RMetS
Royal Meteorological Society

Characteristics of Ensemble Transform Kalman Filter adaptive sampling guidance for tropical cyclones

Sharanya J. Majumdar,[a]* Shin-Gan Chen[b] and Chun-Chieh Wu[b]
[a]*University of Miami, Florida, USA*
[b]*National Taiwan University, Taipei, Taiwan*
*Correspondence to: Dr S. J. Majumdar, RSMAS Division of Meteorology and Physical Oceanography, 4600 Rickenbacker Causeway, Miami 33149, Florida, USA. E-mail: smajumdar@rsmas.miami.edu

The Ensemble Transform Kalman Filter (ETKF) is one candidate strategy for targeting observations around tropical cyclones. The characteristics of target areas identified by a revised version of the ETKF are investigated using the THORPEX Interactive Grand Global Ensemble. In order to emphasize sensitivity to the environmental flow, the axisymmetric circulation associated with the tropical cyclone is removed from each ensemble member. The guidance is found to differ markedly from the ensemble variance, and the guidance products based on ensembles from different operational centres often disagree. For the pre-recurvature stage of typhoon *Sinlaku* (2008), the targets correspond to asymmetries close to the typhoon, neighbouring features including the adjacent subtropical ridge and a nearby upper-level trough, and far-field features including the monsoon trough and areas of flow convergence. During and after recurvature, targets exist in the midlatitude jet upstream and downstream of the typhoon. The midlatitude targets are discernible for forecasts beyond 3 days. The sensitivity to hypothetical observations of wind is more pronounced than those of temperature or specific humidity. The targets based on upper-level observations correspond to the outflow of the tropical cyclone and nearby and remote jets, while the lower-level targets correspond to confluent flow near the cyclone. Multiple remote targets are evident, particularly for longer forecast times, and some of these may be spurious. When ensemble forecasts initialized at different times are used for the same case, the characteristics of the guidance differ, although not dramatically. Future priorities include the evaluation of the effects of assimilating observations sampled in target areas *versus* those in non-target areas, and the improvement of the data assimilation methodology assumed in the ETKF.

Key Words: targeted observations; data assimilation; typhoon *Sinlaku*

Received 26 March 2010; Revised 6 November 2010; Accepted 19 November 2010; Published online in Wiley Online Library 28 February 2011

Citation: Majumdar SJ, Chen S-G, Wu C-C. 2011. Characteristics of Ensemble Transform Kalman Filter adaptive sampling guidance for tropical cyclones. *Q. J. R. Meteorol. Soc.* **137**: 503–520. DOI:10.1002/qj.746

1. Introduction

Short-range (1–3 d) forecasts of tropical cyclone (TC) track, structure and intensity are prone to errors in the initial conditions, in the cyclone and its environment. These errors may be initially small and growing rapidly, or initially large in an area of high uncertainty. In order to reduce initial condition errors, aircraft-borne dropwindsonde observations have been 'targeted' in the synoptic environment of TCs. Following twenty

'Synoptic Flow' experiments conducted by the National Oceanographic and Atmospheric Administration (NOAA) during 1982–1996 (Burpee *et al.*, 1996), NOAA procured a Gulfstream IV-SP jet aircraft (G-IV) and has conducted synoptic surveillance missions since 1997. The strategy for targeting observations around the TC is based on a combination of sampling in all quadrants of the TC, and sampling in neighbouring areas of high variance in the deep-layer (850–200 hPa) mean wind predicted by the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS; Aberson, 2010). Since 2003, the Dropwindsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR; Wu *et al.*, 2005, 2007a) programme has flown frequent surveillance missions around TCs threatening Taiwan, using a similar strategy to NOAA while incorporating new strategies described below. In 2008, The Observing-system Research and Predictability EXperiment (THORPEX) Pacific Asian Regional Campaign (T-PARC) allowed for the use of multiple aircraft to target TCs in the western North Pacific basin, based on a consensus decision using guidance products for targeted observations (e.g. Harnisch and Weissmann, 2010). The current state of the science in targeted observations is summarized well by Langland (2005): 'The primary scientific challenges for targeting include the refinement of objective methods that can identify optimal times and locations for targeted observations, as well as identify the specific types of satellite and *in situ* measurements that are required for the improvement of numerical weather forecasts'.

Since the utilization of ensemble variance (Aberson, 2003), several strategies for targeted observations have been employed. Targeted singular vectors (SVs), first proposed by Palmer *et al.* (1998), have been extended for use in TCs (Peng and Reynolds, 2006). In a case-study, Yamaguchi *et al.* (2009) demonstrated that data from a DOTSTAR mission around typhoon *Conson* (2004) were particularly effective at improving the track forecast if they were collected in a region of SV sensitivity. A relatively new method, the Adjoint-Derived Sensitivity Steering Vector (ADSSV; Wu *et al.*, 2007b), seeks sensitive areas in which initial perturbations would modify the steering flow. Another technique that has proven to be effective for short-range (Majumdar *et al.*, 2002) and medium-range (Sellwood *et al.*, 2008) forecasts of winter weather is the Ensemble Transform Kalman Filter (ETKF), which has also recently been extended for use with TCs. While all these strategies are based on reducing error in a forecast metric, detailed intercomparison studies have shown that they disagree frequently, for TCs in the Atlantic basin (Majumdar *et al.*, 2006; Reynolds *et al.*, 2007) and in the western North Pacific basin (Wu *et al.*, 2009a). A recent evaluation by Aberson *et al.* (2010) using a limited number of cases demonstrated that sampling based on guidance from both SV and ETKF methods may improve the track forecast. Given the lack of a clear consensus in many cases, it is imperative for the mission planner to provide a skilful interpretation of the targets produced by these methods. Such an interpretation has been provided for SVs (Peng and Reynolds, 2006; Reynolds *et al.*, 2009) and the ADSSV (Wu *et al.*, 2009a,b). However, no clear description yet exists for the ETKF, which is different from the 'dynamical sensitivity' SV and ADSSV methods since it combines ensemble-based data assimilation with perturbation evolution to identify the sensitivity of a TC forecast to *observations*. The purpose of this paper is to introduce a revised version of the ETKF pertaining specifically to TCs, and to illustrate the characteristics of ETKF targets through the TC life cycle, for variables at different levels, over a range of forecast times.

The new version of the ETKF documented in this paper is available for use in forthcoming field campaigns such as DOTSTAR. It is also timely to comment on the potential implications of targeting new types of data, such as activating rapid-scan mode aboard geostationary satellites to yield higher-density atmospheric motion vectors (Velden *et al.*, 2005; Langland *et al.*, 2009). Furthermore, only a very basic version of the ETKF has been employed to date, with vertically averaged measures of horizontal wind and temperature being employed in the analysis and forecast-error covariance matrices. The sensitivity to different variables at various levels has not yet been examined. Similarly, the dependence of the ETKF on the ensemble has not been illustrated. And crucially, ETKF targets to date have been dominated by the large ensemble spread of the wind field that is associated with different positions of the TC at the analysis time. While it is logical that reducing the spread in the wind field of the TC would likely lead to a reduction in the associated spread at a later time, the targeting of the synoptic environment is also important. The new version of the ETKF seeks to emphasize the environment, by removing the axisymmetric component of the TC from each ensemble member. The justification is that the ensemble spread of locations at the targeted observing time is likely to be much smaller than that predicted by the ensemble, which is typically initialized two days prior to the observing time. This article also represents an opportune time to improve the ETKF, given that Ensemble Kalman Filter data assimilation schemes with similar flow-dependent error statistics that exhibit strong spatial variability in TCs are being developed in research (Torn and Hakim, 2009a; Zhang *et al.*, 2009) and also as potential operational schemes with global models (Szunyogh *et al.*, 2008; Whitaker *et al.*, 2008). A quantitative evaluation of the effectiveness of this version of the ETKF in improving TC forecasts will be necessary in the future.

The paper is organized as follows. In section 2, the basic methodology and modifications to the ETKF are described. In section 3, the ensemble variance and ETKF targets after the removal of the axisymmetric vortex are illustrated. In section 4, the ETKF sensitivities to observing variable, level, lead-time, and stages of life cycle are shown. In section 5, a discussion of the characteristics of ETKF targets is provided in the context of available targeted observations and other methodologies, followed by concluding remarks in section 6.

2. Methodology

2.1. TIGGE ensembles

The THORPEX Interactive Grand Global Ensemble (TIGGE; Bougeault *et al.*, 2010) was recently established as a pathway towards establishing a Global Interactive Forecast System. Currently, ensemble forecasts from several operational global models are archived in a common format on servers such as http://tigge.ecmwf.int/ and are freely available to researchers. Examples of TC research using TIGGE include Majumdar and Finocchio (2010) who explored the ability of ensembles to predict TC track probabilities, and Yamaguchi and Majumdar (2010)

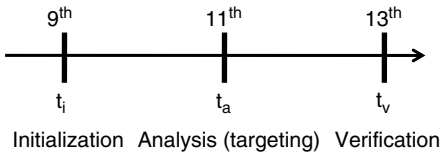

Figure 1. Typical timeline for preparation of ETKF guidance, using ensembles initialized at time t_i. The goal is to identify target areas in which the forecast error variance is maximally reduced at time t_v, due to the targeting of observations at time t_a.

who investigated the structure and evolution of ensemble perturbations in the TC steering flow. In this study, the NCEP GEFS, European Centre for Medium-range Weather Forecasts (ECMWF) and Canadian Meteorological Centre (CMC) outputs at 1° resolution, using 20, 50 and 20 ensemble members respectively (Wei *et al.*, 2008; Buizza *et al.*, 2003; Pellerin *et al.*, 2003, respectively), are used to prepare a matrix $\mathbf{Z}^{\mathrm{i}}(t|H^{\mathrm{i}})$ of linearly independent forecast perturbations, initialized at time t_i with operator H^{i}. Ensembles from each individual model and a combined NCEP+ECMWF+CMC ensemble comprising 89 linearly independent members are used in the ETKF computations. The perturbation for each member is computed about the mean of the ensemble from which that member is drawn, in order to produce realistic samples from the true distribution of dynamically evolving initial condition errors.

2.2. ETKF theory and ensemble perturbation modification

The basic formulation of the ETKF follows that of Majumdar *et al.* (2010), to which the reader is referred for further details and justification. The version of the ETKF ('Var-ETKF') used during T-PARC and in this article uses operational estimates of analysis-error variance to constrain ensemble perturbations at the analysis time. The ETKF then uses these ensemble perturbations to predict the reduction in forecast-error variance within a given verification region at verification time t_v, based on the assimilation of a group of adaptive observations at analysis time $t_a (t_i < t_a < t_v)$. In this study, $t_a - t_i = 2$ d, and $t_v - t_a = 2$ d unless otherwise stated (Figure 1). While $t_a - t_i = 2$ d is a relatively long time for ensemble perturbations to evolve in data assimilation, it is the minimum lead time that is presently necessary for timely aircraft planning or satellite rapid-scan activation. As for all ensemble filters, the error covariance matrix is estimated as the matrix outer product of $\mathbf{Z}^{\mathrm{i}}(t|H^{\mathrm{i}})$ and its transpose:

$$\mathbf{P}^{\mathrm{i}}(t|H^{\mathrm{i}}) = \mathbf{Z}^{\mathrm{i}}(t|H^{\mathrm{i}})\mathbf{Z}^{\mathrm{iT}}(t|H^{\mathrm{i}}). \tag{1}$$

Serial assimilation theory (Bishop *et al.*, 2001) is then used to divide the observational network at time t_a into 'routine' (operator $\mathbf{H}^{\mathrm{r}}$, error covariance $\mathbf{R}^{\mathrm{r}}$) and 'adaptive' (operator $\mathbf{H}^{\mathrm{q}}$, error covariance $\mathbf{R}^{\mathrm{q}}$) components. Analysis- and forecast-error covariance matrices $\mathbf{P}(t|\mathbf{H})$ are computed via a linear transformation $\mathbf{T}$ of the ensemble that is based on these subsets of the observational network. For the routine observational network, an ensemble transform is used together with the monthly mean root mean square analysis-error variance $\mathbf{P}^{\mathrm{i}}{}_{\mathrm{OPER}}(t_a|H^{\mathrm{i}})$ provided by the Naval Research Laboratory (NRL) Atmospheric Variational Data Assimilation System (NAVDAS) to solve for the appropriate transformation matrix $\mathbf{T}^{\mathrm{r}}$, via

$$\mathbf{T}^{\mathrm{rT}}\,\mathbf{Z}^{\mathrm{i}}\{\mathbf{P}^{\mathrm{i}}{}_{\mathrm{OPER}}(t_a|H^{\mathrm{i}})\}^{-1}\mathbf{Z}^{\mathrm{i}}\mathbf{T}^{\mathrm{r}} = \mathbf{I}. \tag{2}$$

The routine analysis-error covariance matrix valid at time t_a is then given by

$$\mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}}) = \mathbf{Z}^{\mathrm{i}}(t_a|H^{\mathrm{i}})\mathbf{T}^{\mathrm{r}}\mathbf{T}^{\mathrm{rT}}\mathbf{Z}^{\mathrm{iT}}(\mathrm{t_a}|H^{\mathrm{i}}). \tag{3}$$

To produce the ETKF guidance maps, the analysis-error covariance due to each hypothetical deployment of targeted observations (denoted by superscript q) is then computed as

$$\begin{aligned}\mathbf{P}^{\mathrm{r+q}}(t_a|\mathbf{H}^{\mathrm{r+q}}) = {}& \mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}}) \\ & -\mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}})\mathbf{H}^{\mathrm{qT}}\{\mathbf{H}^{\mathrm{q}}\mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}})\mathbf{H}^{\mathrm{qT}} + \mathbf{R}^{\mathrm{q}}\}^{-1} \\ & \times\mathbf{H}^{\mathrm{q}}\mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}}).\end{aligned} \tag{4}$$

Extending (4) to time t_v yields the forecast-error covariance

$$\mathbf{P}^{\mathrm{r+q}}(t_v|\mathbf{H}^{\mathrm{r+q}}) = \mathbf{P}^{\mathrm{r}}(t_v|\mathbf{H}^{\mathrm{r}}) - \mathbf{S}^{\mathrm{q}}(t_v|\mathbf{H}^{\mathrm{q}}),$$

where

$$\begin{aligned}\mathbf{S}^{\mathrm{q}}(t_v|\mathbf{H}^{\mathrm{q}}) = {}& \mathbf{Z}^{\mathrm{i}}(t_v|\mathbf{H}^{\mathrm{r}})\mathbf{Z}^{\mathrm{iT}}(t_a|\mathbf{H}^{\mathrm{r}})\mathbf{H}^{\mathrm{qT}} \\ & \times\{\mathbf{H}^{\mathrm{q}}\mathbf{P}^{\mathrm{r}}(t_a|\mathbf{H}^{\mathrm{r}})\mathbf{H}^{\mathrm{qT}} + \mathbf{R}^{\mathrm{q}}\}^{-1} \\ & \times\mathbf{H}^{\mathrm{q}}\mathbf{Z}^{\mathrm{i}}(t_a|\mathbf{H}^{\mathrm{r}})\mathbf{Z}^{\mathrm{iT}}(t_v|\mathbf{H}^{\mathrm{r}}).\end{aligned} \tag{5}$$

The diagonal of $\mathbf{S}^{\mathrm{q}}(t_v|\mathbf{H}^{\mathrm{q}})$ localized within the verification region is rapidly produced and plotted as a function of the qth targeted observation on the ETKF guidance used throughout this paper. The qth targeted observation is of wind and/or temperature and/or specific humidity, sampled at either 200, 500 or 850 hPa, with adjacent observations sampled at 1° resolution. The summary map represents a mosaic of reduction in forecast-error variance within the fixed verification region, as a function of the observation location. The optimal target location is the value of q for which the diagonal of $\mathbf{S}^{\mathrm{q}}(t_v|\mathbf{H}^{\mathrm{q}})$ localized within the verification region is largest. All ETKF guidance presented in this paper is normalized with respect to the maximum value on the guidance map. In the future, the examination of the un-normalized value of signal variance would be useful, in order to quantitatively estimate day-to-day changes in the potential for forecast-error reduction. The ETKF has shown promise in this regard in the midlatitudes (Majumdar *et al.*, 2001, 2002; Sellwood *et al.*, 2008). For tropical cyclones, it is not yet clear to the authors how to interpret quantitatively the reduction in forecast-error variance, given that it is largely a function of both the TC wind field and its position.

Theoretically, the reduction in forecast error covariance $\mathbf{S}^{\mathrm{q}}(\mathrm{t_v}|\mathbf{H}^{\mathrm{q}})$ is equal to the covariance of 'signals', where a signal is defined as the difference between two numerical forecasts that are identical in every respect except that one forecast includes the targeted data in the assimilation, while the second withholds the targeted data (Bishop *et al.*, 2001). Therefore, the ETKF attempts to predict the variance of the propagation of the effect of any set of targeted observations, whose variables and levels at the observing time are explored through the article. The signal variance at the verification

time is based on a vertically averaged 'difference kinetic energy' (DKE) norm between pressure levels p_0 and p_t:

$$E(x, y, t) = \frac{1}{p_t - p_0} \int_{p_0}^{p_t} ({u_s}^2 + {v_s}^2) \frac{dp}{2}. \quad (6)$$

The quantities u_s and v_s represent the signal in the zonal and meridional wind fields respectively. The vertical average is computed over three pressure levels: 850, 500 and 200 hPa. We elect to use the DKE norm, given that it is the variance in the wind forecast errors in and around the TC that are most suitable for reduction in global models, and it is at least partially related to track forecast-error variance. Other perturbation variables such as temperature and specific humidity were found to produce a secondary contribution to the signal variance at the verification time.

As with all ensemble filters, sampling errors due to a limited ensemble size lead to spurious, long-distance error covariance structures. This in turn yields targets that may be very remote to the verification region, which is the TC in this case. Unlike data assimilation in which the sampling error is mitigated via a distance-dependent covariance localization around either the model grid point (Szunyogh *et al.*, 2008) or the observation location (Whitaker *et al.*, 2008), such a localization is not employed in the ETKF adaptive sampling strategy, since no known method exists to propagate the localized error covariance information forward in time. Potential dynamical localization methods for use with targeted observations are under development (Craig Bishop, personal communication), and one exploratory method to reduce the size of ensemble perturbations far from the TC is discussed in section 5.

2.3. *Removal of axisymmetric component of TC*

The versions of the ETKF reported in Majumdar *et al.* (2006), Reynolds *et al.* (2007) and Wu *et al.* (2009a) all contain ensemble perturbations of horizontal wind that are dominant near the TC at time t_a. These large perturbations are associated with both the variability in the high winds of the primary circulation and the spreading in the track forecast, which can be considerable when the ensemble is initialized two or more days prior to the targeted observations being collected (as is normally the case for aircraft mission planning or satellite rapid-scan activation). Accordingly, the overwhelming ETKF target is most often associated with the TC itself (e.g. Figures 6(c,d) of Reynolds *et al.*, 2007). It is intuitive that targeting observations in the TC will reduce the error variance in its evolved wind field at later forecast times, so no new insight is gained via guidance that simply emphasizes the TC. In order to identify the guidance associated with the TC asymmetry and near- and far-field environment, we modify each ensemble member by removing the axisymmetric component of the flow associated with the TC. The following is done for each ensemble member. First, the stream function is computed at each of the three pressure levels (200, 500, 850 hPa), from which the TC centre is deduced as the local minimum. (If a centre does not exist, as is often the case at 200 hPa, the centre at 850 hPa is used. The centre location does not vary significantly with height, given the 1° resolution of the ensemble output.) The average stream function at a specified radius R_c from the TC centre is then computed, followed by the axisymmetric component and resulting azimuthal wind field. This axisymmetric wind field is then subtracted from the original wind field.

3. Typhoon *Sinlaku*: Ensemble variance, vortex removal, ETKF guidance

3.1. *Synoptic overview of Sinlaku*

Typhoon *Sinlaku* (15 W) was the most heavily observed TC during the Tropical Cyclone Structure (TCS-08)/T-PARC field campaign in 2008. After forming as a depression east of the Philippines at 0000 UTC on 8 September, it was upgraded to a tropical storm by the Japan Meteorological Agency (JMA) at 1800 UTC on the same day. By 1200 UTC on 9 September, *Sinlaku* had already become a 65 kt (33 m s^{-1}) typhoon, and it underwent rapid intensification into a 100 kt (51 m s^{-1}) typhoon by 1800 UTC on 10 September. During the period of rapid development and subsequent slight weakening due to eyewall replacement, the motion of *Sinlaku* was slow and meandering, prior to its landfall in Taiwan on 13 September. *Sinlaku* continued to decay to a tropical storm as it moved slowly over northern Taiwan and recurved on 15 September. During its northeastward track towards Japan, *Sinlaku* surprisingly re-intensified into a typhoon on 18 September.

The western North Pacific basin commonly possesses several systems that modify TC motion, such as the subtropical high, the midlatitude trough, the subtropical jet, and the southwesterly monsoon (Wu, 2006). The environment of *Sinlaku* exhibited all these features, which played different roles during the life cycle of the storm. Shortly before 0000 UTC on 11 September, *Sinlaku* was situated directly to the south of an upper-level low pressure system (labelled 'A' on Figure 2(b)), and between two subtropical high pressure systems, one over mainland China ('B') and the other over the western North Pacific ('C'). Among the more remote systems was the monsoon trough over Vietnam ('D'), a zone of convergence on the northwestern side of the China subtropical high ('E') and a tropical depression, which was sampled by five TCS-08/T-PARC aircraft missions for TC formation and targeting ('F'). The depression did not develop further, while *Sinlaku* became organized rapidly. Finally, two sections of the midlatitude jet are labelled 'G' and 'H'.

Sinlaku's motion towards Taiwan as it intensified was very slow, with neither subtropical ridge ('B' or 'C') nor the cut-off low ('A') possessing a dominant influence. There was considerable uncertainty in the track forecasts of *Sinlaku*, thereby justifying six coordinated targeted observation missions by the DOTSTAR and DLR Falcon aircraft during 10–16 September 2008 (with numerous other missions focusing on TC structure and extratropical transition). The influence of assimilating dropwindsonde data from these aircraft on numerical forecasts of *Sinlaku* is reported by Harnisch and Weissmann (2010) and Weissmann *et al.* (2010).

3.2. *Ensemble variance*

Prior to investigating the ETKF guidance, it is instructive to analyze the variance in each of the individual model ensembles, and the modification to the distribution of variance that is achieved by removing the axisymmetric component of the vortex. First, the distribution of the

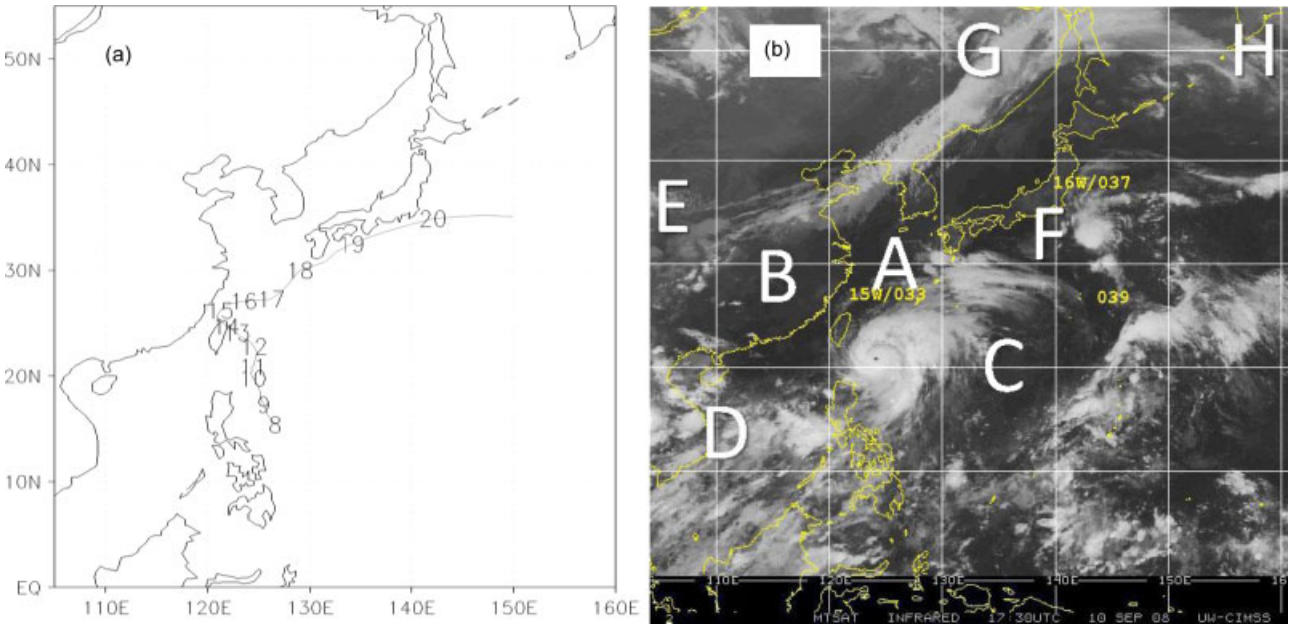

Figure 2. (a) Japan Meteorological Agency (JMA) best track of typhoon *Sinlaku*, with the numbers indicating the position at 0000 UTC on that date. (b) MTSAT infrared satellite image of typhoon *Sinlaku* at 1730 UTC on 10 September 2008 This figure is available in colour online at wileyonlinelibrary.com/journal/qj

48h ensemble forecast variance in the horizontal wind speed components averaged over 850, 500 and 200hPa is illustrated for each of the NCEP, ECMWF, CMC, and combined 89-member ensembles, valid at 0000 UTC on 11 September when *Sinlaku* was a category 2 typhoon to the southeast of Taiwan (Figure 2). The features depicted in Figure 2(b) are evident in the ensemble mean streamlines in Figure 3. It is immediately apparent that the ECMWF and CMC ensembles possess high variance associated with *Sinlaku* (Figures 3(b,c)), while the NCEP values are relatively small (Figure 3(a)). The TC in the NCEP ensemble is weak due to its coarse horizontal resolution (T126) in 2008, and the variance is low since the TC is relocated to the same position in each ensemble member at the initial time. The NCEP ensemble variance values are also relatively small throughout the domain, meaning that the ensemble members are relatively tightly clustered compared with ECMWF and CMC. The local variance maxima correspond to distinctive features, including the upper-level low to the north of *Sinlaku* ('A'), the convergence zone over China ('E'), and a broad cloudy area in the midlatitude storm track ('G'). In contrast to NCEP, the ECMWF ensemble variance is dominated by the variance in the track and wind field of *Sinlaku*, with secondary maxima in the upper-level low ('A'), midlatitude jet ('G'), and the tropical depression southeast of Japan which did not develop ('F'). Areas of low variance exist in the Tropics, associated with the monsoon trough ('D') and an easterly wave to the southeast of *Sinlaku*. There is little variance in the area of convergence over China ('E'). The distribution of variance in the CMC ensemble is considerably broader than that of NCEP and ECMWF, with multiple local maxima. When all three ensembles are combined, the structure of variance appears to be dominated by ECMWF (Figure 3(d)). The magnitudes of variance maxima in the combined ensemble are slightly smaller than those corresponding maxima in the ECMWF ensemble. In order to determine whether the dominant contribution by ECMWF to the variance (and ETKF guidance) produced by the combined ensemble was due to the large number of ECMWF ensemble members, equivalent plots were produced by randomly sampling 20 of the 50 ECMWF members and repeating the computations. Little difference was found in either the variance or the ETKF guidance when 20 ECMWF members were used rather than 50 (figures not shown). We therefore interpret the dominance of the variance (and ETKF guidance) by ECMWF as a consequence of the variance of ECMWF perturbations exhibiting higher spatial variability than NCEP or CMC, which is not surprising given ECMWF's diabatic singular vector method for constructing ensemble perturbations around TCs (Barkmeijer *et al.*, 2001; Puri *et al.*, 2001), and the consequent high perturbation growth in the ECMWF ensemble documented by Yamaguchi and Majumdar (2010).

3.3. *Vortex removal*

The effect of removing the axisymmetric wind field associated with the vortex is not pronounced in the NCEP ensemble, comparing Figure 3(a) against Figure 4(a). The weak circulation associated with *Sinlaku* is replaced by a saddle point with low variance. The adjacent upper-level low is now more distinct in the mean flow streamlines and variance field. As expected, the effect of the vortex removal is more distinct in the ECMWF, CMC and combined ensembles (Figures 4(b,c,d)). It is worth emphasizing that the streamlines and variance fields are left unaltered in the remote environment of *Sinlaku*. A radius $R_c = 1000$ km was necessary to ensure that the entire circulation associated with *Sinlaku* was removed.

The eigenvalues (listed in the diagonal matrix Γ) and eigenvectors (listed as the columns of matrix $\mathbf{E}$) of the routine analysis-error covariance matrix were computed for each ensemble, via the singular value decomposition

$$\mathbf{Z}^{\mathrm{i}\,\mathrm{T}}(t_\mathrm{a}|H^\mathrm{i})\mathbf{Z}^\mathrm{i}(t_\mathrm{a}|H^\mathrm{i}) = \mathbf{C}\,\Gamma\,\mathbf{C}^\mathrm{T},$$

followed by

$$\mathbf{E} = \mathbf{Z}^\mathrm{i}(t_\mathrm{a}|H^\mathrm{i})\mathbf{C}\,\Gamma^{-1/2}.$$

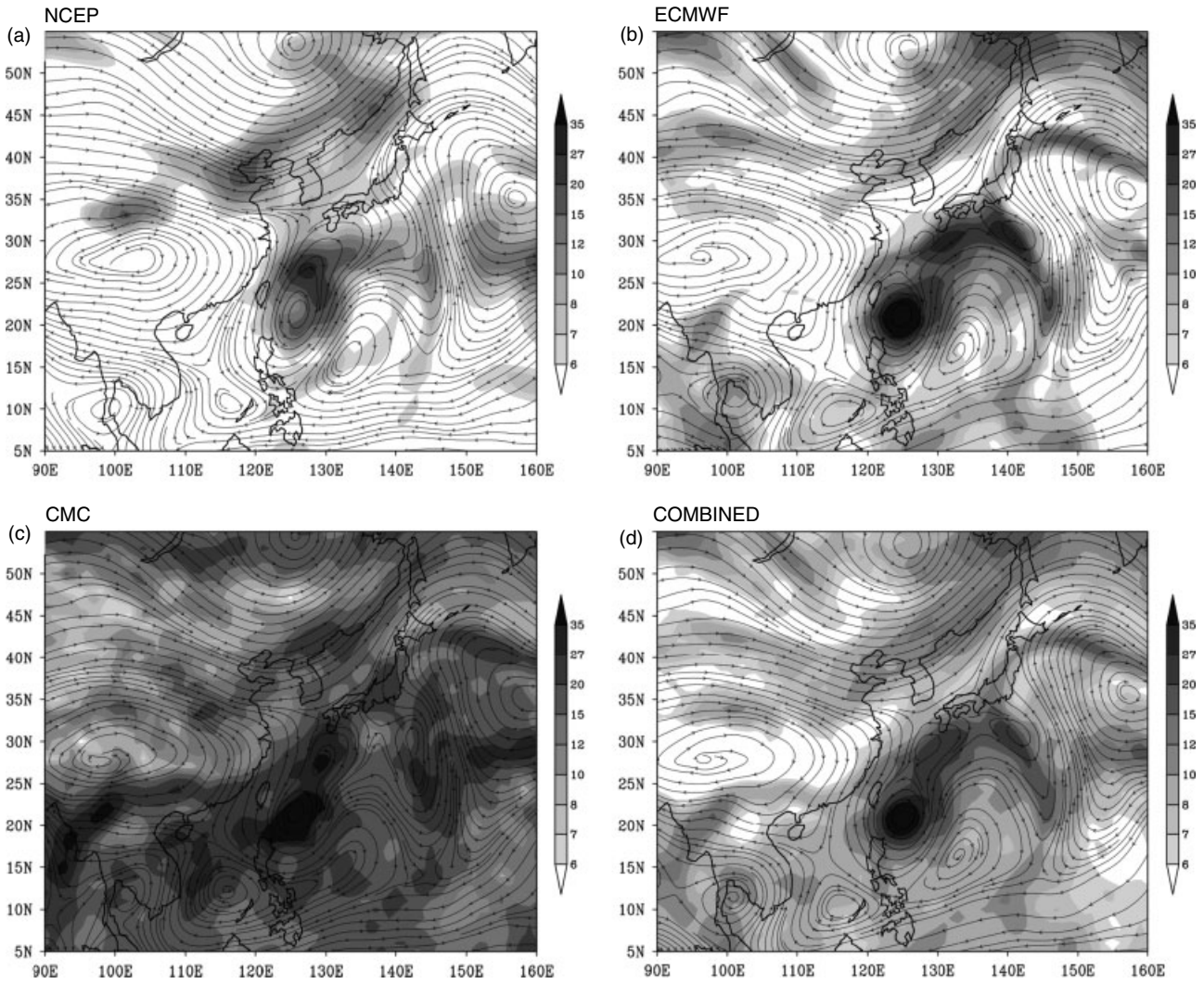

Figure 3. 48 h ensemble forecast variance (shading, $m^2\ s^{-2}$) of horizontal wind averaged over 850, 500, and 200 hPa for (a) NCEP, (b) ECMWF, (c) CMC and (d) combined NCEP+ECMWF+CMC ensembles, valid at 0000 UTC on 11 September 2008. The streamlines correspond to the ensemble mean averaged over the same three levels, valid at the same time.

The ECMWF ensemble produces a more dominant leading eigenvector and a steeper spectrum than NCEP or CMC (Figures 5(a–c)), with the eigenvector structures being larger in the midlatitudes than in the Tropics. The eigenvalue spectrum of the combined ensemble (Figure 5(d)) is more similar to that of ECMWF than to those of NCEP or CMC, although with added variance in the trailing eigenvalues.

3.4. ETKF

The ETKF targets aimed at sampling at 0000 UTC on 11 September 2008 to reduce error variance in a 2 d forecast as *Sinlaku* approached Taiwan are now presented, and compared against the ensemble variance. In a recent study by Aberson *et al.* (2010), who assimilated subsets of dropwindsonde data in areas selected by the ensemble variance, ETKF and SVs into the operational NCEP GFS system, it was found that the ensemble variance and ETKF targets local to the TC were usually similar, leading to only minor differences between the track forecasts when the variance and ETKF were used. However, an earlier version of the NCEP GFS ensemble was used to create the targets in that study, and no vortex removal had been considered. Here, the ETKF guidance is first illustrated in Figure 6 for each ensemble, with the axisymmetric component preserved and a $10^\circ \times 10^\circ$ verification region centred on the ensemble mean forecast position at $t_v = t_i + 96$ h. The ETKF guidance based on the NCEP ensemble is dominated by a broad region north of the actual position of *Sinlaku* (Figure 6(a)), which is likely associated with the ensemble variance in the flow between *Sinlaku* and the upper-level low ('A') in Figure 3(a). For ECMWF, the dominant target is the typhoon itself, with an area of confluence between the upper-level low and the subtropical ridge over the western Pacific ('C') as a secondary target (Figure 6(b)). The main target using the CMC ensemble is also centred on *Sinlaku*, with several secondary targets (Figure 6(c)). For the combined NCEP+ECMWF+CMC ensemble, the primary target is focused on *Sinlaku*, with no secondary targets discernible on the chosen scale (Figure 6(d)). The dominance of the TC as a target in the ETKF guidance is evident for most mature TCs (e.g. Figure 3(b) of Harnisch and Weissmann, 2010), and is normally similar to the maxima of ensemble variance although more localized around the TC. The ETKF guidance therefore presents limited useful information on the optimal target area, given that the TC is an obvious target in which the assimilation of additional observations within

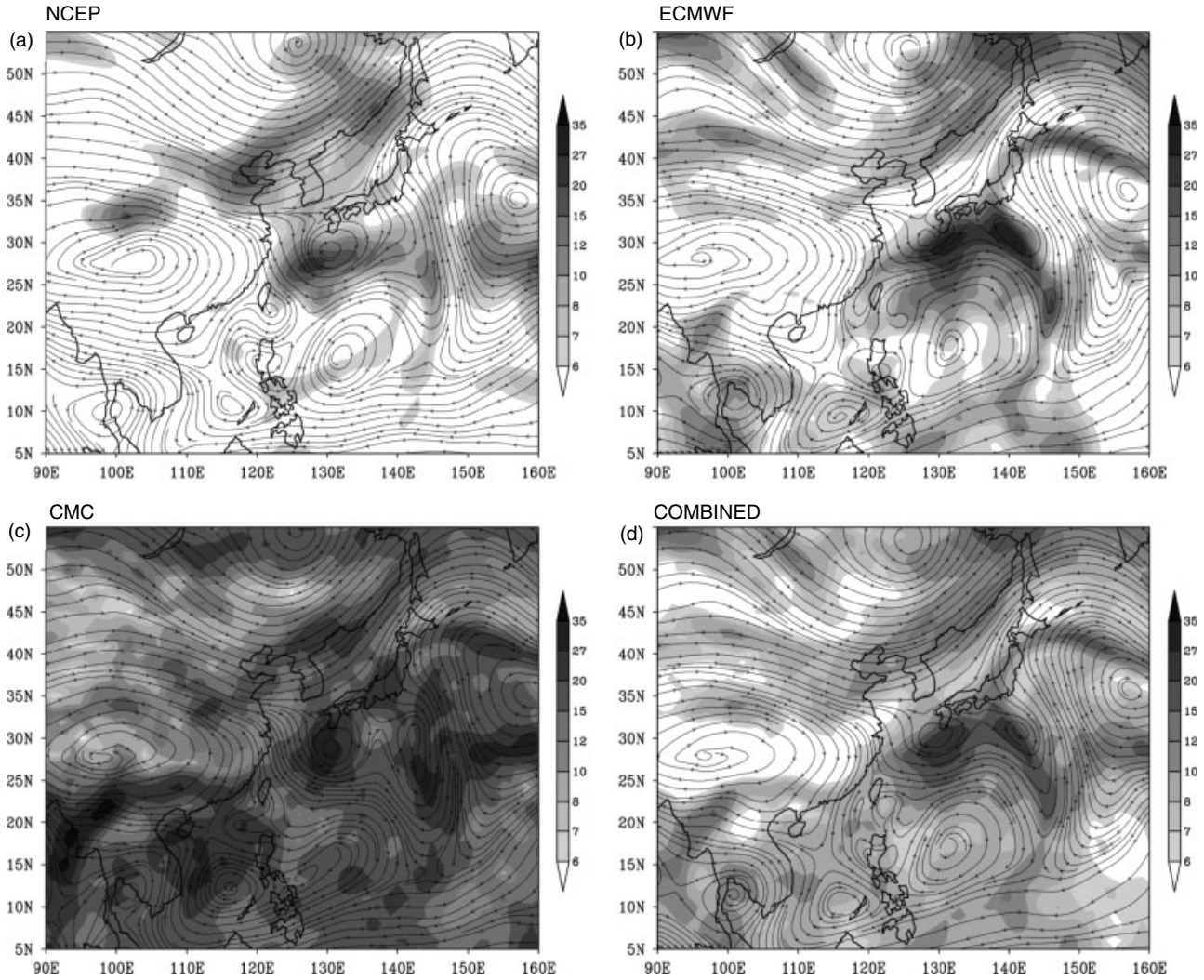

Figure 4. As Figure 3, but with the axisymmetric component of the flow removed within 1000 km of typhoon *Sinlaku*.

the TC would be likely to reduce forecast errors within the TC at a later time.

Next, the corresponding ETKF guidance with the axisymmetric component of the TC vortex removed is illustrated in Figure 7. It is evident that the ETKF guidance in Figure 7 is less localized on the TC than when the axisymmetric component is included (Figure 6). After the axisymmetric component has been removed, the remaining flow corresponds to the mean background flow and the asymmetries associated with the TC. The targets in these areas indicate that the uncertainties in the background flow and/or the asymmetries would be important to sample, as they may ultimately influence the TC motion.

The guidance with the axisymmetric component removed is different from the corresponding ensemble variance in Figure 4. For example, the ETKF guidance based on the NCEP ensemble places emphasis on the upper-level low and trough to the north of *Sinlaku* ('A'), the subtropical ridge to the southeast of *Sinlaku* ('C'), and the upstream area of convergence over China ('E') (Figure 7(a)). The wind fields in all these features may be important in modifying the track and wind field of *Sinlaku*. Some remote locations, such as the midlatitude jet ('G') in which the variance is large are not evident in the ETKF guidance. The primary targets in the ECMWF ensemble are associated with *Sinlaku*, the northwestern side of the subtropical ridge ('C') over the western Pacific, and a region of flow confluence between the upper-level low ('A') and the subtropical ridge ('C') (Figure 7(b)). Secondary targets are evident in the deep Tropics and midlatitudes. The ETKF guidance based on the ECMWF ensemble emphasizes features closer to the storm than the distribution of ensemble variance. The ETKF guidance using the CMC ensemble is broader than that based on ECMWF or NCEP, but with discernible targets in the vicinity of *Sinlaku* and the upstream region of convergence over China ('E') (Figure 7(c)). The ETKF guidance for the combined NCEP+ECMWF+CMC ensemble resembles that of ECMWF (Figures 7(b,d)), although more localized around *Sinlaku*. In order to provide a more even weighting to each ensemble, a rescaling of the perturbations from the respective ensembles is worth considering in the future.

4. Sensitivity to level, variable, TC stage, forecast and lead-time

4.1. Pressure level of targeted wind and temperature observations

It is useful to determine situations in which forecasts would benefit from the targeted deployment of platforms or sensors that preferentially sample at a particular level or layer. ETKF guidance maps for the same 2 d forecasts as section 3.4

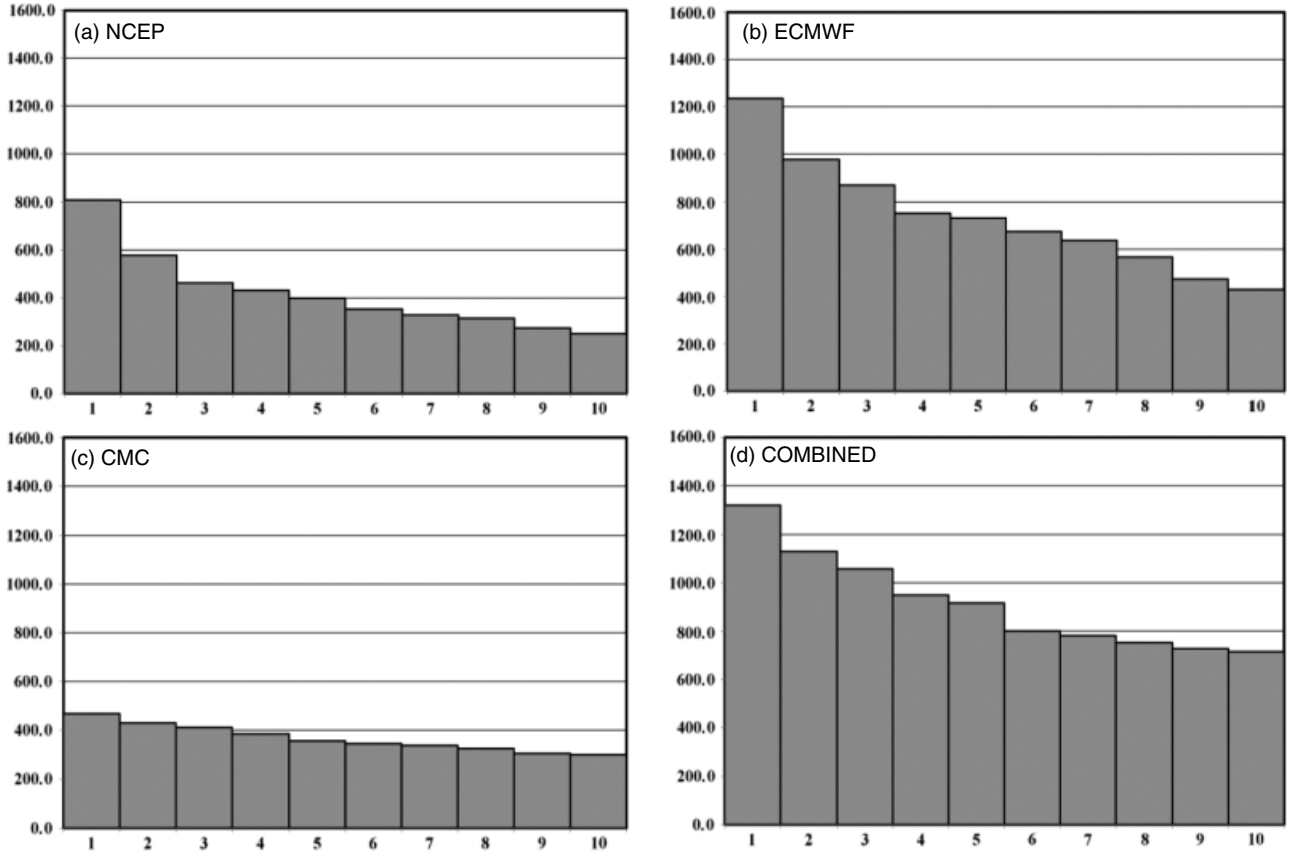

Figure 5. Spectra of the leading ten eigenvalues of the ensemble-based routine analysis error covariance matrix for (a) NCEP, (b) ECMWF, (c) CMC and (d) combined NCEP+ECMWF+CMC ensembles, valid at 0000 UTC on 11 September 2008. The axisymmetric component of the flow has been removed within 1000 km of typhoon *Sinlaku* in each ensemble member.

are illustrated in Figure 8, targeting horizontal winds and temperature at 0000 UTC on 11 September 2008 at the 850, 500 and 200 hPa levels, using the combined 89-member ensemble with the axisymmetric component of the vortex removed.

Several differences between the guidance products for targeting horizontal wind are evident at different levels (Figures 8(a–c)). First, at 200 hPa, the targets correspond to the trough to the northeast of *Sinlaku*, the confluence of the subtropical jet, and the anticyclonic outflow of *Sinlaku*. There is little sensitivity in the midlatitudes, even though significant wind variance exists (Figure 4(d)). When the axisymmetric component of *Sinlaku* is not removed, *Sinlaku* is the dominant target, though the other targets mentioned here are still significant (figure not shown). At 500 hPa, the main target is more local around *Sinlaku*, and the western periphery of the mid-level subtropical ridge ('C') east of *Sinlaku* (Figure 8(b)). On the preceding day, when the cold-core low ('A') north of *Sinlaku* is more prominent at the middle levels, this cold-core low is more of a significant target. At 850 hPa, *Sinlaku* and its immediate environment including an area of low-level confluence are primary targets (Figure 8(c)). The guidance over all levels (Figure 7(d)) logically represents a combination of the guidance at the respective individual levels. The corresponding guidance without the vortex removal reveals a near-circular target corresponding to *Sinlaku* at 850 and 500 hPa, with no remote targets at the scale of Figure 8.

The ETKF sensitivity to temperature observations is in distinct contrast to the guidance based on wind observations alone. While the primary upper-level target is the trough to the northeast of *Sinlaku* in an area of high temperature gradient (Figure 8(d)), which is collocated with one of the primary targets for wind, significant targets for upper-level temperature exist in the midlatitudes. The lower-level targets for temperature appear unstructured, lying mostly in baroclinic regions over central China and towards the north (Figure 8(f)). The targets for specific humidity (not shown) are weak and highly localized, sometimes near to but often remote from the TC in areas of high gradient. It is possible that the remote midlatitude targets for temperature and specific humidity are spurious for a cyclone in the Tropics. Overall, the wind sensitivity in Figures 8(a–c) dominates the temperature and humidity sensitivities, given that the guidance with the combined targeting variables closely resembles those for wind observations. The contribution of temperature, and to a lesser extent specific humidity, is to introduce new targets in the midlatitudes. It is worth noting that, during this stage of *Sinlaku*'s life cycle, there is no significant temperature or moisture advection that is affecting the TC, unlike the later stages.

4.2. *Stages of life cycle*

The ETKF guidance products for 2 d forecasts throughout the life cycle of *Sinlaku* are now explored, using the combined 89-member ensemble. Prior to genesis, the guidance is widespread, likely with multiple spurious targets (Figure 9(a)). Compared with later stages in the life cycle,

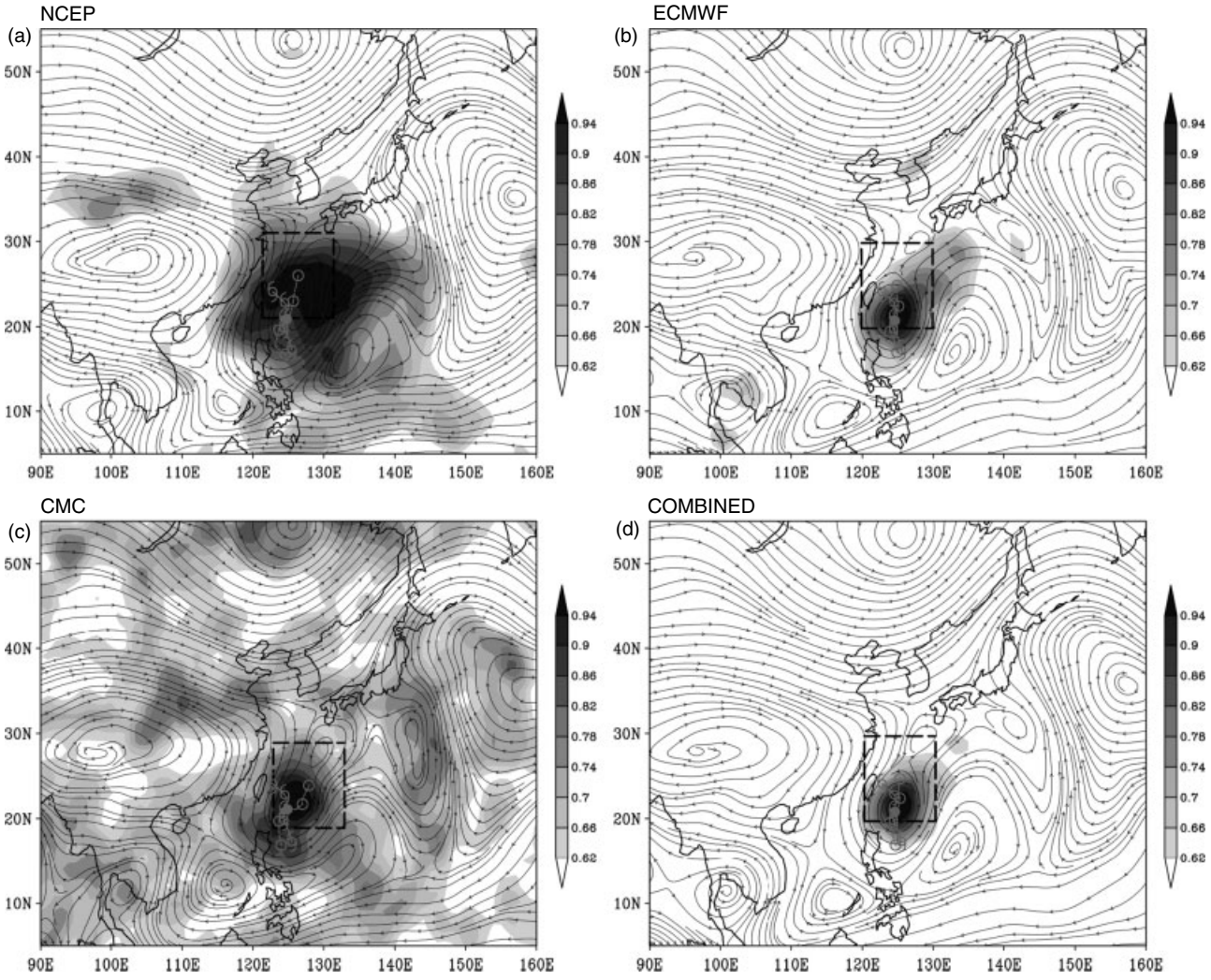

Figure 6. ETKF summary map guidance for a 2 d forecast of typhoon *Sinlaku*, for a targeted observing time of 0000 UTC on 11 September 2008, using (a) NCEP, (b) ECMWF, (c) CMC and (d) combined ensembles. The (shaded) values are normalized by the maximum value in the domain. The streamlines correspond to the ensemble mean averaged over 850, 500 and 200 hPa at the targeting time. The best track (tropical cyclone symbols) and ensemble mean track (circles) are shown every 24 h. The verification region is represented by the dashed box. No component of the flow has been removed in any of the ensemble members.

there is higher emphasis on targeting in the deep Tropics. An area of sensitivity exists in the broad monsoon trough ('D') to the southwest of the developing depression. Another target is the area associated with the wave that developed into *Sinlaku*. ETKF guidance for developing TCs has not been examined in any detail to date. It is also worth noting for this case that the ECMWF ensemble initialized on 0000 UTC on 6 September 2008 did not include diabatic singular vectors targeted on the area, since *Sinlaku* was not a tropical cyclone at this time. Moving forward to two days later, when *Sinlaku* was intensifying rapidly but moving slowly, the targets were now associated with the asymmetric field of *Sinlaku*, and the upper-level cold-core low to the north ('A') (Figure 9(b)). Multiple other targets existed in the Tropics and midlatitudes, and it is again likely that those targets in midlatitudes ('G' and 'H') are spurious. The targets for the next day are presented in Figures 7 and 8, and the corresponding targets for a day later, when *Sinlaku* is still moving slowly towards Taiwan but with little or no change in intensity, are presented in Figure 9(c). The guidance is more local to *Sinlaku* than in Figures 9(a,b), with the outflow and westerly winds on the northern side of the subtropical ridge east of *Sinlaku* ('C') being the main target region outside the storm itself. The cold-core low ('A') has evolved into a deep trough downstream, and is only a minor target. At this time, the subtropical ridges 'B' and 'C' were more prevalent, with a mostly zonal jet to the north of *Sinlaku*. Moving forward to targeting at the time at which *Sinlaku* begins its recurvature, the targets become slightly better defined (Figure 9(d)). A weak upstream short wave to the northwest of the recurving TC is identified as one target, together with the midlatitude jet, outflow and the western periphery of the Pacific subtropical ridge which has extended to the north and expanded in size. During the middle period post-recurvature, the targets become concentrated on the storm, and areas of the mostly zonal midlatitude jet upstream and particularly downstream over Japan (Figure 9(e)). Finally, as *Sinlaku* began to undergo extratropical transition, there is no longer a target over *Sinlaku*. The targets are concentrated primarily in the midlatitude jet, including an upstream shortwave over China and a trough downstream (Figure 9(f)). The Pacific subtropical ridge has extended westward towards China, and any targets remaining in the Tropics, such as the monsoon trough and the wave that develops into typhoon *Hagupit* are likely spurious.

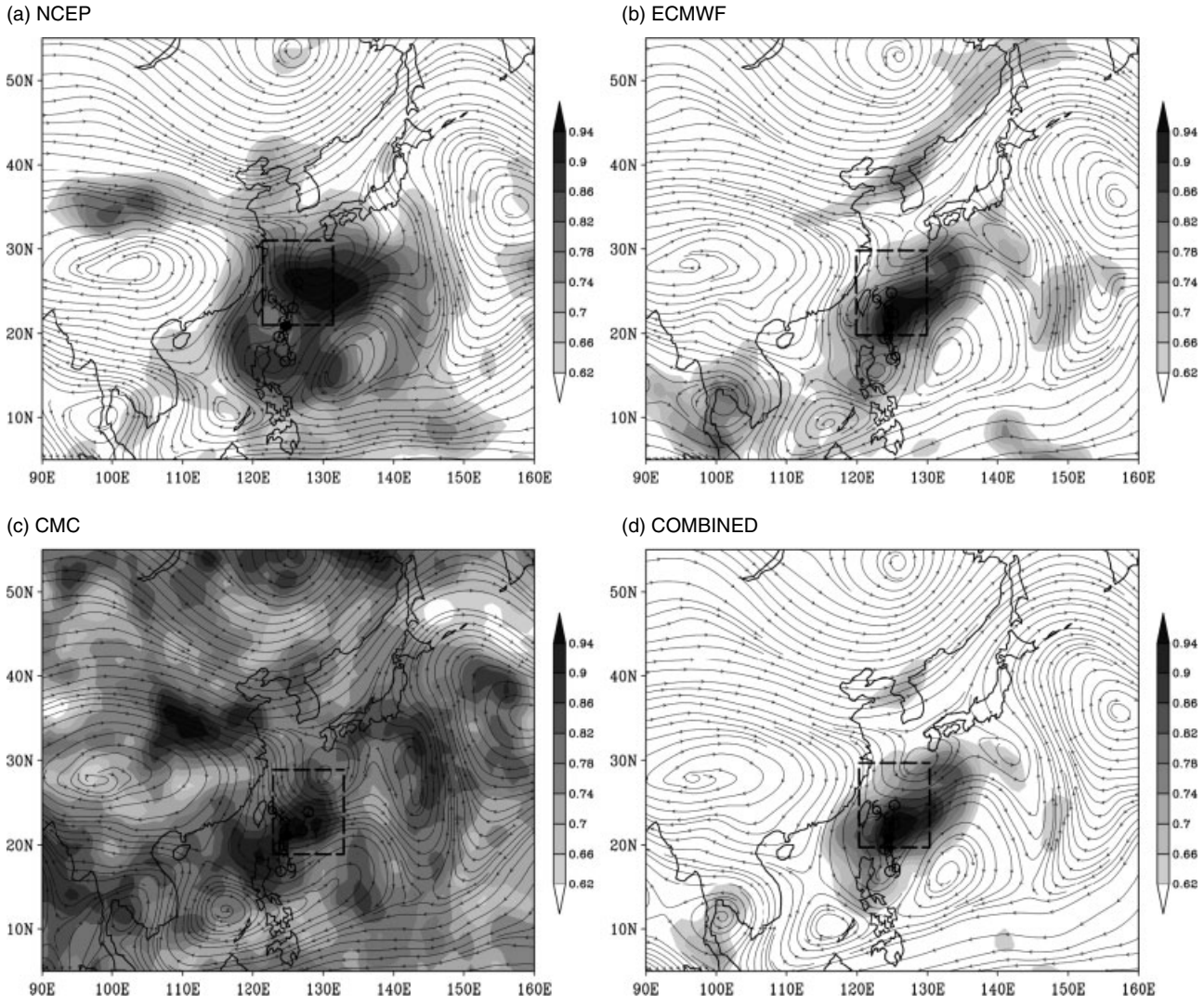

Figure 7. As Figure 6, but with the axisymmetric component of the flow removed within 1000 km of typhoon *Sinlaku* in each ensemble member.

4.3. *Varying forecast times* ($t_v - t_a$)

Given that the error statistics in the ETKF rely on linear dynamics, the theoretical assumptions may be violated as the forecast time is extended beyond 2 d. Nevertheless, as demonstrated by Sellwood *et al.* (2008) and Majumdar *et al.* (2010), the ETKF guidance may be plausible out to seven days in the midlatitude storm track. No quantitative evaluations have yet been performed for TCs, and will be left for future research. Here, the qualitative nature of the ETKF targets as the forecast time is extended out to +84 h is examined, for a verification time at which *Sinlaku* was expected to reach southern Japan. As the 'optimization' time interval $t_v - t_a$ is changed, the targets are also expected to change, with different dynamical processes dominating depending on the optimization time. The guidance is presented here in reverse chronological order, beginning from when the targeted observation and verification times are identical ($t_a = t_v$). In this case, a single target lies mostly within the verification region. As the forecast time is increased to 12 h, the target area expands slightly upstream from the verification region (Figure 10(a)). Extending the forecast time to 36 h, the primary target outside *Sinlaku* is further upstream, over mainland China. A target extends downstream in the westerly jet through Korea and southwestern Japan (Figure 10(b)). At 60 h, *Sinlaku* remains a target, and the targets in the westerly jet have become better defined further upstream and downstream (Figure 10(c)). However, multiple targets, possibly spurious, have also become increasingly evident. For a forecast time of 84h, the targets are still discernible and have extended further along the jet, west of China (Figure 10(d))

Similar guidance was computed for earlier stages in *Sinlaku*'s life cycle. Though the primary target always resides in the verification region when the forecast time is 0 h, the targets spread out in multiple directions even with a forecast time of 1 d. The evolution of the targets with increasing forecast time is often ambiguous. The most convincing cases for targeting beyond 2 d are generally when the TC is interacting with a trough upstream, with other synoptic systems playing a negligible role.

4.4. *Varying lead-time* ($t_a - t_i$)

As mentioned earlier, a lead-time (or ensemble age), $t_a - t_i$ of at least 2 d is normally required for aircraft mission planning or satellite rapid-scan activation. The ETKF guidance therefore needs to be computed using an ensemble that is initialized at least 2 d prior to the targeted observing time. The accuracy of the ETKF guidance may be

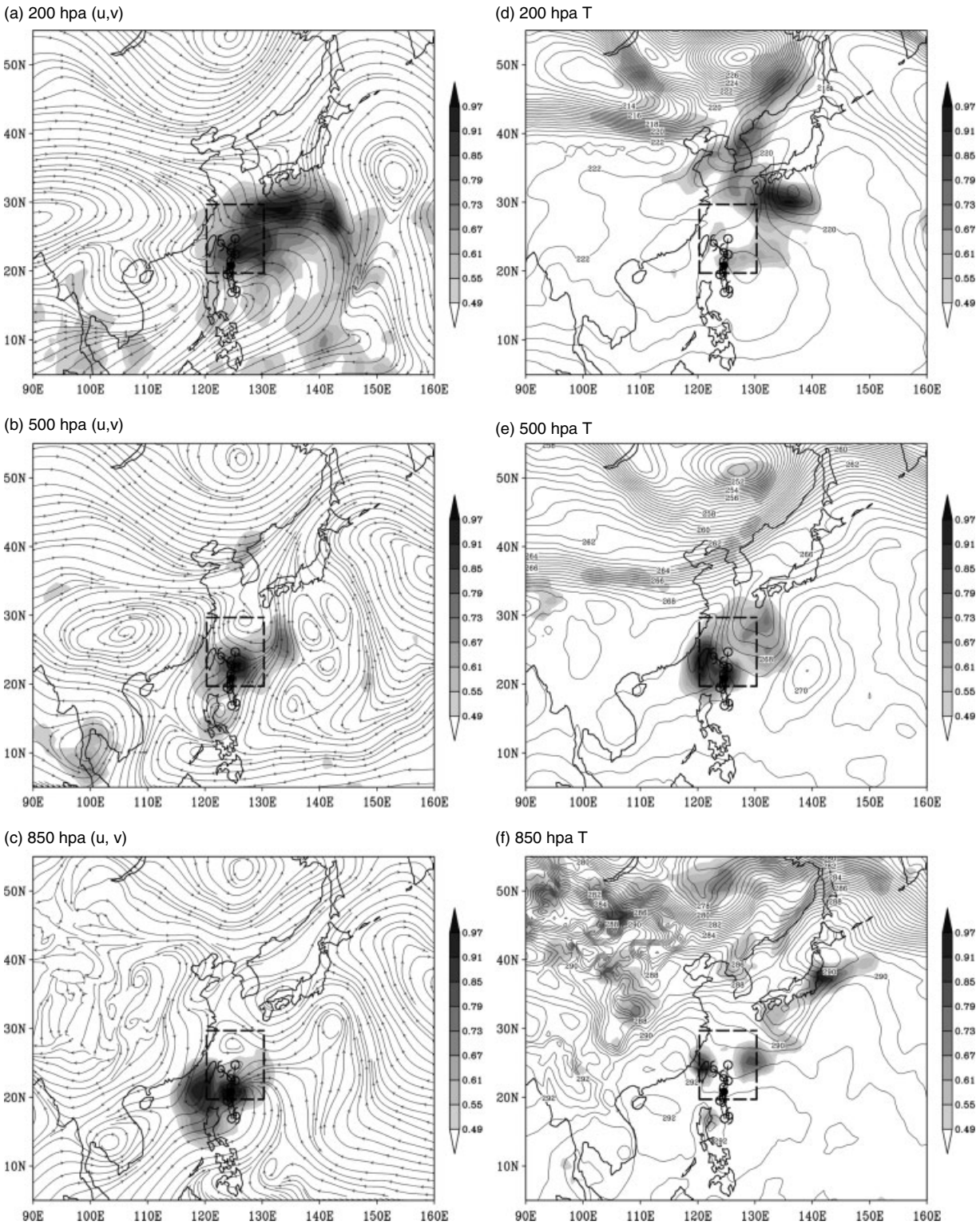

Figure 8. ETKF guidance for the same case as Figure 7, for targeted observations of horizontal wind at (a) 200 hPa, (b) 500 hPa and (c) 850 hPa (streamlines at each level). (d)–(f) are as (a)–(c), but for temperature observations (contours with interval 0.5 K). The best track (tropical cyclone symbols) and ensemble mean track (circles) are shown every 24 h.

compromised as the ensemble gets older, given the inevitable poor specification of error covariance with perturbations that evolve far longer than those assumed in an ensemble Kalman filter data assimilation scheme. For the same case as in Figure 7, the sensitivity to ensemble lead time is evident in Figure 11. If the lead time is 3 d, as is common for DOTSTAR mission planning, the target is dominated by the cold-core low ('A') and the northern periphery of the subtropical ridge ('C') (Figure 11(a)). However, if the lead time is reduced to 1 d (Figure 11(b)), the cold-core low is now predicted to be significantly weaker than that in the 2 d (Figure 7(d)) or 3 d ensemble forecasts. For the 1 d lead time, the target is focused more on *Sinlaku* and the western periphery of the subtropical ridge ('C') (Figure 11(b)). Therefore, a flight path from a mission such as DOTSTAR may have been modified, although not drastically, if the track were able

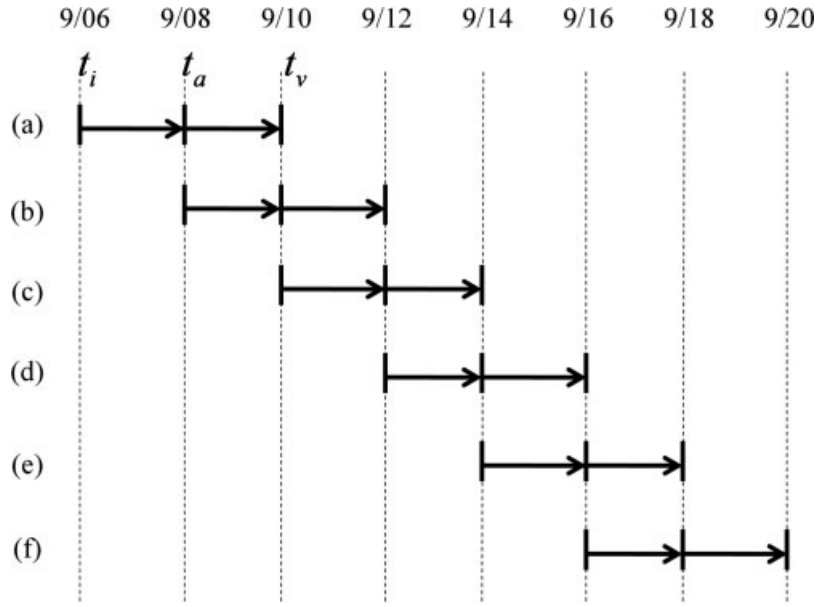

Figure 9. ETKF guidance for targeting (horizontal wind, temperature and specific humidity) for the combined ensemble, for two-day forecasts through the life cycle of typhoon *Sinlaku*, as illustrated in the timeline above. The (shaded) values are normalized by the maximum value in the domain. The streamlines correspond to the ensemble mean averaged over 850, 500, and 200 hPa at the targeting time. The best track (tropical cyclone symbols) and ensemble mean track (circles) are shown every 24 h.

to be designed less than a day before the mission. Overall, in contrast to the sensitivity of the target regions to the optimization time $t_v - t_a$, there is generally lower sensitivity to the lead time $t_a - t_i$.

5. Discussion

In the planning of special aircraft- and satellite-based deployments to improve TC forecasts, it is necessary to have a firm grasp of the available methodologies used to 'target' the observations. In order to progress beyond the obvious interpretation of the primary ETKF guidance, the axisymmetric component of the vortex has been removed from all ensemble members in this study, and the comparison with ensemble variance, and sensitivity to forecast variable, targeting level, stage of development and forecast time has been assessed for one typhoon case. Similar computations have been performed for a large number of cases (Wu *et al.*, 2009a), albeit with the axisymmetric vortex included.

The main consequence of removing the axisymmetric part of the vortex is to give emphasis to targets associated with asymmetries in the TC environment, and local and remote environmental features that may influence the track. Such targets are often masked (and therefore ignored) when the full vortex is present in the ensemble members. In contrast to Aberson *et al.* (2010), who determined that targets local to the TC were similar in the ensemble variance and the ETKF, this study has demonstrated that considerable differences exist between guidance provided by the upgraded ETKF and the ensemble variance. In addition to the TC asymmetries, the ETKF identifies the outflow and adjacent features such as the subtropical ridge and nearby upper-level troughs or low pressure systems as targets. Multiple remote targets exist in the deep Tropics, such as the monsoon trough and nearby easterly waves. As the TC undergoes recurvature, the primary targets move to the midlatitude storm track, both upstream and downstream of the TC. While extratropical transition (ET) has not been explored here, it has recently been shown that ET events are associated with plumes of large forecast uncertainty (Harr *et al.*, 2008; Anwender *et al.*, 2008), and ensemble sensitivity methods offer further insights (Torn and Hakim, 2009b).

The results in this article are in distinct contrast to those of Petersen *et al.* (2007), who found that ETKF targets for short-range forecasts of winter weather were generally unique and corresponded to an area of high gradient in the mass-momentum field, such as an upstream winter storm system or baroclinic zone. Also in contrast to our findings, Petersen *et al.* (2007) found little change to the structure of the targets in the vertical, and their conclusions did not vary even if the ensemble size was reduced to ten members. We suggest that the differences between our results and those of Petersen *et al.* (2007) are primarily due to the relatively low ensemble-based dimensionality of the flow in middle latitudes, as illustrated in Oczkowski *et al.* (2005), thereby producing more unique, and less ambiguous targets than in the Tropics, even when the ensemble size is small. A result consistent with Petersen *et al.* (2007) was that the ETKF targets exhibited sensitivity to the lead-time of the ensemble, demonstrating that the necessary operational lead-times may affect the accuracy of the guidance.

The results of Sellwood *et al.* (2008) and Majumdar *et al.* (2010), which demonstrated that ETKF targets in the midlatitude storm track could be traced upstream out to 7 d in zonal flow regimes, are unlikely to be corroborated for TCs except in cases of a distinct and unambiguous interaction with an approaching midlatitude trough. This higher ambiguity in ETKF targets for TCs heightens the need for careful interpretation by field programme planners. The effects of spurious correlations appear to be more acute in the Tropics than in midlatitudes, where the ensemble-based correlations can be interpreted dynamically (Ancell and Hakim, 2007; Hakim and Torn, 2008). The multiple targets, some of them spurious, render it challenging to provide common conclusions about the dynamical meaning of the

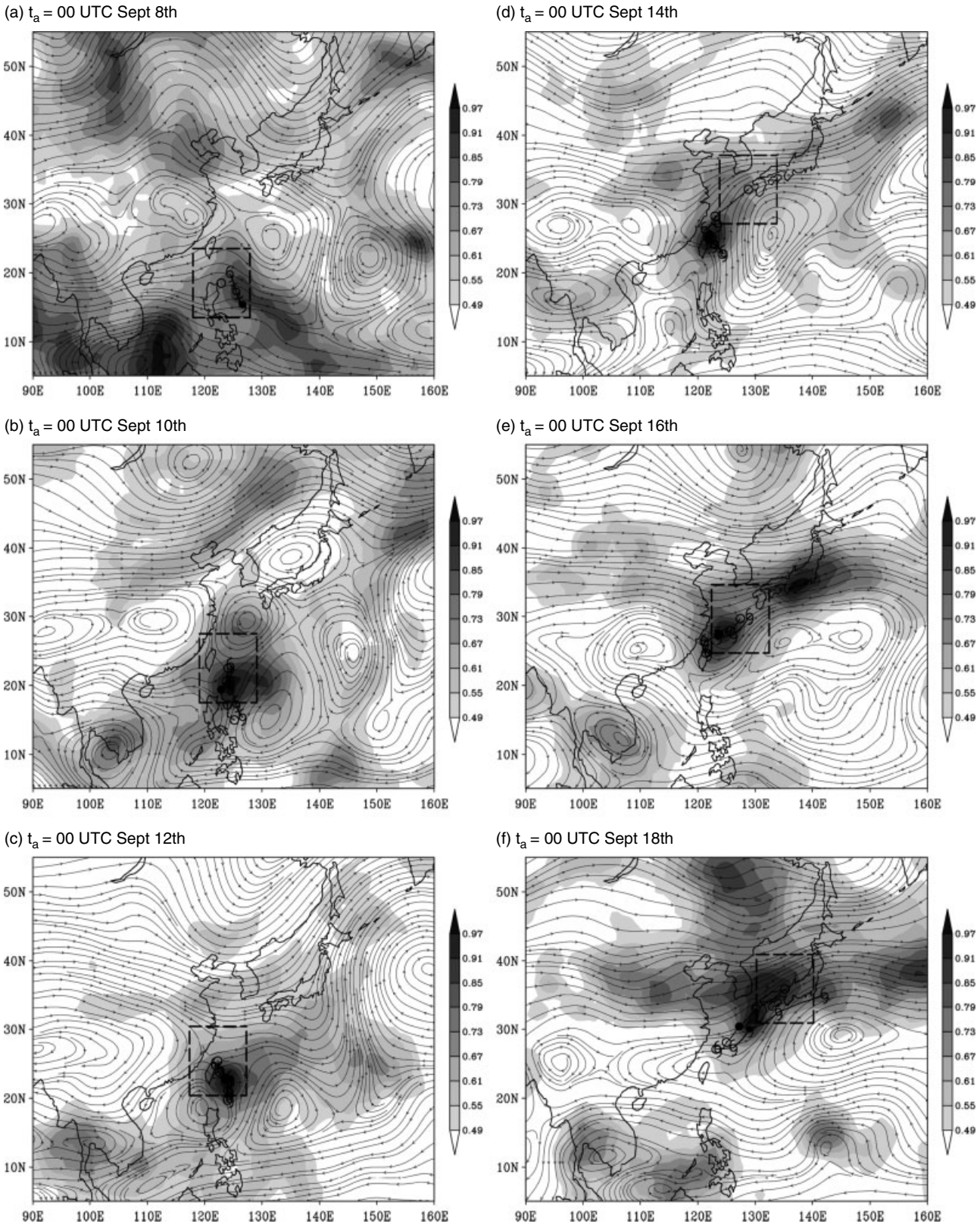

Figure 9. (continued)

targets over multiple cases, hence the descriptive nature of this article.

As reported in previous articles, some other targeting methods applied to TCs yield a clearer interpretation. For example, targets based on total-energy singular vectors often correspond to an annulus around the TC in which the gradient of potential vorticity is changing sign, and also remote locations towards the northwest of the TC (Peng and Reynolds, 2006; Reynolds *et al.*, 2007; Chen *et al.*, 2009; Kim and Jung, 2009; Reynolds *et al.*, 2009). The total-energy singular vectors, which exhibit these common characteristics regardless of the forecast model (Majumdar *et al.*, 2006; Wu *et al.*, 2009a), reflect *dynamical analysis sensitivity*, and therefore no assumption about the

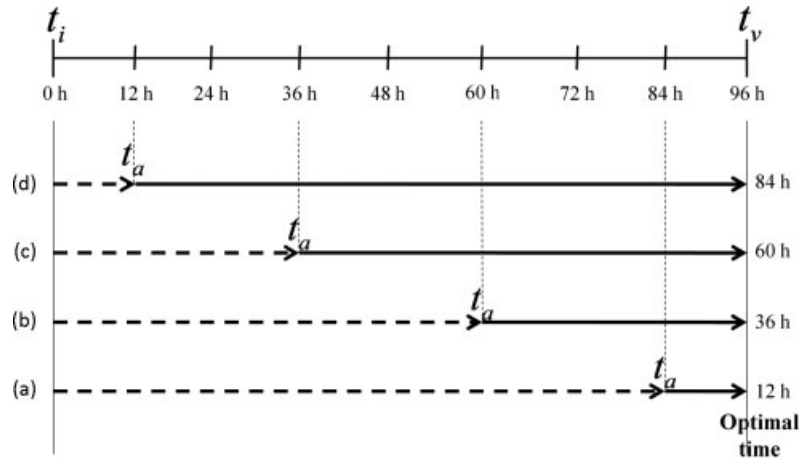

Figure 10. ETKF guidance for targeting (horizontal wind, temperature and specific humidity) for the combined ensemble, varying the time of targeted observations. The values are normalized by the maximum value in the domain. The streamlines correspond to the ensemble mean averaged over 850, 500, and 200 hPa at the targeting time. The ensemble initialization time is 0000 UTC on 15 September 2008 in all cases, and the verification time is 0000 UTC on 19 September 2008 in all cases. The time indicated on the figure is the difference between the targeting and verification times, as also illustrated on the timeline. The best track (tropical cyclone symbols) and ensemble mean track (circles) are shown every 24 h.

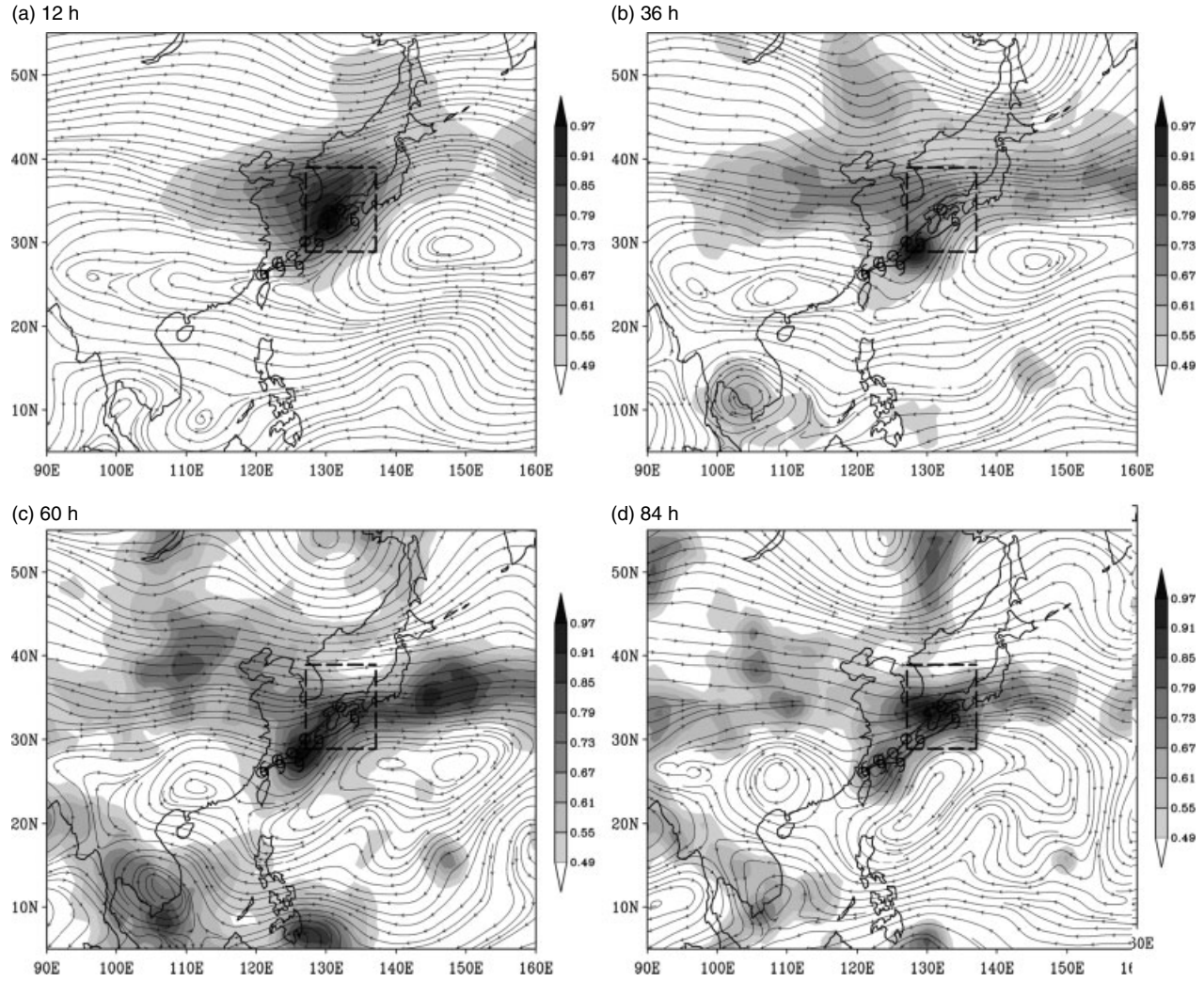

Figure 10. (continued)

data assimilation is included. However, if the singular vectors are constrained by an ensemble-based analysis-error covariance metric, their targets become more scattered, including targets downstream of the TC ('VARSV' targets in Figure 11 of Reynolds *et al.*, 2007). Since the effects of assimilating targeted observations on forecasts rely heavily on the data assimilation scheme, it is prudent to account for data assimilation in the targeting strategy, therefore reflecting *observation sensitivity* (e.g. Langland and Baker, 2004). This can lead to a less intuitive dynamical

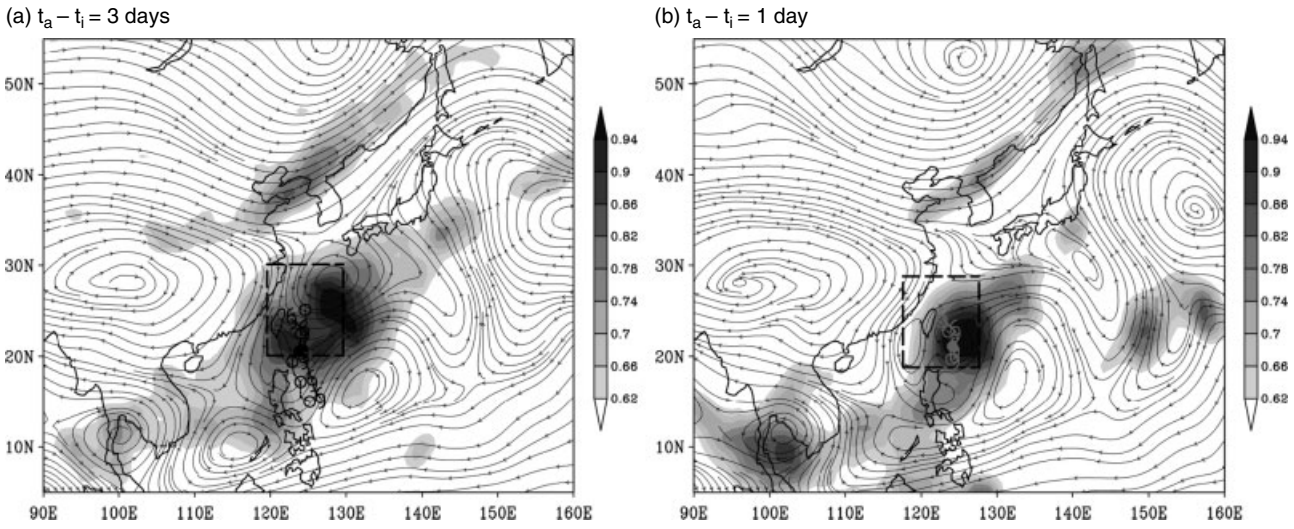

Figure 11. ETKF summary map guidance for a 2 d forecast of typhoon *Sinlaku*, for a targeted observing time of 0000 UTC on 11 September 2008, using the combined ensemble. The ensemble is initialized (a) 3 d and (b) 1 d prior to the targeted observing time. The corresponding plot for 2 d is Figure 7(d). The values are normalized by the maximum value in the domain. The streamlines correspond to the ensemble mean averaged over 850, 500, and 200 hPa at the targeting time. The best track (tropical cyclone symbols) and ensemble mean track (circles) are illustrated for every 24 h.

interpretation of the guidance, as is evident in the VARSVs and the ETKF. It should also be noted that some common interpretations of the ETKF and SV targets do exist, such as the confluence of the flow associated with the TC and also neighbouring synoptic systems such as the subtropical high and midlatitude jet (Chen *et al.*, 2009). As with the SVs, the ADSSV method focuses on dynamical features that may affect the TC motion, and it has been shown to capture the TC, the subtropical ridge adjacent to the TC and upstream (but not downstream) locations in midlatitudes (Wu *et al.*, 2009b). Again, some common interpretations with the ETKF exist, but some significant differences between the targets are also evident. Theoretical relationships and connections between adjoint- and ensemble-based methods have been proposed by Leutbecher (2003), Majumdar *et al.* (2006) and Ancell and Hakim (2007). Comparisons between ETKF and ensemble sensitivity guidance (Torn and Hakim, 2008), which possess some theoretical commonality, are under way.

The ETKF differs from other targeting strategies in its identification of targets north or northeast of the TC as it undergoes recurvature. While it is commonly accepted that the track of a TC is modified dynamically by the immediate environment or an approaching feature from the northwest, the errors in long-wave patterns and the upstream modification of such patterns by initial disturbances far downstream may still be significant. In the midlatitude storm track, ongoing work by the lead author suggests that an initial perturbation created in a region such as a short wave trough can propagate upstream and downstream, consistent with the findings of Simmons and Hoskins (1979) who showed that upstream and downstream developments occur in nonlinear simulations of an initially local perturbation in zonal baroclinic flows. Furthermore, McTaggart-Cowan *et al.* (2003) found that a TC may undergo reintensification or extratropical transition due to the influence of a zonal jet slightly downstream, suggesting that downstream sensitivity is possible. From a Kalman filtering perspective, a high ensemble-based covariance can exist between the TC and a wave trough downstream, due to the relatively large amplitude and ensemble spread in both regions. Further investigations are required to determine whether ETKF targets far downstream of a TC are genuine or spurious.

As was evident in the comparison between ETKF targets produced using the NCEP, ECMWF and CMC ensembles in section 3, the targets can differ significantly depending on the ensemble used. The dynamics of ensemble perturbation evolution and growth can be very different in the respective ensembles (Yamaguchi and Majumdar, 2010), leading to a vastly different distribution of variance and covariance in each ensemble. The construction of an ensemble perturbation method that produces an appropriate magnitude and distribution of spread in both the midlatitudes and Tropics has proven to be challenging (McLay *et al.*, 2008). The differences between the respective ensembles compound the difficulty in making general conclusions about the characteristics of ETKF targets.

As noted above, the effectiveness of a targeting strategy is dependent upon the data assimilation scheme. An analysis increment produced by a scheme such as that of NCEP is quasi-isotropic (Kleist *et al.*, 2009), while an equivalent increment that exploits the flow-dependent and anisotropic error covariance of an Ensemble Kalman Filter possesses a structure based on the dynamics of the flow (Torn and Hakim, 2009a). The ETKF produces increments similar to the latter, although the lack of covariance localization in the ETKF permits larger-scale correlations which may be spurious. An exploratory method to reduce the likelihood of spurious remote targets is via an artificial reduction of the magnitude of the ensemble perturbations at large distances from the TC. An example is illustrated in Figure 12, in which each element of the matrix of analysis ensemble perturbations pertaining to the routine observational network $[\mathbf{Z}^{\mathrm{r}}(\mathrm{t_a}|H^{\mathrm{i}})]_{\mathrm{mn}}$ is substituted by

$$[\mathbf{Z}^{\mathrm{r}}(t_{\mathrm{a}}|H^{\mathrm{i}})]_{mn} \exp\left\{\frac{-\alpha(R-1000)}{\max R}\right\},$$

where R is the distance between the location of the (mn)th element and the ensemble mean position of the TC. The factor α is chosen to ensure that the perturbations decay to half their original magnitude at R = 5000 km. Since the

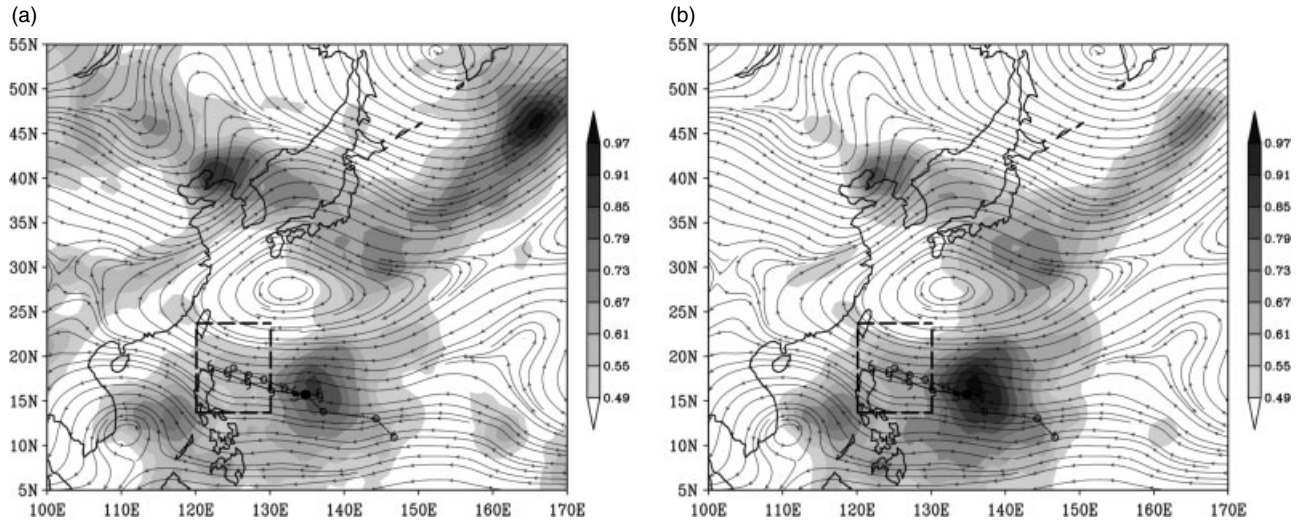

Figure 12. ETKF summary map guidance for a 2 d forecast of typhoon *Nuri*, for a targeted observing time of 0000 UTC on 18 August 2008, using the combined ensemble. In (a), the axisymmetric component of *Nuri* is removed. In (b), the same plus the distance-dependent rescaling of routine analysis ensemble perturbations is applied. The (shaded) values are normalized by the maximum value in the domain. The streamlines correspond to the ensemble mean averaged over 850, 500, and 200 hPa at the targeting time. The best track (tropical cyclone symbols) and ensemble mean track (circles) are shown every 12 h.

axisymmetric component of the TC has been removed, the primary target may be very remote in some cases, such as for typhoon *Nuri* (Figure 12(a)). With the reduction of the magnitude of the ensemble perturbations in this remote target area, the long-distance correlations are reduced, and the covariance of errors in the flow closer to the TC is thereby emphasized via a new transformation matrix $\mathbf{T}^{q}$, leading to a more local target (Figure 12(b)). Note that the ensemble perturbations at the verification time are modified linearly via the same transformation matrix, but there is no according nonlinear modification of the perturbation structures themselves. Nevertheless, such a method, with some refinement, may be practical in emphasizing those error covariance structures that are most likely to be exploited by the targeted observations, giving less ambiguous guidance to mission planners.

The horizontal and vertical structures of analysis and forecast errors in the Tropics also require further investigation, along the lines of the investigation by Hakim (2005) for the midlatitudes. It is also worth noting that data-denial experiments to evaluate the influence of assimilating a group of observations on a forecast may be contaminated by initially tiny differences that amplify and spread rapidly over 1–2 d, and this may be particularly acute in convective areas (Zhang *et al.*, 2003; Hodyss and Majumdar, 2007; Aberson, 2008). All present targeting strategies require improvement in being able to predict error structures in operational models, and accordingly represent the propagating effect of targeted observations based on the data assimilation scheme.

Several recommendations may be suggested via this study. First, local environmental features such as the upper-level outflow, low-level confluent flow near a TC or a neighbouring ridge or trough are proposed to be suitable for targeting, as is already being performed in field programmes. Additional observations, such as those from an airborne Doppler wind lidar (Weissmann *et al.*, 2005) in the near environment of the TC may provide an improved representation of local asymmetries in the wind structure. A second recommendation would be to consider the targeted sampling of remote synoptic features beyond the range of conventional aircraft. The targeted sampling of broad areas by satellite-borne instruments may reduce the error variance in such locations upstream and downstream, thereby improving the representation of the synoptic scale flow, as argued by Langland (2005). Third, evaluations of signal variance structures and the predicted influence of observations on a TC forecast are required.

6. Conclusions

The properties of target areas selected by a revised version of the ETKF used in tropical cyclone (TC) field programs have been investigated in this article, for typhoon *Sinlaku* (2008). Guidance was constructed from ensemble forecasts available in the TIGGE database, including NCEP, ECMWF and CMC, and a combination of the three individual ensembles. In order to reduce the heavy emphasis on the targets associated with the TC and to highlight sensitivity to the synoptic flow and asymmetries in the near environment, the primary axisymmetric circulation of the TC was removed from each ensemble member. It was found that the ETKF guidance for 2 d forecasts differed markedly from the associated ensemble variance, and different characteristics of the targets were exhibited for each ensemble. The ECMWF perturbations tended to overwhelm their NCEP and CMC counterparts, rendering the ETKF targets produced by the combined ensemble similar to those produced using ECMWF alone, even with a reduced ECMWF ensemble. For all ensembles, the targets were found to correspond to asymmetries in or near the TC, neighbouring features such as the adjacent subtropical ridge and a nearby trough, and far-field features such as the monsoon trough and waves in the deep Tropics, and areas of flow convergence and broad portions of the midlatitude jet. For forecast times beyond a day, multiple targets were evident, and it is unclear from this study which targets are genuine or spurious. The targets spread out further at longer forecast times. Additionally, the characteristics of the targets differ, although not dramatically, when ensembles initialized at different times are used for the same forecast case.

The commonly used ETKF guidance that assumed the assimilation of horizontal wind, temperature and specific humidity bore most resemblance to the guidance produced using wind alone. The upper-level targets corresponded more to the TC outflow and nearby and remote jets, while the lower-level targets corresponded to inflow near the TC. The guidance based on observing temperature produced remote targets, often in the midlatitudes. In the early stages of the TC life cycle, targets in the deep Tropics, such as the monsoon trough and easterly waves, were present. As the storm approached recurvature, targets in the midlatitude jet became increasingly evident, both upstream and downstream. The targets in the midlatitudes were traceable beyond a 3 d forecast time, offering the potential for using the ETKF beyond 2 d forecasts in clear scenarios of recurvature. However, the above results are difficult to generalize, given the complexity of multiple flows interacting with the TC. These results are in contrast to previous studies that revealed less ambiguous ETKF targets for winter weather events.

Further improvements to and evaluations of the ETKF are necessary. First, the spurious correlations associated with the TC can be reduced by including more ensemble members, and potentially by modifying the perturbations to produce a superior estimate of analysis-error covariance. The localization of signal variance around the observation sites, and new methods to propagate these structures forward in time requires investigation. Basic evaluations on the relative efficacy of assimilating observations sampled in ETKF target areas *versus* non-target areas are necessary. If these evaluations demonstrate that the ETKF can provide accurate quantitative estimates of signal variance and rank target areas in order of importance, the technique would be a suitable candidate for operational targeting of aircraft and satellite data to improve tropical cyclone forecasts.

Acknowledgements

The authors acknowledge funding from Ron Ferek and Chris Velden via the Office of Naval Research grant number N00014-08-1-0251. Funds to support a visiting scholarship for Shin-Gan Chen were provided by the National Science Council of Taiwan by grant NSC97-2917-I-002-147. The authors are grateful to Munehiko Yamaguchi for assistance with the TIGGE data, and to the THORPEX TIGGE team for creating and maintaining the TIGGE database. Detailed comments by two anonymous reviewers helped improve parts of the paper.

References

Aberson SD. 2003. Targeted observations to improve operational tropical cyclone track forecast guidance. *Mon. Weather Rev.* **131**: 1613–1628.

Aberson SD. 2008. Large forecast degradations due to synoptic surveillance during the 2004 and 2005 hurricane seasons. *Mon. Weather Rev.* **136**: 3138–3150.

Aberson SD. 2010. Ten years of hurricane synoptic surveillance (1997–2006). *Mon. Weather Rev.* **138**: 1536–1549.

Aberson SD, Majumdar SJ, Reynolds CA, Etherton BJ. 2010. An observing system experiment for tropical cyclone targeting techniques using the Global Forecast System. *Mon. Weather Rev.* In press. DOI: 10.1175/2010MWR3397.1.

Ancell B, Hakim GJ. 2007. Comparing adjoint- and ensemble-sensitivity snalysis with applications to observation targeting. *Mon. Weather Rev.* **135**: 4117–4134.

Anwender D, Harr PA, Jones SC. 2008. Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Case studies. *Mon. Weather Rev.* **136**: 3226–3247.

Barkmeijer J, Buizza R, Palmer TN, Puri K, Mahfouf J-F. 2001. Tropical singular vectors computed with linearized diabatic physics. *Q. J. R. Meteorol. Soc.* **127**: 685–708.

Bishop CH, Etherton BJ, Majumdar SJ. 2001. Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**: 420–436.

Bougeault P, Toth Z, Bishop CH, Brown B, Burridge D, Chen DH, Ebert B, Fuentes M, Hamill TM, Mylne K, Nicolau J, Paccagnella T, Park Y-Y, Parsons D, Raoult B, Schuster D, Silva Dias P, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S. 2010. The THORPEX Interactive Grand Global Ensemble (TIGGE). *Bull. Amer. Meteorol. Soc.* **91**: 1059–1072.

Buizza R, Richardson DS, Palmer TN. 2003. Benefits of increased resolution in the ECMWF ensemble system and comparison with poor-man's ensembles. *Q. J. R. Meteorol. Soc.* **129**: 1269–1288.

Burpee RW, Franklin JL, Lord SJ, Tuleya RE, Aberson SD. 1996. The impact of Omega dropwindsondes on operational hurricane track forecast models. *Bull. Amer. Meteorol. Soc.* **77**: 925–933.

Chen J-H, Peng MS, Reynolds CA, Wu C-C. 2009. Interpretation of tropical cyclone forecast sensitivity from the Singular Vector perspective. *J. Atmos. Sci.* **66**: 3383–3400.

Hakim GJ 2005. Vertical structure of midlatitude analysis and forecast errors. *Mon. Weather Rev.* **133**: 567–578.

Hakim GJ, Torn RD. 2008. Ensemble synoptic analysis. In *Synoptic-dynamic meteorology and weather analysis and forecasting: A tribute to Fred Sanders*. *Meteorol. Monogr.* **33**: 147–162.

Harnisch F, Weissmann M. 2010. Sensitivity of typhoon forecasts to different subsets of targeted dropsonde observations. *Mon. Weather Rev.* **138**: 2664–2680.

Harr PA, Anwender D, Jones SC. 2008. Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Methodology and a case study of typhoon *Nabi* (2005). *Mon. Weather Rev.* **136**: 3205–3225.

Hodyss D, Majumdar SJ. 2007. The contamination of 'data impact' in global models by rapidly growing mesoscale instabilities. *Q. J. R. Meteorol. Soc.* **133**: 1865–1875.

Kim HM, Jung B-J. 2009. Singular vector structure and evolution of a recurving tropical cyclone. *Mon. Weather Rev.* **137**: 505–524.

Kleist DT, Parrish DF, Derber JC, Treadon R, Wu WS, Lord S. 2009. Introduction of the GSI into the NCEP global data assimilation system. *Weather Forecasting* **24**: 1691–1705.

Langland RH. 2005. Issues in targeted observing. *Q. J. R. Meteorol. Soc.* **131**: 3409–3425.

Langland RH, Baker NL. 2004. Estimation of observation impact using the NRL atmospheric variational data assimilation system. *Tellus.* **56A**: 189–201.

Langland RH, Velden C, Pauley PM, Berger H. 2009. Impact of satellite-derived rapid-scan wind observations on numerical model forecasts of hurricane *Katrina*. *Mon. Weather Rev.* **137**: 1615–1622.

Leutbecher M. 2003. A reduced-rank estimate of forecast-error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.* **60**: 729–742.

Majumdar SJ, Finocchio PM. 2010. On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities. *Weather Forecasting* **25**: 659–680.

Majumdar SJ, Bishop CH, Etherton BJ, Toth Z. 2002. Adaptive sampling with the Ensemble Transform Kalman Filter. Part II: Field program implementation. *Mon. Weather Rev.* **130**: 1356–1369.

Majumdar SJ, Aberson SD, Bishop CH, Buizza R, Peng MS, Reynolds CA. 2006. A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Weather Rev.* **134**: 2354–2372.

Majumdar SJ, Bishop CH, Etherton BJ, Szunyogh I, Toth Z. 2001. Can an Ensemble Transform Kalman Filter predict the reduction in forecast error variance produced by targeted observations? *Q. J. R. Meteorol. Soc.* **127**: 2803–2820.

Majumdar SJ, Sellwood KJ, Hodyss D, Toth Z, Song Y. 2010. Characteristics of target areas selected by the Ensemble Transform Kalman Filter for medium-range forecasts of high-impact winter weather. *Mon. Weather Rev.* **138**: 2803–2824.

McLay JG, Bishop CH, Reynolds CA. 2008. Evaluation of the Ensemble Transform Analysis Perturbation Scheme at NRL. *Mon. Weather Rev.* **136**: 1093–1108.

McTaggart-Cowan R, Gyakum JR, Yau MK. 2003. The influence of the downstream state on extratropical transition: Hurricane *Earl* (1998) case study. *Mon. Weather Rev.* **131**: 1910–1929.

Oczkowski M, Szunyogh I, Patil DJ. 2005. Mechanisms for the development of locally low-dimensional atmospheric dynamics. *J. Atmos. Sci.* **62**: 1135–1156.

Palmer TN, Gelaro R, Barkmeijer J, Buizza R. 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **55**: 633–653.

Pellerin G, LeFaivre L, Houtekamer PL, Girard C. 2003. Increasing the horizontal resolution of ensemble forecasts at CMC. *Nonlin. Proc. Geophys.* **10**: 463–468.
Peng MS, Reynolds CA. 2006. Sensitivity of tropical cyclone forecasts as revealed by singular vectors. *J. Atmos. Sci.* **63**: 2508–2528.
Petersen GN, Majumdar SJ, Thorpe AJ. 2007. The properties of sensitive area predictions based on the ensemble transform Kalman filter (ETKF). *Q. J. R. Meteorol. Soc.* **133**: 697–710.
Puri K, Barkmeijer J, Palmer TN. 2001. Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Q. J. R. Meteorol. Soc.* **127**: 709–734.
Reynolds CA, Peng MS, Majumdar SJ, Aberson SD, Bishop CH, Buizza R. 2007. Interpretation of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Weather Rev.* **135**: 4006–4029.
Reynolds CA, Peng MS, Chen J-H. 2009. Recurving tropical cyclones: Singular vector sensitivity and downstream impacts. *Mon. Weather Rev.* **137**: 1320–1337.
Sellwood KJ, Majumdar SJ, Mapes BE, Szunyogh I. 2008. Predicting the influence of observations on medium-range forecasts of atmospheric flow. *Q. J. R. Meteorol. Soc.* **134**: 2011–2027.
Simmons AJ, Hoskins BJ. 1979. The downstream and upstream development of unstable baroclinic waves. *J. Atmos. Sci.* **36**: 1239–1254.
Szunyogh I, Kostelich EJ, Gyarmati G, Kalnay E, Hunt BR, Ott E, Satterfield E, Yorke JA. 2008. A local Ensemble Transform Kalman Filter data assimilation system for the NCEP global model. *Tellus.* **60**: 113–130.
Torn RD, Hakim GJ. 2008. Ensemble-based sensitivity analysis. *Mon. Weather Rev.* **136**: 663–677.
Torn RD, Hakim GJ. 2009a. Ensemble data assimilation applied to RAINEX: Observations of hurricane *Katrina* (2005). *Mon. Weather Rev.* **137**: 2817–2829.
Torn RD, Hakim GJ. 2009b. Initial condition sensitivity of Western Pacific extratropical transitions determined using ensemble-based sensitivity analysis. *Mon. Weather Rev.* **137**: 3388–3406.
Velden C, Daniels J, Stettner D, Santek D, Key J, Dunion J, Holmlund K, Dengel G, Bresky W, Menzel P. 2005. Recent innovations in deriving tropospheric winds from meteorological satellites. *Bull. Amer. Meteorol. Soc.* **86**: 205–223.
Wei M, Toth Z, Wobus R, Zhu Y. 2008. Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus.* **60A**: 62–79.
Weissmann M, Busen R, Dörnbrack A, Rahm S, Reitebuch O. 2005. Targeted observations with an airborne wind Lidar. *J. Atmos. Oceanic Technol.* **22**: 1706–1719.
Weissmann M, Harnisch F, Wu C-C, Lin P-H, Ohta Y, Yamashita K, Kim Y-H, Jeon E-H, Nakazawa T, Aberson SD. 2010. The influence of assimilating dropsondes on typhoon track and midlatitude forecasts. *Mon. Weather Rev.* In press. DOI: 10.1175/2010MWR3377.1.
Whitaker JS, Hamill TM, Wei X, Song Y, Toth Z. 2008. Ensemble data assimilation with the NCEP Global Forecast System. *Mon. Weather Rev.* **136**: 463–482.
Wu C-C, Lin P-H, Aberson SD, Yeh T-C, Huang W-P, Hong J-S, Lu G-C, Hsu K-C, Lin I-I, Chou K-H, Lin P-L, Liu C-H. 2005. Dropwindsonde observations for typhoon surveillance near the Taiwan region (DOTSTAR): An overview. *Bull. Amer. Meteorol. Soc.* **86**: 787–790.
Wu C-C. 2006. 'Targeted observation and data assimilation for tropical cyclone track prediction'. *Proceedings of 6th International Workshop on Tropical Cyclones*, WMO/CAS/WWW, San Jose, Costa Rica, 21–28 November, 409–423.
Wu C-C, Chou K-H, Lin P-H, Aberson SD, Peng MS, Nakazawa T. 2007a. The impact of dropwindsonde data on typhoon track forecasts in DOTSTAR. *Weather Forecasting* **22**: 1157–1176.
Wu C-C, Chen J-H, Lin P-H, Chou K-H. 2007b. Targeted observations of tropical cyclone movement based on the adjoint-derived sensitivity steering vector. *J. Atmos. Sci.* **64**: 2611–2626.
Wu C-C, Chen J-H, Majumdar SJ, Peng MS, Reynolds CA, Aberson SD, Buizza R, Yamaguchi M, Chen S-G, Nakazawa T, Chou K-H. 2009a. Intercomparison of targeted observation guidance for tropical cyclones in the northwestern Pacific. *Mon. Weather Rev.* **137**: 2471–2492.
Wu C-C, Chen S-G, Chen J-H, Chou K-H, Lin P-H. 2009b. Interaction of typhoon *Shanshan* (2006) with the midlatitude trough from both adjoint-derived sensitivity steering vector and potential vorticity perspectives. *Mon. Weather Rev.* **137**: 852–862.
Yamaguchi M, Iriguchi T, Nakazawa T, Wu C-C. 2009. An observing system experiment for typhoon *Conson* (2004) using a singular vector method and DOTSTAR data. *Mon. Weather Rev.* **137**: 2801–2816.
Yamaguchi M, Majumdar SJ. 2010. Using TIGGE data to diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Mon. Weather Rev.* **138**: 3634–3655.
Zhang F, Snyder S, Rotunno R. 2003. Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.* **60**: 1173–1185.
Zhang F, Weng Y, Sippel JA, Meng Z, Bishop CH. 2009. Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an Ensemble Kalman Filter. *Mon. Weather Rev.* **137**: 2105–2125.

