Simple Sensitivity Analysis Using Ensemble Forecasts

Takeshi ENOMOTO

Disaster Prevention Research Institute, Kyoto University, Uji, Japan
Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Shozo YAMANE

Faculty of Science and Engineering, Doshisha University, Kyotanabe, Japan

and

Wataru OHFUCHI

Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Abstract

Simple methods are formulated using an ensemble forecast to identify the sensitive initial perturbations that grow in a specified region at the verification time. These methods do not require the tangent-linear or adjoint models, but use an ensemble forecast to obtain approximated solutions. Input to the sensitivity calculation can be any ensemble forecast integrated from initial conditions perturbed with the bred vector, singular vector, or ensemble Kalman filter methods. Two formulations are presented here to approximate the adjoint and singular vector methods using an ensemble forecast. The ensemble singular vector sensitivity, which has already been applied in previous studies, is obtained with a single eigenvector calculation. The ensemble adjoint sensitivity only requires an even simpler matrix-vector multiplication.

To validate the formulations, ensemble-based sensitivity analysis has been conducted in a few cases. First, the two methods were applied to identify the sensitive initial perturbations that grow in the verification region over Japan in January and August 2003. The first singular vector mode indeed achieves the largest amplitude at the verification time, but that is not necessarily true after the verification time. Both methods can identify the sensitive regions more specifically than the regions with large ensemble spread in cases with a mid-latitude cyclone and with a tropical cyclone. The monthly-mean sensitivity in January 2003 indicates the effect of Rossby waves and synoptic disturbances in upstream sensitive regions over Siberia, Tibet, and a downstream sensitive region in the north-western Pacific; the sensitivity in August 2003 suggests the influence of the Asian summer monsoon. Next, for an August 2002 storm case in Europe, global 20-km resolution simulations were conducted from the initial conditions perturbed by the ensemble singular vector method to compare with the unperturbed simulation. In the perturbed simulation, the cyclone is deeper by a few hPa in its north-east sector with more precipitation north of the Alps more consistently with observations. These results indicate that reasonable sensitive regions can be identified with our methods.

Corresponding author: Takeshi Enomoto, Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, Kyoto, Japan
E-mail: enomoto.takeshi.3n@kyoto-u.ac.jp

1. Introduction

Sensitive regions are the geographical distribution of initial perturbations that influence the growth of the disturbances within a given region at the verification time. Sensitivity analysis is employed to identify areas for adaptive observations (Snyder 1996) and to interpret dynamics (Fujii et al. 2008). Sensitivity has typically been computed with the adjoint (Rabier et al. 1996; Langland et al. 2000) or singular-vector (Gelaro et al. 1998; Buizza and Montani 1999) methods, which require the tangent-linear or adjoint models. Alternatively, sensitivity can also be obtained by ensemble-based methods (Bishop and Toth 1999; Bishop et al. 2001; Hamill and Snyder 2002; Ancell and Hakim 2007; Torn and Hakim 2008; Liu et al. 2009; Kalnay et al. 2012). Mu et al. (2003) proposed the conditional nonlinear optimal perturbation (CNOP) that is applicable on a longer time scale for nonlinear phenomena, such as El Niño Southern Oscillation. For tropical cyclones (TCs), Wu et al. (2007) proposed the adjoint-derived sensitivity steering vector (ADSSV) method to identify not only the locations but also the critical direction of the steering flow. Ito and Wu (2013) further proposed to use the TC position as a metric in their typhoon-position-oriented sensitivity analysis (TyPOS).

The purpose of this paper is to derive two formulations of ensemble-based sensitivity analysis and to clarify their mathematical backgrounds. The first formulation approximates the adjoint method using an ensemble forecast, which is referred to as the ensemble adjoint sensitivity analysis (EnASA). The second formulation approximates the singular-vector (SV) method using an ensemble forecast, which is referred to as the ensemble singular-vector sensitivity analysis (EnSVSA). The two formulations are not necessarily new, but derived in consistent with the linear perturbation theory and with adjoint and SV methods. Each formulation measures forecast sensitivity in a different manner. Adjoint sensitivity is the ratio of the norm at the initial time to that of the verification time at each grid point, whereas singular-vector sensitivity is the growth rate of growing modes.

Unlike some other methods in the literature, (pseudo) observations are not considered for the sake of simplicity. For example, EnSVSA is essentially identical to the method of Bishop and Toth (1999) in that the singular vectors are obtained from the forecast error covariance matrix but without the transformation matrix that corresponds to the deployment of an additional observation. Our simplified methods use only simple matrix operations and do not require a data assimilation system in the sensitivity calculation. For example, the method proposed by Liu et al. (2009) and simplified by Kalnay et al. (2012) requires components of the data assimilation system using an ensemble Kalman filter. Their method computes the forecast sensitivity to observations, whereas our methods calculate the forecast sensitivity to the initial perturbations. Yet the perturbations obtained by our methods provide geographical distributions of sensitive perturbations and may be used in targeted observations.

EnSVSA, introduced in Enomoto et al. (2006), has already been applied in a few case studies of a cut-off cyclone (Enomoto et al. 2007), the propagation of Rossby wave packets (Nishii and Nakamura 2010), and a blocking anticyclone (Matsueda et al. 2011). EnSVSA has successfully identified the sensitive regions for the development of the disturbances in each case.

Gelaro et al. (1998) constructed perturbations optimized for the 48-h forecast error from the linear combination of the SV. They showed that the forecast error was indeed reduced in the forecast from the perturbed initial conditions. Along the line of Gelaro et al. (1998), the cut-off cyclone that occurred in August 2002, in Europe, which was discussed in Enomoto et al. (2007), is revisited to examine if the perturbations generated with the EnSVSA improve the forecast by comparing the unperturbed and perturbed experiments.

The layout of this paper is as follows. Section 2 defines the mathematical problem for the sensitivity analysis, reviews the adjoint and SV sensitivity, and formulates the ensemble-based approximations. Section 3 describes the step-by-step procedures and demonstrates a few applications. Finally, a summary and a discussion are presented in Section 4.

2. Formulation

Prior to the formulation of the sensitivity analysis using an ensemble forecast, the adjoint and singular-vector methods are briefly reviewed.

2.1 Linear growth of perturbations

Let $y \in R^n$ be the initial perturbations at $t = 0$ that are linearized about the state x. Assuming linearity in the time evolution, the perturbations at the verification time t denoted by z may be expressed by a projection of y as follows:
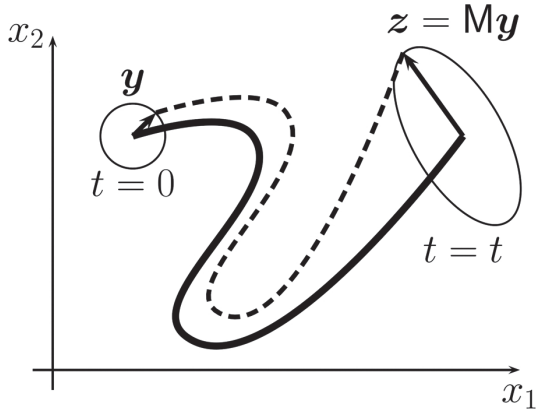
$$z = My, \tag{1}$$

Fig. 1. Schematic illustration of the perturbation growth. The perturbations at the initial time $t = 0$ and at the verification time t are denoted by a circle and an oval, respectively. The thick solid and dotted curves represent the control forecast and the forecast with the largest perturbation growth at t, respectively. The vector $\boldsymbol{y}$ represents the direction of the perturbations at $t = 0$ that corresponds to the vector $\boldsymbol{z}$ that represents the direction of the largest growth at t. M represents the time integration with the tangent linear model.

where M is an $n \times n$ square matrix representing time integration with the tangent linear model (Fig. 1).

To measure the growth of disturbances, the norm must be defined. The choice of the norm depends on the aspect of the disturbances of interest. In the present study, the norm is expressed by the inner product as follows:

$$\|\boldsymbol{y}\|_0 \equiv \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle_0} = \sqrt{\boldsymbol{y}^\mathsf{T} \mathsf{G}_0 \boldsymbol{y}},$$
$$\|\boldsymbol{z}\|_t \equiv \sqrt{\langle \boldsymbol{z}, \boldsymbol{z} \rangle_t} = \sqrt{\boldsymbol{z}^\mathsf{T} \mathsf{G}_t \boldsymbol{z}}, \tag{2}$$

at the initial and verification times, respectively, where G_0 and G_t are arbitrary non-negative symmetric matrices that denote the weight in their respective definitions of norm. For example, G_t may be used to define the verification region. Note that the inner product (2) is defined in a general form and that they reduce to a natural (Euclidian) inner product if G_0 or G_t is an identity matrix.

a. Sensitivity analysis using the adjoint method

Assume that a scalar variable $\Theta = \Theta(\boldsymbol{x})$ varies with $\boldsymbol{z}$. For example, Θ may be any scalar of interest such as the area-averaged temperature or total dry energy in a given verification region. Θ is often referred to as

a cost function. Let θ be small changes from Θ due to $\boldsymbol{z}$. The relationship between θ and $\boldsymbol{z}$ is as follows:

$$\theta = \left[\frac{\partial \Theta}{\partial \boldsymbol{x}_t} \right]^\mathsf{T} \boldsymbol{z}. \tag{3}$$

By substituting (1) into (3), the relationship between θ and $\boldsymbol{y}$ is obtained.

$$\theta = \left[\frac{\partial \Theta}{\partial \boldsymbol{x}_t} \right]^\mathsf{T} \mathsf{M}\boldsymbol{y} = \left[\mathsf{M}^\mathsf{T} \frac{\partial \Theta}{\partial \boldsymbol{x}_t} \right]^\mathsf{T} \boldsymbol{y} = \left[\frac{\partial \Theta}{\partial \boldsymbol{x}_0} \right]^\mathsf{T} \boldsymbol{y}. \tag{4}$$

When θ is the maximum (or minimum), Θ is most sensitive to the initial perturbation $\boldsymbol{y}$. Therefore, the task of the sensitivity analysis using the adjoint method is to find $\boldsymbol{y}$ that maximizes θ with $\|\boldsymbol{y}\|_0 = 1$. Note that the squared sum of unity here refers to the perturbation amplitude that is an order of magnitude (or more) smaller than the unperturbed state. Based on the definition of the norm (2), $\boldsymbol{y}$ that maximizes θ with $\boldsymbol{y}^\mathsf{T} \mathsf{G}_0 \boldsymbol{y} = 1$ may be obtained by finding the extremum of a function,

$$F(\boldsymbol{y}, \lambda) \equiv \theta + \lambda (1 - \boldsymbol{y}^\mathsf{T} \mathsf{G}_0 \boldsymbol{y}), \tag{5}$$

where λ is the Lagrange multiplier. Substituting (4) into (5) and calculating $\partial F / \partial \boldsymbol{y} = \boldsymbol{0}$, it follows that

$$2\lambda \boldsymbol{y} = \mathsf{G}_0^{-1} \mathsf{M}^\mathsf{T} \frac{\partial \Theta}{\partial x_t} = \mathsf{G}_0^{-1} \frac{\partial \Theta}{\partial x_0}. \tag{6}$$

This equation states that θ is an extremum when $\boldsymbol{y}$ is parallel to the vector of the right hand side of (6). Integration with the adjoint model (M^T) backward in time is required to obtain the most sensitive initial perturbations $\boldsymbol{y}$.

b. Sensitivity analysis using the singular vectors

In the singular vector method, the task is to find the initial perturbations that optimally grow at the verification time i.e., to maximize $\|\boldsymbol{z}\|_t$ subject to a condition $\|\boldsymbol{y}\|_0 = 1$.

The solution may be obtained by finding the extremum of a Lagrange function

$$F(\boldsymbol{y}, \lambda) \equiv \boldsymbol{z}^\mathsf{T} \mathsf{G}_t \boldsymbol{z} + \lambda (1 - \boldsymbol{y}^\mathsf{T} \mathsf{G}_0 \boldsymbol{y})$$
$$= \boldsymbol{y}^\mathsf{T} \mathsf{M}^\mathsf{T} \mathsf{G}_t \mathsf{M}\boldsymbol{y} + \lambda (1 - \boldsymbol{y}^\mathsf{T} \mathsf{G}_0 \boldsymbol{y}), \tag{7}$$

where λ is the Lagrange multiplier. The derivative of F is calculated as follows:

$$\frac{\partial F(\boldsymbol{y}, \lambda)}{\partial \boldsymbol{y}} = 2\boldsymbol{y}^\mathsf{T} \mathsf{M}^\mathsf{T} \mathsf{G}_t \mathsf{M} - 2\lambda \boldsymbol{y}^\mathsf{T} \mathsf{G}_0 = 0. \tag{8}$$

The equation reduces to an eigenvalue problem

$$G_0^{-1}M^\mathsf{T}G_t My = \lambda y. \tag{9}$$

The solutions are identified as the eigenvectors of $G_0^{-1}M^\mathsf{T}G_t M$ using the SV method. The singular value, the square root of the eigenvalue λ, corresponds to the growth rate of each mode.

The number of degrees of freedom in an atmospheric general circulation model (AGCM) is so large that it is computationally too expensive to calculate the exact solution. In practice, the approximate solution is obtained by the Lanczos method where the tangent-linear (M) and adjoint (M^T) models are integrated recursively (Buizza et al. 1993).

2.2   Use of ensemble forecast

In this section, the adjoint and SV methods are approximated by the ensemble forecast of m members. The assumption of linear time evolution (1) is replaced with an ensemble forecast

$$z_i = M(x + y_i) - M(x), i = 1, 2, \cdots, m, \tag{10}$$

where x is the mean of the ensemble forecast or control forecast and M is the nonlinear model.

Assume that an initial perturbation can be expressed in a linear combination of the ensemble initial perturbations

$$y = p_1 y_1 + p_2 y_2 + \cdots + p_m y_m, \tag{11}$$

where

$$p^\mathsf{T} = (p_1, p_2, \cdots, p_m) \tag{12}$$

are coefficients.

Assuming linear time evolution, the corresponding perturbations at the verification time t are

$$z = p_1 z_1 + p_2 z_2 + \cdots + p_m z_m. \tag{13}$$

Note that the same coefficients p are used at $t = 0$ and t. Ensemble sensitivity analysis looks for the optimal y and z by choosing p in the sense explained in the following subsections.

Using short-hand notations

$$Y = (y_1, y_2, \cdots, y_m), \ Z = (z_1, z_2, \cdots, z_m) \tag{14}$$

the linear combinations (11) and (13) are expressed as follows:

$$y = Yp, \tag{15}$$

$$z = Zp, \tag{16}$$

respectively.

a.   Ensemble adjoint sensitivity analysis

Assume that from each ensemble member,

$$\theta^\mathsf{T} = (\theta_1, \theta_2, \cdots, \theta_m) \tag{17}$$

is known, for example, as the deviations from the ensemble mean. If (3) holds for each evolved member, the θ may also be expressed in the linear combination as in (11),

$$\theta = \theta^\mathsf{T} p. \tag{18}$$

With (15) and (17), the task of the sensitivity analysis is to find the p that maximizes $\theta^\mathsf{T} p$ with $y^\mathsf{T}G_0 y = p^\mathsf{T}Y^\mathsf{T}G_0 Yp = 1$.

The Lagrange function (5) is replaced by

$$F(p, \lambda) \equiv \theta^\mathsf{T} p + \lambda(1 - p^\mathsf{T}Y^\mathsf{T}G_0 Yp). \tag{19}$$

The derivative of (19) gives

$$\theta^\mathsf{T} - 2\lambda p^\mathsf{T}Y^\mathsf{T}G_0 Y = 0, \tag{20}$$

or

$$2\lambda p = (Y^\mathsf{T}G_0 Y)^{-1}\theta, \tag{21}$$

where

$$Y^\mathsf{T}GY = \begin{pmatrix} \langle y_1, y_1 \rangle_0 & \langle y_1, y_2 \rangle_0 & \cdots & \langle y_1, y_m \rangle_0 \\ \langle y_2, y_1 \rangle_0 & \langle y_2, y_2 \rangle_0 & \cdots & \langle y_2, y_m \rangle_0 \\ \vdots & \vdots & \ddots & \vdots \\ \langle y_m, y_1 \rangle_0 & \langle y_m, y_2 \rangle_0 & \cdots & \langle y_m, y_m \rangle_0 \end{pmatrix}. \tag{22}$$

Therefore, the optimal initial perturbations are parallel to $Y(Y^\mathsf{T}G_0 Y)^{-1}\theta$. Ancell and Hakim (2007) derived (21) (their (11)) by multiplying (4) by y^T and taking the expected value. Thus, they considered ensemble of adjoint sensitivity. We first assumed the linear combination of the initial perturbation (11) or (15) and used the Lagrange multiplier as in 2.1a.

For a special case that for all m the norms of y_i are equal

$$\|y_1\|_0 = \|y_2\|_0 = \cdots = \|y_m\|_0 \tag{23}$$

and that y_i are orthogonal one another

$$\langle y_i, y_j \rangle_0 = 0, i \neq j, \tag{24}$$

$$Y^\mathsf{T}G_0 Y \propto I, \tag{25}$$

where I is the identity matrix, the optimal initial perturbations become

$$y = Y\theta = \theta_1 y_1 + \theta_2 y_2 + \cdots + \theta_m y_m. \tag{26}$$

Thus, p is simply proportional to θ from (21) and (25) (See Ito and Wu 2013 for a more generalized expression). Here, the EnASA is conducted in the $\mathbf{R}^m$ subspace generated by y_i ($i = 1, 2, ..., m$). As $m \to n$, the solution asymptotically approaches the sensitivity analysis (6) using the adjoint model in full $\mathbf{R}^n$.

b.   Ensemble singular-vector sensitivity analysis

With (15) and (16), the task of the sensitivity analysis is to find the p that maximizes $p^\mathsf{T} Z^\mathsf{T} G_t Z p$ with $y^\mathsf{T} G_0 y = p^\mathsf{T} Y^\mathsf{T} G_0 Y p = 1$.

The Lagrange function (7) is replaced by

$$F(p, \lambda) = p^\mathsf{T} Z^\mathsf{T} G_t Z p + \lambda (1 - p^\mathsf{T} Y^\mathsf{T} G_0 Y p), \tag{27}$$

where

$$Z^\mathsf{T} G_t Z = \begin{pmatrix} \langle z_1, z_1 \rangle_t & \langle z_1, z_2 \rangle_t & \cdots & \langle z_1, z_m \rangle_t \\ \langle z_2, z_1 \rangle_t & \langle z_2, z_2 \rangle_t & \cdots & \langle z_2, z_m \rangle_t \\ \vdots & \vdots & \ddots & \vdots \\ \langle z_m, z_1 \rangle_t & \langle z_m, z_2 \rangle_t & \cdots & \langle z_m, z_m \rangle_t \end{pmatrix}. \tag{28}$$

Its derivative

$$\frac{\partial F(p, \lambda)}{\partial p} = 2 p^\mathsf{T} Z^\mathsf{T} G_t Z - 2\lambda p^\mathsf{T} Y^\mathsf{T} G_0 Y = \mathbf{0} \tag{29}$$

yields an eigenvalue problem

$$(Y^\mathsf{T} G_0 Y)^{-1} Z^\mathsf{T} G_t Z p = \Lambda p, \tag{30}$$

where Λ is the matrix whose diagonal elements are eigenvalues.

If y_i is normalized (23) and orthogonal (24), matrix $Y^\mathsf{T} G_0 Y$ is proportional to the identity matrix (25); thus, the eigenvectors of $Z^\mathsf{T} G_t Z$ provide the solution. EnSVSA is conducted in the $\mathbf{R}^m$ subspace generated by y_i ($i = 1, 2, ..., m$). However, as $m \to n$, the eigenvectors of $Z^\mathsf{T} G_t Z$ asymptotically approach those of (9) using the tangent-linear and adjoint models in the full $\mathbf{R}^n$ space.

3.   Applications

In this section, the formulations derived in the previous section are validated with an operational one-week ensemble forecast from the Japan Meteorological Agency (JMA). Note that the $(Y^\mathsf{T} G_0 Y)^{-1}$ factor is ignored in the following applications in favour of

simplicity although (25) does not usually hold true.

3.1   JMA One-week ensemble forecast

For the period used here, the initial perturbations are generated with the breeding method (Toth and Kalnay 1993, 1997) in JMA's ensemble prediction system (EPS) for medium-range weather forecasting (Kyouda 2002, 2006). The ensemble size is 25: a control run without perturbation and 24 runs with perturbations produced from 12 normalized- and orthogonalized-bred vectors with positive and negative signs. Perturbations are added in the Northern Hemisphere and in the Tropics (90°N–20°S). The ensemble forecast for 9 days is run daily at resolution T106L40 (truncation wave number of 106 using the triangular truncation, which corresponds to 1.125° horizontal resolution, and 40 vertical levels). The amplitudes of the perturbations are adjusted with the geopotential height at 500 hPa. The perturbations are not strictly orthogonal although the Gram-Schmidt orthogonalization is applied (Kyouda 2006).

The JMA one-week ensemble is available online in the General Regulary-distributed Information in Binary form (GRIB) from the Research Institute for Sustainable Humanosphere (RISH) data server, Kyoto University, and from the Grid Point Value (GPV) data archive at Kitsuregawa Laboratory, Institute of Industrial Science, the University of Tokyo. The pressure-level data consist of the geopotential height, winds, and temperature at 300-, 500- and 850-hPa levels at 2.5° horizontal resolution for an 8-day forecast with a 1-day interval. The relative humidity is only available at the 850-hPa surface. The surface data provide the sea-level pressure, 10-m winds, and accumulated precipitation.

3.2   Procedures of sensitivity analysis

In the following applications, sensitivity analysis is performed with the dry total energy norm (Buizza et al. 1993)

$$\begin{aligned} e = \frac{1}{2Ap_\mathrm{r}} \int_{p_\mathrm{r}}^0 \int_A u'^2 + v'^2 + \frac{c_p}{T_\mathrm{r}} T'^2 \\ + RT_\mathrm{r} \left(\frac{p_\mathrm{s}'}{p_\mathrm{r}} \right)^2 \mathrm{d}A\mathrm{d}p, \end{aligned} \tag{31}$$

where superscript $'$ denotes the perturbation from the control run; A is the unit area; u, v, T, p_s are the zonal wind, meridional wind, temperature, and surface pressure, respectively; c_p and R are the specific heat at constant pressure and the gas constant of dry air, respectively; T_r and p_r are the reference temperature

and pressure, respectively. Here, we use $T_r = 270$ K and $p_r = 1000$ hPa.

First, the ensemble forecast is preprocessed using the following procedure.

1. Calculate perturbations of dry prognostic variables (u, v, T, p_s) by subtracting the value in the control forecast from the value in each ensemble member in the verification region at the verification time t and store the perturbation values in an $n \times m$ array Z, where n is the degrees of freedom (the number of grid points in the verification region multiplied by the number of variables) and m is the ensemble size.

2. Multiply the perturbations by the areal factor (cos ϕ, where ϕ is the latitude or the Gaussian weights) and the layer thickness factor.

3. Multiply T' and p_s' by the square roots of the coefficients in (31), i.e., $\sqrt{c_p/T_r}$ and $\sqrt{RT_r}/p_r$, respectively.

Using the prepared forecast perturbations, ensemble sensitivity can be obtained with either adjoint or singular-vector formulations as follows.

a. Ensemble adjoint sensitivity analysis

The perturbation dry total energy norm for each member e_i is calculated from the forecast perturbation z_i as

$$e_i = \frac{1}{2} z^\mathsf{T} z. \tag{32}$$

Here θ is chosen to be the normalized energy norm calculated as follows:

$$\theta_i = \frac{e_i}{\sum e_i}, \tag{33}$$

which can be used to obtain the distribution of the initial perturbations.

b. Ensemble singular-vector sensitivity analysis

The eigenanalysis (30) is conducted with the singular value decomposition of Z

$$\mathsf{Z} = \mathsf{U}\Sigma\mathsf{V}^\mathsf{T}. \tag{34}$$

The eigenvectors of the covariance $\mathsf{Z}\mathsf{H}\mathsf{Z}^\mathsf{T}m^{-1}$ are given by the left singular vector U. Due to the m^{-1} factor in the covariance, the eigenvalues are related to the square of the singular values by

$$\Lambda = \frac{\Sigma\Sigma^\mathsf{T}}{m}, \tag{35}$$

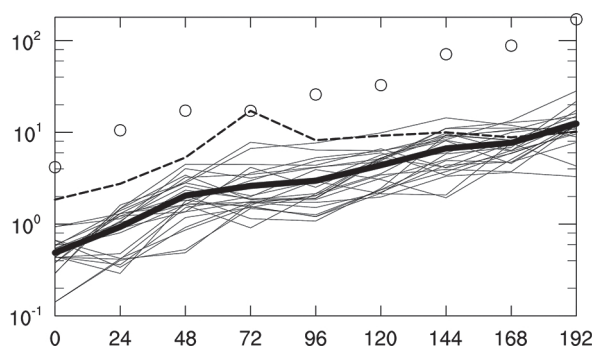(Hartmann 2014). V^T represent the normalized prin-

Fig. 2. Time evolution of the total energy norm calculated from the JMA one-week ensemble forecast from 1200 UTC on 2 January 2003 in the verification region near Japan (120°E–150°E, 20°N–50°N). The horizontal and vertical axes represent the forecast (verification) time (FT, h) and the vertically integrated dry total energy (J kg^{-1}) in the unit area, respectively. The thin solid lines are the ensemble members, the thick solid line is the ensemble mean, and the broken line is the first mode at the FT = 72 h. The circles represent the first mode at each forecast time.

cipal components and the i-th contribution rate $\sigma_i/\mathrm{tr}(\Sigma\Sigma^\mathsf{T})$, where σ_i is a singular value in Σ. Each mode of the initial perturbations can be expressed as an eigenvector multiplied by the square root of the eigenvalue,

$$\mathsf{E}\sqrt{\Lambda} = \frac{\mathsf{U}\Sigma}{\sqrt{m}} = \frac{\mathsf{Z}\mathsf{V}}{\sqrt{m}}, \tag{36}$$

which can be regarded as regression (Hartmann 2014).

3.3 Sensitivity analysis over Japan

EnSVSA and EnASA were applied to the JMA one-week ensemble forecast for every day in January and August 2003, except for 1 January due to data corruption. The verification region was set to Japan (125°E–150°E, 25°N–50°N).

First, the perturbation growth of the first mode of EnSVSA is compared with the ensemble members and the ensemble mean. Figure 2 shows the time evolution of the total energy norm calculated at the forecast from 1200 UTC on 2 January in the verification region. The first mode of the EnSVSA with various verification times (circles) is always larger than the ensemble members (thin solid) and the ensemble mean (thick solid) at the respective forecast time (FT). The first mode of the EnSVSA for FT =

(a)

TE J/kg mode=1-10 contrib=92.1% FT=72hr INIT=20030121

(b)

TE J/kg Adjoint sensitivity FT=72hr INIT=20030121

(c)

TE spread J/kg FT=0hr INIT=20030121
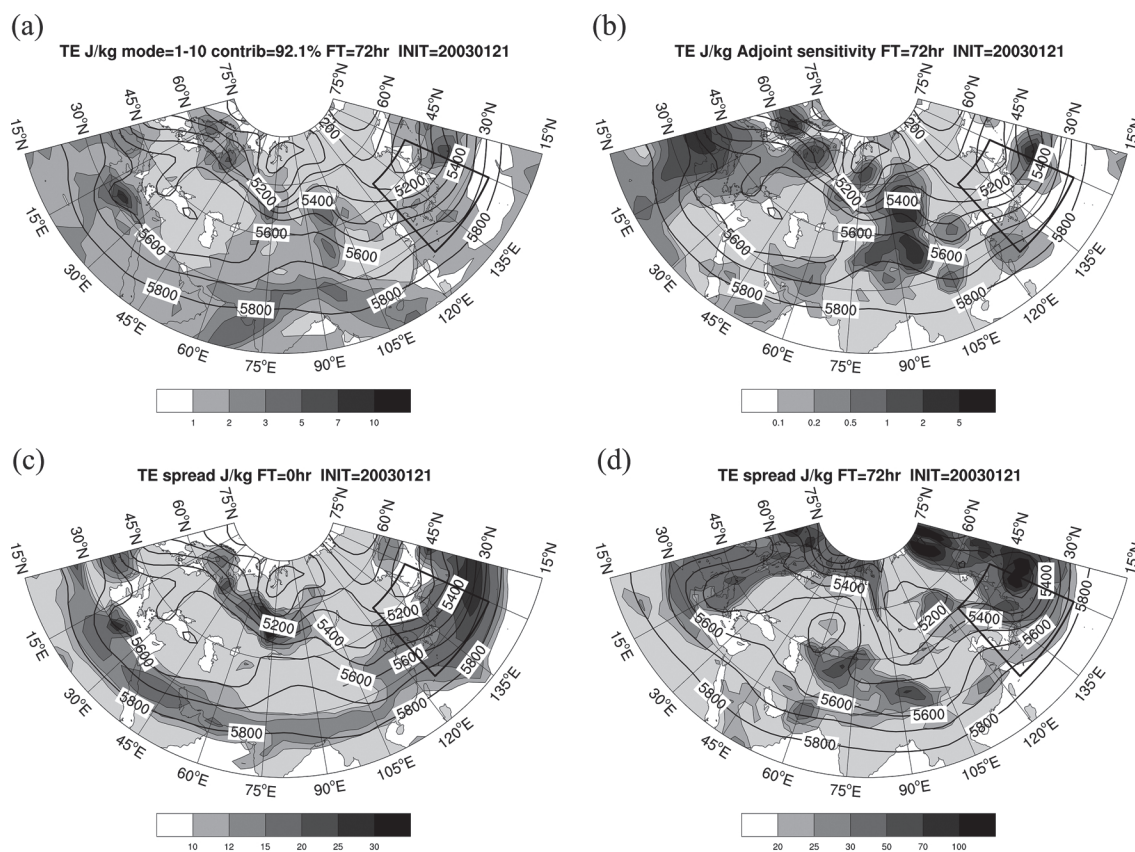
(d)

TE spread J/kg FT=72hr INIT=20030121

Fig. 3.   Distribution of (a) ensemble SV (the sum of the 10-leading modes) and (b) adjoint sensitivity in terms of the vertically integrated dry total energy calculated from the 72-h forecast in the JMA one-week ensemble forecast at 1200 UTC on 21 January 2003. Panels (c) and (d) show the initial and forecast ensemble spreads. The contours show the 500-hPa geopotential height at the initial (a, b, c) and at the verification time (d). The verification region is marked by a rectangle.

72 h (broken) increases rapidly between FT = 48 h and 72 h and is significantly larger than the ensemble members at FT = 72 h, but the first mode is not necessarily the largest in after FT = 96 h. This example indicates that the EnSVSA successfully extracts the mode that is maximized at the verification time in the verification region.

Next, the horizontal distributions of the sensitive regions are shown in terms of the total energy norm (31). Figures 3a and 3b display SV (the sum of the 10-leading modes) and adjoint sensitivity regions, respectively, as calculated from the 72-h forecast at 1200 UTC on 21 January, which is a few days prior to the rapid development of a mid-latitude cyclone near Japan. Both EnSVSA and EnASA indicate large sensitivity near the upstream ridge (90°E–105°E, 30°N–60°N). A train of sensitivity further upstream implies the propagation of Rossby waves along the polar jet. Some sensitivity is also found along the subtropical jet. Some of the remote sensitivities may be false signals due to the ignorance of the $(Y^T G_0 Y)^{-1}$ factor because these sensitivities are insensitive to the valid time and remain even in the calculation from the analysis. The downstream sensitivity is probably due to a cyclone matured to the east of Japan. The initial ensemble spread (Fig. 3c) is large along the polar and subtropical jets and does not highlight the most sensitive ridge discussed previously. The perturbations that correspond to the forecast ensemble spread (Fig. 3d) are used in the ensemble sensitivity analysis.

The EnSVSA and EnASA are calculated from the 72-h forecast at 1200 UTC on 5 August, which is a few days prior to the landfall of Typhoon Etau in Japan (Figs. 4a, b). High sensitivity is found near

(a)

TE J/kg mode=1-10 contrib=90.3% FT=72hr INIT=20030805

(b)

TE J/kg Adjoint sensitivity FT=72hr INIT=20030805

(c)

TE spread J/kg FT=0hr INIT=20030805
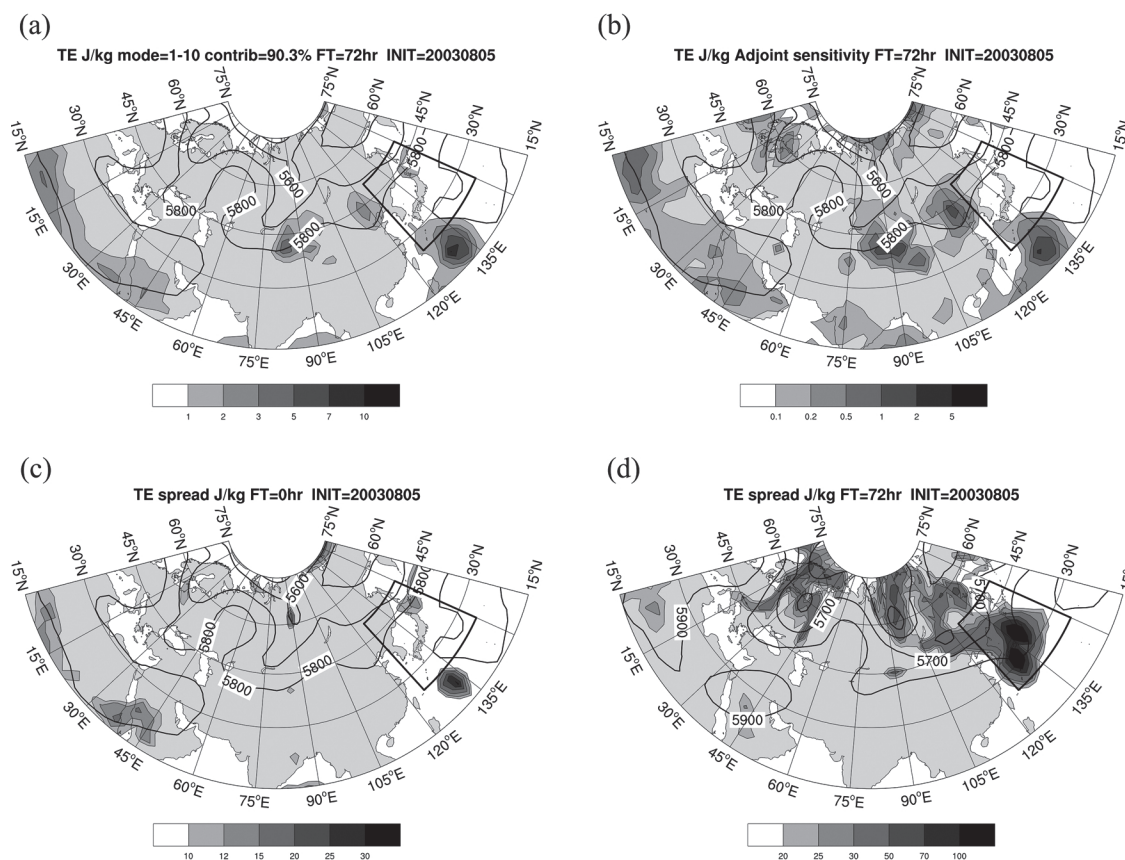
(d)

TE spread J/kg FT=72hr INIT=20030805

Fig. 4. As in Fig. 3, but from 5 August, 2003.

Etau, similar to that in EnSVSA and EnASA of the previous few days. On 5 August, both EnSVSA and EnASA find other sensitive regions along the subtropical jet near (120°E, 45°N) and (90°E, 45°N). The initial ensemble spread is not particularly large in those regions (Fig. 4c). At the verification time, the typhoon and the mid-latitude cyclones merge into a trough (Fig. 4d).

Figure 5 shows the monthly average for January (a, b) and August (c, d) using EnSVSA (a, c) and EnASA (b, d). In January, both EnSVSA (Fig. 5a) and EnASA (Fig. 5b) indicate that the upstream and downstream sensitive regions influence Japan 72 h later. Both EnSVSA and EnASA detect sensitive regions along the subtropical jet centred near (70°E, 35°N) and (80°E, 35°N), respectively. EnSVSA finds a meridionally extended sensitive region at (75°E–60°E, 15°–60°). The northward extension of the SV sensitive regions could be interpreted as the sensitivity associated with the polar jet. The corresponding peak is detected by EnASA near (75°E, 45°N) and near

(95°E, 60°). These sensitivity regions may correspond to Rossby-wave propagations and associated intensification of the Tibetan anticyclone (Takaya and Nakamura 2005). The downstream sensitive region to the east of Japan near (150°E, 40°N) probably corresponds to matured cyclones (Yoshida and Asuma 2004). These cyclones could influence the perturbation growth over Japan because they intensify the northerly flow. In August, the sensitive regions are found to the south of 30°N. These sensitive regions suggest the influence of the monsoon activity.

3.4 The August 2002 storm in Europe

In this subsection, a cut-off cyclone that occurred in August 2002, which was discussed in Enomoto et al. (2007), is revisited. The cut-off was initiated in the upper troposphere on 10 August, 2002 and developed over the Mediterranean Sea. This cyclone that caused historical floods in middle Europe was poorly forecast: the forecast skill of this cyclone was limited to three days (Grazzini and van der Grijn 2002).

(a)

TE J/kg mode=1-10 72hr 200301

(b)

TE J/kg Adjoint Sensitivity FT=72hr 200301

(c)

TE J/kg mode=1-10 72hr 200308

(d)

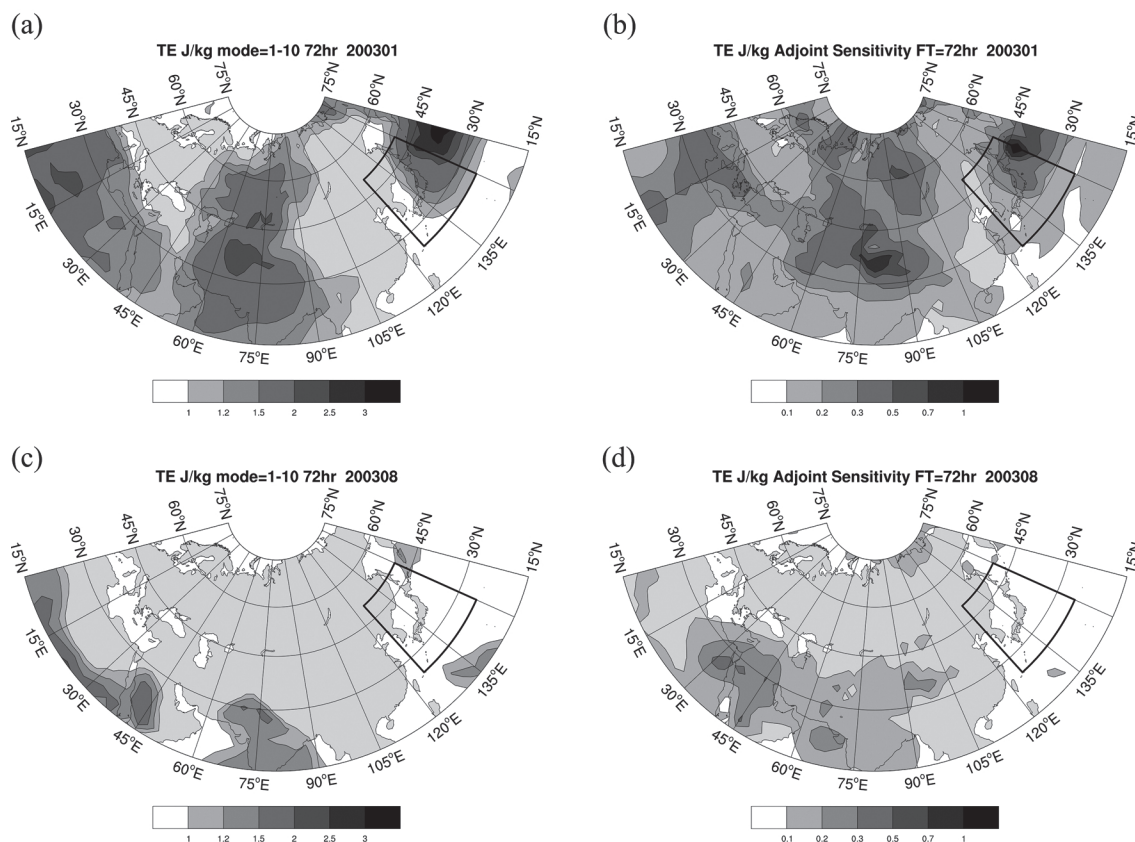TE J/kg Adjoint Sensitivity FT=72hr 200308

Fig. 5.   As in Fig. 3a and 3b, but for the monthly averages of ensemble SV (a, c) and the adjoint (b, d) sensitivity in January (a, b) and in August (c, d).

Figure 6 shows 72-h forecast in the JMA one-week ensemble forecast valid at 1200 UTC on 11 August. Although the cut-off cyclone develops with reasonable intensity at the correct location in some members, it is underestimated in others. In the inaccurate forecast, the cyclone near Ireland and Britain (near the top-left in each panel) is typically excessively developed, whereas the cut-off over middle Europe (centre panel) is too weak.

Enomoto et al. (2007) investigated the predictability of the cut-off in hind cast simulations from the daily JMA analysis at 0000 UTC from 5 to 8 August using AFES (AGCM for the Earth Simulator) (Ohfuchi et al. 2004; Enomoto et al. 2008) with T639L48 (triangular truncation wave number 639, which corresponds to a 0.1875° horizontal resolution, and 48 levels in the vertical). They noticed that Tropical Storm Cristobal off Florida near (72°W, 30°N, Fig. 7) is misplaced in the Gulf of Mexico when the cut-off over Europe is not forecast correctly. The influence from Critobal to the cut-off was investi-

gated in the hindcasts and EnSVSA. Here, we conduct experiments with perturbations from EnSVSA to validate EnSVSA.

a. Influence from a tropical storm

The remote influence of Cristobal can be detected with EnSVSA. In the sensitivity analysis, the JMA one-week ensemble forecast, provided by the Numerical Prediction Division, JMA, is used. Winds, temperature, humidity, geopotential height are provided at five levels (1000, 850, 700, 200, 100 hPa) with a 2.5° horizontal resolution. The sensitivity regions are detected along the east coast of North America by EnSVSA using the JMA one-week ensemble forecast from 1200 UTC on 8 August 2002 with the verification region over western Europe (Fig. 8). The sensitive region is collocated with the trough and associated precipitation; one of the sensitive peaks corresponds to the position of Cristobal centred near (75°W, 30°N, Figs. 7, 8). Among three modes with the contribution rates of 42.6, 21.7, and
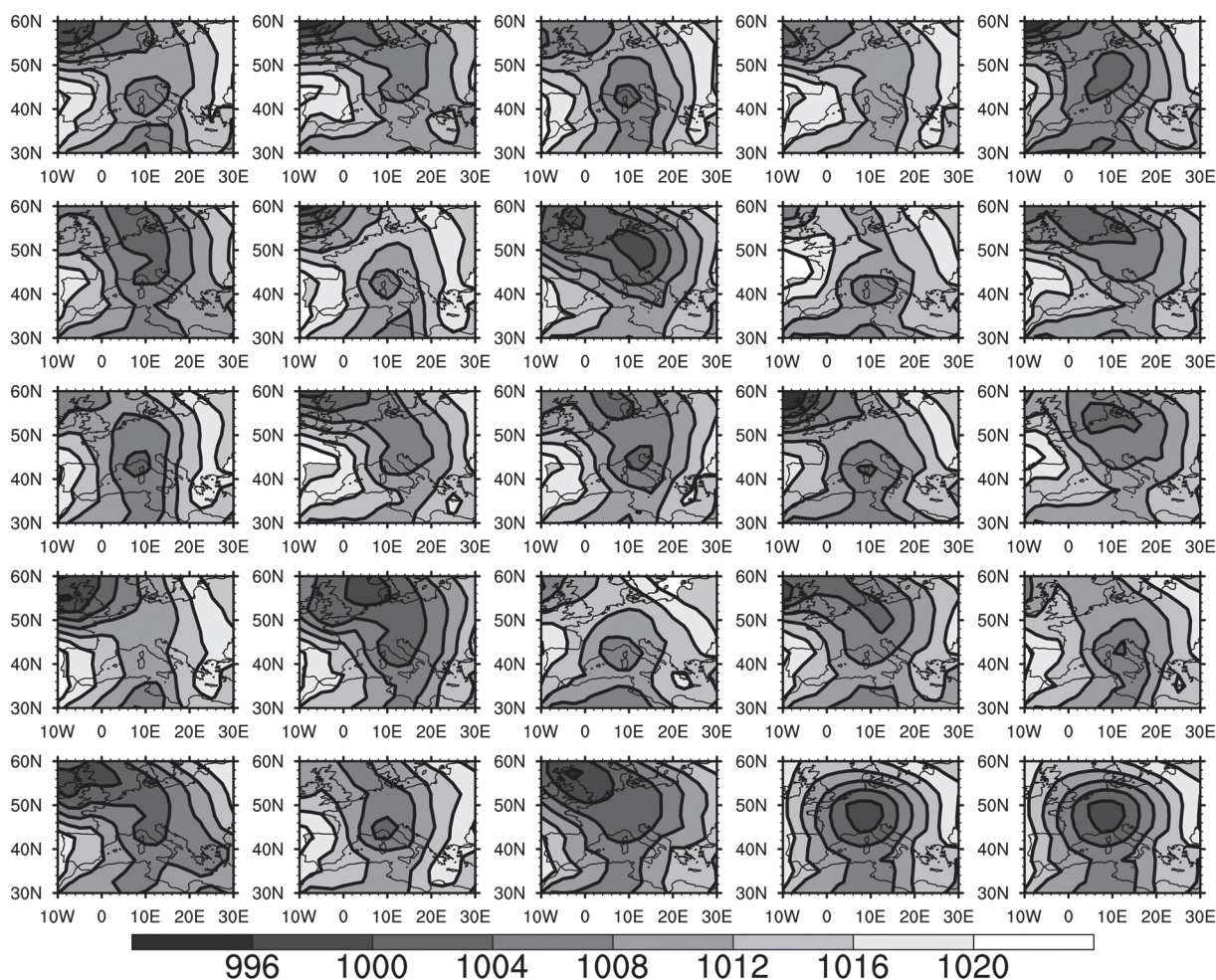
Fig. 6.   Distribution of the sea-level pressure (hPa) within western Europe valid at 1200 UTC on 11 August 2002 in the JMA one-week ensemble forecast from 1200 UTC on 8 August 2002.

8.51 %, respectively, the first SV mode (SV1) is sensitive near Cristobal (Fig. 9a). At the verification time, the cyclone near Ireland and Britain is significantly dampened and the cut-off in the middle Europe is intensified by a few hPa (Fig. 9b).

b.   Perturbed hindcast simulations

The evolved perturbations in SV1 (Fig. 9b) would correct the typical error found in the ensemble forecast (Fig. 6). To test if SV1 grows as suggested by the final distribution in the nonlinear model, perturbed simulations were conducted. The initial conditions prepared from the JMA analysis at 1200 UTC on 8 August 2002 were positively and negatively perturbed with SV1. Figure 10 displays the results of the perturbed simulations of the 20-km resolution

simulation using AFES. The cut-off is produced in both the unperturbed and positively perturbed simulations (Figs. 10a, b), but it is somewhat weaker than the analysis (Fig. 10c). The central pressures are 1002 and 1001 hPa in the unperturbed and positively perturbed simulations, respectively, and 998 hPa in the analysis. Although the pressure change is as small as 1 hPa at the centre, the cyclone in the positively perturbed simulation is deeper by 3–4 hPa near the centre and to the northeast of the centre consistent with SV1 at the verification time (Fig. 9b). The difference (positive minus negative) indicates that increased rainfall occurred to north of the Alps.

The consistency between the EnSVSA and the hindcast simulations can also be confirmed from the vertically integrated total energy norm per unit area
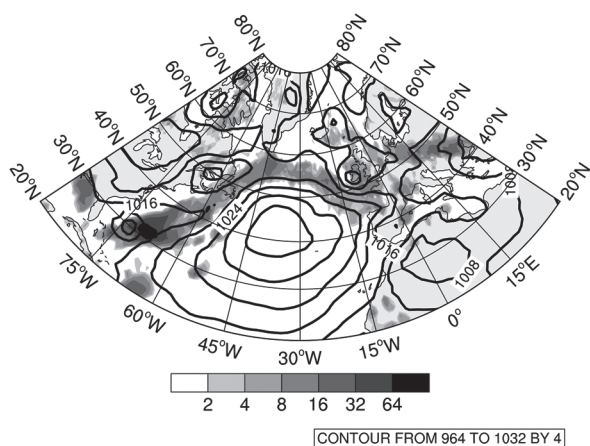
Fig. 7.   Distribution of the sea-level pressure (hPa, JMA analysis) (contoured) at 1200 UTC and daily precipitation (mm day^{-1}, GPCP data) (shaded) on 8 August 2002.
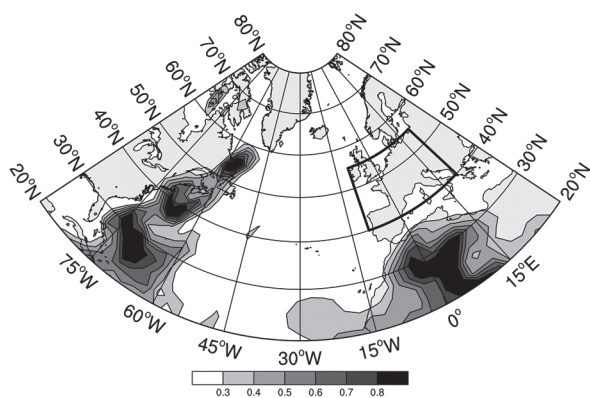

Fig. 8.   Distribution of the sum of the dry total energy norm (J kg^{-1}) of the EnSVSA leading three modes weighted with the contribution rates calculated with the 72-h forecast in the JMA one-week ensemble forecast from 1200 UTC on 8 August 2002. The rectangle marks the verification region over western Europe (10°W–20°E, 40°N–55°N).

(a)

(b)

Fig. 9.   Distribution of the surface pressure perturbation of the first mode computed from the JMA one-week ensemble forecast by the EnSVSA at (a) FT = 0 h and at (b) FT = 72 h from the initial time (1200 UTC on 8 August 2002).

(31) of the positively perturbed simulation and that of SV1 in the verification region (Fig. 11 and Table 1). The perturbation is calculated as the value of the positively perturbed run minus that of the control simulation. In spite of the differences in the forecast models and the data resolution, the time evolution of the total energy norm is similar to each other. The growth in the positively perturbed simulations is faster between
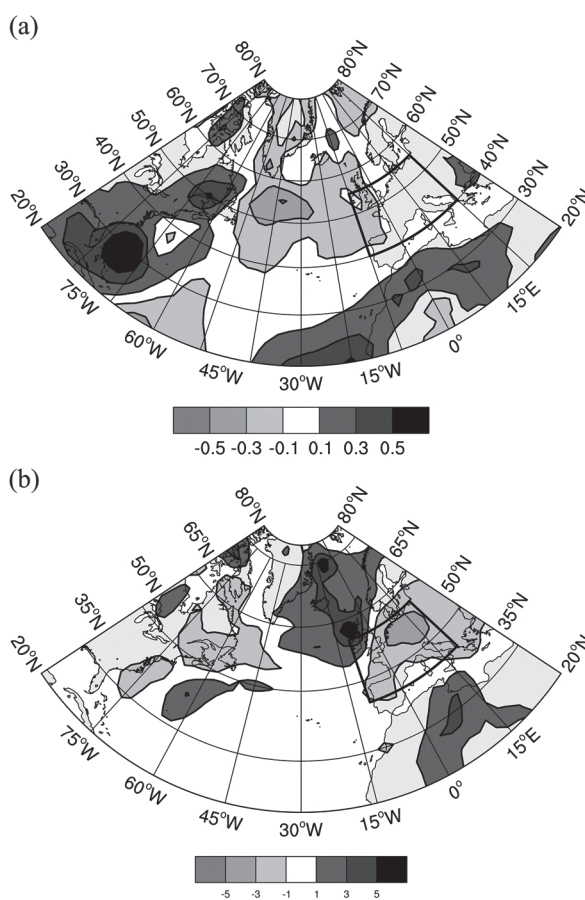
the initial time and 10 August and slower toward the verification time, but the values at the verification time are similar. These results indicate the robustness of EnSVSA.

4.   Summary and discussion

Ensemble-based adjoint and singular vector sensitivity analysis methods were formulated as finite member approximations. These methods are simple because they only require previously run ensemble forecasts as input and a few matrix operations and because they do not consider observations explicitly. Adjoint sensitivity is calculated as the deviation of the norm from the control forecast (26) without an adjoint model M^T. Singular vector sensitivity is computed from the singular vector analysis of the evolved
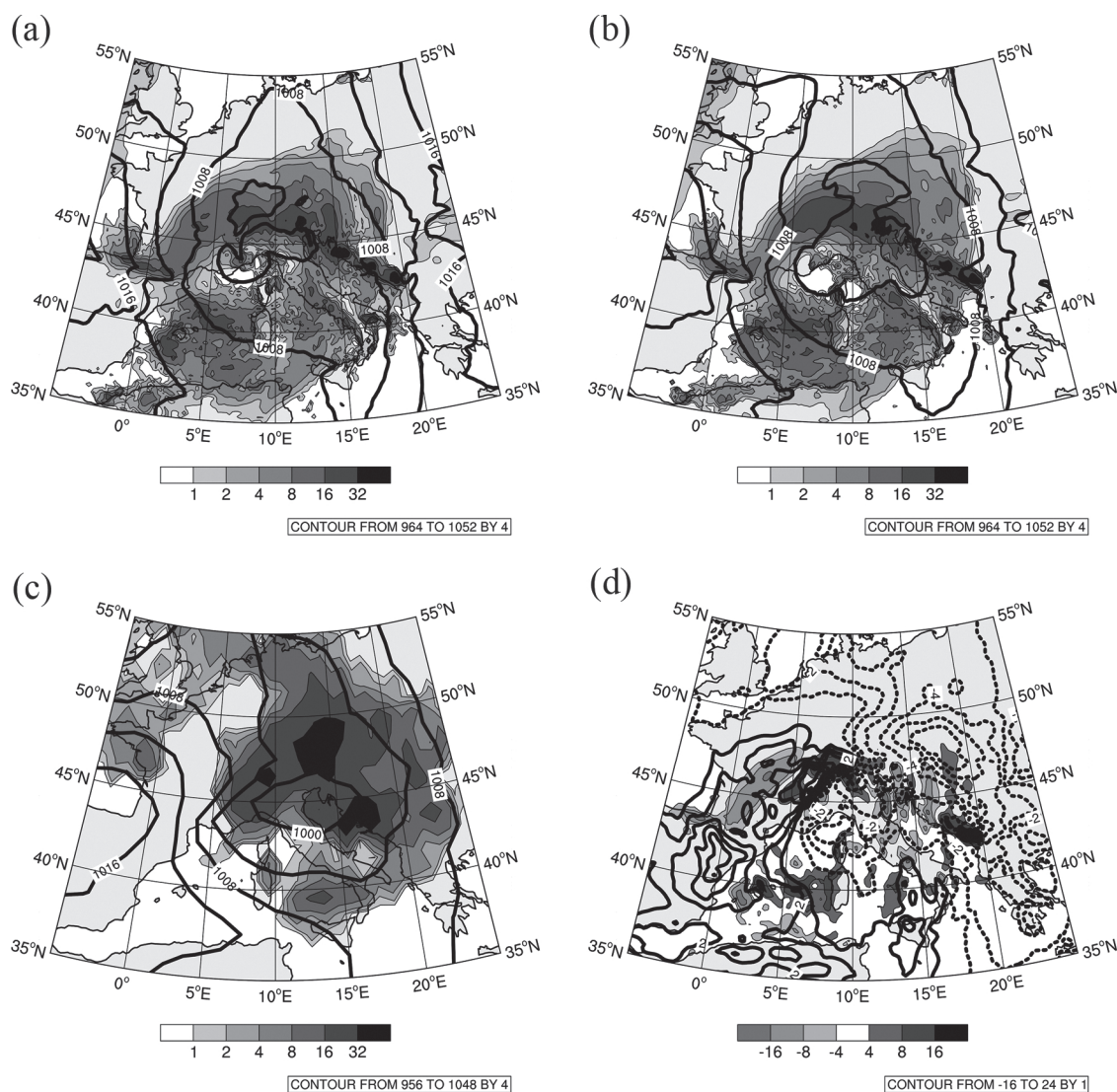
Fig. 10.   Distribution of the sea-level pressure (hPa) and precipitation (mm day^{-1}) at 1200 UTC on 11 August, 2002 in (a) unperturbed and (b) perturbed experiments using AFES, (c) the analysis, and (d) the difference between positively perturbed and negatively perturbed simulations. The sea-level pressure and precipitation in (c) are prepared from the JMA analysis and GPCP data.

perturbation matrix Z (30) without forward or backward time integration of tangent linear M or adjoint models M^T, respectively. Because the ensemble size is significantly smaller than the degrees of the freedom of the model, our ensemble-based methods are computationally inexpensive.

Despite their simplicity our methods are able to identify regions sensitive to the initial perturbations that grow in the verification region at the verification time (Fig. 2) due to the presence of a mid-latitude

cyclone (Fig. 3) and a tropical cyclone (Fig. 4); the regions identified by these methods are more specific than the regions with large forecast ensemble spread. The monthly means of sensitivity in boreal winter and in boreal summer reveal potentially important regions for enhanced observations (Fig. 5). The growth of detected sensitive initial perturbations is confirmed in nonlinear simulations of the August 2002 storm in Europe (Figs. 8, 9, 10).
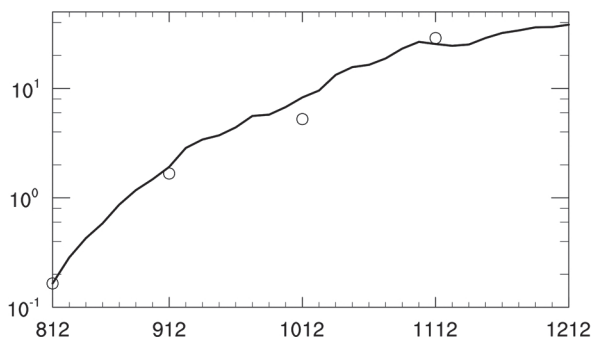
For practical purposes, such as finding the candi-

Fig. 11.   Time evolution of the total energy norm (J kg^{-1}) of the positively perturbed simulation (solid line) and the first mode of ensemble SV (circles) calculated from the JMA one-week ensemble forecast from 1200 UTC on 8 August 2002 in the verification region over western Europe (10°W–20°E, 40°N–50°N).

Table 1.   The total energy norm (J kg^{-1}) of the positively perturbed simulation (AFES, middle column) and the first mode of ensemble SV (SV1, right column) calculated from the JMA one-week ensemble forecast from 1200 UTC on 8 August 2002 in the verification region over western Europe (10°W–20°E, 40°N–50°N) at 1200 UTC on each day (date, left column) in August 2002.

date	AFES	SV1
8	0.165	
9	2.85	1.67
10	9.56	5.24
11	24.5	28.8

dates for targeted observations and investigating remote effects (Enomoto et al. 2007; Nishii and Nakamura 2010; Matsueda et al. 2011), our ensemble-based sensitivity analysis would be a useful tool. For more quantitative applications, three issues that were not addressed here may require careful examinations. First, we did not orthogonalize the perturbations for simplicity: the initial perturbations y used as input may not always be orthogonal, depending on how the perturbations are generated. In addition, we retained both positively and negatively perturbed members because the evolved perturbations have nonlinear components. Obviously, the number of independent modes is not equal to the ensemble size of 25, but rather equal to 12. Therefore, at most, the leading 10 modes were investigated in this study. We believe the orthogonalization does not change the results of the sensitivity analysis presented here, at least qualitatively.

Second, similar to other ensemble-based methods, including ensemble-based data assimilation, sampling error is an open issue. Some physical interpretation is required to judge sensitive regions, particularly for remote sensitive regions. For example, we could filter noise by applying the localization. We did not filter noise in this paper because some remote signals, such as Rossby wave trains, cannot be simply dropped based on distance. Although more objective methods are preferable, we can still use information from the simplified sensitivity analysis. In the practical situation of field campaigns, we could ignore sensitive regions that are outside of the extent of the targeted observations.

Third, it was not examined how the ensemble SV and adjoint sensitivity can represent the SV and adjoint sensitivity calculated using the adjoint model. We discussed that the same solution should be obtained in the limit when the ensemble size matches the degree of freedom in the model. It is of practical and theoretical interest how fast the results become similar with increasing ensemble size. It is also interesting to examine the difference in the structure and distribution of noise due to sampling error when the ensemble size is limited.

Finally, it is still unclear to what extent our methods are valid for the verification time. The initial perturbations are assumed to be small and grow linearly, although the ensemble forecast itself is nonlinear. Thus, our methods cannot be applied to a long verification time. The precise length depends on the time scale of each phenomenon. In the mid-latitudes, for a synoptic scale motion, the scale is of the order of several days. For a convective scale motion, the time scale can be a day or less. In practice, with mixed scale motions under the influence of synoptic and convective scales, determining the suitability of the sensitivity analysis is not straightforward. One of the methods to validate the sensitivity analysis can be the perturbed simulations conducted in Gelaro et al. (1998) and in Subsection 3.4b.

and improve the manuscript significantly. The authors also thank Dr. Takemasa Miyoshi for inviting us for contribution and for editing. The NCAR Command Language (NCL) was used for data analysis and visualization.

References

Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sesitivity analysis with applications to observation targeting. *Mon. Wea. Rev.*, **135**, 4117–4134.

Bishop, C. H., and Z. Toth, 1999: Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, **56**, 1748–1765.

Bishop, C. H., J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Buizza, R., and A. Montani, 1999: Targeting observations using singular vectors. *J. Atmos. Sci.*, **56**, 2965–2985.

Buizza, R., J. Tribbia, F. Molteni, and T. Palmer, 1993: Computation of optimal unstable structures for a numerical weather prediction model. *Tellus A*, **45**, 388–407.

Enomoto, T., S. Yamane, and W. Ohfuchi, 2006: Simple sensitivity analysis using ensemble forecast. *Proceedings of Third Workshop on Mechanisms of Climate Variation and its Predictability*, 40–43 (in Japanese).

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. *Meteor. Atmos. Phys.*, **96**, 29–42.

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

Fujii, Y., H. Tsujino, N. Usui, H. Nakano, and M. Kamachi, 2008: Application of singular vector analysis to the Kuroshio large meander. *J. Geophys. Res.,* **113**, C07026, doi:10.1029/2007JC004476.

Gelaro, R., R. Buizza, T. N. Palmer, and E. Klinker, 1998: Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. *J. Atmos. Sci.*, **55**, 1012–1037.

Grazzini, F., and G. van der Grijn, 2002: Central European floods during summer 2002. *ECMWF newsletter*, **96**, 18–28.

Hamill, T. M., and C. Snyder, 2002: Using improved background-error co-variances from an ensemble Kalman filter for adaptive observations. *Mon. Wea. Rev.*, **130**, 1552–1572.

Hartmann, D. L., 2014: EOF/PC Analysis. [Available at http://www.atmos.washington.edu/~dennis/552_Notes_4.pdf.]

Ito, K., and C.-C. Wu, 2013: Typhoon-position-oriented sensitivity analysis. Part I: theory and verification. *J. Atmos. Sci.*, **70**, 2525–2546.

Kalnay, E., Y. Ota, T. Miyoshi, and J. Liu, 2012: A simpler formation of forecast sensitivity to observations: application to ensemble Kalman filters. *Tellus A*, **64**, 1–9.

Kyouda, M., 2002: The ensemble prediction system for medium-range weather forecasting at JMA. *Research Activities in Atmospheric and Oceanic Modelling.* Ritchie, H. (ed.), Report No. 32 WMO/TD-No. 1105.

Kyouda, M., 2006: The one-week EPS. *Numerical prediction technical report.* **52**, Forecast Department, the Japan Meteorological Agency, chapter 3, 23–33 (in Japanese).

Langland, R. H., M. A. Shapiro, and R. Gelaro, 2000: Initial condition sensitivity and error growth in forecasts of the 25 January 2000 East Coast snowstorm. *Mon. Wea. Rev.*, **130**, 957–974.

Liu, J., E. Kalnay, T. Miyoshi, and C. Cardinali, 2009: Analysis sensitivity calculation in an ensemble Kalman filter. *Quart. J. Roy. Meteor. Soc.*, **135**, 1842–1851.

Matsueda, M., M. Kyouda, Z. Toth, H. L. Tanaka, and T. Tsuyuki, 2011: Predictability of an atmospheric blocking event that occurred on 15 December 2005. *Mon. Wea. Rev.*, **139**, 2455–2470.

Mu, M., W. S. Duan, and B. Wang, 2003: Conditional nonlinear optimal perturbations and its applications. *Nonlin. Processes Geophys.*, **10**, 493–501.

Nishii, K., and H. Nakamura, 2010: Three-dimensional evolution of ensemble forecast spread during the onset of a stratospheric sudden warming event in January 2006. *Quart. J. Roy. Meteor. Soc.*, **136**, 894–905.

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.

Rabier, F., E. Klinker, P. Courtier, and A. Hollingsworth, 1996: Sensitivity of forecast errors to initial conditions. *Quart. J. Roy. Meteor. Soc.*, **112**, 121–150.

Snyder, C., 1996: Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Amer. Meteor. Soc.*, **77**, 953–965.

Takaya, K., and H. Nakamura, 2005: Mechanism of intraseasonal amplification of the cold Siberian high. *J. Atmos. Sci.*, **62**, 4423–4440.

Torn, R. D., and G. J. Hakim, 2008: Ensemble-based sensitivity analysis. *Mon. Wea. Rev.*, **136**, 663–677.

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.

Wu, C.-C., J.-H. Chen, P.-H. Lin, and K.-H. Chou, 2007: Targeted observations of tropical cyclone movement based on the adjoint-derived sensitivity steering vector. *J. Atmos. Sci.*, **64**, 2611–2626.

Yoshida, A., and Y. Asuma, 2004: Structures and environment of explosively developing extratropical cyclones in the northwestern Pacific region. *Mon. Wea. Rev.*, **132**, 1121–1142.