Ensemble Forecasting of Tropical Cyclone Motion Using a Barotropic Model. Part II: Perturbations of the Vortex

KEVIN K. W. CHEUNG AND JOHNNY C. L. CHAN

Department of Physics and Materials Science, City University of Hong Kong, Kowloon, Hong Kong, China

(Manuscript received 30 April 1998, in final form 15 December 1998)

ABSTRACT

In Part I of this study, the technique of ensemble forecasting is applied to the problem of tropical cyclone motion prediction by perturbing the environmental flow. In this part, the focus is shifted to perturbation of the vortex structure. The same barotropic model as in Part I is used to predict 66 cases from the Tropical Cyclone Motion (TCM-90) experiment. Two series of experiments are performed. The first consists of the Monte Carlo forecast for the vortex (MCFV), lagged-average forecast for the vortex (LAFV), and breeding of growing modes for the vortex (BGMV). These are applied to the original analyses without adding any synthetic vortex, and their techniques of generating perturbations are similar to those in Part I. The second series adopts a technique that simulates uncertainties in estimating the synthetic vortex structure. The effect of adding an initial position error (IPER) is first explored. Then a set of four experiments (BETA) is carried out to study the consequence of perturbing the persistence vector and/or various parameters used to generate the β gyres in a spun-up vortex.

Response of the model forecasts to random noise added in the experiment MCFV is found to be low, and the ensemble mean is thus always close to the control forecast. The situation is similar when the rms size of the noise is slightly varied, or when its characteristic length scale is changed. The skill of the IPER experiment also differs little from the control forecast. The remaining experiments other than the MCFV and IPER show a similar average skill to one another when verified both under the perfect model assumption (PMA) and by the best tracks. In the PMA verification, potential improvement over the control forecast can be obtained by the ensemble mean in the LAFV, BGMV, and some sets of the BETA experiments. However, some failure cases are found in the LAFV and BGMV experiments when the vorticity center cannot be well identified during the model integration.

When compared with the best tracks, a portion of the cases can still outperform the control in the LAFV, BGMV, and BETA experiments. Since different control forecasts are used in different experiments, the forecast errors are scaled to the same benchmark before they are compared. It is found that among the BETA experiments, the configuration with the best performance is to perturb the parameters for generating the β gyres and persistence vector simultaneously. It can outperform LAFV after 48 h and has comparable performance with BGMV. This set of BETA ensemble may therefore be suitable for substituting BGMV when a synthetic vortex is necessary. Another implication is that since the persistence vector represents an improved large-scale flow, potential skill should exist for combining the perturbed β gyres with the environmental perturbations used in Part I.

1. Introduction

In the first part of this study, Cheung and Chan (1999, hereafter referred to as Part I), three perturbation methodologies, the Monte Carlo forecast (MCF), the lagged-average forecast (LAF), and the breeding of growing modes (BGM), were applied to the environmental flow associated with tropical cyclones (TCs) to examine the usefulness of the ensemble forecasting technique in TC motion prediction. The effectiveness of perturbing the vortex structure, which is one of the two major factors (the other being the environmental flow) that determine TC motion, is evaluated in this paper.

Of these two factors, environmental steering is a first-order effect because it can explain a large part of the TC motion especially when the flow is strong. Theoretical studies (e.g., Fiorino and Elsberry 1989) have demonstrated that the motion is not very sensitive to the intensity and inner-core structure, but depends more on the outer-core velocity profile. This has a great impact on TC forecasting because the structure of a TC vortex cannot be determined accurately from the sparse observations over the tropical oceans. Furthermore, the interaction between the TC vortex and the environment is a highly complicated feedback process. In the barotropic view, the result of interaction is a systematic deviation of the motion from the environmental steering (Chan and Gray 1982) due to the formation of an azi-

Corresponding author address: Prof. Johnny C.-L. Chan, Department of Physics and Materials Science, City University of Hong Kong, 83 Tat Chee Ave., Kowloon, Hong Kong, China.
E-mail: apjcchan@cityu.edu.hk

muthally asymmetric flow within the vortex (Fiorino and Elsberry 1989), which has been referred to as the beta effect. Because of the relatively simple barotropic dynamics and its computational efficiency, a large amount of work has been performed on the barotropic modeling of TC motion. Some of this concentrated on the effect of vortex structure (Peng and Williams 1990; Smith and Ulrich 1990).

A method common to almost all numerical weather prediction (NWP) centers to improve the representation of the TC vortex in a model is to add a synthetic vortex as a part of the initialization process. The design of the synthetic vortex by different centers can vary substantially. Experiments using different tangential wind profiles in a barotropic model have also been performed (Leslie and Holland 1995). The results show a large sensitivity to the structure of the synthetic vortex. Furthermore, it has been generally recognized recently that the addition of an axisymmetric vortex is not adequate since a certain period is required for the vortex to adapt to the environment (Carr and Elsberry 1992). The consequence is an incorrect motion of the modeled TC. Thus, most of the NWP centers have started to introduce some form of asymmetry in their synthetic vortices. The asymmetry can be a simple persistence vector to force the vortex in a right direction (DeMaria et al. 1992; Heming et al. 1995), barotropic β-gyre structure (Kurihara et al. 1993), or some analytically derived asymmetries (Ueno 1989).

In this study, several experiments are first performed using simple perturbations of the vortex structure represented in the analyses (as was applied in Part I on the environmental flow). Because of the importance of a synthetic vortex and the associated asymmetries, another series of experiments is also performed. This second series involves substituting the original vortex by a synthetic one. The idea is that if the addition of such a vortex is viewed as part of the model configuration, an ensemble formed by systematically perturbing the synthetic vortex can probably sample the vortex uncertainties better when compared with the methods in the first series of experiments. The design philosophy here is to consider how errors in the observed quantities (e.g., position and intensity) are transferred to the various components of the vortex. Recently, Krishnamurti et al. (1997) studied an ensemble of hurricane forecasts using different analyses, which represent a combination of estimates of analysis error and perturbations to the model physics (from different assimilation centers) at the initial time. On the other hand, Zhang and Krishnamurti (1997) used a combination of perturbations of the position and an empirical orthogonal function analysis of error growth in the large-scale flow to form a hurricane ensemble. The methodologies here differ from those two studies in that perturbations to the environmental flow and vortex circulation are separated, and those from dynamically constrained methods and purely stochastic ones are also isolated.

The flow of this paper is as follows. The ideas and operational details of the experiments are given in section 2. Similar to Part I, the ensembles are verified both under the perfect model assumption and using the best tracks. The verification results are presented in section 3. The implications of these results are further discussed in section 4. A summary and some remarks concerning the relationship of the present study to Part I together with further possible extensions are given in section 5.

2. Experimental design

a. Configuration of the model

The configuration of the model is identical to that used in Part I. It is a semi-Lagrangian barotropic model (Holland et al. 1991) integrated on a 50-km-resolution Mercator grid covering approximately the domain 7°S–48°N, 100°–162°E. The input to the model is the deep-layer-mean (850–300 hPa) winds of the final analyses from the Tropical Cyclone Motion experiment (Rogers et al. 1993). The large-scale steering derived from this layer has been shown to have a systematic relationship with TC motion (Chan and Gray 1982), and similar definitions of ''layer mean'' have also been used in barotropic modeling (e.g., DeMaria et al. 1992; Aberson and DeMaria 1994). The way of boundary-data nesting also remains unchanged; that is analysis data or those from their time interpolation (for 0600 or 1800 UTC) are blended to the domain during the integration.

This study is based on 66 cases taken from nine TCs that are the same as those in Part I. The selection criteria can also be found there. Among the 66 cases, 41 allow forecasts up to 72 h, 15 stop at 60 h, and the remaining 10 at 48 h. A complete listing of the cases is given in the appendix.

b. Ensembles without a synthetic vortex

The three types of perturbation methods described in Part I, that is, the Monte Carlo method, the lagged-average forecast, and the breeding of growing modes, are also employed here. Each of these methods as applied to the vortex flow is described in this section together with the control forecasts.

1) CONTROL FORECASTS

Since the experiments do not involve the addition of a synthetic vortex, the control runs are taken to be those forecasts made directly from the analyses. Two major problems are readily apparent from the tracks for all the 66 cases (Fig. 1). First, the initial position identified by the model (a vorticity maximum) usually deviates from the best-track location by around 100 km and sometimes more. Second, the tracks in some cases are irregular and move in a zigzag manner (e.g., in some cases of Yancy and Zola). However, the overall performance is not

much lower than that in Part I, where an axisymmetric synthetic vortex was used to replace that in the analysis. A skill score relative to the climatology and persistence (CLIPER) forecast (Xu and Neumann 1985) was also used there to measure the control performance. The mean score shows that the model can outperform the CLIPER after 48 h slightly but the case-to-case variability is large. Readers are referred to Part I for a quantitative evaluation.

2) MONTE CARLO FORECAST FOR THE VORTEX

In Monte Carlo forecast for the vortex (MCFV), Gaussian random noise is added to the zonal and meridional wind components within the vortex domain, which is taken as that of the so-called analyzed vortex determined by the vortex specification algorithm from the Geophysical Fluid Dynamics Laboratory (GFDL) (Kurihara et al. 1995). In the corresponding environmental perturbations experiment (MCF in Part I), the random noise added is controlled at a root-mean-square (rms) value of 3 m s^{-1}. This restriction is relaxed here and the sensitivity to the magnitude of perturbations will be examined. Initially no spatial correlation exists between the random noise at any two grid points. Additional experiments are also performed to add spatially filtered perturbations to the vortex that are generated as follows. Random errors are first created on the whole model domain. Then, a filtering operator with a characteristic length scale is applied successively to the errors. By varying the parameter controlling the length scale, noise shorter than a particular wavelength can be largely removed and in this way correlation among the errors on different grid points is introduced. Finally, the errors in the vortex region are added to the analysis. In the experiments, the characteristic length scale is varied from about 4° to 11° latitude, which should include the size of all the TCs considered. Details of the filtering algorithm can be found in Kurihara et al. (1993). Another result also found in the MCF in Part I is that the variance of such an ensemble remains almost the same when the ensemble size does not go below 50. Thus, 50 members are also used in the present MCFV ensemble. Nine experiments (one from each TC) are performed for verification.

3) LAGGED-AVERAGE FORECAST FOR THE VORTEX

The background idea of lagged-average forecast for the vortex (LAFV) is exactly the same as the methodology LAF in Part I except that the lagged error (short-range forecast error prior to the initial forecast time) is restricted to the vortex. The integration schedule (Fig. 2) is identical to that for the environmental perturbation. A maximum lag period of 24 h is used, with two successive lagged integrations separated by 6 h. In total, 16 perturbations can be generated. Four (L1, . . . , L4) belong to the simple-lag (SL) members. They are the difference between the forecast vortex (as isolated by the Kurihara filtering algorithm, denoted by F) and the original vortex in the analysis (denoted by A) at T00. In practice, the forecast vortex is simply moved to the best-track position at T00 before the prognostic run. The other 12 (S12 . . . , S43) belong to the short-range forecast difference (SRFD) type (denoted by D), in which the differences between the forecast vortices from the lagged integrations are computed and then added to the wind values only within the vortex domain at T00. Together with the control forecast, the whole ensemble consists of 17 members.

4) BREEDING OF GROWING MODES FOR THE VORTEX

In breeding of growing modes for the vortex (BGMV), the breeding time for the vortex (24 h) is again the same as that in BGM for the environmental perturbation (Fig. 3) and four 6-h cycles are performed. The first cycle starts from a pair of forecasts: one starting from the analysis (denoted by A) and another from one to which random noise (denoted by R) is added within the vortex domain (i.e., just like a member from the MCFV ensemble). The difference (denoted by D) between the two forecast vortices (6 h later) is added to (the P member) or subtracted from (the N member) the analysis vortex at that time. Integrations from the unperturbed and perturbed analyses are performed again. Three similar cycles are carried out until the initial time (T00) is reached and a pair of perturbations generated. Eight pairs of such bred errors are generated to equate the total number of members in LAFV. Members P1–P8 denote those from the cycles where perturbations are added, while N1 to N8 represent members in which the perturbations are subtracted.

c. Ensembles with a synthetic vortex

1) CONTROL FORECASTS

The preparation of the environmental flow for adding a synthetic vortex is the same as in Part I. The original vortex in the analysis is filtered using the Kurihara vortex specification technique. The synthetic vortex depends on the experiment performed. In the initial position error experiment (IPER) below, only an axisymmetric vortex (Chan and Williams 1987) is placed at the best-track position. Its tangential wind profile is restated here:

$$v_t = v_{\max}\left(\frac{r}{r_{\max}}\right)\exp\left\{\frac{1}{b}\left[1 - \left(\frac{r}{r_{\max}}\right)^b\right]\right\}, \qquad (1)$$

where $v_{\max}$ is the maximum wind speed, $r_{\max}$ the radius at which $v_{\max}$ occurs, and b the size factor ($0 < b \leq 1$). The value of $v_{\max}$ is chosen to be 0.8 times the warning intensity of the cyclone to account for the lower value of the vertically integrated mean than the observed

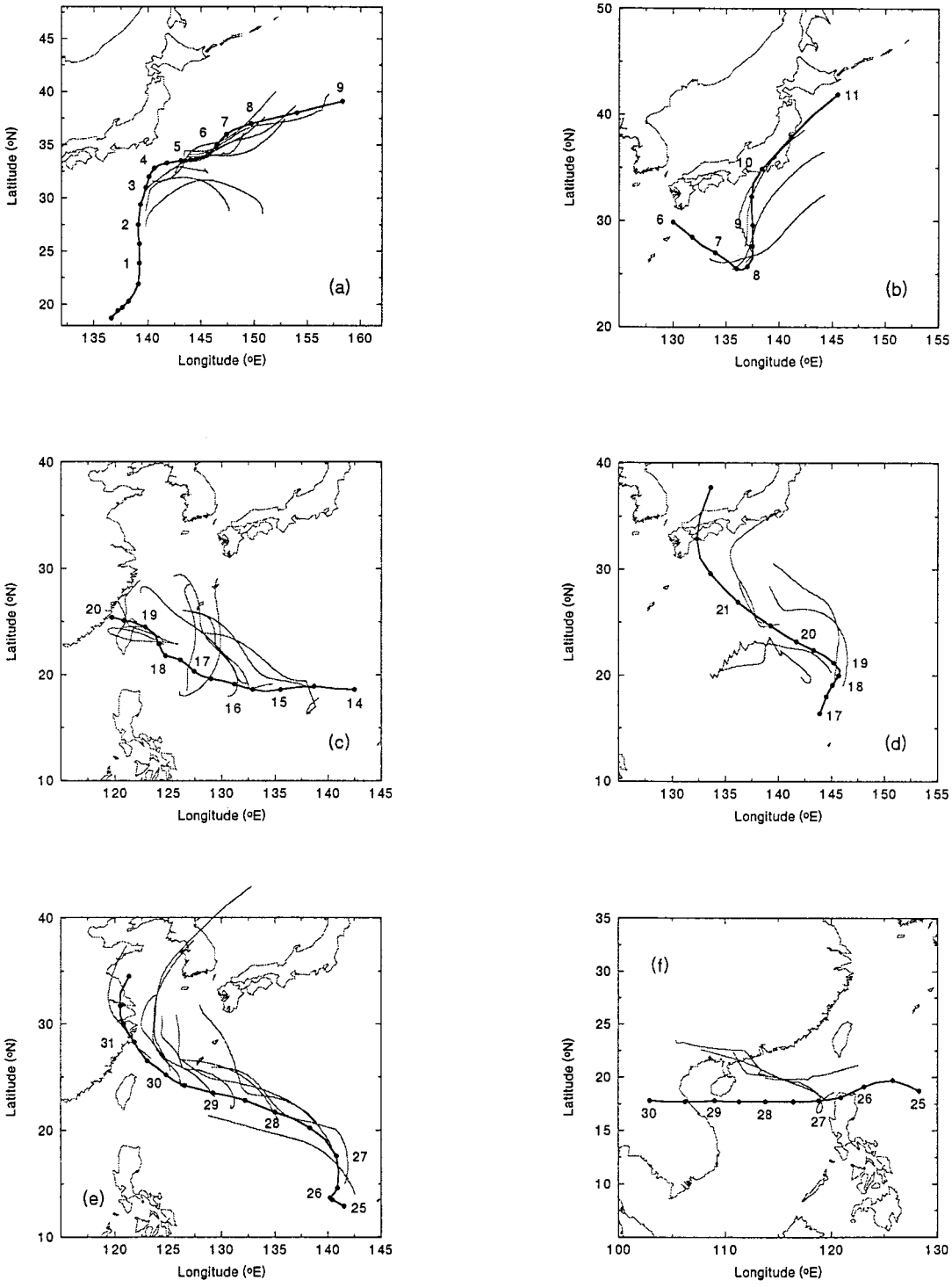

FIG. 1. Control forecasts run directly from analyses for TC (a) Vernon, (b) Winona, (c) Yancy, (d) Zola, (e) Abe, (f) Becky, (g) Dot, (h) Ed, and (i) Flo. The numbers indicated are the dates in August 1990 for (a)–(e) and in September 1990 for (f)–(i).

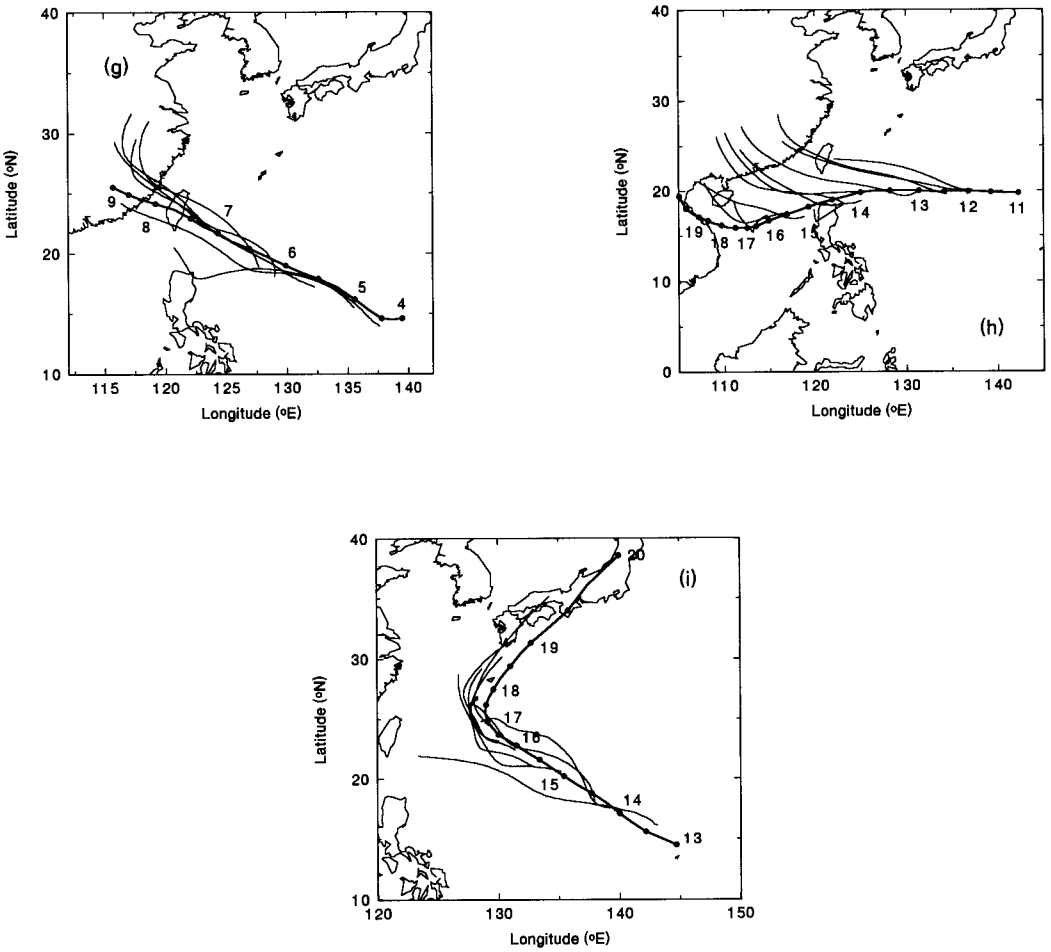

FIG. 1. (*Continued*)

surface wind (DeMaria et al. 1992), and b is obtained indirectly from the radius of the vortex boundary r_f (at which the azimuthally averaged tangential wind is 6 m s^{-1} and its gradient is less than 4×10^{-6} s^{-1}) defined in the Kurihara filtering algorithm. Because the radius of maximum wind is generally not known, r_{max} is set at 100 km.

In Part I, the entire axisymmetric vortex profile in (1) was simply added to the environmental flow without any vortex blending. This has been mentioned as being a possible source of error for the control forecasts. Since the focus there was the effect of environmental perturbations (and the control vortex was the same in all ensemble members), this was not considered to be a serious problem. However, because the perturbations are added specifically to the vortex structure in this part of the study, a clear definition of the radial extent of the perturbations is needed. To meet this requirement, a slight modification to the control forecasts is made in that a vortex blending scheme is adopted. This is accomplished by defining a radial blending zone (bounded by two radii r_{in} and r_{out}) in which the contribution from the environment and the vortex varies linearly. Peng et al. (1993) adopted the constant values of 630 km and 840 km, respectively, for western North Pacific TCs. In our study, instead, the following variable values are taken:

$$r_{in} = 0.85 \times r_f \qquad r_{out} = 1.15 \times r_f. \tag{2}$$

The values of r_{in} and r_{out} from these formulas not only agree roughly with those of Peng et al. (1993) for medium-size TCs, but have the additional advantage of an adjustable blending location for midget or very large TCs (e.g., Typhoon Yancy in our study). The blending (also termed transplantation in some literature) is performed as

$$f(r, \theta) = f_B(r, \theta)w(r) + f_E(r, \theta)[1 - w(r)], \tag{3}$$

where $w(r)$ is the weighting function that controls the relative contribution from the synthetic vortex $[f_B(r, \theta)]$ and the environmental flow $[f_E(r, \theta)]$:

$$w(r) = \begin{cases} 1 & r < r_{in} \\ \dfrac{r_{out} - r}{r_{out} - r_{in}} & r_{in} \le r \le r_{out} \\ 0 & r > r_{out}. \end{cases} \tag{4}$$

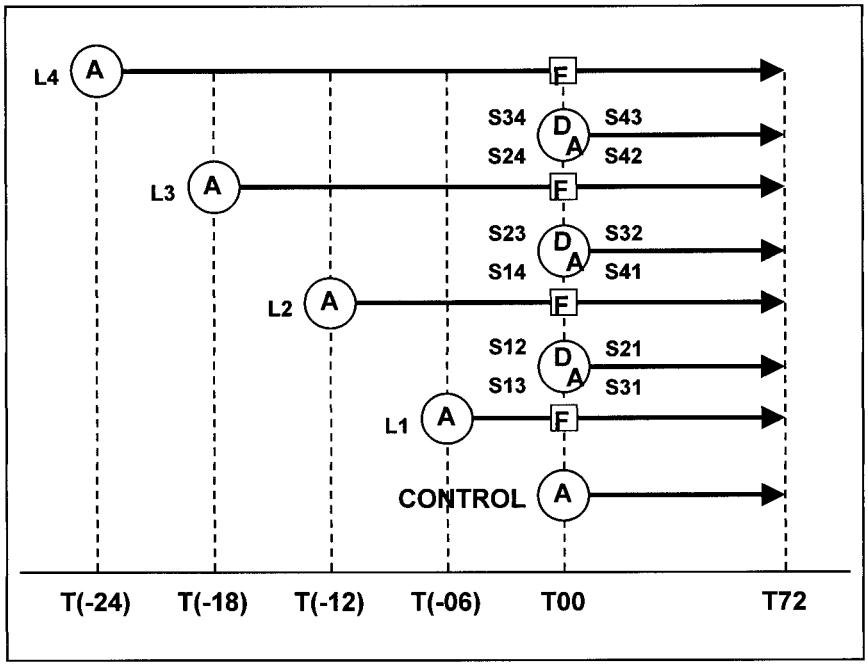

FIG. 2. Schematic diagram of the perturbation methodology LAFV. See text for the explanation of the symbols and labels.

In general, the effect of this blending scheme on the control forecasts is not great. Some fast bias in the previous control forecasts without blending can be reduced, as the blending weakens the winds in the outer-core region. Compared with the forecasts run directly from the analyses, the present ones have a more steady motion due to the improved representation of the vortex circulation (not shown).

The modified control forecasts for the other set of experiments, BETA, will be described in section 2c(3). The blending in those cases includes the asymmetric structure of the vortex as well [see section 2c(3) for details].

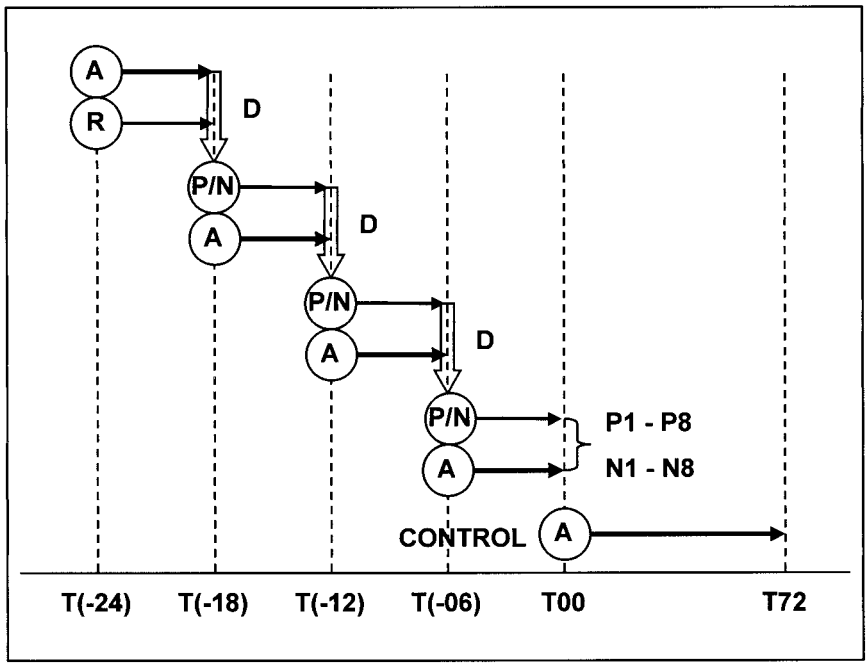

FIG. 3. As in Fig. 2 except for experiment BGMV.

2) INITIAL POSITION ERROR

One of the major difficulties in TC motion prediction is to identify the vortex location accurately. The TC position in an NWP analysis can deviate from the true location by 100 km or more (Curry et al. 1987; Chan and Kwok 1997). The application of satellite and radar imagery techniques in this decade greatly improves this deficiency. However, in real-time forecast situations, the initial position error can still be a serious source of error. This experiment thus tries to simulate the effect of changing the initial position of a synthetic vortex. The distribution of the IPER depends on a particular assimilation system, which is not included in our model. To obtain such information, 1322 cases from the U.K. Meteorological Office (UKMO) global model in the period 1991–95 are drawn to form a statistic. A binormal distribution is fitted to the IPER of these cases (Fig. 4). The standard deviation is found to be 133.7 km (112.4 km) in the zonal (meridional) direction. [The UKMO has adopted a new bogusing scheme since 1994 that can reduce the IPER drastically (Heming et al. 1995). However, as the main concern is the usefulness of perturbing the IPER only, data from the entire 1991–95 period are used. The conclusion of this experiment should not be changed.] The correlation coefficient between the zonal and meridional error is 0.14. Because the resolution of the UKMO global model is 2° latitude, which is about four times that of the present model, the standard deviations of the errors are reduced by a factor of 4 (i.e., values of 33.4 and 28.1 km are used) to take into account the higher resolution of our model. In addition, the mean position error in the statistics shown in Fig. 4 is set to zero because this is only a characteristic of the UKMO model. For each ensemble member, errors in the zonal and meridional directions are selected from the derived binormal distribution and then used to perturb the best-track position. As in MCFV where the nature of perturbations is purely random, an ensemble size of 50 is also adopted here for each of the nine TCs.

3) SYNTHETIC VORTEX WITH β GYRES AND PERSISTENCE

Numerical studies of TC motion in a barotropic framework (e.g., Fiorino and Elsberry 1989) have shown that an initially axisymmetric vortex on a β plane will develop in time an azimuthal wavenumber-1 asymmetry (the so-called β gyres). The flow associated with these gyres at the center of the TC (termed the ventilation flow) generally points toward the northwest in the Northern Hemisphere. Such asymmetries can occasionally be identified in the real atmosphere (Chan and Cheung 1995, 1998; Franklin et al. 1996).

The set of experiments BETA attempts to simulate uncertainties in the structure of the asymmetric flow. The basic methodology is to spin up an initially axisymmetric vortex in a quiescent environment in a high-resolution model for a period long enough for the β gyres to develop and become quasi-steady. The spun-up vortex is then blended into the analysis in the same way as the control forecasts in section 2c(1). The model used for generating the asymmetric vortex is the barotropic model of Chan and Williams (1987) that has a 20-km resolution with a domain of 4000 km^2. Based on the concept discussed by Kurihara et al. (1993) in the GFDL vortex specification scheme, the spinup time is specified as follows: if the TC at 18 h prior to the initial forecast time has not yet reached tropical storm intensity, an 18-h integration period is used; or if the central pressure depression at that time is less than half of the value at the initial time, rapid intensification is assumed and again the integration period is 18 h; for all other cases the integration period is 36 h. After a spun-up vortex is used in our model, the forecasts can generally be improved slightly (not shown).

To create the BETA ensemble forecasts, the axisymmetric part is first perturbed before initiating the spinup procedure by adding random Gaussian noise to the three basic properties of the axisymmetric tangential wind profile: the intensity [i.e., $v_{\max}$, the maximum sustained wind (MSW)], the radius of maximum wind $r_{\max}$, and the size parameter b (that determines how the wind speed decreases to the environmental value in the outer core). The values of the standard deviation of the Gaussian noise added to the three parameters (namely, $\delta v_{\max}$, $\delta r_{\max}$, and δb) are estimated as follows. Operationally, the intensity of a TC is generally estimated using the Dvorak technique. Information on the eye pattern, cloud imagery, and past intensity change are gathered and represented by a current intensity (CI) number. A change of CI by 1.0 corresponds to a change in MSW by 25 kt (12.5 m s^{-1}). The value of $\delta v_{\max}$ is thus estimated to be the possible error in specifying the CI, that is, 12.5 kt (6.25 m s^{-1}). To allow for the reduction of wind speed in a deep-layer mean (by a factor of 0.8, as in our determination of $v_{\max}$), the actual value of $\delta v_{\max}$ is taken to be 5 m s^{-1}. The value of $\delta r_{\max}$ is taken to be consistent with the resolution of the model (i.e., 25 km). The value of b in the control forecasts is calculated by specifying $v_t = 6$ m s^{-1} at $r = r_f$. If this boundary value of the tangential wind is altered by 1 m s^{-1}, the corresponding change in b will be about 0.1. Thus, δb is taken to be 0.1. Extreme values of $v_{\max}$, $r_{\max}$, and b are avoided by rejecting errors deviating from the original values by more than two standard deviations of their corresponding error distributions.

To go one step further, a forecast may be improved by adding an asymmetric flow to the vortex in the form of a persistence vector, which forces the cyclone to move initially in the correct direction. The vector is defined both in direction and magnitude as the past 6-h motion vector of the TC, and added to every grid point within the vortex domain. It is known that addition of such an asymmetry can improve a 24-h forecast, but does not persist for longer than 48 h [Chan and Kwok (1997) for the western North Pacific; the results for the Atlantic

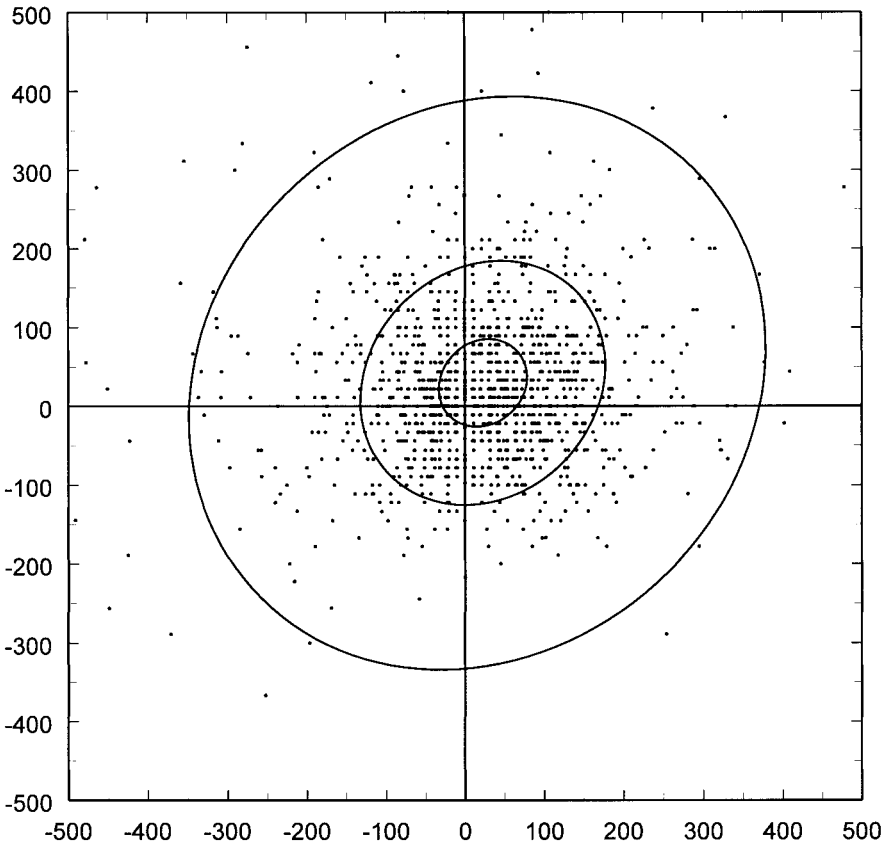

FIG. 4. Initial position errors (*dy* vs *dx* in km) of 1322 cases for the UKMO global model in the period 1991–95. The error ellipses are the 25%, 50%, and 90% probability contours of the fitted binormal distribution.

basin being slightly different as shown in Aberson and DeMaria (1994)]. If the persistence vector is used in the present model (with a spun-up vortex), the forecasts show improvement over those without the vector only for several TCs (Fig. 5). For example, the forecasts for the westward-moving Yancy, Dot, and Ed are much better due to the added persistence vector. However, for cases with rapid recurvature (e.g., Flo), a persistence may not be beneficial. On average, the skill is similar to that using just an axisymmetric vortex. The reason is probably that the combined effect of the β gyres and the persistence provides too large an estimation of the asymmetry in the vortex. Thus, in terms of a single forecast, the spinup procedure and persistence need not be applied at the same time. This combination of spun-up vortex and persistence is used in the ensemble experiments here because the perturbed components may provide a better sampling of the uncertainties in the asymmetry and thus some skillful members in the ensemble. To perturb the persistence vector, the uncertainty in the past 6-h position is used. This is changed by using the same positional error statistic as in experiment IPER. The perturbed position is then used to calculate the altered persistence vector.

Four experiments with different perturbations are performed (see Table 1). BETA(0) consists of perturbed spun-up vortices only. The β gyres and the persistence vector are then perturbed, respectively, in BETA(1) and BETA(2), with a control persistence also present in the former. In BETA(3), both kinds of asymmetry are perturbed simultaneously. In all four experiments the number of members is equal to that in the LAFV and BGMV (i.e., 17). The objective here is to seek an optimized perturbation scheme with both usefulness and efficiency. More implications from these experiments will be given when their relative performance is discussed.

The properties of all the experiments performed are summarized in Table 1. Note that the column describing the application of a synthetic vortex gives the configuration of the control forecast in each set of experiments.

3. Verification of the ensembles

a. Statistical measures and skill scores

The statistical measures and skill scores are identical to those in Part I. They are summarized here for further discussion. If F_i denotes a parameter of the ith ensemble member, the ensemble centroid or ensemble mean can be calculated as

$$\overline{F} = \frac{1}{N}\sum_{i=1}^{N} F_i, \tag{5}$$

where N is the total number of members.

Two measures of the error of an ensemble can be defined. One is the error of the ensemble mean E when verified by the corresponding parameter A in the analysis:

$$E = |\overline{F} - A|. \tag{6}$$

In our case of TC motion forecast, the quantity of primary interest is the location of the vortex center. Hence the quantities F_i and A are position vectors, and the norm (denoted by $|\cdot|$) in (6) is simply taken as the great-circle distance between two center positions. Another measure is the mean square error e^2:

$$e^2 = \frac{1}{N}\sum_{i=1}^{N} |F_i - A|^2. \tag{7}$$

The quantity used to measure the spread of an ensemble is the mean-square distance δ^2 among the members:

$$\delta^2 = \frac{1}{N(N-1)}\sum_{i,j=1}^{N} |F_i - F_j|^2. \tag{8}$$

Note that δ is independent of the verification. In a perfect model, δ should be a good estimation of the potential rms error e. This relation is used to study the spread–error (between δ and e) correlation of different perturbation schemes.

Two relative skill scores (RSS's) are calculated for the two types of verification: one under the perfect model assumption (PMA) and the other verified by the best tracks. In the former, one of the ensemble members is chosen as the ''perfect'' forecast for verifying the control and the mean forecast of the remaining members. As in Part I, an averaging is performed over all possible realizations (i.e., each member except the control is selected once as the verifying member). The purpose is to reduce the sensitivity on the verifying member chosen. The RSS calculated is

$$\mathrm{RSS}_{\mathrm{PMA}} = \frac{E_{\mathrm{control}} - E_{\mathrm{ensemble}}}{E_{\mathrm{control}} + E_{\mathrm{ensemble}}} \times 100\%, \tag{9}$$

where E_{control} is the error of the control run, E_{ensemble} the error of the ensemble mean, both verified by the chosen ''reality.'' A positive value of $\mathrm{RSS}_{\mathrm{PMA}}$ indicates an outperformance of the ensemble mean over the control. When verified by the best tracks, the score (termed $\mathrm{RSS}_{\mathrm{best}}$) to be calculated is the same as (9) except the errors are with respect to the best track.

b. Skill of the MCFV and IPER experiments

The ensembles for MCFV and IPER are verified directly by the best tracks. It is found that in the former experiment the response to the random noise added to the vortex is small, and the ensemble mean differs little from the control forecast. This is because the cyclonic shear in the vortex is very strong and regions of local vorticity maximum/minimum generated by the noise can easily be smoothed out. The rms size of the noise is varied between 1 and 3 m s^{-1}, and the response is still similar. As described in section 2b(2), some additional experiments are performed in which correlation is introduced among the random noise by removing some of the shortwave components using a specified characteristic length scale. It is found that when this length scale is small, both the skill and spread of the ensemble usually differ little from those for the uncorrelated errors. Further, the spread of the ensemble shows a decreasing trend when the characteristic length scale increases. A plausible reason is that the vorticity distribution is modified only slightly by the perturbations as their scales exceed the typical size of a TC vortex. Again, the ensemble mean is almost identical to the control. Therefore, it can be concluded performing such a perturbation methodology to the vortex is not useful.

The skill of the IPER ensembles is also found to be close to the control forecasts and thus the RSS calculated is not far from zero (Table 2). Sometimes, slight improvement can be obtained up to 24 h. Examples are the cases E91400 (TC Ed at 0000 UTC 14 September) and F91500. This is in accord with the result that a better short-term forecast can be produced by an improved initial position (Chan and Kwok 1997). However, this effect vanishes beyond 24 h and the RSS simply fluctuates about zero. Thus, an ensemble in which only the initial position is perturbed is not useful.

The conclusion from these two experiments is that, if no synthetic vortex is used, then purely stochastic perturbations cannot generate a more skillful ensemble mean than the control forecast (as in the MCFV). Even when a synthetic vortex is added, the initial position is not an essential system parameter to be perturbed.

c. Skill of the experiments LAFV, BGMV, and BETA

The remaining six sets of experiments: LAFV, BGMV, BETA(0)–BETA(3) are performed on all the 66 available cases. The ensembles are verified first under the PMA and then by the best tracks. The scores of the six sets of experiments are presented in this and the following subsections, while detailed discussions and implications of the results are given in the next section.

For some cases in LAFV and BGMV, it is found that when perturbations are added to the vortex in the analysis, it is not possible to identify the vorticity center in the model integrations. The number of these ''failure'' cases is 33 for both LAFV and BGMV (but not necessarily the same cases). These include not only weak TCs but also some strong ones. For consistency, these failure cases are neglected in the verification. Therefore, the scores presented for these two experiments are based on the remaining 33 cases, respectively. Diagnostics and implications for the failure cases will be given in section 4.

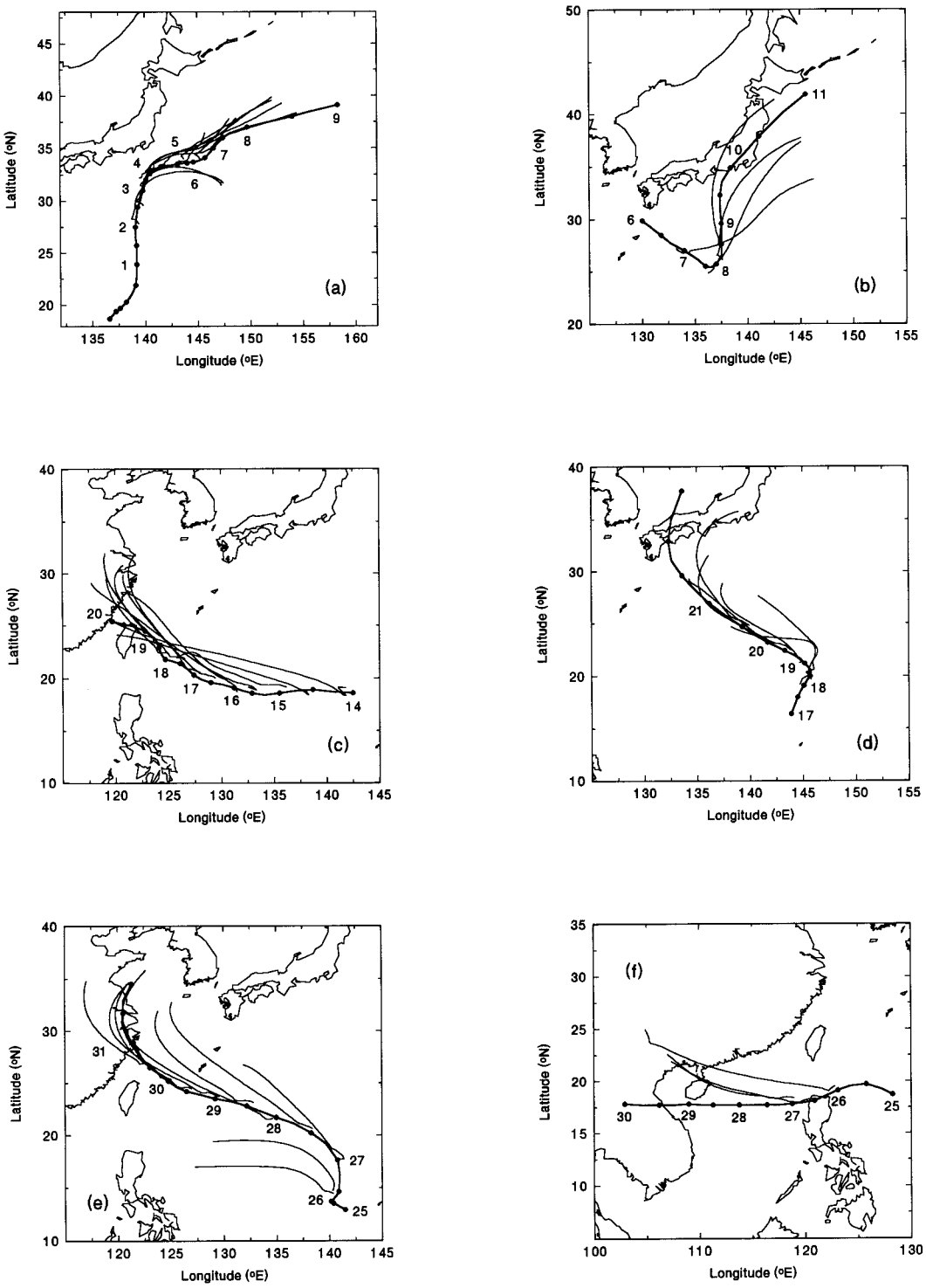

FIG. 5. As in Fig. 1 except a spun-up vortex and persistence vector are added in the control forecasts.

Under PMA, the performance of LAFV and BGMV in terms of the range and temporal evolution are quite similar (Figs. 6a and 6b). Most cases can attain a positive score up to a lead time of 72 h, and the median score for BGMV is generally higher than that of LAFV for most of the forecast times (Table 3). This is also true for the percentage of cases with positive score. This agrees with the result in Part I that the breeding algorithm can provide better samples from the analysis errors than the simple lagged-average approach.

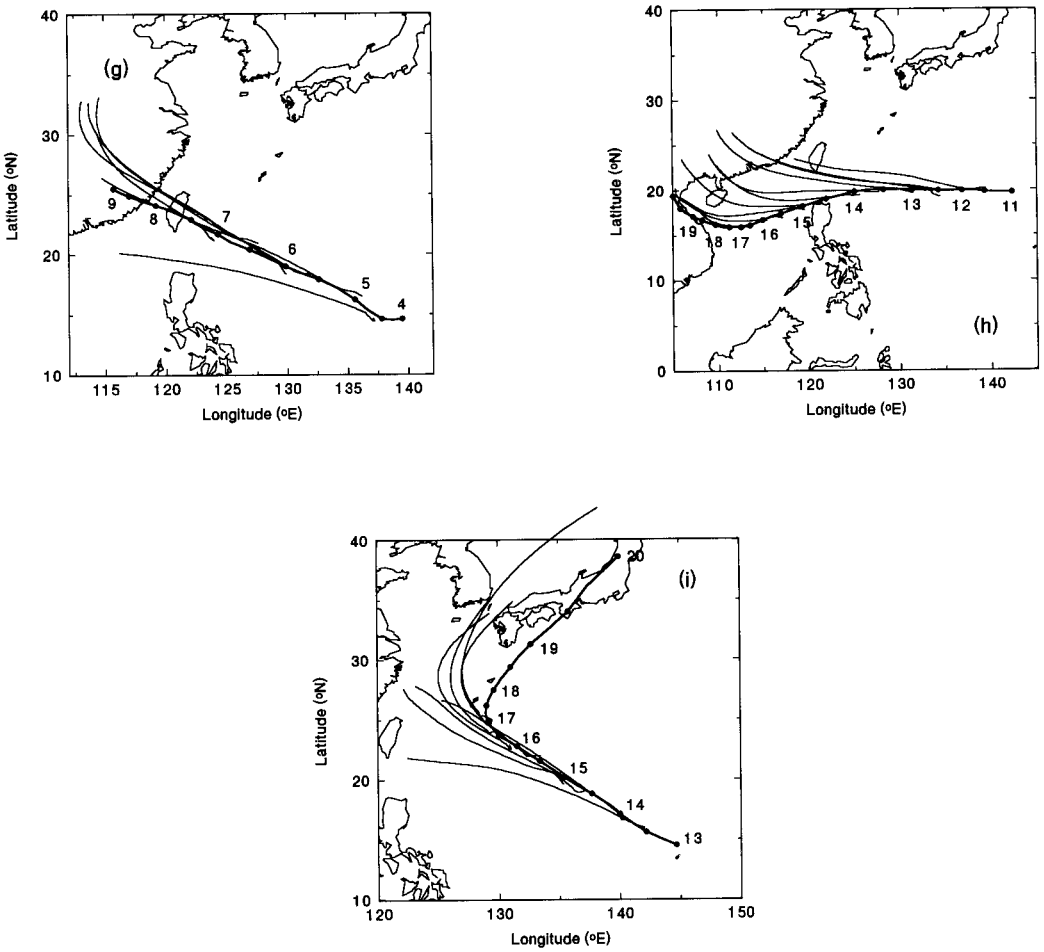

FIG. 5. (*Continued*)

Among the BETA experiments, the outcomes for BETA(0) and BETA(1) are better than for the other two under PMA verification (Figs. 6c and 6d). Similar to the performance of LAFV and BGMV, the median scores at all times can be positive and most cases lie above the zero line (Table 4, first two columns). This reconfirms the β gyres as an important factor in determining the motion of a TC. In other words, if the dynamics governing a TC is purely barotropic and the uncertainties in the vortex asymmetry are well specified, an ensemble approach can be used to improve the forecast provided that the large-scale flow is accurately mod-

TABLE 1. Summary of the eight sets of experiments performed in this study.

Expt	No. of members	Synthetic vortex applied	Nature of perturbation	Components perturbed
MCFV	50	No	Random	Vortex flow
LAFV	17	No	Dynamical	Lagged-error + short-range forecast difference within vortex as perturbation
BGMV	17	No	Dynamical	Bred short-range forecast error within vortex as perturbation
IPER	50	Axisymmetric vortex	Random	Initial position
BETA(0)	17	Spun-up vortex	Random	Parameters to generate the β gyres
BETA(1)	17	Spun-up vortex + persistence	Random	Parameters to generate the β gyres
BETA(2)	17	Spun-up vortex + persistence	Random	Persistence vector
BETA(3)	17	Spun-up vortex + persistence	Random	Parameters to generate the β gyres + persistence

TABLE 2. Relative skill scores (RSS_{best}) for the nine cases performed in IPER.

Lead time (h)	V80200	W80700	Y81500	Z81800	A82600	B82600	D90600	E91400	F91500
12	−13.3	−0.4	0.5	−0.1	0.6	0.1	−0.2	21.1	2.7
24	−4.0	−0.2	0.2	0.0	0.0	−0.2	−0.1	−0.2	0.5
36	−1.9	−0.9	0.4	0.0	−0.2	0.0	−0.1	0.1	−0.5
48	−1.5	−0.8	0.3	−0.1	0.0	0.1	0.2	−0.1	−0.5
60	−1.5	−0.9	0.3	0.1	−0.2	0.0	0.1	−0.1	0.3
72	−1.8	−1.1	0.5	0.0	0.1	−0.1	0.0	0.0	0.3

eled. When the persistence vector is perturbed in BETA(2), the ensemble always spreads symmetrically with respect to the control. The consequence is that the ensemble mean is close to the control forecast for all cases (as in IPER examined earlier), and thus the RSS's are all close to zero (not shown). The median scores usually fluctuate about a small value and few cases can attain a positive score (Table 4, third column). The effect of perturbing the β gyres and the persistence vector simultaneously [BETA(3)] is not great, and the score distribution remains about the same as in BETA(2) (Table 4, fourth column). In general, the magnitude of the persistence vector should be greater than that of the propagation effect of the β gyres. Therefore, it is be-

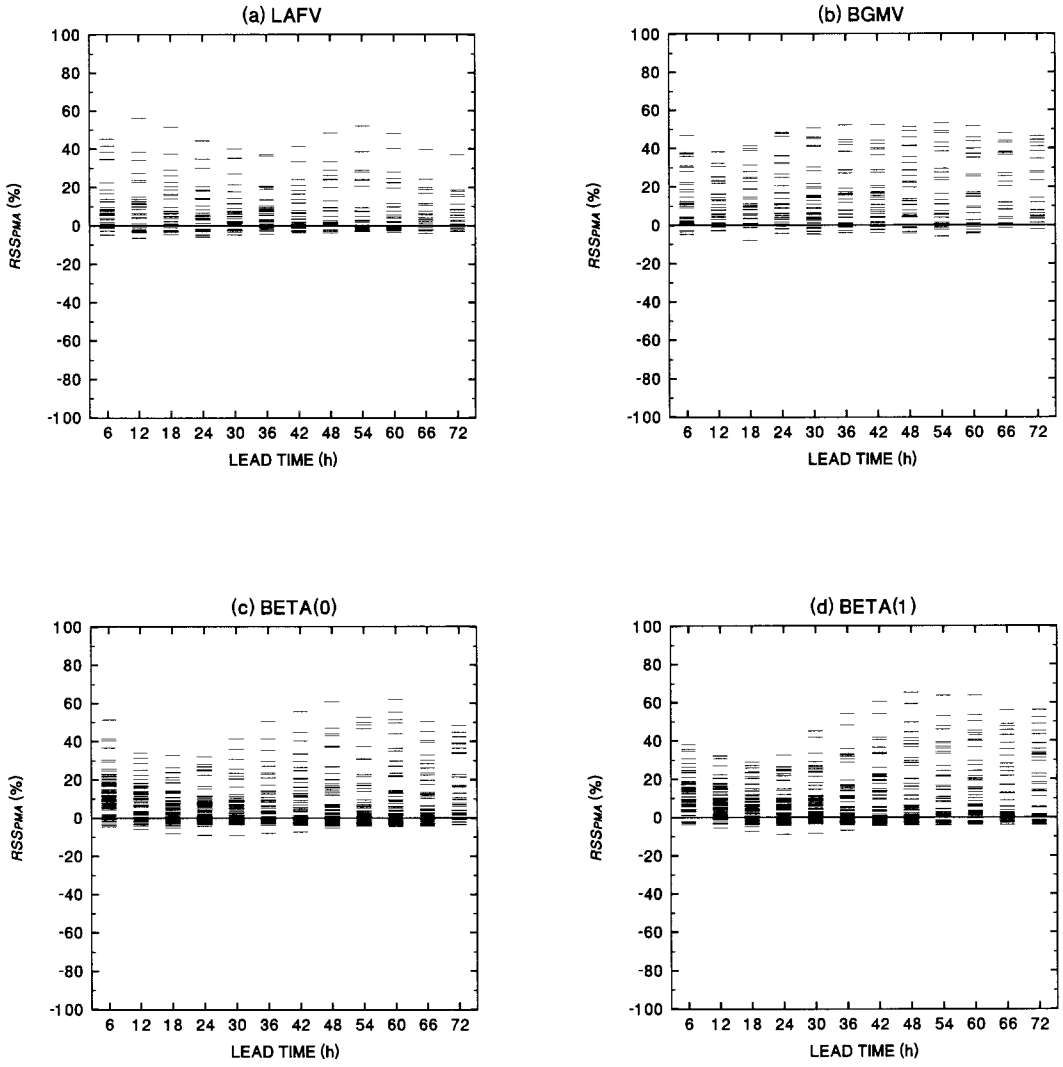

FIG. 6. Relative skill score under PMA of all the cases and lead times for experiments (a) LAFV, (b) BGMV, (c) BETA(0), and (d) BETA(1). Each horizontal bar represents one case.

TABLE 3. Median values of RSS_{PMA} for the two experiments without a synthetic vortex, LAFV and BGMV. The values in parentheses are the percentage of positive scores at a particular time.

Time (h)	LAFV	BGMV
6	6.6 (76)	10.6 (91)
12	8.7 (73)	8.2 (82)
18	5.1 (79)	9.8 (94)
24	4.1 (70)	9.4 (85)
30	3.6 (76)	9.3 (85)
36	5.5 (73)	7.9 (88)
42	3.8 (76)	13.1 (82)
48	2.0 (67)	12.4 (82)
54	2.0 (61)	11.0 (82)
60	2.3 (57)	14.8 (79)
66	3.1 (65)	11.8 (87)
72	4.2 (74)	11.8 (91)

TABLE 4. As in Table 3 except for the set of experiments BETA.

Time (h)	BETA(0)	BETA(1)	BETA(2)	BETA(3)
6	11.0 (85)	9.2 (88)	1.5 (61)	4.4 (79)
12	6.4 (76)	6.3 (77)	−2.1 (23)	−0.1 (47)
18	4.7 (71)	4.7 (70)	−1.9 (20)	−1.5 (38)
24	4.0 (67)	4.1 (68)	−2.1 (18)	−1.6 (33)
30	3.4 (64)	3.6 (68)	−2.0 (24)	−1.9 (36)
36	2.6 (61)	2.9 (65)	−1.9 (26)	−1.6 (36)
42	1.5 (61)	3.3 (61)	−1.5 (35)	−1.4 (36)
48	2.5 (67)	3.0 (67)	−0.6 (45)	−0.9 (42)
54	2.5 (66)	2.9 (68)	−1.0 (52)	−0.2 (45)
60	4.2 (71)	6.2 (73)	1.4 (55)	0.4 (54)
66	4.8 (76)	5.9 (76)	3.3 (61)	1.1 (61)
72	7.7 (76)	13.4 (73)	3.9 (71)	4.1 (68)

lieved that the contribution from the latter has possibly been overwhelmed by the former in this experiment. The generally low usefulness of perturbing the persistence vector does not contrast our previous understanding of adding such a component. Since the persistence vector is perturbed in a random manner here (in both magnitude and direction), the vortex must be forced to move initially in the direction of the perturbed persistence vector. Therefore, the ensemble tracks should be expected to spread symmetrically about the control when the forecast starts. When the persistence effect fades out, the behavior of the tracks is not observed to have any substantial change. Nevertheless, the persistence vector is designed using the actual past motion and its role in BETA(3) will be discussed in section 4.

When the verification is by the best tracks, the scores for the experiments depicted earlier in Fig. 6 degrade substantially. In LAFV and BGMV, spread of score among the cases is large during the first 24 h (Figs. 7a and 7b). There are two possible reasons for this. The first one is that when the errors are still small, the score is more sensitive to their changes. Another probable reason is that this is the period when the vortex and also the environmental flow are adjusting to the perturbations, and the effect on different cases can vary to a large extent. When this adjustment period is over, the spread of scores among the cases can be seen to decrease. The general performance of LAFV and BGMV are again similar, and the median scores of BGMV show a slightly increasing trend (Table 5). After 48 h, it can attain a positive median score (though the magnitude is only 1%–2%), and in more than half the cases improvement can be obtained. This indicates that though the large-scale environmental flow (which largely determines the motion) remains unchanged, on average the addition of vortex perturbations (such as the breeding method in BGMV) can possibly improve a forecast by the end of a 72-h period.

For the BETA experiments, the scores decrease substantially so that the performances of the four are about the same. The situations for BETA(0) and BETA(1) are shown in Figs. 7c and 7d, respectively. Similar to that in LAFV and BGMV, the case sensitivity (spread among the cases) is also large in the first 12 h and then decreases, indicating the environment also needs a certain period to adapt to the spun-up vortex. However, the overall low skill of BETA(0) and BETA(1) (Table 6 gives the median scores) suggests that this kind of perturbation (changing the asymmetric flow in the synthetic vortex) is not quite satisfactory in indicating the essential uncertainties in the specification of a TC vortex. [The low skill of BETA(2) is to be expected from its PMA verification.] At the least, they should not be better than the dynamical approach in LAFV and BGMV. BETA(3) shows a slight trend of improvement by the end of the forecast. This is probably because the variability due to the combined modified large-scale flow (represented by the persistence vector) and β gyres is larger than when only one of them is changed.

To conclude this subsection, changing the persistence vector alone has been eliminated as a useful means of perturbation. For a barotropic flow, perturbing the asymmetric flow (β gyres) has potential skill if the large-scale flow is predicted accurately. However, this may not work in a realistic situation. On the other hand, dynamically constrained perturbations are still the more promising methodology. Therefore, the following discussion will focus on the LAFV, BGMV, and the three BETA experiments [except BETA(2)]. Further, three points should be stressed. First, the PMA score is a test of the potential of the perturbation methodology, and the absolute skill cannot reflect the applicability of the experiments when actual forecasts are to be made. Second, because different control forecasts are used for the experiments, conclusions can be drawn on the performance of an individual experiment but their intercomparison needs further analyses to take into account the slightly different skill of the three sets of control forecast (directly from the analysis, addition of a synthetic vortex with and without persistence). This is true if the LAFV/BGMV is to be compared with the BETA experiments, or BETA(0) is to be compared with BETA(1)/BETA(3). Finally, it should be cautioned that the conclusions drawn here may not necessarily be extended to full baroclinic models because even if each horizontal layer is

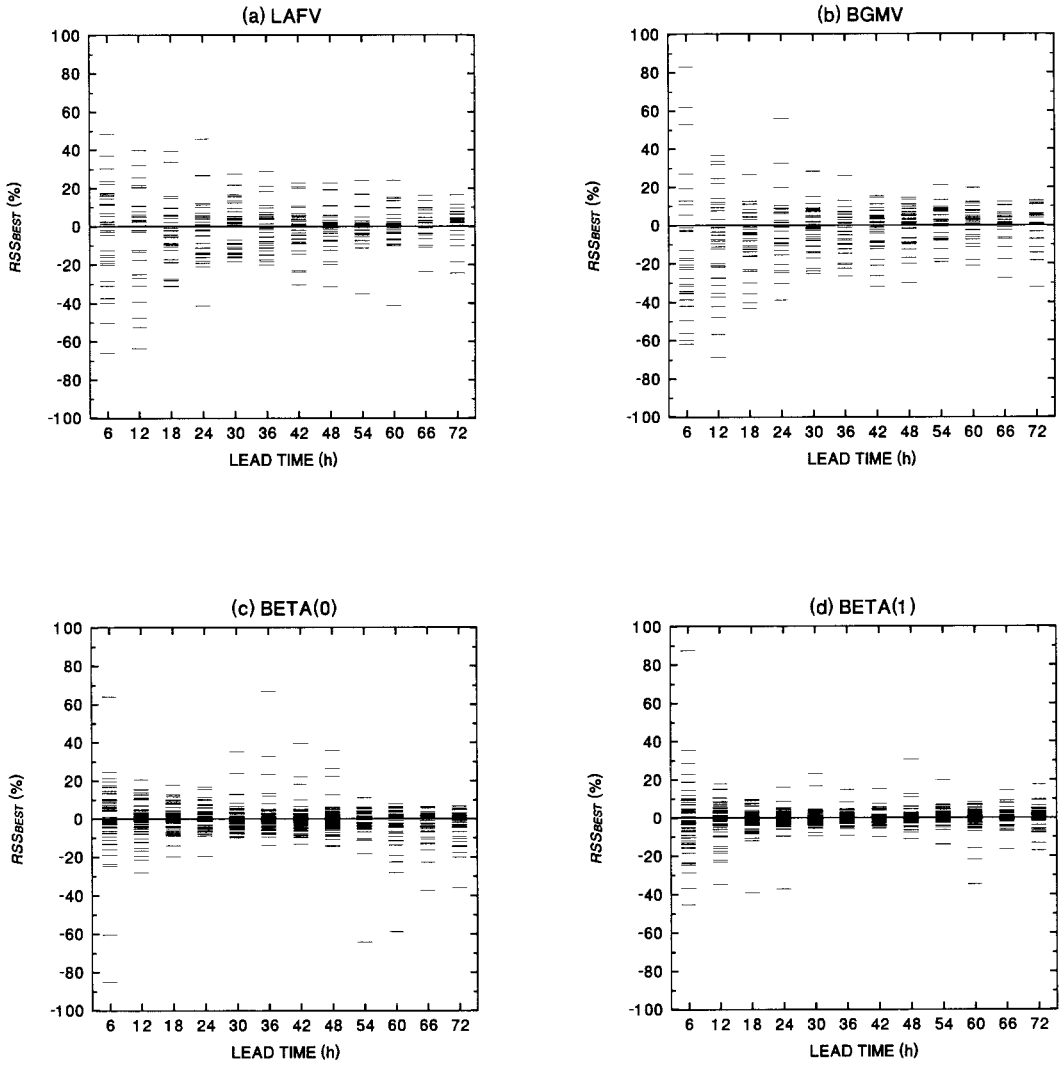

FIG. 7. As in Fig. 6 except verification is by the best track.

treated as a barotropic flow, vertical coupling between one another will probably alter the results.

d. Spread–error relationship

As in Part I, the correlation between the spread and the error measures of the ensembles is studied. The background idea is that the degree of spread among the members in an ensemble is a measure of the response with respect to the perturbation, and a situation with smaller spread (i.e., the ensemble members agree well with one another) may result in a more skillful forecast. Thus if the perturbation methodology can effectively simulate the underlying statistics of the error sources of the system (from observations and model), the spread and error should be related.

Under PMA, values of the spread (as measured by the quantity δ) and the rms error e from a single realization selected randomly are used because the averaging over realizations will automatically make the two quantities almost identical. The correlation between them is high for all four of the selected experiments [LAFV, BGMV, BETA(0), and BETA(1); see Fig. 8 where δ and e are scaled by their respective standard deviations with respect to the 66 cases to give a better illustration of the relationship]. The lowest value of the correlation coefficient, ρ, is about 0.6 and the highest one is 0.94. This high correlation can be extended out to 72 h. Therefore, in terms of the ability to indicate the predictability of a forecast using the response to perturbations (when no model deficiency is present), these four methodologies work well. A detailed comparison among them suggests that LAFV and BGMV can be said to be better than the two BETA experiments because the relationship between spread and potential error size is clearer and they can indicate some large errors correctly by the end of the 72-h forecast.

When the verification is by the best tracks, the spread–

TABLE 5. As in Table 3 except the scores are RSS_{best}.

Time (h)	LAFV	BGMV
6	2.3 (55)	−17.2 (24)
12	2.2 (55)	−8.1 (33)
18	−4.5 (39)	−4.8 (36)
24	−1.8 (55)	−3.2 (39)
30	1.5 (58)	−1.5 (45)
36	−0.9 (48)	−0.3 (45)
42	0.3 (52)	1.1 (52)
48	−0.4 (48)	1.6 (58)
54	−0.5 (43)	1.3 (61)
60	0.5 (54)	1.8 (61)
66	1.6 (61)	1.8 (65)
72	2.2 (61)	1.2 (65)

TABLE 6. As in Table 4 except the scores are RSS_{best}.

Time (h)	BETA(0)	BETA(1)	BETA(2)	BETA(3)
6	−0.6 (48)	−2.1 (39)	−2.6 (33)	−2.7 (41)
12	1.0 (56)	−0.7 (44)	−1.0 (41)	−1.2 (41)
18	0.3 (53)	−1.4 (39)	−0.5 (42)	−1.2 (38)
24	0.2 (53)	−0.9 (44)	−0.4 (44)	−1.0 (38)
30	−0.1 (48)	−0.7 (39)	−0.3 (47)	−0.6 (39)
36	−0.2 (48)	0.0 (45)	−0.4 (45)	−0.5 (42)
42	−0.3 (47)	0.1 (52)	−0.2 (44)	−0.2 (47)
48	−0.6 (44)	0.1 (52)	−0.1 (48)	0.3 (55)
54	−0.4 (48)	0.2 (54)	0.4 (54)	1.0 (57)
60	−0.2 (46)	0.2 (54)	0.2 (52)	1.3 (59)
66	−1.0 (41)	0.4 (54)	0.1 (51)	1.5 (59)
72	−1.2 (41)	0.3 (56)	0.0 (49)	2.0 (59)

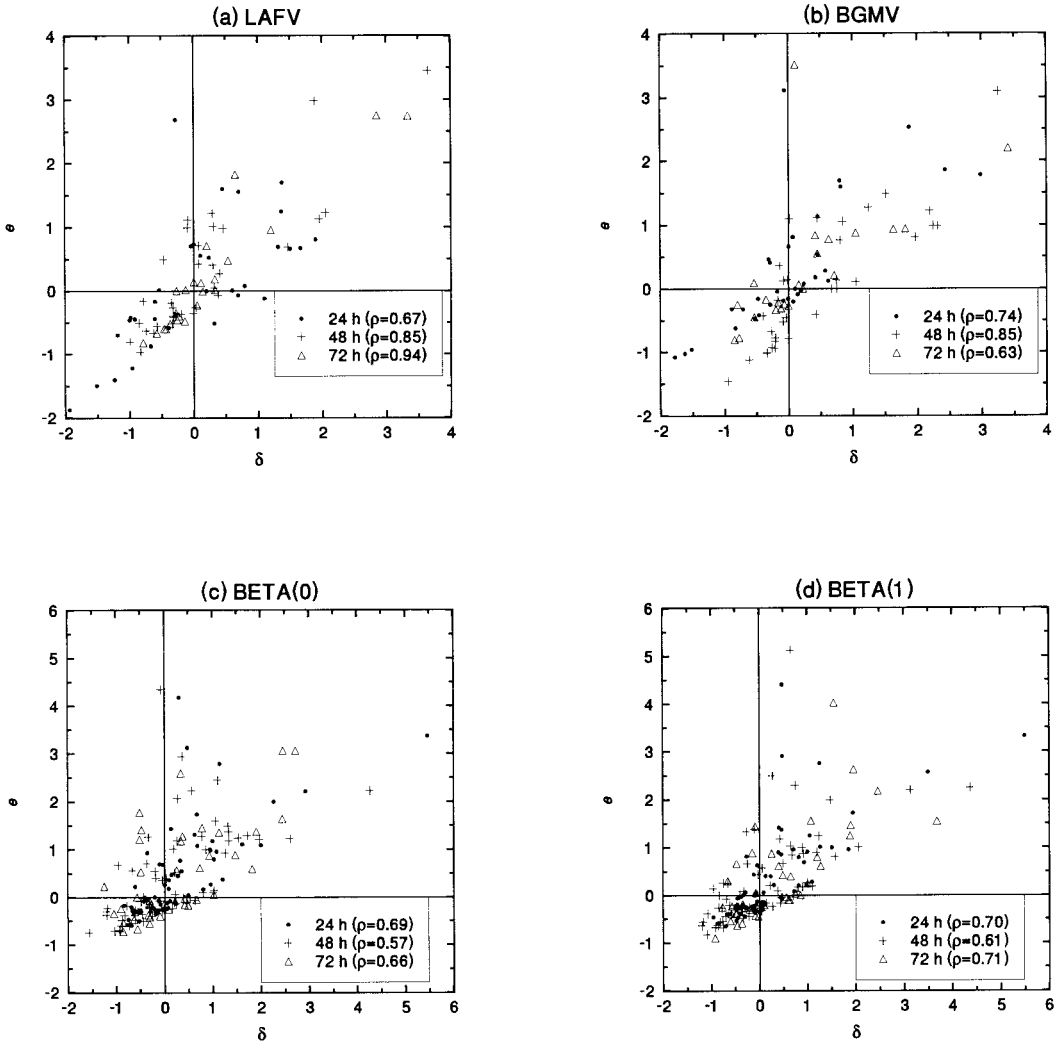

FIG. 8. Scatter diagram of the δ–e pair under PMA for experiments (a) LAFV, (b) BGMV, (c) BETA(0), and (d) BETA(1) at three lead times of 24, 48, and 72 h. The values plotted are deviations from the respective median. The linear correlation coefficient ρ for each time is also given.

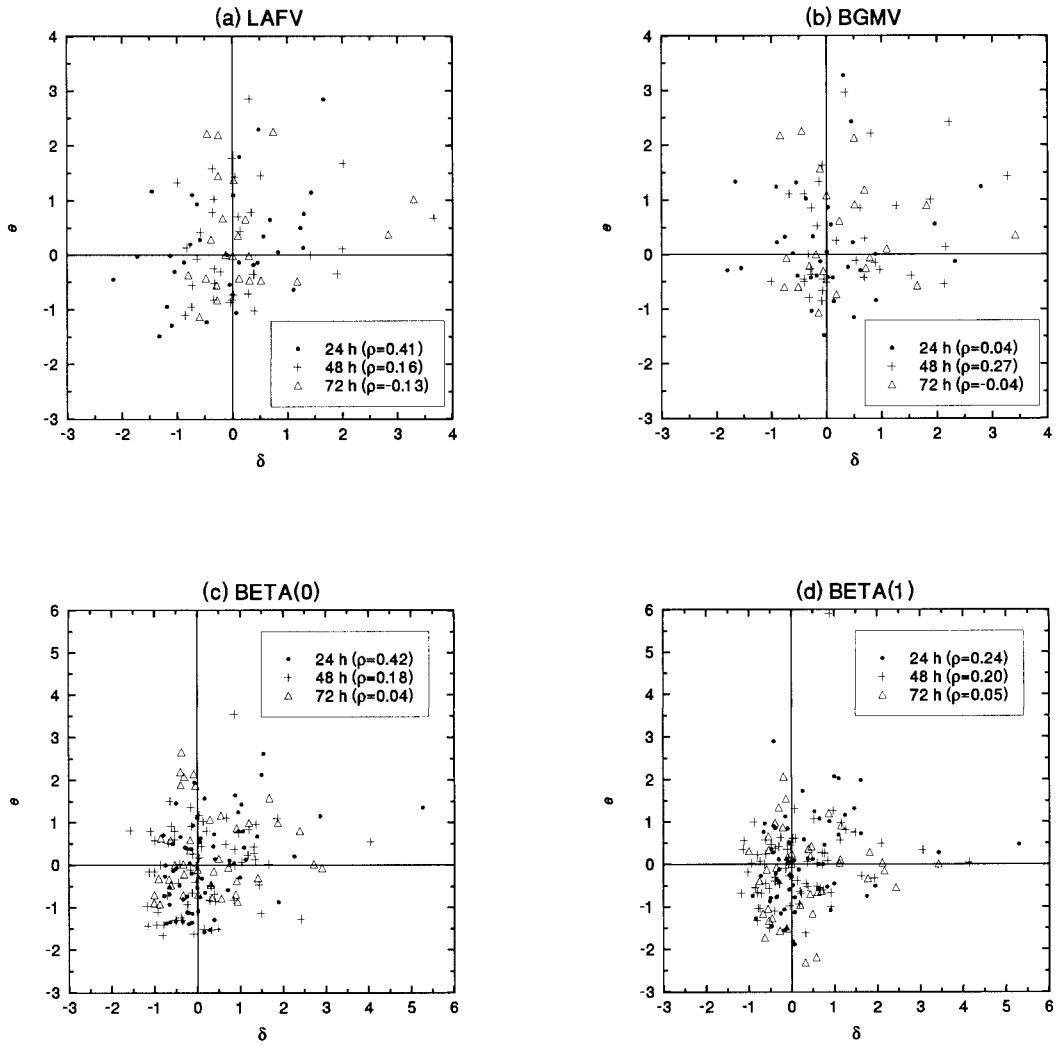

FIG. 9. As in Fig. 8 except verification is by the best track.

skill relationship (again scaled by their respective standard deviations as in Fig. 8) is not as clear (Fig. 9). For each of the four experiments considered here, the spread of the ensemble is too small compared with the corresponding value of the rms error. The values of the correlation coefficient, ρ, are much lower than those in Part I. This may indeed be expected because the possible bias in the forecast of the unperturbed environmental flow is an important source of error. Not treating the variability of errors from this source will certainly underestimate the final forecast error.

To summarize, the underlying idea of forecasting forecast skill using ensemble spread is validated here for TC motion ensembles when perturbations are added only to the vortex circulations and no model deficiencies are present. Also, all four perturbation methodologies considered have similar abilities to indicate the rms error. However, in a realistic situation, it appears to be more practical to use the ensembles of Part I because the error in forecasting the large-scale flow can dominate the total error.

4. Further analyses of the results

a. Failure cases in LAFV and BGMV

Half of the cases in both LAFV and BGMV fail during the integration because the perturbed vortex cannot provide a clear vorticity center (refer to section 3c), which may be due to two possible reasons. One is that the original representation of the vortex circulation in the analysis is already unrealistic, and the other is the distortion of the original circulation by the perturbation. An example of the former is the case W80800 in LAFV. The cyclonic circulation in the original analyzed vortex is highly asymmetric (Fig. 10a). The flow to the east and north is much stronger than that to the south. In one of the members (S13), it can be seen that a small

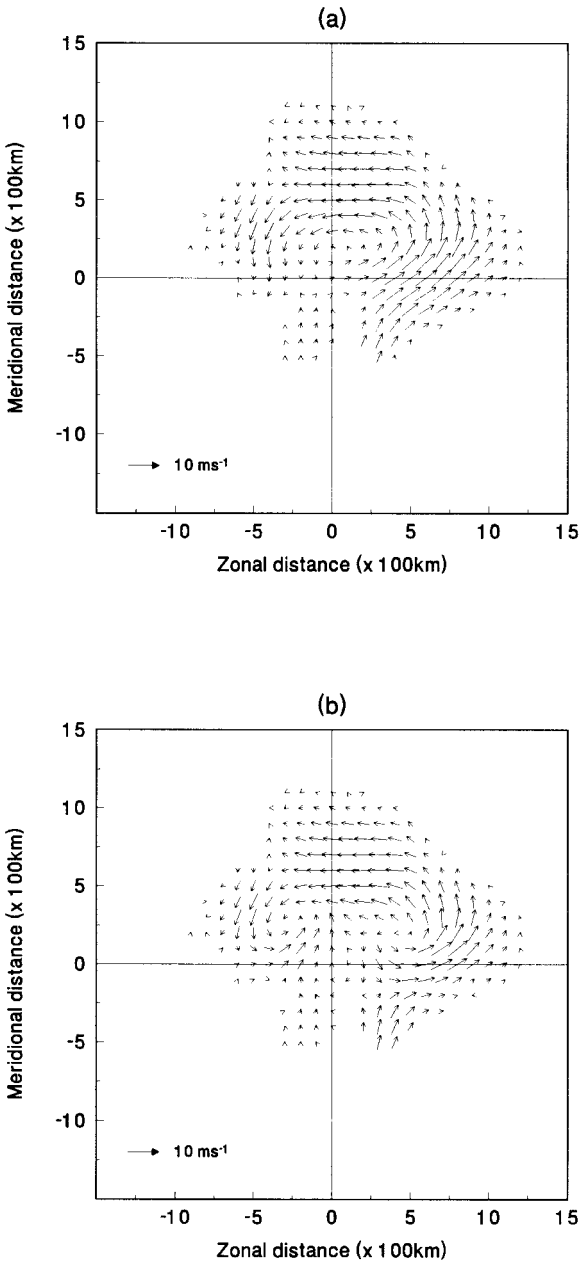

FIG. 10. (a) Analyzed vortex from the original analysis for the case W80800 (Winona at 0000 UTC 8 August) and (b) one of the perturbed vortices (S13) from the LAFV ensemble for the same case.

cyclonic center is generated at the western side and another one at the eastern side (Fig. 10b). Thus the initial vorticity center is expected to be shifted and subsequently the two subcenters will probably further evolve so that a definite vorticity center cannot be identified. Sometimes those ''unqualified'' members can be distinguished from the normal ones by examining their azimuthally averaged tangential wind profiles. In the same case W80800 (Fig. 11), the ''normal'' cases (denoted by solid lines) possess profiles with radii of maximum wind (RMWs) closer to the center but the failed ones (denoted by dashed lines) are seen to have shifted RMWs. Some of them even have anticyclonic flow near the center. In general, the SL members in the LAFV experiments usually behave well but problems similar to that illustrated by the above example frequently occur among the SRFD members.

For the BGMV cases, the contrast between the normal members and the failed ones is not as clear as in the LAFV, probably because of the random nature of the perturbation when starting a breeding cycle. However, some perturbed vortex circulation can be readily identified to be unrealistic. Many of them occur for the TCs Yancy, Zola, Abe, and Becky (Fig. 12 gives two examples). Either the cyclonic flow is incomplete and/or a smaller ''subvortex'' is generated.

What conclusions can be drawn from these failed cases in the LAFV and BGMV? If the control forecasts of these cases are examined (Fig. 1), it is found that their skill is generally low compared with the others. The reason is obvious when the circulation of the original analyzed vortex at the initial time is identified to be unrealistic. On the other hand, if the vortex circulation can be distorted easily by the perturbations, we can conclude that its sensitivity to errors (from both the analyses and model deficiencies) is large. This also implies that the possible skill of the forecast is low. Thus, the presence of these failed cases can help identify forecasts that are likely to be unskillful. Some irregular forecast tracks can easily be recognized, but the perturbation methodologies here provide a dynamical basis for their identification.

b. Scaled measure of forecast skill

In section 3c an overview of the performance of the experiments based on the verification results is provided, but a quantitative comparison of the RSSs between one method and another is not fair because different control forecasts were used in these experiments. If the forecast errors of the two control forecasts with a synthetic vortex (spun-up vortex with or without persistence, described qualitatively in section 2c) are scaled by that of the control forecast run directly from the analyses (described qualitatively in section 2b), their relative performance can be depicted clearly. The time series of the median values of these scaled relative errors (Fig. 13) shows that adding a synthetic vortex can reduce the error substantially in the first 24 h, but afterward the skill is about the same as that without one. Further, the addition of persistence to a spun-up vortex makes the forecast worse, a phenomenon already discussed in section 2c. Nevertheless, when no persistence is considered, a spinning-up process is beneficial and able to yield a better forecast than using just an axisymmetric vortex (skill not shown in Fig. 13) throughout the forecast times.

To remove the possible bias in comparing different

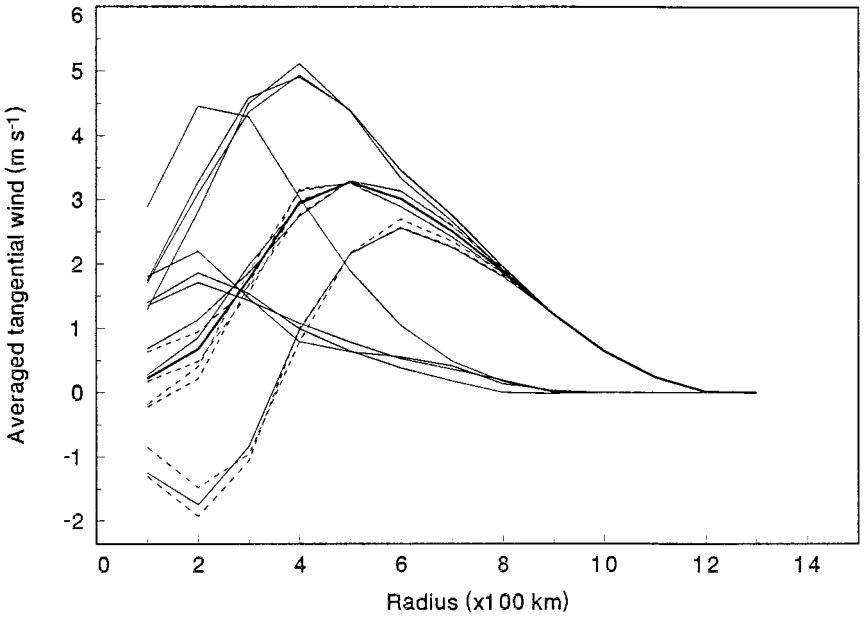

FIG. 11. Azimuthally averaged tangential wind profiles of the LAFV ensemble members for case W80800. The thick solid line indicates the control vortex profile. The thin solid lines and dashed ones are for the SL and SRFD members, respectively.

experiments that have different control forecasts, another quantity is derived here to measure their respective skill. The basic idea is that if the control forecast is already a good one, it is more difficult to improve the forecast further. Therefore, a higher rating should be assigned to an improvement of the ensemble mean over the control forecast. On the other hand, if the control is a poor one so that improvement is easily accomplished, a lower rating should be given. Note that a positive RSS has already taken this into account because the denominator in its definition involves the control forecast error. For example, for an E_{ensemble} of 50 km, it receives an RSS of 50% if the E_{control} is 150 km, but only 33% for an E_{control} of 100 km. However, the above arguments have to be reversed if the ensemble mean is a degradation instead of an improvement and the RSS fails to include this situation. The new quantity to be examined is not the RSS in (9) because it is already nonlinear and not suitable for further scaling. Rather, the error difference (ED) between that of the ensemble mean and the control forecast (i.e., the numerator of the RSS), scaled by a correction factor (CF) depicting the above arguments about the quality of the control forecast, is used [termed scaled error difference (SED)]:

$$\text{SED} = (E_{\text{control}} - E_{\text{ensemble}}) \times \text{CF} = \text{ED} \times \text{CF}. \quad (10)$$

The CF is simply taken as the ratio of a certain reference error (termed E_{ref}) and the control error if the ED is positive and the reciprocal of it if the ED is negative. That is,

$$\text{CF} = \begin{cases} \dfrac{E_{\text{ref}}}{E_{\text{control}}} & (\text{ED} > 0) \\ \dfrac{E_{\text{control}}}{E_{\text{ref}}} & (\text{ED} < 0). \end{cases} \quad (11)$$

c. Intercomparison of the BETA experiments

When the three BETA experiments under the best-track verification are compared [with BETA(2) neglected], the control forecast errors with persistence [those used in BETA(1) and BETA(3)] are used as the reference (with all the 66 cases) and those for BETA(0) are scaled. Their general relative performances as measured by SED are not far from one another for the first 42 h but begins to diverge afterward (Fig. 14). The median value of BETA(0) becomes negative, as presented in the previous subsection. That of BETA(1) becomes positive after 36 h although the value of SED remains small. The increasing trend of BETA(3) is more apparent and by 72 h the median value is about 25 km. Note that this trend is actually similar to their median scores (Table 6). However, greater confidence can be given to the present result based on SED because the differential performance of the control forecasts has been taken into account.

It is interesting to note that, in the first 36 h, the median SED of BETA(0) is close to or slightly higher than that of BETA(1). After 36 h, the situation reverses and BETA(1) takes over. The only difference between the two experiments is the presence of a persistence

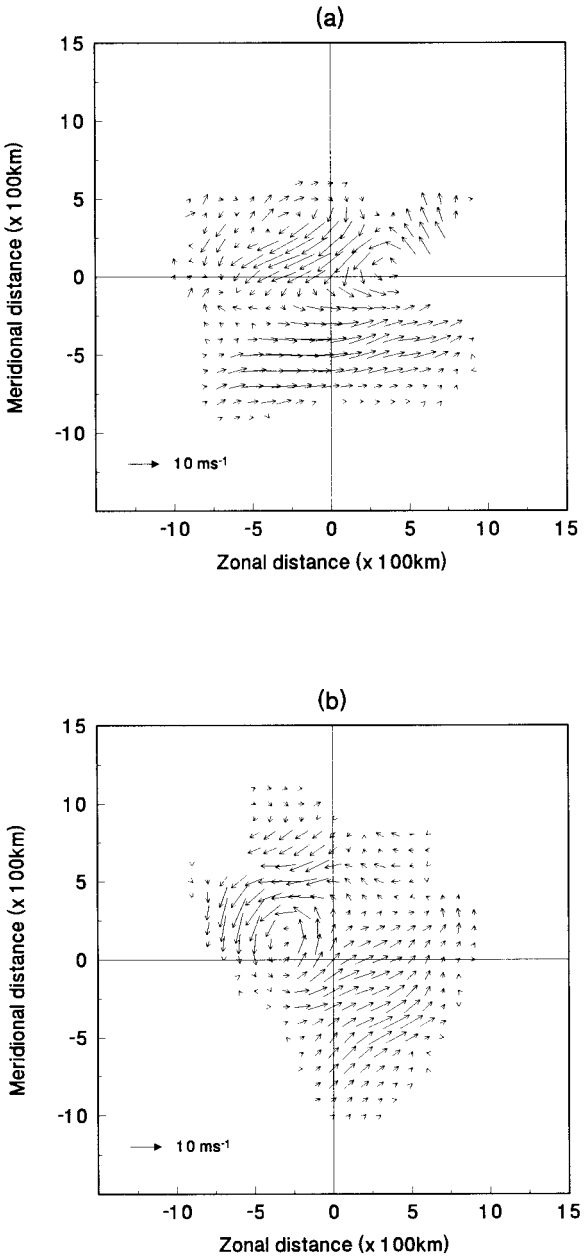

FIG. 12. As in Fig. 10 except for members (a) N06 in Y81612 and (b) B06 in B82700 from the BGMV ensembles.

vector in every member of BETA(1) (including the control). A persistence vector can be thought of as providing an appropriate steering flow (as inferred from the past motion) to the vortex. In a single model run, such a persistence vector can improve the short-term forecast by imposing the right direction of motion at the beginning (section 2c). Therefore, it is likely that because of the addition of the control persistence vector in all the BETA(1) members, the motion of the vortices is determined largely by the persistence. Comparatively, the ensemble effect of perturbing the β gyres in the synthetic vortex is overwhelmed in the first 36 h. Afterward, the persistence effect fades out and the skill of the ensemble averaging is reflected. Therefore, the median score of BETA(1) gradually increases to a higher level. The reason for it being able to outperform BETA(0) is, however, not very clear. One possibility is that the interaction between the perturbed synthetic vortex and the imposed persistence vector (representing a better steering flow) can provide more skillful forecasts among the ensemble members.

Finally, a point is to be addressed concerning BETA(3). If only the persistence vector is perturbed as in BETA(2), the potential skill has been shown to be low. However, the skill is increased in BETA(3) when the variability of the asymmetric flow within the vortex is included. Therefore, the role of the perturbations added to the persistence vector as providing alternative (and possibly true) steering flows cannot work by itself (due to the symmetric stochastic nature of the method), but can be a useful background for another way to change the vortex structure. In other words, if the large-scale flow in the analysis is not modeled well in a case but can be corrected by a perturbed persistence vector at the initial time, the spun-up vortex in some of the members will have greater probability to follow the right track and thus the ensemble mean forecast can be improved. One further implication is the possible potential of combining the vortex perturbations with the environmental ones. The breeding of a growing modes experiment in Part I has been found to be a better methodology for perturbing the environmental flow than a pure Monte Carlo method and lagged-average approach. It has good ability to explore different possible scenarios of the motion, and the degree of spread–skill correlation allows an estimation of the predictability of individual case. Thus, it can be a candidate for combining with the BETA vortices here. If all combinations are used (17×17), the total number of possible members is huge. Further experiments are needed to find out the optimal way of combination, which will be reported in a later study.

d. LAFV and BGMV versus BETA

The experiment BETA(0) is first selected to be compared with the two without synthetic vortices, LAFV and BGMV, because the other three experiments in the series [BETA(1)–BETA(3)] involve the additional effect of persistence vector. When LAFV and BGMV are to be compared with BETA(0), the control errors for the former two are used as reference. That is, the EDs of the LAFV and BGMV are unchanged while those for BETA(0) are scaled according to (10). Since the 33 ''normal'' cases in LAFV and BGMV are different, only the corresponding ones in BETA(0) are chosen for comparison. The variation of the median value of SED for LAFV and BETA(0) is similar in the first 36 h (Fig. 15a) although the fluctuation in LAFV is larger. The

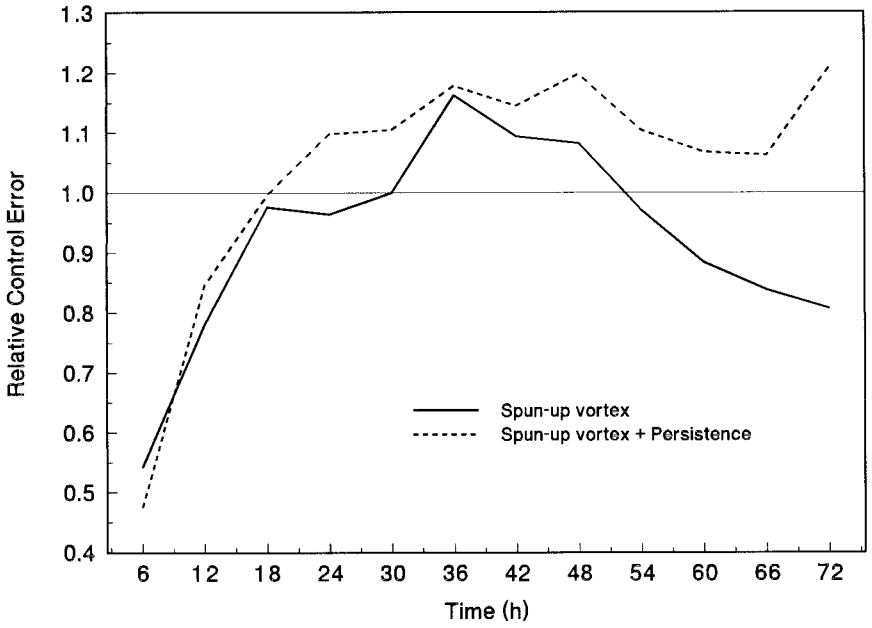

FIG. 13. Time series of the ratio of the median forecast error of the two control forecasts with a synthetic vortex to that run directly from the analyses.

median SED of the LAFV becomes positive from 60 h onward but that of BETA(1) remains negative. As in the previous comparison among the BETA experiments, this trend is also similar to their median scores (Tables 5 and 6). The relationship between SED of the BGMV and BETA(0) is about the same (Fig. 15b). The performance of BGMV is not very good in the first 36 h but becomes better afterward. The median SED of BETA(0) with respect to the 33 BGMV cases is again close to or below zero for all times, and that of BGMV also takes over after 30 h. The general better performance of the two dynamically constrained perturbation methodologies compared with that of BETA(0) can probably be explained by their respective methods in generating the perturbations. In LAFV and BGMV, a lagged period of 24 h is used to generate the perturbations. During this 24 h, interaction between the environment and the TC vortex is allowed. In the perturbed β gyres in BETA(0), however, no information from the environment is available at all. It is only after the prog-

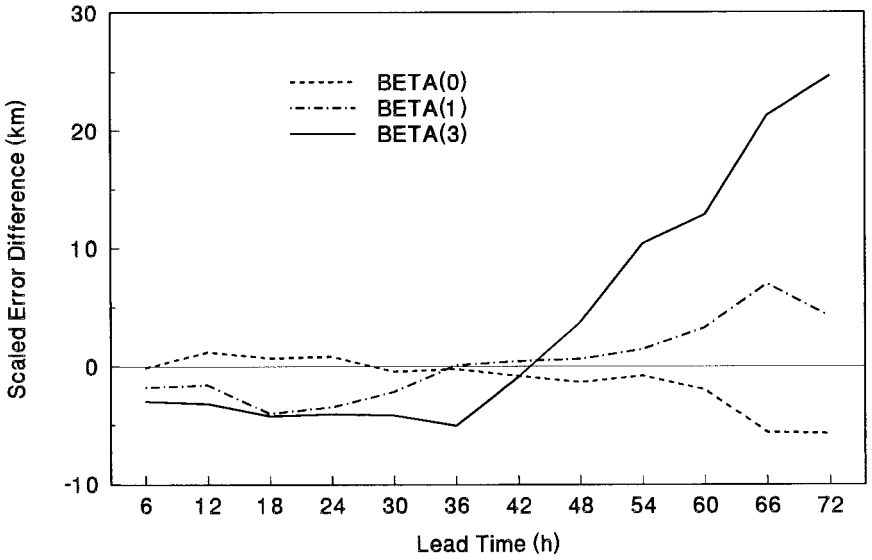

FIG. 14. Time series of the median value of SED (see text for definition) for BETA(0), BETA(1), and BETA(3).

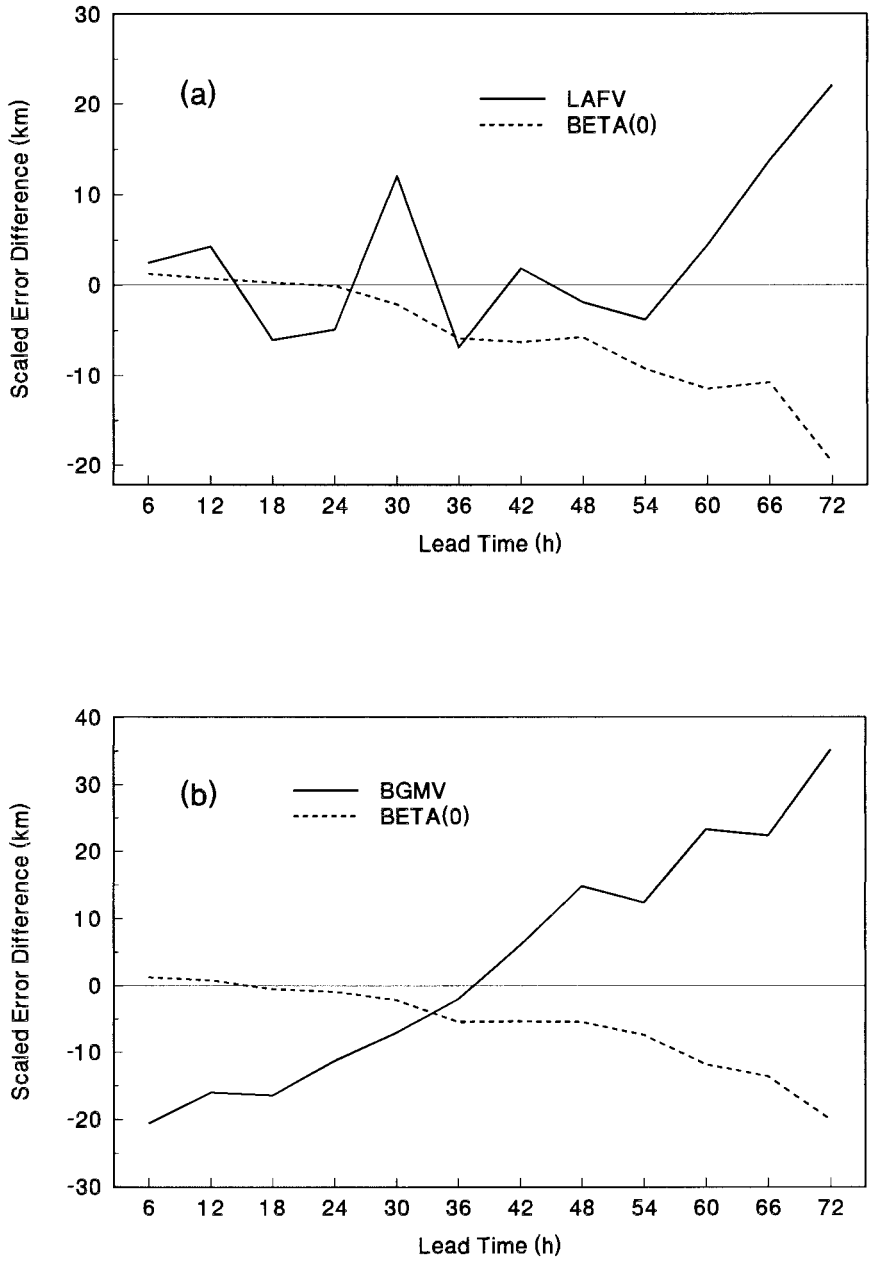

FIG. 15. As in Fig. 14 except for (a) LAFV and BETA(0), and (b) BGMV and BETA(0).

nostic process has started that the environmental flow tries to interact with the perturbed synthetic vortex.

Since BETA(3) possesses the best SED among the BETA experiments after 48 h as shown in the last subsection, it is also compared with LAFV and BGMV using the SED concept. The procedure is the same as the previous comparison of BETA(0) with the two dynamically constrained methodologies. That is, the control errors of LAFV and BGMV are used as the reference and the EDs of BETA(3) are then scaled. For the LAFV–BETA(3) pair, the median SEDs for them are both low in the first 48 h, and begin to rise afterward (Fig. 16a). The actual value of SED for BETA(3) is about double that for LAFV in these later times. When BGMV is compared with BETA(3), their resemblance is again apparent (Fig. 16b). In the first 36 h, BGMV is not very successful, and SED for BETA(3) is close to zero (due to the stochastic nature of perturbing the persistence vector as discussed in the last section). Then their SEDs begin to increase in a steady and similar manner. By the end of the 72-h forecast, an improvement (median value) of about 40 km can be obtained by both methods.

Therefore, the BETA(3) configuration should be an ideal alternative to the dynamically constrained methods

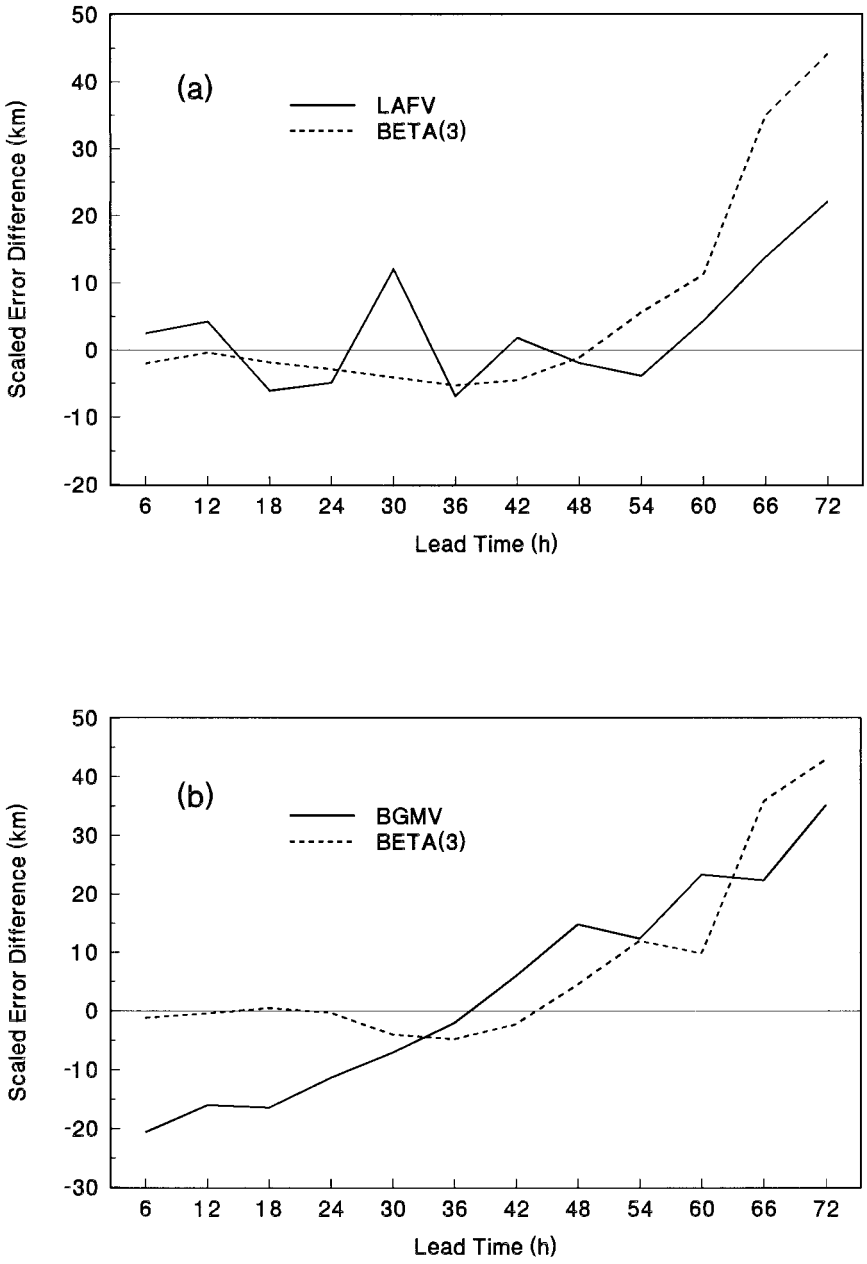

FIG. 16. As in Fig. 14 except for (a) LAFV and BETA(3), and (b) BGMV and BETA(3).

in terms of perturbations to the vortex. On the practical side, all three experiments have their own merits. LAFV and BGMV can be processed directly and conveniently from the analyses. If they are to be combined with the corresponding experiments in Part I, the filtering algorithm for isolating the forecast vortex only needs to be performed once to generate perturbations for both the environmental flow and the vortex. In addition, these methods can identify cases in which the vortex circulation is not well represented in the analysis, as explained in section 4a. For those cases, BETA(3) may play a complementary role to substitute the LAFV or BGMV perturbations.

5. Summary and concluding remarks

a. Summary

This is the second part of a study to examine the usefulness of ensemble forecasting methodology to tropical cyclone motion. The first part (Part I) deals with perturbations added to the environmental flow. In this

part, two series of eight experiments in total are performed. The first series tries to perturb the TC vortex represented in the analysis. The three experiments involved, the Monte Carlo forecast for the vortex (MCFV), the lagged-average forecast for the vortex (LAFV), and the breeding of growing modes for the vortex (BGMV), are modified from Part I to perturb the vortex. The second series intends to simulate uncertainties of a synthetic vortex as inferred from those of the observed quantities of a cyclone. Experiment IPER perturbs the initial position only, while the set BETA of four other experiments perturbs the β gyres in a spun-up vortex with the option of also changing the persistence vector. The same barotropic model as in Part I is used. Verifications are based on 66 cases selected from the 1990 season in the western North Pacific. The main findings of this part of the study are as follows.

1) The response to the random noise added in the experiment MCFV is found to be low, and the ensemble mean is thus always close to the control forecast. The situation is similar when the rms size of the noise is slightly varied, and characteristic length scale of it changed. Therefore, the utility of this perturbation methodology is low.
2) In the nine cases performed for experiment IPER, the ensemble mean is close to the control forecast. Thus, the potential of perturbing the initial position alone to increase the skill is also not promising.
3) Under the perfect model assumption (PMA), the performance of LAFV, BGMV, and the first two experiments in the set BETA is relatively better. If a detailed comparison is made, BGMV indeed performs slightly better than LAFV in agreement with the result in Part I. These results indicate generally the importance of perturbing the asymmetric component in the vortex.
4) When the verification is by the best track, the relative performance of the above four experiments is not far from one another except their actual average skill degrades substantially. An increasing trend in the skill of the BETA(3) (where perturbations to the β gyres and persistence vector are combined) by the end of the 72-h forecast is observed.
5) A high degree of spread–skill correlation is demonstrated under the PMA for the above four experiments. When verified by the best track, the spread is too small compared with the forecast error. The correlation coefficients at various lead times are even smaller than the corresponding ones in Part I.
6) A scaled error difference is designed to take into account the different control forecasts used in the experiments and intercomparisons are made among them. In general, it can be concluded that dynamically constrained perturbations are more promising than a simple change to the synthetic vortex. Potential skill to combine with the environmental perturbations is also demonstrated when the various BETA experiments are compared.

b. Concluding remarks

As an overview of the work performed in Part I and that in the present study, the role of perturbations added to the vortex should be clarified and addressed. In general, the performance of all the methodologies here is lower than those in Part I (except the Monte Carlo forecast). This is measured by our examination of the skill of the ensemble mean and the average score over the available cases. In terms of spread–skill correlation, it is also clear that only perturbing the vortex cannot adequately account for the variability of motion resulted from the vortex–environment interaction.

Thus, we conclude that none of the experiments here are suitable at present for an operational application. Rather, in our discussion, emphasis is placed on their relative performance. For example, potential problems have been identified if the original analysis (including the vortex circulation) is used for the basis of perturbation, and in the choice of appropriate system parameters for a perturbation based on a procedure to add a synthetic vortex. The importance of perturbing the large-scale flow (as represented by the persistence vector) is also highlighted. However, potential candidates still need to be sought for a subsequent combination of the two parts of study to form a complete picture of ensemble forecasting of TC motion in a barotropic framework.

Acknowledgments. The authors would like to thank Dr. Pat Harr of the U.S. Naval Postgraduate School for providing the Final Analyses dataset, Dr. Y. Kurihara of GFDL for providing the hurricane filtering code, and the U.K. Meteorological Office for providing the TC analysis and forecast positions. Prof. L. Leslie of the University of New South Wales (UNSW) provided the semi-Lagrangian barotropic model and supported a trip for the first author to visit UNSW to perform part of this study. The provision of the necessary computing facility by UNSW is also much appreciated. The extensive reviews from three anonymous reviewers have significantly helped improve the manuscript. Their efforts are gratefully acknowledged.

This research is sponsored by the Research Grants Council of the Universities Grants Committee of Hong Kong Grant 9040194 and partially by the U.S. Office of Naval Research Grant N00014-94-1-0824.

APPENDIX

Forecast Cases Considered

This appendix resembles the corresponding one in Part I of the paper and is reproduced here for reference. All the forecast cases are listed in Table A1. They are

TABLE A1. Forecast cases used in this study.

Vernon	Winona	Yancy	Zola	Abe	Becky	Dot	Ed	Flo
V80200*	W80700*	Y81400	Z81800*	A82600*	B82600*	D90412	E91112	F91312
V80212	W80712	Y81412	Z81812	A82612	B82612†	D90500	E91200	F91400
V80300	W80800†	Y81500*	Z81900	A82700	B82700#	D90512	E91212	F91412
V80312	W80812#	Y81512	Z81912	A82712		D90600*	E91300	F91500*
V80400		Y81600	Z82000†	A82800		D90612†	E91312	F91512†
V80412		Y81612†	Z82012#	A82812		D90700#	E91400*	F91600†
V80500†		Y81700†		A82900			E91412	F91612#
V80512†		Y81712†		A82912			E91500†	F91700#
V80600†		Y81800#		A83000†			E91512†	
		Y81812#		A83012#			E91600#	

*Used in the MCFV and IPER.
† 60-h forecast.
# 48-h forecast.

grouped with respect to individual TC and ordered chronologically. The total number is 66. Each forecast case is denoted as Xmddtt where X is the first letter of the TC name, m is the month, dd is the day, and tt is the UTC hour. All cases are run up to a lead time of 72 h except those specified. The selection criteria have been described in section 2b of Part I.

REFERENCES

Aberson, S. D., and M. DeMaria, 1994: Verification of a nested barotropic hurricane track forecast model (VICBAR). *Mon. Wea. Rev.*, **122,** 2804–2815.

Carr, L. E., III, and R. L. Elsberry, 1992: Analytical tropical cyclone asymmetric circulation for barotropic model initial conditions. *Mon. Wea. Rev.*, **120,** 644–652.

Chan, J. C. L., and W. M. Gray, 1982: Tropical cyclone movement and surrounding flow relationship. *Mon. Wea. Rev.*, **110,** 1354–1374.

——, and R. T. Williams, 1987: Analytical and numerical studies of the beta-effect in tropical cyclone motion. Part I: Zero mean flow. *J. Atmos. Sci.*, **44,** 1257–1265.

——, and K. K. W. Cheung, 1995: Fourier components of the circulation associated with tropical cyclone motion. Preprints, *21st Conf. on Hurricanes and Tropical Meteorology*, Miami, FL, Amer. Meteor. Soc., 12–14.

——, and R. H. F. Kwok, 1997: A diagnostic study on the improvement in tropical cyclone motion prediction by the UK Meteorological Office global model. *Meteor. Appl.*, **4,** 1–9.

——, and K. K. W. Cheung, 1998: Characteristics of the asymmetric circulation associated with tropical cyclone motion. *Meteor. Atmos. Phys.*, **65,** 183–196.

Cheung, K. K. W., and J. C. L. Chan, 1999: Ensemble forecasting of tropical cyclone motion using a barotropic model. Part I: Perturbations of the environment. *Mon. Wea. Rev.*, **127,** 1229–1243.

Curry, W. T., R. L. Elsberry, and J. C. L. Chan, 1987: An objective technique for estimating present tropical cyclone locations. *Mon. Wea. Rev.*, **115,** 1073–1082.

DeMaria, M., S. D. Aberson, K. V. Ooyama, and S. J. Lord, 1992: A nested spectral model for hurricane track forecasting. *Mon. Wea. Rev.*, **120,** 1628–1643.

Franklin, J. L., S. E. Fuer, J. Kaplan, and S. D. Aberson, 1996: Tropical cyclone and surrounding flow relationships: Searching for beta gyres in Omega dropwindsonde datasets. *Mon. Wea. Rev.*, **124,** 64–84.

Fiorino, M., and R. L. Elsberry, 1989: Some aspects of vortex structure related to tropical cyclone motion. *J. Atmos. Sci.*, **46,** 975–990.

Heming, J. T., J. C. L. Chan, and A. M. Radford, 1995: A new scheme for the initialisation of tropical cyclones in the UK Meteorological Office global model. *Meteor. Appl.*, **2,** 171–184.

Holland, G. J., L. M. Leslie, E. A. Ritchie, G. S. Dietachmayer, P. E. Powers, and M. Klink, 1991: An interactive analysis and forecast system for tropical cyclone motion. *Wea. Forecasting*, **6,** 415–424.

Krishnamurti, T. N., R. Correa-Torres, G. Rohaly, and D. Oosterhof, 1997: Physical initialization and hurricane ensemble forecasts. *Wea. Forecasting*, **12,** 503–514.

Kurihara, Y., M. A. Bender, and R. J. Ross, 1993: An initialization scheme of hurricane models by vortex specification. *Mon. Wea. Rev.*, **121,** 2030–2045.

——, ——, R. E. Tuleya, and R. J. Ross, 1995: Improvements in the GFDL hurricane prediction system. *Mon. Wea. Rev.*, **123,** 2791–2801.

Leslie, L. M., and G. J. Holland, 1995: On the bogussing of tropical cyclones in numerical models: A comparison of vortex profiles. *Meteor. Atmos. Phys.*, **56,** 101–110.

Peng, M. S., and R. T. Williams, 1990: Dynamics of vortex asymmetries and their influence on vortex motion on a β-plane. *J. Atmos. Sci.*, **47,** 1987–2003.

——, B. F. Jeng, and C. P. Chang, 1993: Forecast of typhoon motion in the vicinity of Taiwan during 1989–90 using a dynamical model. *Wea. Forecasting*, **8,** 309–325.

Rogers, E., S. J. Lord, D. G. Deaven, and G. J. DiMego, 1993: Data assimilation and forecasting for the tropical cyclone motion experiment at the National Meteorological Center. Preprints, *20th Conf. on Hurricanes and Tropical Meteorology*, San Antonio, TX, Amer. Meteor. Soc., 329–330.

Smith, R. K., and W. Ulrich, 1990: An numerical study of tropical cyclone motion using a barotropic model. I: The role of vortex asymmetries. *Quart. J. Roy. Meteor. Soc.*, **116,** 337–362.

Ueno, M., 1989: Operational bogussing and numerical prediction of typhoon in JMA. JMA/NPD Tech. Rep. 28, Japan Meteorological Agency, 48 pp. [Available from Department of Physics and Materials Science, City University of Hong Kong, 83 Tat Chee Ave., Kowloon, Hong Kong, China.]

Xu, Y., and C. J. Neumann, 1985: A statistical model for the prediction of Western North Pacific tropical cyclone motion (WPCLPR). NOAA Tech. Memo. NWS NHC 28, 30 pp. [Available from Department of Physics and Materials Science, City University of Hong Kong, 83 Tat Chee Ave., Kowloon, Hong Kong, China.]

Zhang, Z., and T. N. Krishnamurti, 1997: Ensemble forecasting of hurricane tracks. *Bull Amer. Meteor. Soc.*, **78,** 2785–2795.