, L. Houtekamer, and G. Pellerin, 2007: Verification of an Ensemble Prediction System against Observations. *Mon. Wea. Rev.*, **135**, 2688–2699.
- Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998a: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395.
- Côté, J., J. -G. Desmarais, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. Staniforth, 1998b: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part II: Results. *Mon. Wea. Rev.*, **126**, 1397–1418.
- Descamps L. and O. Talagrand, 2007: On some aspects of the definition of initial conditions for ensemble prediction. *Mon. Wea. Rev.*, **135**, 3260–3272.
- Duchon, C. E., 1979: Lanczos filtering in one and two dimensions. *J. Appl. Meteor.*, **18**, 1016–1022.
- Ehrendorfer, M., and R. M. Errico, 1995: Mesoscale predictability and the spectrum of optimal perturbation. *J. Atmos.*, **52**, 3475–3500.
- Enomoto T., W. Ohfuchi, H. Nakamura, and M. A. Shapiro, 2007: Remote effects of tropical storm Cristobal upon a cut-off cyclone over Europe in August 2002, *Meteorol. Atmos. Phys.*, **96**, 29–42.
- Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteorol.*, **8**, 985–987.
- Errico, R. M., 1997: What is an adjoint model?. *Bull. Amer. Meteor. Soc.*, **78**, 2577–2592.
- Evans, R. E., M. S. J. Harrison, R. J. Graham, and K. R. Mylne, 2000: Joint Medium-Range Ensembles from The Met. Office and ECMWF Systems. *Mon. Wea. Rev.*, **128**, 3104–3127

- Evensen G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99C5**, 10143–10162.
- Evensen G., 2003: The ensemble Kalman filter: theoretical formulation and practical implementation., *Ocean Dynamics*, **53**, 343–367.
- Evensen, G. 2007. *Data Assimilation: The Ensemble Kalman Filter*. Springer, Berlin, 279 pp.
- Ferreira, N., 2007: The use of medium range and seasonal forecast at CPTEC. *Presentation, ECMWF 11th Workshop on Meteorological Operational Systems*, Reading, United Kingdom, ECMWF.
- Froude, L. S. R., L. Bengtsson, and K. I. Hodges, 2007: The Prediction of Extratropical Storm Tracks by the ECMWF and NCEP Ensemble Prediction Systems. *Mon. Wea. Rev.*, **135**, 2545–2567.
- Graham, R. J., A. D. L. Evans, K. R. Mylne, M. S. J. Harrison, and K. B. Robertson, 2000: An assessment of seasonal predictability using atmospheric general circulation models. *Quart. J. Roy. Meteor. Soc.*, **126**, 2211–2240.
- Gray, M. E. B. and Shutts G. J., 2002: A stochastic scheme for representing convectively generated vorticity sources in general circulation model, *APR Turbulence and Diffusion Note. No.285, Met Office, UK*.
- Hamill, T. M. 2003: Ensemble-based atmospheric data assimilation: A tutorial. NOAA-CIRES Climate Diagnostics Center, 46 pp.
- Hamill, T. M., C. Snyder and R. E. Morss 2000: A Comparison of Probabilistic Forecasts from Bred, Singular-Vector, and Perturbed Observation Ensembles. *Mon. Wea. Rev.*, **128**, 1835–1851.
- Hoffman, R. N. and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, **35A**, 100–118.

- Houtekamer, P. L., and H. L. Mitchell, 2005: Ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **131**, 3269–3289.
- Houtekamer, P. L., and J. Derome, 1995: Methods for Ensemble Prediction. *Mon. Wea. Rev.*, **123**, 2181–2196.
- Houtekamer, P. L., L. Lefaiivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.
- Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133**, 604–620.
- Hunt, B. R., E. J. Kostelich and I. Szunyogh, 2007: Efficient Data Assimilation for Spatiotemporal Chaos: A Local Ensemble Transform Kalman Filter. *Physica D*, 230, 112-126.
- Hunt, B. R., E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J. A. Yorke, and A. V. Zimin, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277.
- JMA, 2007: *Outline of the operational numerical weather prediction at the Japan Meteorological Agency*. Japan Meteorological Agency, 196 pp.
- Johnson, C. 2006: On the benefits of multi-model ensembles: idealized studies using the Lorenz 1963 model. *UKMO Forecasting Research Technical Report*, **492**, 47pp.
- Kalnay, E. 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, Cambridge, 341 pp.
- Kimoto, M., H. Mukougawa, and S. Yoden, 1992: Medium-range forecast skill variation and blocking transition: A case study. *Mon. Wea. Rev.*, **120**, 1616–1627.
- Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548–1550.

- Krishnamurti, T. N., C. M. Kishtawal, D. W. Shin, and C. E. Williford, 2000a: Improving Tropical Precipitation Forecasts from a Multianalysis Superensemble. *J. Climate*, **13**, 4217–4227.
- Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, 2000b: Multimodel Ensemble Forecasts for Weather and Seasonal Climate. *J. Climate*, **13**, 4196–4216.
- Krishnamurti, T. N., S. Surendrab, D. W. Shin, R. J. Correa-Torres, T. S. V. V. Kumar, E. Williford, C. Kummerow, R. F. Adler, J. Simpson, R. Kakar, W. S. Olson, and F. J. Turk, 2001: Real-Time Multianalysis-Multimodel Superensemble Forecasts of Precipitation Using TRMM and SSM/I Products. *Mon. Wea. Rev.*, **129**, 2861–2883.
- Lin, J. W. B. and J. D. Neelin, 2000: Influence of a stochastic moist convective parameterization on tropical climate variability, *Geophys. Res. Lett.*, **27**, 3691–3694.
- Lefaiivre, L., P. L. Houtekamer, A. Bergeron, and R. Verret, 1997: The CMC ensemble prediction system. *Proc. Sixth Workshop on Meteorological Operational Systems*, Reading, United Kingdom, ECMWF, 31–44.
- Legras, B. and R. Vautard, 1996: A guide to Liapunov vectors. *Proc. ECMWF Seminar on Predictability*, Reading, United Kingdom, ECMWF, 143–156.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- Lorenz, E. N., 1996: A problem partly solved. *Proc. Workshop on Predictability*, Vol. 1, Reading, United Kingdom, ECMWF, 1–18.
- Matsueda M., M. Kyouda, H. L. Tanaka and T. Tsuyuki, 2007: Daily forecast skill of Multi-Center Grand Ensemble. *SOLA*, **3**, 29–32.
- Matsueda M., M. Kyouda, H. L. Tanaka and T. Tsuyuki, 2006: Multi-Center Grand Ensemble using three operational ensemble forecasts. *SOLA*, **2**, 33–36.
- Mitchell, H. L., and P. L. Houtekamer., 2002: Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **130**, 2791–2808.

- Miyoshi T. and K. Aranami, 2006: Applying a Four-dimensional Local Ensemble Transform Kalman Filter (4D-LETKF) to the JMA Nonhydrostatic Model (NHM). *SOLA*, **2**, 128–131.
- Miyoshi T., S. Yamane, and T. Enomoto, 2007: The AFES-LETKF Experimental Ensemble Reanalysis: ALERA. *SOLA*, **3**, 45–48.
- Miyoshi T., and S. Yamane, 2007: Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution. *Mon. Wea. Rev.*, **135**, 3841–3861.
- Miyoshi T. and Y. Sato, 2007: Assimilating Satellite Radiances with a Local Ensemble Transform Kalman Filter (LETKF) Applied to the JMA Global Model (GSM). *SOLA*, **3**, 37–40.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis., 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119.
- Mullen, S. L., and R. Buizza, 2002: The impact of Horizontal resolution and ensemble size on probabilistic forecasts of precipitation by the ECMWF ensemble prediction system. *Wea. Forecasting*, **17**, 173–191.
- Murphy, A. H., 1971: A note on the ranked probability score. *J. Appl. Meteorol.*, **10**, 155–156.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772.
- Mylne, K. R., R. E. Evans, and R. T. Clark, 2002: Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quart. J. Roy. Meteor. Soc.*, **128**, 361–384.
- Ott, E, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, J. A. Yorke, 2004: A local ensemble Kalman filter for atmospheric data assimilation., *Tellus*, **56A**, 415–428.

- Palmer, T. N., 2001: A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parameterization in weather and climate prediction models. *Quart. J. Roy. Meteor. Soc.*, **127**, 279–304.
- Palmer, T. N., F. J. Doblas-Reyes, R. Hagedorn, A. Alessandri, S. Gualdi, U. Andersen, H. Feddersen, P. Cantelaube, J.-M. Terres, M. Davey, R. Graham, P. Delecluse, A. Lazar, M. Deque, J. -F. Gueremy, E. Diez, B. Orfila, M. Hoshen, A. P. Morse, N. Keenlyside, M. Latif, E. Maisonnavé, P. Rogel, V. Marletto and M. C. Thomson, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor.*, **85**, 853–872.
- Parrish, D. F., and Derber, J., 1992: The National Meteorological Center’s spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.
- Park, Y.-Y., 2007: Numerical weather forecast and digital forecast at KMA. *Presentation, ECMWF 11th Workshop on Meteorological Operational Systems*, Reading, United Kingdom, ECMWF.
- Pellerin, G., L. Lefaivre., P. Houtekamer, and C. Girard, 2003: Increasing the horizontal resolution of ensemble forecasts at CMC. *Non. Proc. Geophys.*, **10**, 463–468.
- Pellerin, G., M. Charron, P. L. Houtekamer, L. Lefaivre, and H. L. Mitchell, 2005: Reviewing the Ensemble Prediction System at the Canadian Meteorological Centre. *Proceedings, ECMWF Tenth Workshop on Meteorological Operational Systems*, Reading, U.K., ECMWF, 63–71.
- Richardson, D. S., 2001a: Ensembles using multiple models and analyses. *Quart. J. Roy. Meteor. Soc.*, **127**, 1847–1864.
- Richardson, D. S., 2001b: Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size. *Quart. J. Roy. Meteor. Soc.*, **127**, 2473–2489.
- Ritchie, H. 1991: Application of the semi-Lagrangian method to a multilevel spectral primitive-equations model. *Quart. J. Roy. Meteor. Soc.*, **117**, 91–106.

- Ritchie, H., and C. Beaudoin, 1994: Approximations and sensitivity experiments with a baroclinic semi-Lagrangian spectral model. *Mon. Wea. Rev.*, **122**, 2391–2399.
- Simmons, A., G. Kelly, J. -N. Thépaut, D. Tan, and E. Andersson, 2006: The ECMWF forecasting system: data impacts and prospects for ADM-Aeolus. *Presentation, The ADM-Aeolus Workshop*. 26–28 September 2006, Noordwijk, Netherlands.
- Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. F. Frame, J. A. Kettleborough, S. Knight, A. Martin, J. M. Murphy, C. Pian, D. Sexton, L. A. Smith, R. A. Spicer, A. J. Thorpe, and M. R. Allen, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature*, **433**, 403–406.
- Stensrud, D. J., H. E. Brooks, J. D. M. S. Tracton, and E. Rogers, 1999: Using Ensembles for Short-Range Forecasting. *Mon. Wea. Rev.*, **127**, 433–446.
- Szunyogh, I., and Z. Toth, 2002: The Effect of Increased Horizontal Resolution on the NCEP Global Ensemble Mean Forecasts. *Mon. Wea. Rev.*, **130**, 1125–1143.
- Szunyogh, I., E. Kalnay, and Z. Toth, 1997: A comparison of Lyapunov vectors and optimal vectors in a low resolution GCM. *Tellus*, **49A**, 200–227.
- Talagrand, O., 1981: A study of the dynamics of four-dimensional data assimilation. *Tellus*, **33**, 43–60.
- Toth, Z. and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330.
- Toth, Z. and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319.
- Tracton, M. S. and E. Kalnay, 1993: Operational Ensemble Prediction at the National Meteorological Center: Practical Aspects. *Wea. Forecasting*, **8**, 379–398.
- Uppala, S. M., and Coauthors 2005: The ERA-40 re-analysis. *Quart. J. R. Meteorol. Soc.*, **131**, 2961–3012.

- Wang, X. and C. H. Bishop, 2003: A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes. *J. Atmos. Sci.*, **60**, 1140–1158.
- Wang, X., C. H. Bishop, and S. J. Julier, 2004: Which is better, an ensemble of positive/negative pairs or a centered spherical simplex ensemble? *Mon. Wea. Rev.*, **132**, 1590–1605.
- Wei, M., Z. Toth, R. Wobus, Y. Zhu, and C. H. Bishop, 2005a. Initial perturbations for NCEP Ensemble Forecast System. First THORPEX Internal Science Symposium. 6–10 December 2004, Montreal, Canada. The Symposium Proceedings in a WMO Publication 2005, *WMO/TD-No.1237*, WWRP THORPEX No. 6, 2005. p227–230.
- Wei, M., Z. Toth, R. Wobus, Y. Zhu, D. Hou, and coauthors. 2005b. NCEP Global Ensemble: recent developments and plans. 2nd SRNWP Workshop on Short Range Ensemble. 7–8 April 2005, Bologna, Italy. Available at <http://smwp.cscs.ch/LeadCentres/2005Bologna/Agenda.htm>
- Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus*, **60A**, 62–79.
- Wei, M., Z. Toth, R. Wobus, Y. Zhu, C. H. Bishop, and X. Wang, 2006a: Ensemble Transform Kalman Filter-based ensemble perturbations in an operational global prediction system at NCEP. *Tellus*, **58A**, 28–44.
- Wei, M., Z. Toth, R. Wobus, Y. Zhu, and C. H. Bishop, 2006b. The Use of Ensemble Transform Technique for Generating Initial Ensemble Perturbations. NOAA Thorpex PI Workshop at NCEP. 17–19 January 2006, Maryland, America. Available at <http://www.emc.ncep.noaa.gov/gmb/ens/THORPEX/PI-shop-2006.html>
- Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2006c. Initial Perturbations based on the Ensemble Transform (ET) Technique in the NCEP global Ensemble Forecast System. *US Department of Commerce, NOAA/NCEP Office Note*, **453**, 33 pp.
- Whitaker, J. S. and Hamill, T. M. 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

- Wilks, D. S. 2006: *Statistical methods in the atmospheric sciences (2nd Ed)*. International Geophysics Series Vol. 91, Academic Press, 627 pp.
- WMO, 2005: THORPEX. WMO-No. 978, 15pp.
- WMO, 2007: Report of the twenty second session of the CAS/JSC working group on numerical experimentation. National Center for Atmospheric Research, Boulder, USA, 25-27 October 2006. *WMO/TD-No.1411*, 57pp.
- Wu, C. -C., and coauthors, 2005: Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR). *Bull. Amer. Meteor.*, **86**, 787–790.
- Yamaguchi, M., T. Iriguchi, T. Nakazawa, and C. -C. Wu, 2008: An observing system experiment for Typhoon Conson (2004) using a singular vector method and DOTSTAR data. *submitted to Mon. Wea. Rev.*.
- Ziehmann, C., 2000: Comparison of a single-model EPS with a multi-model ensemble consisting of a few operational model. *Tellus*, **52A**, 280–299.