

An Evaluation of the Software System Dependency of a Global Atmospheric Model

SONG-YOU HONG, MYUNG-SEO KOO, AND JIHYEON JANG

Department of Atmospheric Sciences, Yonsei University, Seoul, South Korea

JUNG-EUN ESTHER KIM

Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, and National Oceanic and Atmospheric Administration/Earth System Research Laboratory, Boulder, Colorado

HOON PARK

Department of Atmospheric Sciences, Yonsei University, and Numerical Weather Prediction Center, Korea Meteorological Administration, Seoul, South Korea

MIN-SU JOH AND JI-HOON KANG

Supercomputer Center, Korea Institute of Science and Technology Information, Daejeon, South Korea

TAE-JIN OH

Korea Institute of Atmospheric Prediction Systems, Seoul, South Korea

(Manuscript received 11 December 2012, in final form 8 July 2013)

ABSTRACT

This study presents the dependency of the simulation results from a global atmospheric numerical model on machines with different hardware and software systems. The global model program (GMP) of the Global/Regional Integrated Model system (GRIMs) is tested on 10 different computer systems having different central processing unit (CPU) architectures or compilers. There exist differences in the results for different compilers, parallel libraries, and optimization levels, primarily a result of the treatment of rounding errors by the different software systems. The system dependency, which is the standard deviation of the 500-hPa geopotential height averaged over the globe, increases with time. However, its fractional tendency, which is the change of the standard deviation relative to the value itself, remains nearly zero with time. In a seasonal prediction framework, the ensemble spread due to the differences in software system is comparable to the ensemble spread due to the differences in initial conditions that is used for the traditional ensemble forecasting.

1. Introduction

Numerical atmospheric models for weather forecasting and climate research include a software package that consists of numerically discretized mathematical equations written in programming languages, such as FORTRAN and C. Massively parallel computing cluster computers, which contain many networked processors, are commonly used to achieve superior computational performance

for numerical modeling. The message-passing interface (MPI) is the most commonly used package for massively parallel processing (MPP). The MPI parallel codes of an atmospheric model may not produce the same output if they are ported to a new software system that is defined as the computational platform that includes the parallel communication library, compiler, and its optimization level in this study. Thus, it is important to consider the ability of software system to optimize the codes that consist of an atmospheric model, as well as the tolerance of software system and its impact on model results used in weather forecasting and climate prediction.

Rounding error is another important consideration in atmospheric modeling that can arise because of various

Corresponding author address: Song-You Hong, Department of Atmospheric Sciences, College of Science, Yonsei University, Seoul 120-749, South Korea.
E-mail: songyouhong@gmail.com

characteristics of the computer software implementing the model, including: 1) the rounding method, 2) the order of arithmetic calculations, 3) intrinsic functions, and 4) mathematical libraries, such as the fast Fourier transform (FFT). The binary output from an atmospheric model run on two different machines could be identical only if these characteristics are identical. However, matching output cannot be ensured because two different machines may have different software system as well as hardware system such as unique central processing unit (CPU) architectures. For example, porting a large-eddy simulation (LES) code produces results that are too difficult to compare, and turbulent flows produced by LES are highly sensitive to the number of processors, small changes in initial condition, and floating-point precision (Senoner et al. 2008). Rosinski and Williamson (1997) suggested a necessary condition for successful port that the growth of the difference between two machines should not be worse than the growth of a minimal perturbation imposed on the machine with accurate floating-point representation. Thomas et al. (2002) examined the impact of floating-point optimization, mathematics libraries, and processor configuration on forecast accuracy in a regional model and found that the iterative solver in the dynamical core is most sensitive to processor configuration.

It is important to consider the degree to which computer architecture-based differences in atmospheric models should be accommodated, and under what circumstances. For example, differences in output tolerance could be dependent on model integration time. This study seeks to address this issue by integrating a global atmospheric model on various software systems. Twenty computational platforms are chosen. The global model program (GMP) of the Global/Regional Integrated Model system (GRIMs; Hong et al. 2013; <http://grims-model.org>) is integrated for 10 days to examine the reproducibility of weather forecasts. A seasonal climatology, which is an ensemble seasonal average of the results of a 4-month-long integration, is compared to examine the effect of rounding error on atmospheric simulations. We address the tolerance question using the 500-hPa geopotential height spread for medium-range forecasts and the software system ensemble spread for seasonal climate simulations.

Section 2 describes the experimental setup, section 3 describes the results, and the paper is concluded in section 4.

2. Model and experimental setup

a. Model descriptions

The GRIMs is used in this study. Model physics include longwave and shortwave radiation, cloud–radiation interaction, planetary boundary layer processes, shallow

convection, gravity wave drag, enhanced topography, hydrology, and vertical and horizontal diffusion. For precipitation physics, the GRIMs employs both grid-resolvable precipitation processes and subgrid processes through cumulus parameterization scheme. The model is flexible and can be implemented on multiple platforms from personal to supercomputers in either thread or MPI modes. The dynamical core of the GRIMs-GMP can be either spherical harmonics (SPH) or double Fourier series (DFS; Cheong 2006; Park et al. 2013; Koo and Hong 2013). Within the identical C-preprocessor (cpp) source codes, the regional model program (RMP) and single-column model program (SMP) can be selected for regional downscaling and physics algorithm validation, respectively. The system also has multiple options in each physics parameterization for physics development as well as mechanism studies. The GRIMs-GMP is also coupled with the Modular Ocean Model, version 3 (MOM3; Pacanowski and Griffies 1998) for climate prediction and mechanism studies. The regional model program with the scale-selective spectral nudging (Hong and Chang 2012) has been applied to downscaling of climate changes. The GRIMs version utilized in this study is the GRIMs-GMP with the SPH dynamical core and physics package, version 3.2. See Hong et al. (2013) and the GRIMs official website (<http://www.grims-model.org>) for more details.

b. Experimental design

Table 1 shows the 20 computing environments including FORTRAN compilers, parallel communication libraries, and optimization levels of the compilers. The Yonsei University (YSU) Linux cluster is equipped with 12 Intel Xeon CPUs (model name: X5650) per node and supports the PGI and Intel FORTRAN compilers. The Korea Institute of Science and Technology Information (KISTI; <http://www.kisti.re.kr>) provides a computing environment with high-performance IBM and SUN platforms. Each platform is equipped with different CPU: Intel Xeon X5570 for KISTI-SUN2 platform, Power5 + processor of Power 595 server for KISTI-IBM1 platform, and Power6 dual-core processor of p5 595 server for KISTI-IBM2 platform. Each machine has a different architecture and approximately 500–20 000 CPUs.

For medium-range forecasts, the 10-day simulations are initiated at 0000 UTC 14 July 2001. The selected period represents a heavy rainfall case over Korea. A significant amount of precipitation was recorded in Korea on 15 July 2001, with a local maximum of approximately 371.5 mm near Seoul, South Korea [see Hong and Lim (2006) for further detail]. Two horizontal resolution test beds are chosen, namely T62 and T254, which truncate the horizontal wavenumber at 62 and 254, respectively. The number of grid points in Cartesian coordinates is

TABLE 1. Computing environment including FORTRAN compilers, parallel communication libraries, and optimization levels of the compiler. Identical results are marked by a symbol. Ten ensemble members with different software system are highlighted in boldface.

Name	Machine	FORTRAN compiler	Parallel communication library	Optimization level	Mark
EXP1	KISTI SUN2	INTEL 11.1	openmpi 1.4	O3	□
	KISTI SUN2	INTEL 11.1	mvapich2 1.5	O3	□
EXP2	KISTI SUN2	INTEL 11.1	mvapich1 1.2	O3	○
	KISTI SUN2	INTEL 11.1	openmpi 1.4	O4	□
EXP3	KISTI SUN2	INTEL 11.1	openmpi 1.4	O2	△
EXP4	KISTI SUN2	INTEL 11.1	openmpi 1.4	O1	<
EXP5	KISTI SUN2	INTEL 11.1	openmpi 1.4	O0	>
EXP6	KISTI SUN2	PGI 9.0.4	openmpi 1.4	O2 (-fastsse)	■
	KISTI SUN2	PGI 9.0.4	mvapich2 1.5	O2 (-fastsse)	■
	KISTI SUN2	PGI 9.0.4	mvapich1 1.2	O2 (-fastsse)	■
	KISTI SUN2	PGI 8.0.6	mvapich1 1.2	O2 (-fastsse)	■
	YSU Cluster	PGI 10.6	mvapich1 1.2	O2 (-fastsse)	■
	YSU Cluster	PGI 10.6	mvapich1 1.2	O3 (-fastsse)	■
EXP7	YSU Cluster	PGI 10.6	mvapich1 1.2	O1	●
EXP8	YSU Cluster	PGI 7.1.6	mvapich1 1.2	O2 (-fastsse)	▲
EXP9	KISTI IBM 1	XLF 10.1	—	O3	★
	KISTI IBM 2	XLF 12.1	—	O3	★
	KISTI IBM 1	XLF 10.1	—	O4	★
EXP10	KISTI IBM 1	XLF 10.1	—	O2	▲
	KISTI IBM 1	XLF 10.1	—	O1	▲

192 (360) × 94 (181) for the T62 (T254) test bed, which corresponds to approximately 200 km (50 km) in mid-latitude. The number of vertical layers is 28. The observed precipitation dataset to verify the model results is the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007), which has a 0.25° × 0.25° spatial resolution and 3-hourly time interval over the time range ±90 min within the global latitude belt of 50°S–50°N. The National Centers for Environmental Prediction (NCEP) Final Analysis (FNL; available online at <http://rda.ucar.edu/datasets/ds083.2/>), which employs a 1.0° × 1.0° global grid and 6-hourly time step, is used for the initialization and evaluation of synoptic features.

For seasonal simulations, the boreal summer in June–August (JJA) of 1996, which is the normal year in terms of sea surface temperature (SST) over the Pacific, is chosen. To estimate and filter out the unpredictable part of the flow, 10-member initial condition ensemble runs are performed. These ensemble runs are initialized at 0000 UTC from 1 to 10 May 1996. In addition, 10 software system ensemble runs are done with different computer platforms of EXP1 to 10 experiments as described in Table 1. Initial conditions are obtained from the NCEP/Department of Energy (DOE) Atmospheric Model Intercomparison Project II (AMIP-II) Reanalysis-2 (RA2) data on a 2.5° × 2.5° grid (Kanamitsu et al. 2002), which is also used for validation of the results. The observed SST is updated daily from the optimal interpolation SST weekly dataset (Reynolds and Smith 1994). Observed monthly precipitation data from

the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) data (Xie and Arkin 1997) were used to evaluate the modeled precipitation.

Standard deviation (std), used to investigate equivalence of the results according to the software system, is defined as

$$\text{std} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}, \quad (1)$$

where n is the number of the simulations, x_i is the prognostic variable at i grid point, and \bar{x} is the average of the prognostic variable from the n simulations. The 500-hPa geopotential height and its standing eddy are used to validate the results for medium-range forecasts and seasonal simulation, respectively. The standing eddy is defined as the time-averaged departure from zonal mean at the same latitude and is used to validate the model's general ability (Seol and Hong 2009; Hong et al. 2013).

3. Results

a. General

All the experiments are carried out with the double precision since the default source codes in the GRIMs are compiled only for IEEE double-precision computations with the “-r8” option. Before discussing the software system dependency, it was confirmed that for a given system the GMP produces identical binary output

irrespective of the number of processors in MPI mode. In other words, the simulation results remain the same although the number of processors or nodes is changed for the software combinations presented in Table 1, as well when the GMP is executed using a single processor. The identical binary result irrespective of the processor number is a necessary condition that the parallelized code is bug free. In practice, it was found that the lack of identical binary results allowed us to identify a bug in the parallelized code. In addition, the same results were achieved irrespective of hardware system if the used compiler and parallel library are identical. For example, when PGI compiler with mvapich1 library is used, the YSU cluster produces results that are identical to those from the KISTI-SUN2 system. This result implies that the differences in hardware system can be less important than the kind of software system.

The PGI compiler produces identical results for a given optimization level in versions 8, 9, and 10; however it produces different results with version 7. This result holds for the IBM compiler versions 10 and 12 as well (Table 1). We also tested sensitivity to optimization level of the Intel compiler. The simulation results are the same between the O3 and O4 levels, but are different among the other levels less than O3. The improvement of run-time performance is noticeable only in level-one (O1) optimization, whereas more aggressive optimizations do not significantly affect the speed up (not shown).

In Table 1, the mark indicates consentaneity of the simulation result and implies that the result is exactly identical with all those results achieved using the same compiler irrespective of the computing environment in which the compiler was implemented. Therefore, the number of experiments related to software system is confined to 10 with no inclusion of the experiments that produce the exact same results, as will be discussed in the following subsections.

b. Medium-range forecasts

Figure 1 shows the daily accumulated precipitation and sea level pressure (SLP) over East Asia at 0000 UTC 15 July 2001 at a T254 resolution. The observed precipitation in Fig. 1a was interpolated onto the model grid and shows major precipitation over Korea and the southern part of China, along a summer monsoon front. A subtropical high is centered over the oceans east of Japan. It is apparent that the model reproduces the observed pattern well, although the amount of precipitation along the monsoon front is overestimated (cf. Figs. 1a,b). It was also confirmed that the model reproduces the synoptic-scale features over the globe well (not shown). Discernible discrepancies in precipitation amount exhibit along the monsoon front when its difference fields

between the experiments were plotted (figures not shown). However, there were no systematic patterns in differences from an experiment to another. The resemblance of Figs. 1b,c assures that the model output is not altered by a given software system from a forecast point of view. The smaller standard deviation of the simulated precipitation and SLP over East Asia supports our assurance that the software system would not negatively interfere (Fig. 1d). The maximum 24-h accumulated precipitation over Korea (33°–43°N, 120°–133°E) at T254 resolution, averaged from 10-member software system experiments, is 287.64 mm day⁻¹, and the standard deviation of the maximum precipitations from 10 software-system simulations is 4.52 mm day⁻¹. Since the standard deviation is 1.6% of the average over Korea, the software system would not allow a spurious overestimation near the precipitation core and the results can be considered to be equivalent.

Figure 2 shows the time series of the globally averaged standard deviation for the 500-hPa geopotential height, obtained from the 10 software-system experiments. It is apparent that the standard deviation increases with time at both the T62 and T254 resolutions (Fig. 2a). At the 5-day forecast, the deviations are 3.17 and 2.66 m, respectively, for the T62 and T254 resolutions. The differences at the 10-day forecast are 10.60 and 16.24 m for each resolution. This increasing pattern reflects the chaotic nature of the modeled atmosphere (Lorenz 1963). Lorenz found that small differences in initial conditions can produce different solutions in nonlinear systems such as the atmosphere. Here, slight differences originating from the dissimilar software system induce increasingly unpredictable conditions with time. A rapid growth of the standard deviation is prominent over the first day, and a slower growth continues after day 2. The increase in the deviation exhibited before day 2 may be attributable to the imbalanced initial condition. The standard deviation was reduced when the 24-h forecast results were used for the initial condition (figures not shown), which reflects that the gravity wave oscillation in the initial time is exerted when the FNL data are used for the initial conditions. A higher-resolution model that resolves smaller scales would lead to a larger standard deviation with forecast time. Thus, the standard deviation is larger in the T254 experiment after day 6.

In Fig. 2a, the fractional changes of the globally averaged standard deviations are calculated and compared between the resolutions. The fractional tendency (R_{std}) is defined as

$$R_{\text{std}}(h) = \frac{\overline{\text{std}}_{h+1} - \overline{\text{std}}_h}{\overline{\text{std}}_h}, \quad (2)$$

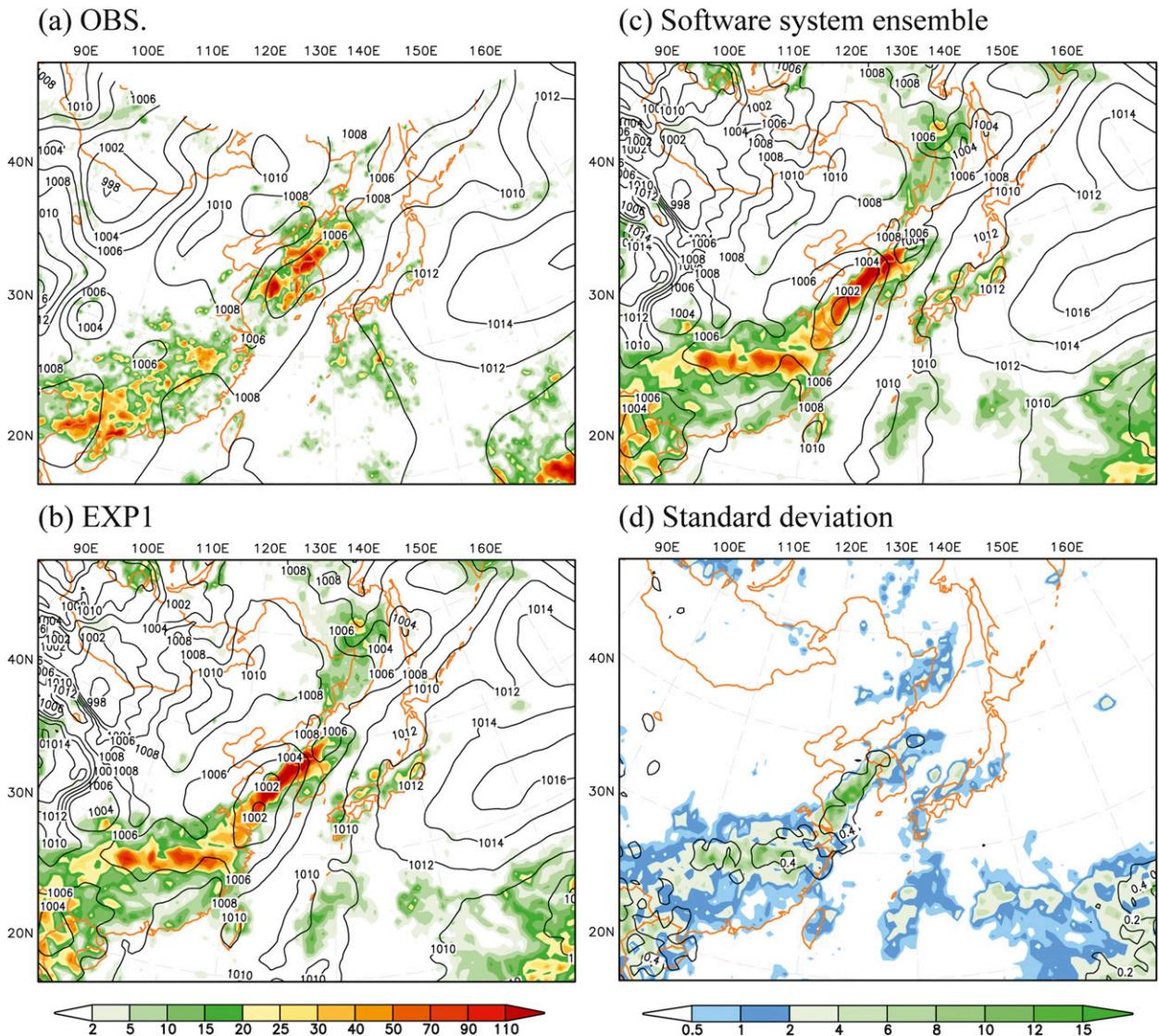


FIG. 1. Daily accumulated precipitation (mm day^{-1} ; shading) and sea level pressure (hPa; contour) over East Asia at 0000 UTC 15 Jul 2001, obtained from the (a) TMPA and FNL data, (b) EXP1 run, (c) ensemble mean, and (d) standard deviation of 10-simulations from the EXP1 to 10 runs at T254 ($\sim 50 \text{ km}$) resolution.

where h is the output interval at 1 h, and $\overline{\text{std}}$ is an area-weighted average over the globe. The fractional changes are almost flat and become comparable with zero tendencies at each resolution after day 2. In Fig. 2b, the fractional standard deviations of RMSEs, the ratio of the standard deviation to the mean of RMSEs for the 10 simulations averaged over the globe, are remarkably small. The RMSE over the analyzed fields increases with time at both resolutions, as expected, but the fractional standard deviation of the RMSE remains less than 0.5% and 2% at day 5 and day 10, respectively. Since the differences of the model results and their forecast skills due to the software system are small and their changes as

time passes are almost flat, the model is considered to simulate the equivalent patterns regardless of the software system.

c. Seasonal simulation

This section focuses on the mean pattern of the 10-member ensemble simulations for the boreal summer of 1996, as the authors intend to evaluate the mean features of the tropical rainfall and atmospheric structure with considering uncertainties related to the initial data and the software system. As shown in Fig. 3, the both ensemble results produce a tropical rainfall pattern comparable to the observation, with the main rainbelt along the intertropical

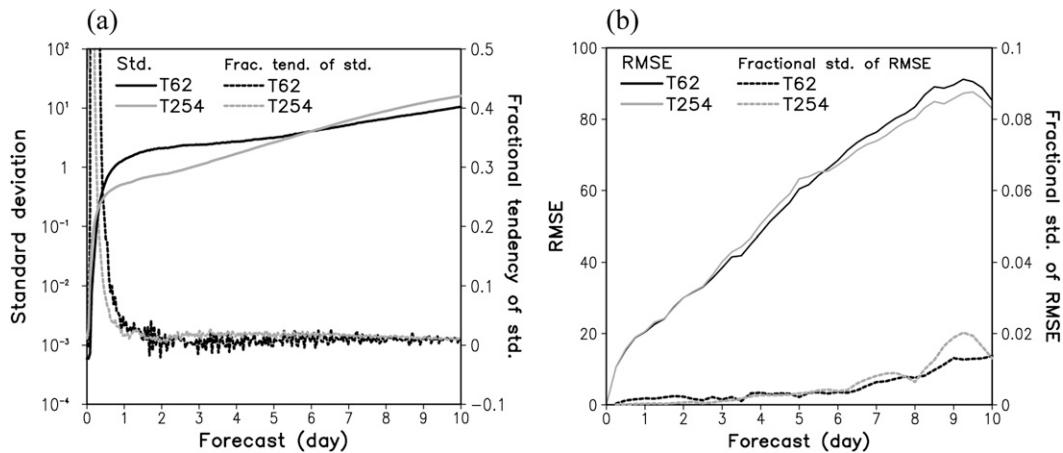


FIG. 2. Time series of (a) globally averaged standard deviation (solid) and fractional tendency of the standard deviation (dotted), and of (b) RMSE of the ensemble mean (solid) and fractional standard deviation of RMSEs (dotted) of the 500-hPa geopotential height (m) simulated from the EXP1 to 10 runs at T62 (black) and T254 (gray) resolutions.

convergence zone (ITCZ; cf. Figs. 3a–c). It confirms the resemblance that 3-month-averaged daily precipitations simulated from the 10-member initial condition ensemble runs are within a narrow range between 3.387 and 3.395 mm day^{-1} . Those from the 10-member software

system ensemble runs are within a similar range between 3.388 and 3.397 mm day^{-1} . In addition, the maximum and minimum pattern correlations between the simulated precipitations with different initial conditions (software systems) are 0.991 (0.992) and 0.987 (0.987), respectively.

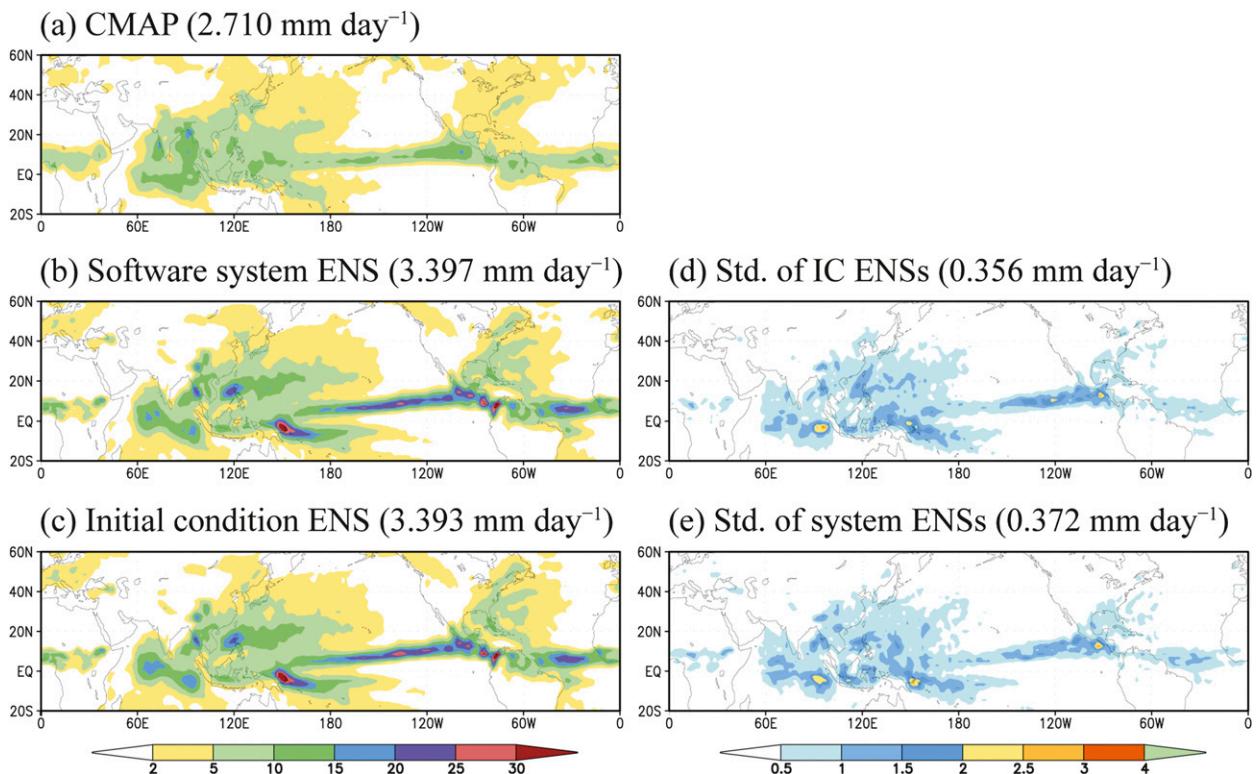


FIG. 3. Three-month-averaged precipitation (mm day^{-1}) during the period of JJA 1996 from the (a) CMAP data, (b) ensemble mean for different software systems with the initial condition of 1 May 1996, and (c) ensemble mean for different initial conditions with the software system of EXP1. Standard deviation among 10 sets of the (d) initial condition ensembles and (e) software system ensembles.

TABLE 2. Globally averaged standard deviation of the 500-hPa geopotential height eddy (m) from the 10-member ensemble with different initial conditions for a given software system (i.e., initial condition ensemble), and the corresponding standard deviation from the 10-member ensemble with different software systems for a given initial condition (i.e., software system ensemble).

Initial condition ensemble		Software system ensemble	
Software system	Std dev	Start date of forecast	Std dev
EXP1	10.65	1 May 1996	11.40
EXP2	10.50	2 May 1996	11.37
EXP3	10.39	3 May 1996	10.20
EXP4	11.17	4 May 1996	10.51
EXP5	11.38	5 May 1996	10.58
EXP6	10.07	6 May 1996	10.70
EXP7	11.35	7 May 1996	10.23
EXP8	11.80	8 May 1996	11.04
EXP9	10.49	9 May 1996	10.57
EXP10	10.78	10 May 1996	11.25
Avg	10.86	Avg	10.78

The standard deviations of ensemble means due to the initial condition and software system are smaller than the corresponding global mean precipitation by a factor of about 1 (cf. Figs. 3d,e).

What level of difference between different system platforms should be allowed? To answer this question, we compute the standard deviations of the 500-hPa geopotential height eddies among the 10 members for each experimental set according to the software system and compare the corresponding value for each ensemble set for different computer platforms (Table 2). It is clear that the system-produced spread is comparable to ensemble spread because of the differences in initial conditions. The horizontal distribution of the 500-hPa standing eddy is comparable to the ensemble average of the 10 software-system experiments (Fig. 4). The distribution of eddies due to the system ensemble reaffirms the climatology's resemblance to the result from the ensemble mean due to initial conditions (cf. Figs. 4b,c), which is also close to the observed (Fig. 4a). In addition, the maximum and minimum pattern correlations of the 500-hPa geopotential height eddies from the RA2 data are 0.725 and 0.659, respectively, whose difference is 0.066. This assures that the seasonal climatology ensemble from the different software platforms produces a realistic modeled climatology.

4. Conclusions

We conducted tests of the software system dependency of simulation results from the GRIMs-GMP run on 10 different software system platforms with different

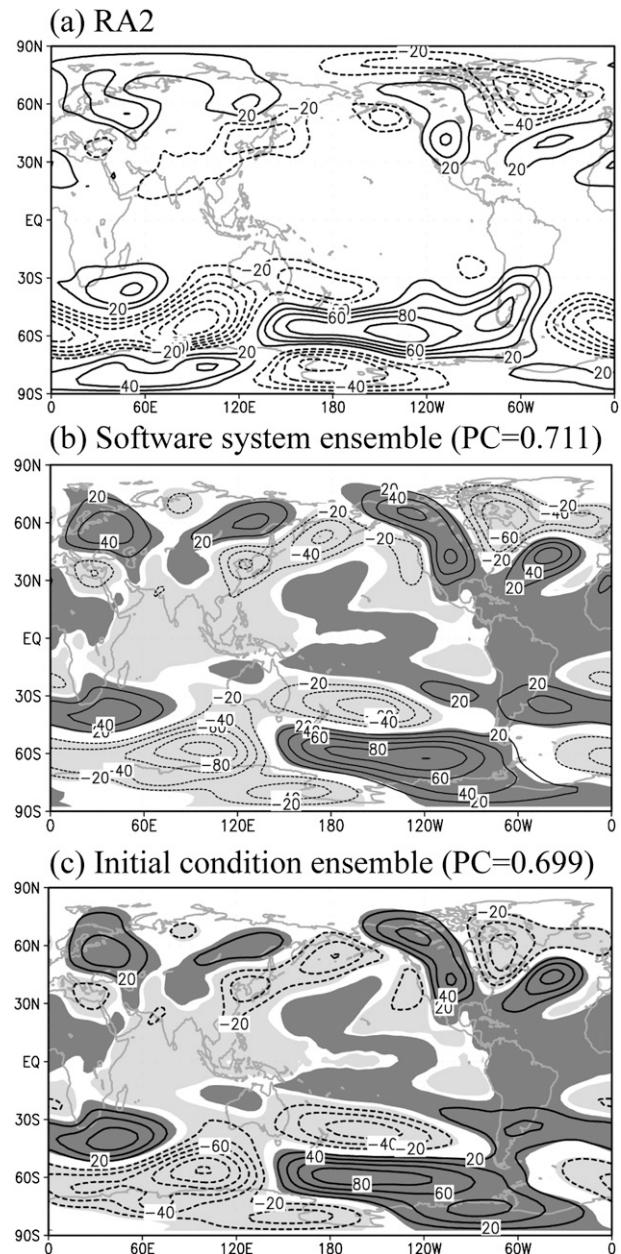


FIG. 4. The 500-hPa geopotential height eddy (m) from the (a) RA2 data, and from the ensemble mean of 10-simulations (b) due to different software systems with the initial condition of 1 May 1996 and (c) due to different initial conditions for the EXP1 run. Shading indicates the significance level greater than 95%. The values in the parentheses indicate the pattern correlation coefficients between the RA2 data and each ensemble mean.

compilers, parallel libraries, and optimization level. We confirm that the model program does not possess the coding errors for the parallel operation in that the model maintains the bitwise-reproducibility irrespective of the number of the processors, as well as when using a single processor. There exist differences in the results for

different compilers and parallel libraries, primarily due to the treatment of rounding errors by the different software systems.

On a weather forecast framework with the GMP having a 200-km horizontal spacing, the system dependency, which is the standard deviation of the 500-hPa geopotential height averaged over the globe, is 3.17 m for the forecast day 5. At higher resolutions, this dependency is smaller in early integration times, but increases with time as a result of greater freedom. However, the change of the globally averaged standard deviation for both resolutions remains nearly zero with time. The RMSE over the analyzed fields increases with time at both resolutions, as expected, but the fractional standard deviation of the RMSE, the ratio of the standard deviation to the mean of RMSEs for the 10 simulations averaged over the globe, remains less than 0.5% and 2% at day 5 and day 10, respectively. The dependency of the simulated heavy rainfall over Korea on the software system is not discernible in terms of its pattern, although the magnitude of peak intensity differs by less than 5%. In a seasonal prediction framework, the spread due to the differences in software system is comparable to the ensemble spread, in that each member uses slightly different initial conditions, which reflects that the influence of uncertainties due to the software system and different initial conditions are comparable.

Realizing that the MPI is the most commonly used package for massively parallel processing on various computer architectures, our study can be adapted to check the implementation of a model to a new computer platform. Although there is no physical background on the quantitative measure of system spread, the values of the standard deviation and the fractional standard deviation over the RMSE we achieved in this study can be used as guidance to the model setup for the short-range forecast. The fractional tendency of standard deviation due to software system stays near zero with time, which can be a measure of adequateness of the model setup. For seasonal simulation, the comparable difference due to different software system to that from the traditional ensemble spread due to different initial conditions is a requirement. Our results also confirm that different software systems can effectively be used to increase the ensemble member for seasonal forecasts, after evaluation of the measures we proposed.

Acknowledgments. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2013-0004008), and by the Korea Meteorological Administration

Research and Development Program under Grant CATER 2012-3084. The authors would like to acknowledge the support from KISTI supercomputing center through the strategic support program for the supercomputing application research (KSC-2012-G3-07).

REFERENCES

- Cheong, H.-B., 2006: A dynamical core with double Fourier series: Comparison with the spherical harmonics method. *Mon. Wea. Rev.*, **134**, 1299–1315.
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- , and E.-C. Chang, 2012: Spectral nudging sensitivity experiments in a regional climate model. *Asia-Pac. J. Atmos. Sci.*, **48**, 345–355.
- , and Coauthors, 2013: The Global/Regional Integrated Model system (GRIMs). *Asia-Pac. J. Atmos. Sci.*, **49**, 219–243.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.
- Koo, M.-S., and S.-Y. Hong, 2013: Double Fourier series dynamical core with hybrid sigma-pressure vertical coordinate. *Tellus*, **65A**, 19851, doi:10.3402/tellusa.v65i0.19851.
- Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- Pacanowski, R. C., and S. M. Griffies, 1998: MOM 3.0 manual. NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 692 pp.
- Park, H., S.-Y. Hong, H.-B. Cheong, and M.-S. Koo, 2013: A Double Fourier Series (DFS) dynamic core in a global atmospheric model with full physics. *Mon. Wea. Rev.*, **141**, 3052–3061.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Rosinski, J. M., and D. L. Williamson, 1997: The accumulation of rounding errors and port validation for global atmospheric models. *SIAM J. Sci. Comput.*, **18**, 552–564.
- Senoner, J.-M., M. García, S. Mendez, G. Staffelbach, and O. Vermorel, 2008: Growth of rounding errors and repetitivity of Large Eddy Simulations. *AIAA J.*, **46**, 1773–1781.
- Seol, K.-H., and S.-Y. Hong, 2009: Relationship between the Tibetan snow in spring and the East Asian summer monsoon in 2003: A global and regional modeling study. *J. Climate*, **22**, 2095–2110.
- Thomas, S. J., J. P. Hacker, M. Desgagné, and R. B. Stull, 2002: An ensemble analysis of forecast errors related to floating point performance. *Wea. Forecasting*, **17**, 898–906.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.