Observing System Research Using Ensemble-based Data Assimilation Methods

Project Representative

Takeshi Enomoto   Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology

Authors

Takeshi Enomoto[*1], Takemasa Miyoshi[*2], Jun Inoue[*3], Qoosaku Moteki[*3], Miki Hattori[*3] and Shozo Yamane[*4]

∗1 Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology

∗2 Department of Atmospheric and Oceanic Science, University of Maryland

∗3 Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology

∗4 Department of Environmental Systems Science, Doshisha University

A data assimilation system for global atmospheric observations has been developed using an ensemble method. This system is composed of the Atmospheric General Circulation model for the Earth Simulator (AFES) and the Local Ensemble Transform Kalman Filter (LETKF). It assimilates global atmospheric observations from public data archives efficiently on the Earth Simulator. The new architecture of the renovated Earth Simulator required re-optimization of AFES and LETKF. The dynamical core of the AFES has been optimized to run twice as fast. The optimized version of LETKF achieved a bump of more than three times. A stream from 1 January 2008 is being conducted to give preliminary results. Smoother fields in the polar regions are achieved by the updated LETKF. Predicted precipitation compares well with satellite observations. Analysis error estimated as analysis ensemble spread is used to evaluate atmospheric observations and to study atmospheric predictability. Observing system experiments are conducted to clarify the influence of pressure observations by Arctic drifting buoys and to identify the planetary-scale propagation of the impact of additional dropsonde observations in the Indian Ocean. Precursory signals are found in various atmospheric phenomena in which the analysis ensemble spread increases prior to the events.

Keywords: Atmospheric General Circulation Model, Ensemble Kalman Filter, Observing System Experiment, Atmospheric Predictability, Optimization

1. Introduction

Japan Agency for Marine-Earth Science and Technology (JAMSTEC) conducts field campaigns in many parts of the world every year. In order to better capture the climate system of the Earth as a whole not only the observations of ocean and atmosphere by research vessels and buoys but also those of land and cryosphere are becoming increasingly important. In addition in-situ and remote sensing observations are investigated by researchers at JAMSTEC. The aim of the project is to apply the simulation techniques to help evaluating the various observations and to develop methodologies for observing system design.

The observing systems are defined by types, distributions, intensity and frequency of observations. Traditionally the observing systems are designed based on the feature of the phenomena in question. The data assimilation system provides an objective tool for designing an optimal observing system.

To achieve our aim an ensemble-based data assimilation system is developed. First it is applied to the assimilation of global atmospheric observations and then it will be applied to other data such as greenhouse gases, land or cryospheric data. Ensemble-based techniques have advantages in assimilating various processes in the climate system since it does not require adjoint models. The climate system, or even a component of it, is so complex that the adjoint models required by the variational techniques are difficult to obtain.

Another important task of the project is evaluation of observations. The value of observations in a particular field campaign is quantified in experiments to test sensitivity to input observations (observing system experiments, OSEs). The analysis error is obtained as the analysis ensemble spread in an ensemble-based data assimilation system. The change in the analysis error represents improvement or degradation of the analysis in OSEs.

The multi-year plan of the project is as follows:

- production of ensemble global atmospheric datasets,

- evaluation of field campaigns by OSEs and
- application of the ensemble-based data assimilation techniques to greenhouse gases, land or cryosphere data.

For the fiscal 2009, 2-year global atmospheric analysis and about 1.5-year equivalent of OSEs and the development of ensemble-based data assimilation techniques were planned. However we were unable to conduct our research and development as scheduled due to a delay is the development of the data assimilation system. After finding a problem related to a compiler bug and applying a workaround, preliminary assimilation experiments from 1 January 2008 were conducted. The delay hindered production runs and OSEs.

The plan of this report is as follows. The overview of the data assimilation system is given and optimization for the Earth Simulator 2 is described in Section 2. Scientific achievements are outlined in Section 3. Concluding remarks are found in Section 4.

2. Development of an ensemble data assimilation system

2.1 System overview

The atmospheric data assimilation system is composed of AFES [1] [2][3] and LETKF [4] [5] [6] [7]. Both AFES and LETKF are updated from those used to produce ALERA (AFES-LETKF experimental ensemble reanalysis) [6]. The planned dataset to be produced with the new system will be called ALERA2. The specifications of ALERA and ALERA2 are summarized in Table 1.

The horizontal resolution is somewhat coarser (about 100 instead of 80 km). The T119L48 (triangular truncation wavenumber at 119, 48 model levels) is used in other simulations such as CFES mini [8] and parameters in physical schemes are determined thorough experiments. The AFES 3.x [9] T119L48 outperforms AFES 2.x T159L48 used in ALERA in forecast experiments of August 2004. Improvements can be mainly attributable to introduction of the new grid condensation scheme. In addition a more sophisticated the land surface scheme MATSIRO [10] is used to obtain land variables.

The ensemble size is increased from 48 to 63. In addition

Table 1 Specifications of ALERA and ALERA2.

	ALERA [6]	ALERA2
Resolution	T159L48	T119L48
Ensemble size	40	63
Covariance localization	21x21x13	400km/0.4 lnp
Spread inflation	0.1	
Observation dataset	JMA	NCEP
SST	NOAA 1° weekly OI [12]	NOAA 1/4° daily [13]

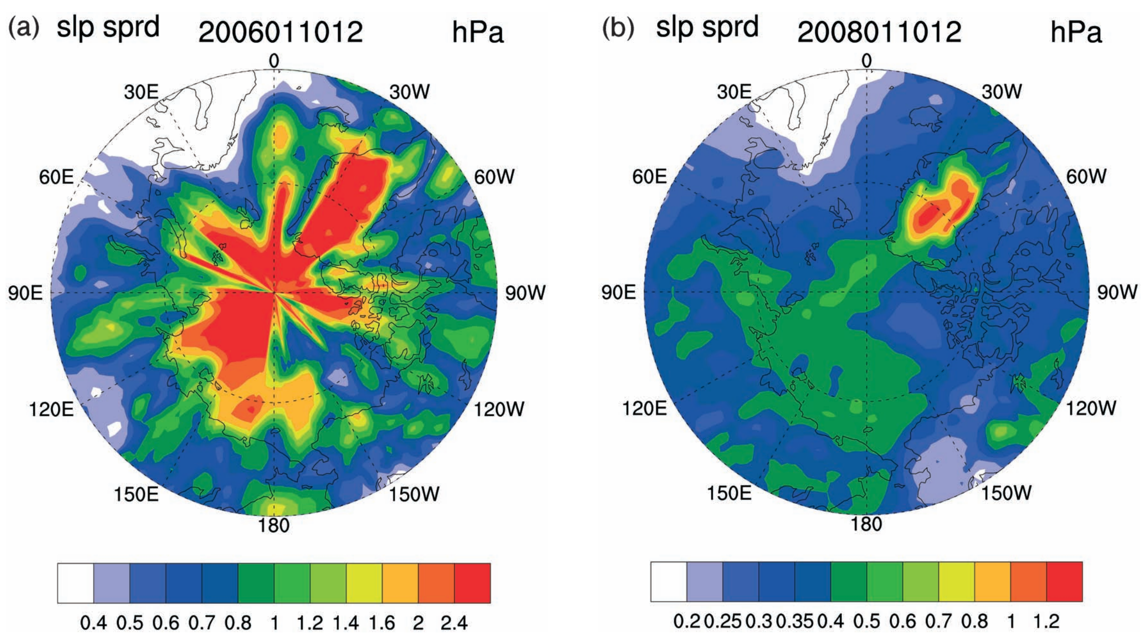

Fig. 1 Snapshots of the analysis ensemble spread of the sea-level pressure (hPa) (a) in ALERA (12 UTC 10 January 2006) and (b) in ALERA2 (12 UTC 10 January 2008). The choice of the date is arbitrary.

the control run from the ensemble mean is conducted. The differences between the guess (6-hr forecast without assimilation) mean and the control indicate the nonlinearity.

The new LETKF assimilates observations with weights based on physical distances without local patches [7] (Fig. 1). ALERA suffers from discontinuities of the analysis ensemble spread of the sea-level pressure due to the local patches converging toward the pole. Such noise is absent in ALERA2 and the analysis ensemble spread is smooth in the polar regions.

ALERA provides the analysis and guess of prognostic variables (wind, temperature, dew point depression) and limited diagnostic variables (geopotential height and sea-level pressure) [6]. As a result the investigation tends to be limited within the dry dynamics. In ALERA2 6-hr forecast of selected variables such as precipitation and surface fluxes are provided. Figure 2 shows daily total (grid-scale and convective) precipitation. Although there are differences in the representation of weak rain, the simulated precipitation associated with extratropical cyclones and tropical convection compares very well with satellite estimates [11] both in terms of distribution and in

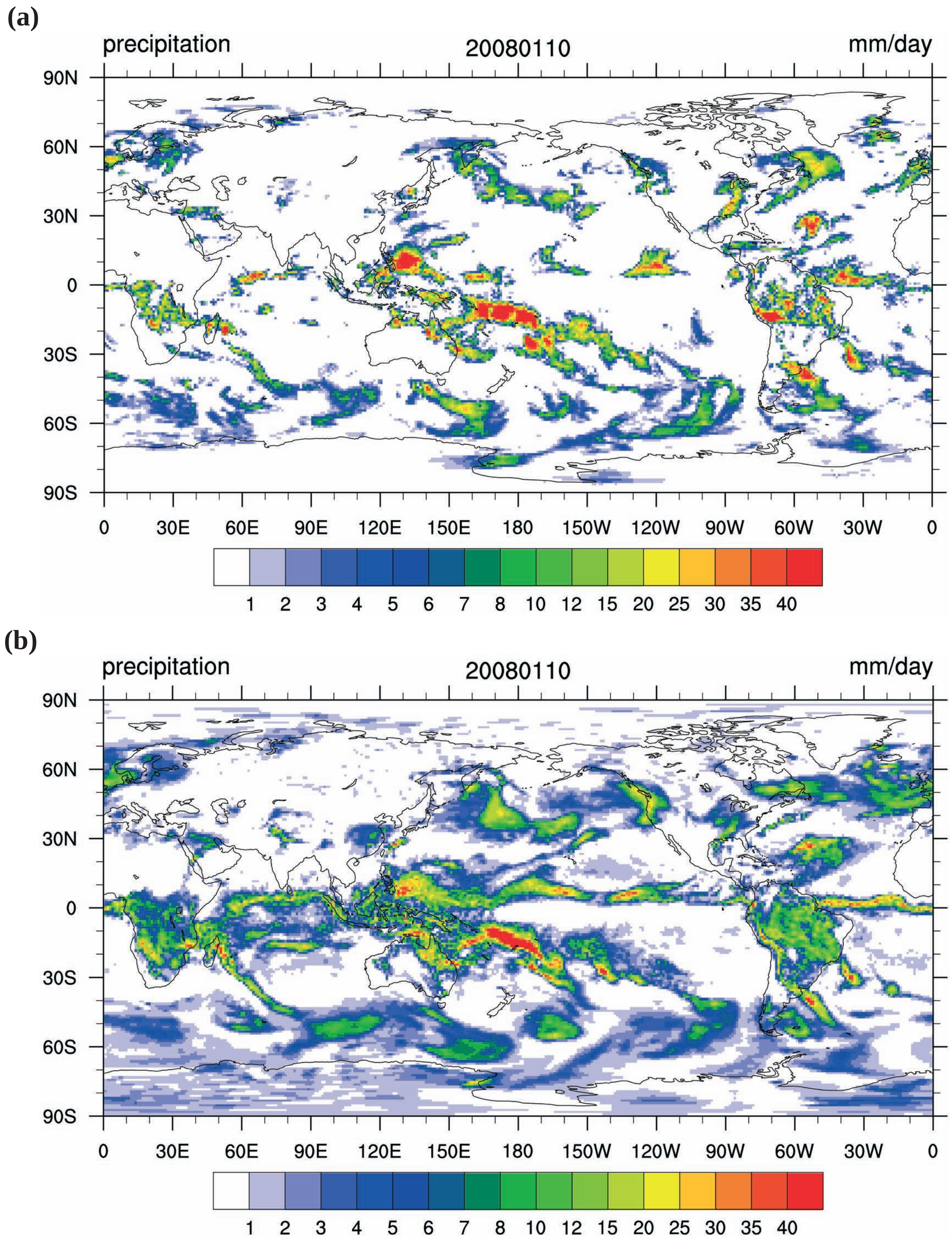

Fig. 2 Daily precipitation of 10 January 2008 from (a) Global Precipitation Climatology dataset [11] and ALERA2.

intensity. Increase of output variables encourages investigation of statistical features of atmospheric phenomena that are represented in physics schemes such as radiation, condensation and convection.

Global atmospheric observations compiled by the National Centers for Environmental Prediction (NCEP) are obtained from the University Corporation for Atmospheric Research (UCAR). This dataset includes most of the observations used for numerical weather prediction at NCEP. ALERA [6] uses the National Oceanic and Atmospheric Administration (NOAA) optimum interpolation (OI) 1° weekly sea-surface temperature (SST) dataset [12]. ALERA2 uses a spatiotemporally higher resolution NOAA 1/4° daily OI SST [13].

2.2 Optimization

Eight nodes are used to produce 64 member ensemble forecast. Thus each processor runs a process of AFES and intra-node parallelism (i.e. microtasking) is not used. AFES is shared by a few projects and optimization has been done collaboratively. In this project the dynamical core has been optimized. In the Legendre analysis some work arrays are eliminated to reduce memory access. The manual unrolls in the Legendre synthesis are replaced by the matrix-vector multiplication calls. The associated Legendre functions now stored by even or odd of n+m and their derivatives are eliminated and replaced by the recurrence formula. Note that other methodologies may help at other resolutions since the optimizations above are adhoc and are found by trial and error. As a result of optimization one-day integration of AFES T119L48 takes 29 s a half of 58 s required by the unoptimized version.

LETKF suffered from a serious problem: LETKF failed produce analysis in several upper layers. The optimization flags were found to cause the problem. Unoptimized LETKF produced reasonable analysis but run significantly slower to retard efficient development. A workaround was then found to use BLAS rather than macro replacements by the compiler. In addition loops of the number of the observations are expanded for vectorization. The optimized LETKF run more than three times faster (1,267 s vs 4,046 s) with real observations.

3. Scientific achievements

3.1 Observing system experiments Observing system experiments

The pressure observation in the Arctic Ocean is mainly conducted by Arctic drifting buoys. An OSE to evaluate pressure observations has been conducted for about a half years to reveal its importance [14] [15]. The influence is not limited to the surface but extends the lower troposphere as shown in Fig. 3 [14]. Information from ALERA has been used to determine the location of POPS (Polar Ocean Profiling System) buoys installed by JAMSTEC. Our findings have attracted attention of the International Arctic Buoy Programme (IABP).

OSEs have been conducted to evaluate sonde observations conducted during MISMO (Mirai Indian Ocean cruise for the Study of the MJO-convection Onset) [16]. These observations act to reduce the analysis error near the observation sites in the central Indian Ocean (Fig. 4a, b). The difference between the two indicates that the improvement extends globally in the form of equatorial waves. The eastern flank of the influence reaches

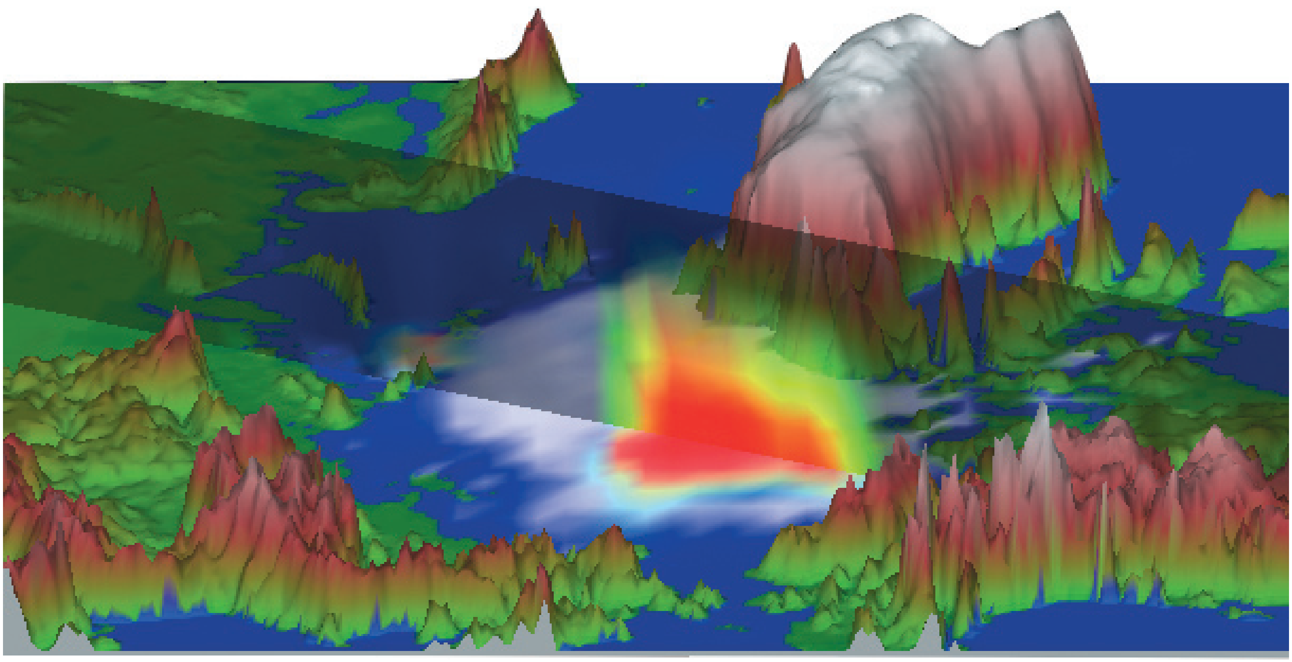

Fig. 3 Three-dimensional visualization of the difference in ensemble spread of the sea-level pressure and in the geopotential height field [15]. In the diagram, the warmer the color, the lower the accuracy of estimates, due to removal of observations. Sea-level pressure observations affect the estimates of geopotential height fields and temperature in the lower troposphere. Sea ice is drawn in white.

the western Pacific and affects the representation of the typhoon genesis.

3.2 Atmospheric predictability

The ensemble analysis is useful in the research on atmospheric predictability. The analysis error is similar to the bred vector and contains information on nonlinear growth. The analysis error is expected to be realistic since it is distributed around the analysis ensemble mean, which is the best estimate of the truth. Investigation of the analysis ensemble spread reveals the increase of the analysis ensemble spread prior to the onset of various atmospheric phenomena: typhoon genesis, westerly bursts, monsoon onset and even stratospheric sudden warming [17].

In the Southern Vietnam the north easterlies and westerlies prevail during the winter and summer monsoons, respectively. Figure 5 shows the 30-day running average of the zonal wind and its analysis error in this region. Prior to the set up of the westerly core with its vertical extension the analysis error increases in the middle troposphere. Although the precursory signals are expected features implied by the linear perturbation theory the precise mechanisms behind the phenomena are not yet known and requires further studies. Increased output variables in ALERA2 will help investigations.

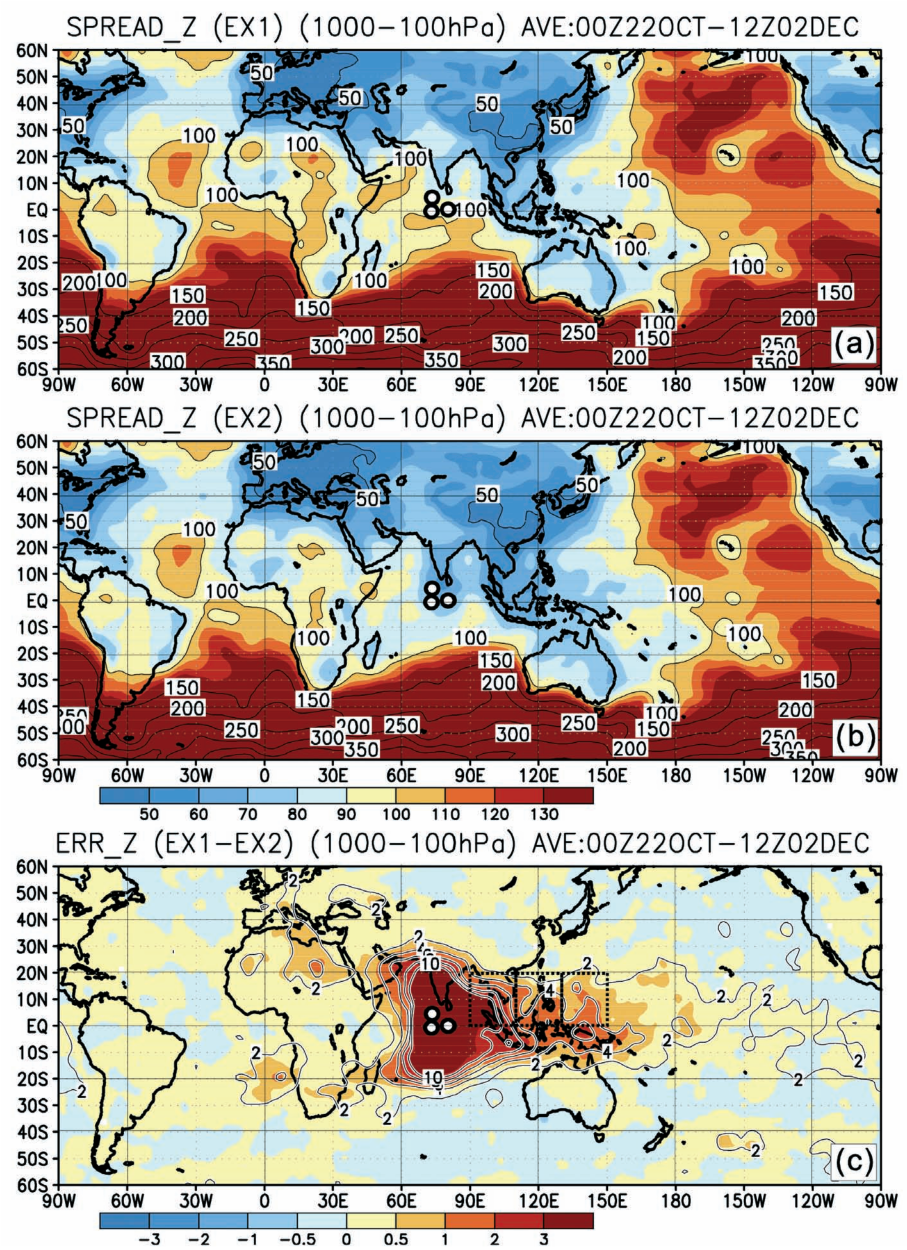

Fig. 4 Distribution of the analysis ensemble spread of the geopotential height vertically integrated from 1000 to 10 hPa and averaged during 0 UTC 22 October and 12 UTC 2 December 2006 (a) in ALERA, (b) in ALERA with additional sonde observations and (c) the difference between the two ((a)–(b)) [16]. The three open circles at the centre of each panel are the observation sites.

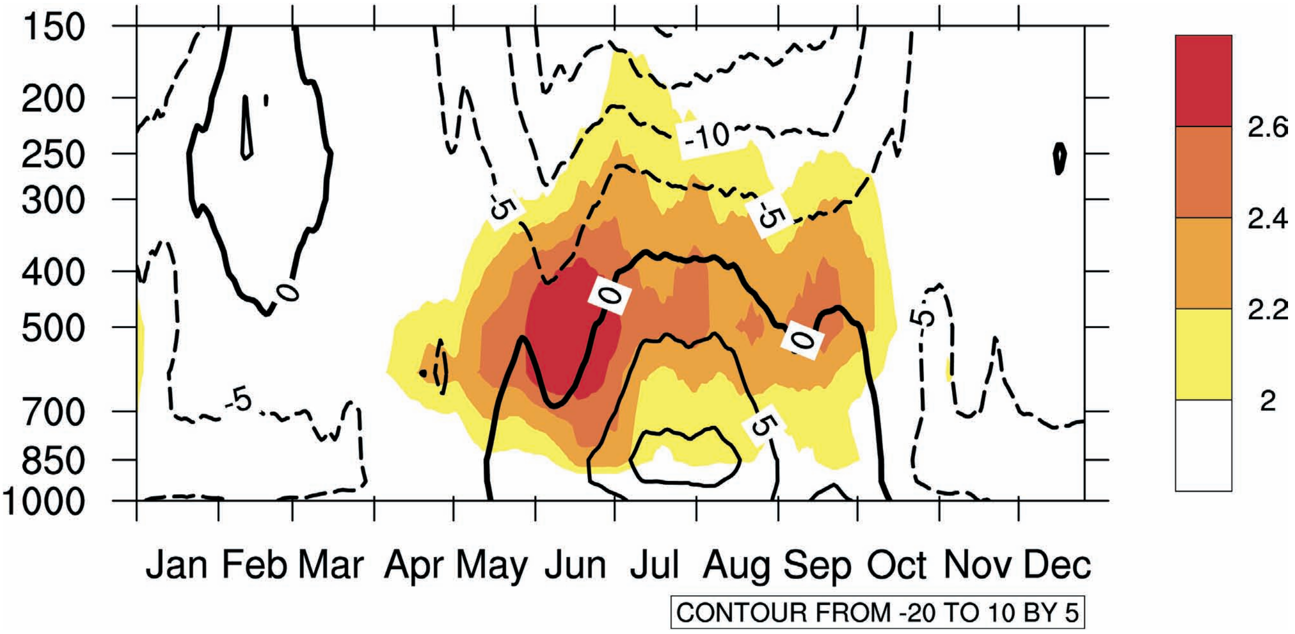

Fig. 5 The 30-day running average of the zonal wind (contour) and its analysis error (colour shades) in the Southern Vietnam (10–15°N, 107.5°–110°E). Produced by M. Hattori [17].

4. Concluding remarks

An ensemble-based data assimilation system has been developed using updated versions of AFES and LETKF. The code has been optimized for the new Earth Simulator to obtain 2x to 3x speed up. Preliminary results show realistic distribution of atmospheric variables including precipitation and much smoother analysis ensemble spread in the polar regions. OSEs conducted with the original Earth Simulator are investigated to obtain new insights into the polar and equatorial dynamics. ALERA has been investigated to identify precursory signals are found in various atmospheric phenomena. Delayed production runs and OSEs will be conducted in the new fiscal year. Applications of the ensemble methods to cryosphere, land and greenhouse gases are being prepared and data assimilation experiments are planned on the Earth Simulator.

Acknowledgement

ALERA was produced under the collaboration among the Japan Meteorological Agency, Japan Agency for Marine- Earth Science and Technology, and Chiba Institute of Science. The atmospheric observation datasets are compiled by the National Centers for Environmental Prediction and obtained from the data archive at the University Corporation for Atmospheric Research (UCAR), USA. The National Oceanic and Atmospheric Administration optimum interpolation 1/4 degree daily sea surface temperature analysis (Reynolds et al. 2007) is also obtained from UCAR.

References

[1] Numaguti, A., M. Takahashi, T. Nakajima, and A, Sumi, 1997: Description of CCSR/NIES atmospheric general circulation model. *CGER'S supercomputer monograph report*, National Institue for Environmental Studies, **3**, 1-48.

[2] Ohfuchi, W., H. Nakamura, M. K. Yoshioka, T. Enomoto, K. Takaya, X. Peng, S. Yamane, T. Nishimura, Y. Kurihara, and K. Ninomiya, 2004: 10-km mesh meso-scale resolving simulations of the global atmosphere on the Earth Simulator: Preliminary outcomes of AFES (AGCM for the Earth Simulator). *J. Earth Simulator*, **1**, 8-34.

[3] Enomoto, T., A. Kuwano-Yoshida, N. Komori, and W. Ohfuchi, 2008:Description of AFES 2: Improvements for high-resolution and coupled simulations. *In High Resolution Numerical Modelling of the Atmosphere and Ocean, K. Hamilton and W. Ohfuchi (eds.)*, chapter 5, pp. 77-97, Springer, New York.

[4] Hunt, B., E. J. Kostelich and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: a local transform Kalman filter. *Physica D*. **230**, 112-126.

[5] Miyoshi, T., S. Yamane and T. Enomoto, 2007: The AFES-LETKF experimental ensemble reanalysis: ALERA. *SOLA*, **3**, 45-48., doi:10.1029/2010GL042723.

[6] Miyoshi, T. and S. Yamane, 2007: Local ensemble transform Kalman filtering with an AGCM at a T159/L48. *Mon. Wea. Rev.*, **135**, 3841-3861.

[7] Miyoshi, T., S. Yamane and T. Enomoto, 2007: Localizing the error covariance by physical distances within a local ensemble transform Kalman filter (LETKF). *SOLA*, 3, 89-92, doi:10.2151/sola.2007-023.

[8] Sasaki, H., T. Enomoto, N. Komori, A. Kuwano-Yoshida, K. Tsuboki, J. Inoue, B. Taguchi and W. Ohfuchi, 2009:

Understanding and forecasting high-impact phenomena in the atmosphere and ocean. *Annual Report of the Earth Simulator Center April 2008–September 2008,* Earth Simulator Center, JAMSTEC, Yokohama, Japan, 23-28.

[9] Kuwano-Yoshida, A., T. Enomoto and W. Ohfuchi: An improved PDF cloud scheme for climate simulations. *Quart. J. Roy. Meteor. Soc.,* in press.

[10] Takata, K., S. Emori and T. Watanabe, 2003: Development of the minimal advanced treatments of surface interaction and runoff, *Global and Planetary Change,* **38**, 209-222.

[11] Huffman, G. J., R. F. Adler, M. M Morrrissey, D. T. Bolvin, S. Curtic, R. Joyce, B. McGavock and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.,* **2**, 36-50.

[12] Reynolds, R. W., N. A. Rayner, T. M. Smith, D. Stokes and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate,* **15**, 1609-1625.

[13] Reynolds, R. W., L. Chunying, T. M. Smith, D. B. Chelton, M. G. Schlax, and K. S. Casey, 2007: Daily high-resolution-blended analyses for sea surface temperature, *J. Climate,* **20**, 5473-5496.

[14] Inoue, J., T. Enomoto, T. Miyoshi and S. Yamane, 2009: Impact of observations from Arctic drifting buoys on the reanalysis of surface fields. *Geophys. Res. Lett.,* **36**, L08501, doi:10.1029/2009GL037380.

[15] Inoue, J. and T. Enomoto, 2009: Reduced Arctic Sea Ice Hinders Accurate Climate Monitoring—Impact of Depleted Arctic Drifting Buoy Network—, press release, Japan Marine-Earth Science and Technology, 2 April, Yokosuka, Japan.

[16] Moteki, Q., K. Yoneyama, R. Shirooka, H. Kubota, K. Yasunaga, J. Suzuki, A. Seki, N. Sato, T. Enomoto, T. Miyoshi and S. Yamane: The influence of observations propagated by convectively coupled equatorial waves, *Q. J. Roy. Meteor. Soc.,* in revision.

[17] Enomoto, T., M. Hattori, T. Miyoshi and S. Yamane, 2010: Precursory signals in analysis ensemble spread. *Geophys. Res. Lett.,* **37**, L08804, doi:10.1029/2010GL042723.

アンサンブル同化手法を用いた観測システムの最適化に関する研究

プロジェクト責任者
榎本　　剛　　海洋研究開発機構　地球シミュレータセンター

著者
榎本　　剛 *[1]，三好　建正 *[2]，猪上　　淳 *[3]，茂木　耕作 *[3]，服部　美紀 *[3]，山根　省三 *[4]
*1　海洋研究開発機構　地球シミュレータセンター
*2　メリーランド大学　大気海洋科学部
*3　海洋研究開発機構　地球環境変動領域
*4　同志社大学　環境システム学科

シミュレーション技術を応用した観測システム研究を実施し、効率的な観測システムの設計に役立てるため、アンサンブル手法を用いた全球大気データ同化システムを開発した。このシステムは、地球シミュレータ用大気大循環モデル（AFES）と局所アンサンブル変換カルマンフィルタ（LETKF）から構成され、公開されている全球大気観測データを地球シミュレータ上で効率的に大気大循環モデルに同化することができる。本年度は、まず AFES 及び LETKF を更新された地球シミュレータに最適化した。AFES の力学過程は最適化の結果実行時間を半減させることに成功した。LETKF は最適化の結果 1/3 以下の実行時間を達成した。このシステムを用いて、2008 年 1 月 1 日からの同化実験を開始し、初期的な結果が得られた。LETKF の改良により、極付近のノイズが一掃された。降水量の 6 時間予報値は、衛星観測とよく対応している。アンサンブル同化手法を用いると、解析誤差がアンサンブルスプレッドとして推定される。解析誤差は、観測を評価し観測システム設計に役立てることができる。本年度は、北極海の漂流ブイによる気圧観測の影響が対流圏下層に広がっていることやインド洋での追加のゾンデ観測の影響が惑星規模に及ぶこと等を明らかにした。また、大気予測可能性に関し、さまざまな大気現象に先行して解析誤差が増大する現象を発見した。

キーワード: 大気大循環モデル, アンサンブル・カルマンフィルタ, 観測システム実験, 大気予測可能性, 最適化