

Assimilation of wind speed and direction observations: a new formulation and results from idealised experiments

By XIANG-YU HUANG^{1*}, FENG GAO², NEIL A. JACOBS³ and HONGLI WANG¹, ¹National Center for Atmospheric Research, Boulder, CO, USA; ²Nanjing University of Information Science & Technology, Nanjing, China; ³AirDat LLC, Morrisville, NC, USA

(Manuscript received 23 October 2012; in final form 22 January 2013)

ABSTRACT

This article presents a new methodology for assimilating wind observations in their observed form of speed and direction, while taking into account both speed and direction error. It ensures the analysed speed and direction will be consistent with their background and observed values. The new formulation is implemented in the Weather Research and Forecasting Data Assimilation system, and idealised experiments are used to demonstrate the potential benefit. The results suggest that analyses from the new formulation are more reasonable when compared to the conventional methodology. The forecasts generated in these idealised experiments also demonstrate the value of this new formulation. Preliminary results from real data experiments are in general agreement with results presented here, and they will be reported in a following article.

Keywords: data assimilation, wind observations, observation errors, WRFDA, OSSE

1. Introduction

Current data assimilation systems, including the Weather Research and Forecasting Data Assimilation system (WRFDA; Barker et al., 2012), do not consider wind direction errors when assimilating wind observations. Despite being collected in the form of speed (*sp*) and direction (*dir*), wind observations are transformed to their longitudinal and latitudinal components, *u* and *v*, prior to assimilation. The observation errors of *u* and *v*, which may be estimated from the wind speed error, are used in the assimilation. For example, the default observation error table used in WRFDA assumes a constant 3.6 m s^{-1} for *u* and *v* at all levels for aircraft data, following the operational practice at National Centers for Environmental Prediction (NCEP) (Kalnay et al., 1996). Although the *dir* errors can impact the uncertainty of *u* and *v*, they cannot be considered during the assimilation process, and therefore have no independent influence on the assimilation results.

In addition, the conventional method assumes the *u* and *v* observation errors to be uncorrelated to simplify the observation error covariance matrix to a diagonal matrix. Since *u* and *v* components are not observed in the real

world, but are derived from observations of *sp* and *dir*, not only are *u* and *v* observation errors correlated, but they are both directly impacted by the *dir* observation errors. For observing systems directly observing *sp* and *dir* (e.g. surface stations), it is safe to assume the *sp* observation errors are not correlated with *dir* observation errors. For example, most operational 10-m synoptic observation stations use a rotating cup anemometer to measure *sp*, while *dir* is observed independently by a vane, thus isolating independent and uncorrelated characteristics of the respective observing systems. Newer sonic anemometers observe *sp*, and then calculate the *dir* angle from the vector difference before reporting time-averaged values for *sp* and *dir*. In this case, the errors are correlated; however, the *dir* error, despite being a function of *sp*, behaves differently (e.g. van der Molen et al., 2004), so treating *sp* and *dir* independently would still be beneficial when considering the error statistics during data assimilation.

Aircraft-based observing systems derive *sp* and *dir* observations from both air speed and heading information. The observed wind vector is the difference between the ground track vector (i.e. aircraft motion with respect to Earth) and the aircraft track vector. The ground track vector is typically determined by a very accurate satellite-based global positioning system (GPS). The aircraft track vector is calculated from the true air speed (TAS) and the heading angle, where the TAS is the difference between the

*Corresponding author.
email: huangx@ucar.edu

dynamic and static pressure. The heading angle is typically determined by a laser gyro heading system. The pitot tube system that determines TAS, and the heading system that determines the track angle, are considered two unrelated observing systems. Both systems provide information to calculate the wind observation, which is then broken down into u and v vector components. Once the observed sp and dir from the aircraft are converted into u and v , both the original sp error and the original dir error are imbedded in the u and v components, and are no longer recoverable.

This derivation, which uses the vector difference between ground speed and air speed, allows for some dir error, which originates from the heading instrumentation, to impact the sp magnitude. Therefore, based on the specific type of observation, it is theoretically possible for sp and dir errors to be correlated, but the degree of correlation is much less than the direct influence sp and dir errors have on the u and v error when u and v are converted from sp and dir .

A new method of directly assimilating sp and dir observations is proposed in this article and is referred to as asm_sd . It takes into account both sp and dir observation error. It ensures that the analysed sp and dir will be consistent with the background sp and dir , as well as the observed sp and dir . The new formulation and results are compared with the conventional method, which is referred to as asm_uv in this article. In section 2, both asm_uv and asm_sd are introduced using the WRFDA framework. Section 3 presents the differences between the asm_uv and asm_sd analyses using a highly simplified approach and WRFDA single observation pair experiments. Idealised WRFDA experiments are presented in section 4. Conclusions and plans for future work are discussed in section 5.

2. WRFDA and new formulation

WRFDA employs the incremental formulation of Courtier et al. (1994). The analysis increment minimises a cost function,

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}), \quad (1)$$

where the superscripts -1 and T denote the inverse and adjoint, respectively, of a matrix or linear operator; $\delta\mathbf{x} = \mathbf{x}^a - \mathbf{x}^b$ is the analysis increment, where $\mathbf{x}^a$ is the analysis and $\mathbf{x}^b$ is the background; $\mathbf{B}$ is the background error covariance; $\mathbf{d} = \mathbf{y} - H(\mathbf{x}^b)$ is the innovation vector; $\mathbf{y}$ is the observation vector; H is the non-linear observation operator; $\mathbf{H}$ is the linearised H ; and $\mathbf{R}$ is the observation error covariance. To minimise eq. (1), WRFDA computes the gradient of J ,

$$\nabla J = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}),$$

where $\mathbf{H}^T$, the adjoint of $\mathbf{H}$, is needed. A more detailed description of WRFDA can be found in Barker et al. (2004)

and Huang et al. (2009). Here, we only discuss the items related to wind observations.

The model state vector, $\mathbf{x}$, contains, among other variables, u and v ,

$$\mathbf{x} = (\dots, u, v, \dots)^T \quad (2)$$

and the observation vector, $\mathbf{y}$, contains the wind observations. In the conventional data assimilation system (asm_uv), wind observations, despite largely being collected in the format of dir^o and sp^o , are transformed to u^o and v^o before being assimilated; therefore, $\mathbf{y}$ can be written as

$$\mathbf{y}_{asm_uv} = (\dots, u^o, v^o, \dots)^T. \quad (3)$$

In the new formulation (asm_sd), wind observations are assimilated in their original form of dir^o and sp^o , and $\mathbf{y}$ is written as

$$\mathbf{y}_{asm_sd} = (\dots, dir^o, sp^o, \dots)^T. \quad (4)$$

Observation operators H , $\mathbf{H}$ and $\mathbf{H}^T$ provide mapping between the model state space and observation space, and can be expressed as a series of operators

$$\begin{aligned} H(\mathbf{x}) &= H_i(H_t(\mathbf{x})) \\ \mathbf{H} &= \mathbf{H}_i\mathbf{H}_t \\ \mathbf{H}^T &= \mathbf{H}_i^T\mathbf{H}_t^T, \end{aligned} \quad (5)$$

where subscript t denotes variable transforms and subscript i denotes interpolation. The difference between the asm_uv formulation and the asm_sd formulation is in the transform-related operators. For asm_uv , H_t , $\mathbf{H}_t$ and $\mathbf{H}_t^T$ are all identity operators:

$$\begin{aligned} H_t(\mathbf{x}) &= \mathbf{x} \\ \mathbf{H}_t &= \mathbf{I} \\ \mathbf{H}_t^T &= \mathbf{I}. \end{aligned} \quad (6)$$

For asm_sd , the non-linear observation operator, H , which computes sp and dir from u and v , is expressed as

$$\begin{aligned} \begin{pmatrix} sp \\ dir \end{pmatrix} &= H_t \begin{pmatrix} u \\ v \end{pmatrix} \\ sp &= \sqrt{u^2 + v^2} \\ dir &= n\pi + \arctan\left(\frac{u}{v}\right), \end{aligned} \quad (7)$$

where n is an integer, which is chosen to confine the dir part of the innovation vector $\mathbf{d}$ to within $[-180^\circ, +180^\circ]$. The tangent linear operator, $\mathbf{H}_t$, which computes

perturbations in sp and dir from perturbations in u and v , is expressed as

$$\begin{pmatrix} \delta sp \\ \delta dir \end{pmatrix} = \mathbf{H}_t \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} \quad (8)$$

$$\mathbf{H}_t = \begin{pmatrix} \frac{u}{\sqrt{u^2+v^2}} & \frac{v}{\sqrt{u^2+v^2}} \\ \frac{v}{u^2+v^2} & \frac{-u}{u^2+v^2} \end{pmatrix}.$$

The adjoint operator $\mathbf{H}_t^T$, which computes the gradients in u and v from the gradients in sp and dir , is expressed as

$$\mathbf{H}_t^T = \begin{pmatrix} \frac{u}{\sqrt{u^2+v^2}} & \frac{v}{u^2+v^2} \\ \frac{v}{\sqrt{u^2+v^2}} & \frac{-u}{u^2+v^2} \end{pmatrix}. \quad (9)$$

The errors of most observation types are assumed to be uncorrelated; thus, the observation error covariance, $\mathbf{R}$, becomes a diagonal matrix with observational error variance on the diagonal. In the `asm_uv` formulation,

$$\mathbf{R}_{asm_uv} = \begin{pmatrix} \cdot & 0 & 0 & 0 \\ 0 & \sigma_u^2 & 0 & 0 \\ 0 & 0 & \sigma_v^2 & 0 \\ 0 & 0 & 0 & \cdot \end{pmatrix}, \quad (10)$$

where σ_u^2 and σ_v^2 are the observational error variances in u and v , respectively. Despite observations being collected in sp and dir , it is common practice to ignore the dir error, and apply a single sp error to both u and v . In the `asm_sd` formulation, the observation error covariance, $\mathbf{R}$, is expressed as

$$\mathbf{R}_{asm_sd} = \begin{pmatrix} \cdot & 0 & 0 & 0 \\ 0 & \sigma_{sp}^2 & 0 & 0 \\ 0 & 0 & \sigma_{dir}^2 & 0 \\ 0 & 0 & 0 & \cdot \end{pmatrix}, \quad (11)$$

where σ_{sp}^2 and σ_{dir}^2 are observation error variances in sp and dir , respectively. In contrast to the approximations and assumptions made for σ_u^2 and σ_v^2 , estimations of σ_{sp}^2 and σ_{dir}^2 may be obtained directly (Gao et al., 2012).

3. Analysis difference between assimilation methods

3.1. A simple demonstration

To illustrate the difference between the two assimilation methods, we start with a scalar analysis by reducing the dimension of $\mathbf{x}$ and $\mathbf{y}$ to 1, and assuming the observation

and background are collocated. The solution of eq. (1) is easily obtained by

$$x^a = \frac{R}{B+R}x^b + \frac{B}{B+R}y, \quad (12)$$

where all the variables are scalar and otherwise defined the same as their vector equivalent.

From eq. (12), we see that the analysis is a weighted average of the background and the observation. The weights are determined by the ratio between the background error variance (B) and the observational error variance (R). In a special case, when $R/B = 1$, the analysis is simply midway between the background and the observation (i.e. $x^a = (x^b + y)/2$).

To further simplify the illustration, we assume u , v , dir and sp can be analysed independently. We will return to this assumption at the end of this subsection. We use eq. (12) to derive analyses for all of them separately. For `asm_uv`,

$$\begin{aligned} u_{uv}^a &= \frac{1}{2}(u^b + u^o) \\ v_{uv}^a &= \frac{1}{2}(v^b + v^o), \end{aligned} \quad (13)$$

and for `asm_sd`,

$$\begin{aligned} sp_{sd}^a &= \frac{1}{2}(sp^b + sp^o) \\ dir_{sd}^a &= \frac{1}{2}(dir^b + dir^o). \end{aligned} \quad (14)$$

where superscripts a , b and o denote analysis, background and observation, and subscripts uv and sd denote solutions by `asm_uv` and `asm_sd`, respectively.

The analyses are shown in Fig. 1. Starting from the same background (**BG**) wind vector and observation (**OBS**) wind vector, we obtain two very different analysis wind vectors using `asm_uv` and `asm_sd`. To derive the wind analysis using `asm_uv`, eq. (13) is used to calculate u^a and v^a , and the final result is the wind vector **ANA_{uv}**. Likewise, to produce

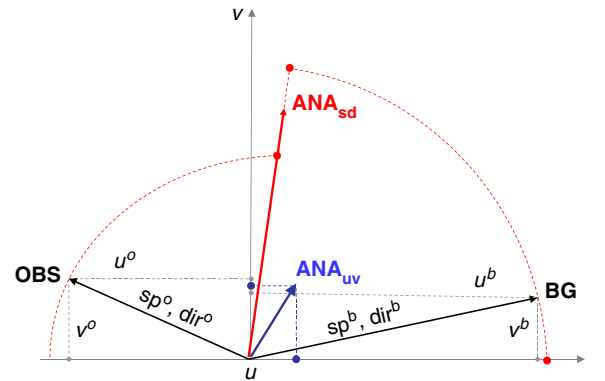


Fig. 1. The background wind vector, the observation wind vector, the analysis wind vector obtained by `asm_uv` (assimilating u and v) and the analysis wind vector obtained by `asm_sd` (assimilating sp and dir). For this illustration, we assume u, v, dir and sp can be analysed independently and $B/R = 1$.

the wind analysis using `asm_sd`, eq. (14) is used to calculate sp^a and dir^a , and the final result is the wind vector $\mathbf{ANA}_{sd}$.

It is evident that `asm_uv` and `asm_sd` could lead to very different analyses. The analysis vectors differ both in directions and in magnitude. The `asm_uv` produces (u^a, v^a) , which is consistent with (u^b, v^b) and (u^o, v^o) , while the `asm_sd` produces (sp^a, dir^a) , which is consistent with (sp^b, dir^b) and (sp^o, dir^o) . In the Appendix, it is shown that under certain conditions, the analysis error covariance of `asm_sd` equals that of `asm_uv`; however, as the wind observations are made with sp and dir , we believe the `asm_sd` analysis is more reasonable.

Under the assumptions stated above, it is easy to show that the analysis wind vector obtained by `asm_sd` is always longer than the analysis wind vector obtained by `asm_uv` when the background wind vector and the observation wind vector do not point in the same direction. The analysis wind speed, as a function of u^b, v^b, u^o and v^o , for both `asm_uv` and `asm_sd`, can be written as

$$\begin{aligned} sp_{uv}^a &= \sqrt{(u_{uv}^a)^2 + (v_{uv}^a)^2} = \frac{1}{2} \sqrt{(u^b + u^o)^2 + (v^b + v^o)^2} \\ sp_{sd}^a &= \frac{1}{2}(sp^b + sp^o) = \frac{1}{2} \left(\sqrt{(u^b)^2 + (v^b)^2} + \sqrt{(u^o)^2 + (v^o)^2} \right). \end{aligned} \quad (15)$$

An example of sp_{uv}^a/sp_{sd}^a is shown in Fig. 2. The background wind vector is chosen to be $(u^b, v^b) = (25.0, 5.0)$, which can

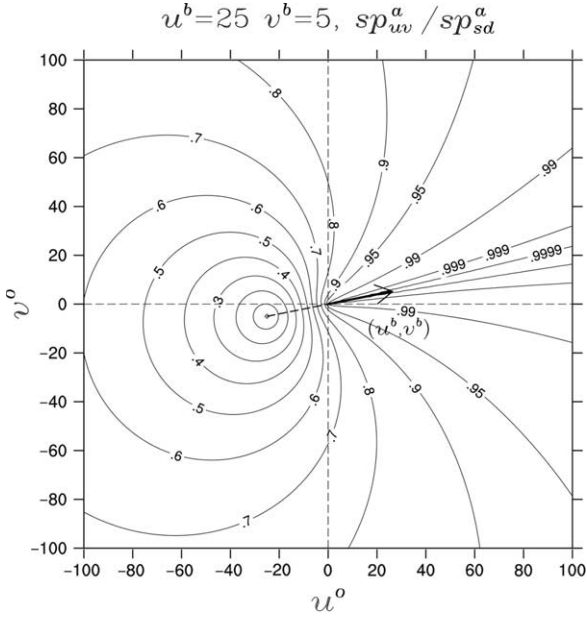


Fig. 2. The ratio of `asm_uv` analysis wind speed sp_{uv}^a to `asm_sd` analysis wind speed sp_{sd}^a as a function of u^o and v^o , which is computed using eq. (15). The background wind vector (u^b, v^b) is also shown.

also be expressed as $(sp^b, dir^b) = (25.5, 258.7)$. The observation wind vectors are varied across the range $(-100:100, -100:100)$, which were produced by 80,802 $[2 \times (201 \times 201)]$ experimental runs. The units used for u, v and sp are m s^{-1} and for dir is $^\circ$ (degree). The ratio is less than 1 with the exception of the case when the observation wind vector is exactly along the background wind vector, and in this case, the ratio equals 1. Thus, `asm_uv` produces a weaker wind analyses than `asm_sd`. In the extreme case of the observed wind vector being exactly opposed to the background wind vector, $(u^o, v^o) = (-25.0, -5.0)$, the `asm_uv` wind analysis vector has 0 magnitude (length), while the `asm_sd` wind analysis vector has the same magnitude (length; 25.5) as the background wind vector or the observed wind vector.

In this particular subsection of the study, an assumption is made that the background error and observational error are equal. In most operational data assimilation systems, the background errors are not typically expressed in terms of u and v . In WRFDA, they are estimated from model data in the form of a stream function and velocity potential. On the other hand, observational errors are expressed in terms of u and v (`asm_uv`). With the new option (`asm_sd`), the observational errors are given in terms of sp and dir . In both situations, it is not straightforward to make the background errors equal to observation errors. Additional experiments, discussed in the next subsection, do not make this same assumption.

The comparisons presented above are based on a univariate analysis performed independently on u, v, sp and dir . Operational data assimilation systems would not analyse u and v separately, so it is highly idealised to make this assumption with `asm_uv`. It is also questionable to assume the observational errors in u and v are uncorrelated, although this assumption has been made in all variational data assimilation systems. It is certainly unreasonable to make the uncorrelation assumption about the background error correlation. Similar arguments can be made for `asm_sd`, and the independent analysis of sp and dir . The motivation for this assumption was based on the simplicity of an idealised case. This assumption is not made in the following subsection and we will illustrate that the results from WRFDA single observation pair experiments are consistent with those we obtained above using the highly idealised approach.

3.2. WRFDA single observation pair experiments

In this subsection, we reproduce qualitatively similar results as the previous subsection by comparing `asm_uv` and `asm_sd` without making the same idealised assumptions as discussed above. Single observation pair experiments are performed. For each experiment, a single

observation pair, (u^o, v^o) or (sp^o, dir^o) , is introduced at a model grid point, which has background values (u^b, v^b) or (sp^b, dir^b) . The experiment is very similar, but not identical, to the common practice of a single observation experiment in data assimilation.

The results from three experiments, with $(u^o, v^o) = (-15.0, 6.0)$ or $(sp^o, dir^o) = (16.2, 111.8)$ and $(u^b, v^b) = (25.0, 5.0)$ or $(sp^b, dir^b) = (25.5, 258.7)$, are shown in Fig. 3. The two wind vectors for background and observation are labelled **BG** and **OBS**, respectively. The WRFDA default background error, in terms of stream function and velocity potential, are employed for all experiments presented in this section. In the *asm_uv* experiment, we use 1.5 m s^{-1} for both u and v observational errors. The analysis wind vector **ANA_{uv}** indicates that the background error in u and v should be reasonably close to the observational error in u and v .

In the first *asm_sd* experiment, we use 1.5 m s^{-1} for the sp observation error and 10° for dir observation error; the latter closely follows the estimated observational errors in Gao et al. (2012). The analysis vector **ANA_{sd}** (solid line) indicates that the background sp error should be similar to the observational sp error, but the background dir error is much smaller than the observational dir error. In the second *asm_sd* experiment, a smaller observational dir error of 2.9° was employed. It is noted that there is an existing approach (Andersson et al., 2000) to show

background errors in observation space, which can help to compare and tune the background and observations errors. The analysis vector **ANA_{sd}** (dot-dashed line) rotated left considerably with a dir 185.2° , which is halfway between the **OBS** dir of 111.8° and **BG** dir 258.7° indicating that the background dir error is close to the observational dir error. As discussed in the previous sections, the observational dir error should influence the analysis, and it is easily demonstrated in this test with *asm_sd*.

It is also evident from Fig. 3 that the same background and observation wind vectors produce a larger magnitude (longer) analysis wind vector when using *asm_sd* versus *asm_uv*. The same $80,802 [2 \times (201 \times 201)]$ experiments from Fig. 2 were repeated using the WRFDA single observation pair configuration, and the results are presented in Fig. 4, which are strikingly similar to those in Fig. 2. The shift of the minimum point simply indicates that the observation errors are not equal to the background errors. Thus, we conclude that even without the assumptions made in the previous subsection, *asm_sd* produces a stronger wind analysis than *asm_uv*.

Through the assimilation of the single wind observation pair at 559 hPa (-15.0 m s^{-1} , 6.0 m s^{-1}) or $(16.2 \text{ m s}^{-1}$, $111.8^\circ)$, the **OBS** vector in Fig. 3, WRFDA produces a 3-D analysis increments of u and v , as well as other variables with its multivariate formulation, using either *asm_uv* or *asm_sd*. The wind analysis increments for *asm_uv* (left),

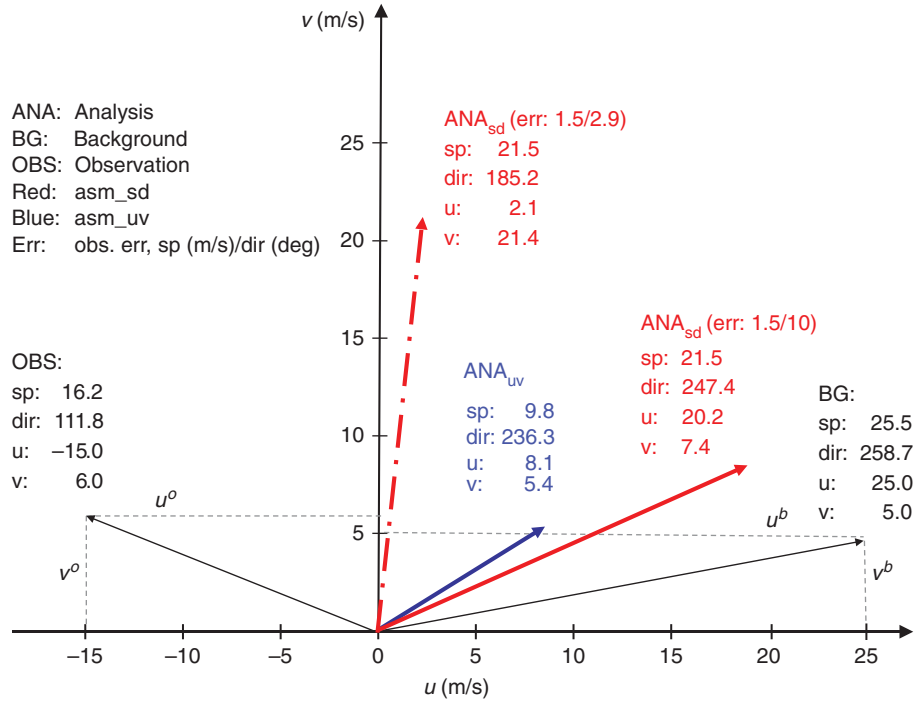


Fig. 3. The background wind vector, the observation wind vector, the analysis wind vector obtained by *asm_uv* (assimilating u and v) and the analysis wind vector obtained by *asm_sd* (assimilating sp and dir). WRFDA is used with the default option (*asm_uv*) and with the new option (*asm_sd*).

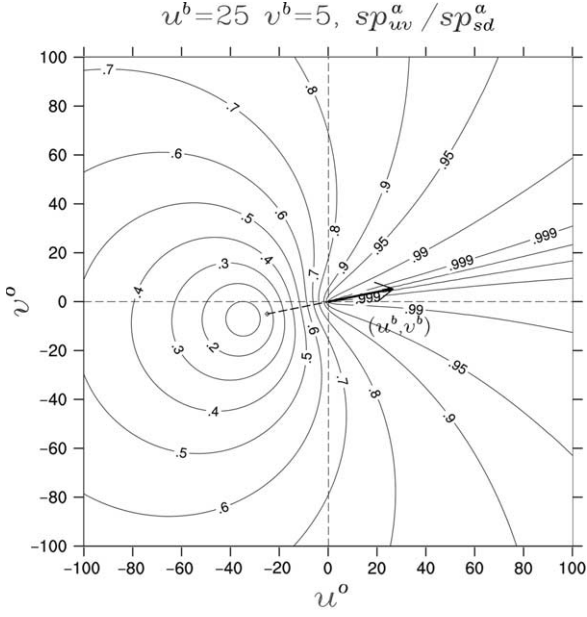


Fig. 4. The ratio of asm_uv analysis wind speed sp_{uv}^a to asm_sd analysis wind speed sp_{sd}^a as a function of u^o and v^o , which is computed using WRFDA. The background wind vector, (u^b, v^b) , is also shown.

and asm_sd (right), are shown in Fig. 5. Due to horizontal correlation of the background errors, the differences are quite large across the entire model domain, as well as at the observation point, which is consistent with Fig. 3. The increment maps in Fig. 5 from asm_uv (5a and 5c) resemble the classic textbook figures (e.g. Daley, 1991) for responses to a single u innovation. This is consistent with Fig. 3, as the v innovation (1.0 m s^{-1}) is much smaller than the u innovation (40.0 m s^{-1}). The increment maps in Fig. 5 from asm_sd (5b and 5d) can be considered a combination of responses to both the u and v innovations, which are comparable in size (cf. Figs. 3 and 5).

The wind increment vectors are also shown in Fig. 5e and 5f. At the observation location, the increment vector in Fig. 5e is consistent with the vector difference between $\mathbf{ANA}_{uv}$ and $\mathbf{BG}$ in Fig. 3. Likewise, the increment vector in Fig. 5f is consistent with the vector difference between $\mathbf{ANA}_{sd}$ (solid) and $\mathbf{BG}$ in Fig. 3. In summary, with an identical background field (i.e. not just a single model point), the same wind observation pair could produce very different analysis wind increments when using asm_uv or asm_sd.

4. Observing system simulation experiments

In the previous section, the difference between asm_uv and asm_sd was demonstrated using both simple examples and WRFDA single observation pair experiments. The primary goal of data assimilation in the context of numerical weather

prediction is to improve the analysis, which leads to a more accurate forecast. In this section, we compare our new formulation (asm_sd) with the conventional formulation (asm_uv) using an observing system simulation experiment (OSSE) framework.

OSSE studies are typically conducted to assess the potential impact of a proposed future observing system (Masutani et al., 2010). The standard OSSE framework consists of four components or steps:

- (1) generate a reference atmosphere by conducting a ‘nature’ run, which is normally done by using a higher resolution than that of the assimilation or forecast model;
- (2) generate simulated observations using the reference atmosphere combined with realistic observational errors;
- (3) assimilate the simulated observations; and
- (4) assess the impact of the simulated observations by comparing the analyses and forecasts to the reference atmosphere.

In this study, we follow the OSSE protocol with the exception of the last step, where instead of assessing the simulated observation impact, we compare the analyses and forecasts to the reference atmosphere to evaluate the performance of the new formulation (asm_sd).

4.1. Nature run

The weather research and forecasting (WRF) modelling system (Skamarock et al., 2008) is used over a domain covering North America and the surrounding oceans (Fig. 6). For the ‘nature run’, a domain with 6-km grid spacing and 57 vertical levels with a model top at 50 hPa are employed. Initial and lateral boundary conditions were obtained from the NCEP final analysis (FNL). The physics configuration is identical to that used by Huang et al. (2009) with the exception of the cumulus parameterisation, which in this case was Kain–Fritsch (Kain, 2004). The reference atmosphere is produced by a 5-day continuous WRF run initialised at 0000 UTC 15 December 2011.

To assess the realism of the nature run, the 500-hPa synoptic flow (e.g. height, temperature and wind vectors) at forecast hour 72 from the nature run is compared to the FNL analysis in Fig. 6, which is valid 0000 UTC 18 December 2011. This is the last cycle time for the outputs of synthetic observations, as well as the assimilation experiments. The nature run provides a solid basis for the following OSSE by successfully simulating the development and propagation of the main synoptic system, where the closed isohypse west of Baja, Mexico, and the deep trough over the eastern United States verified well against the FNL analysis.

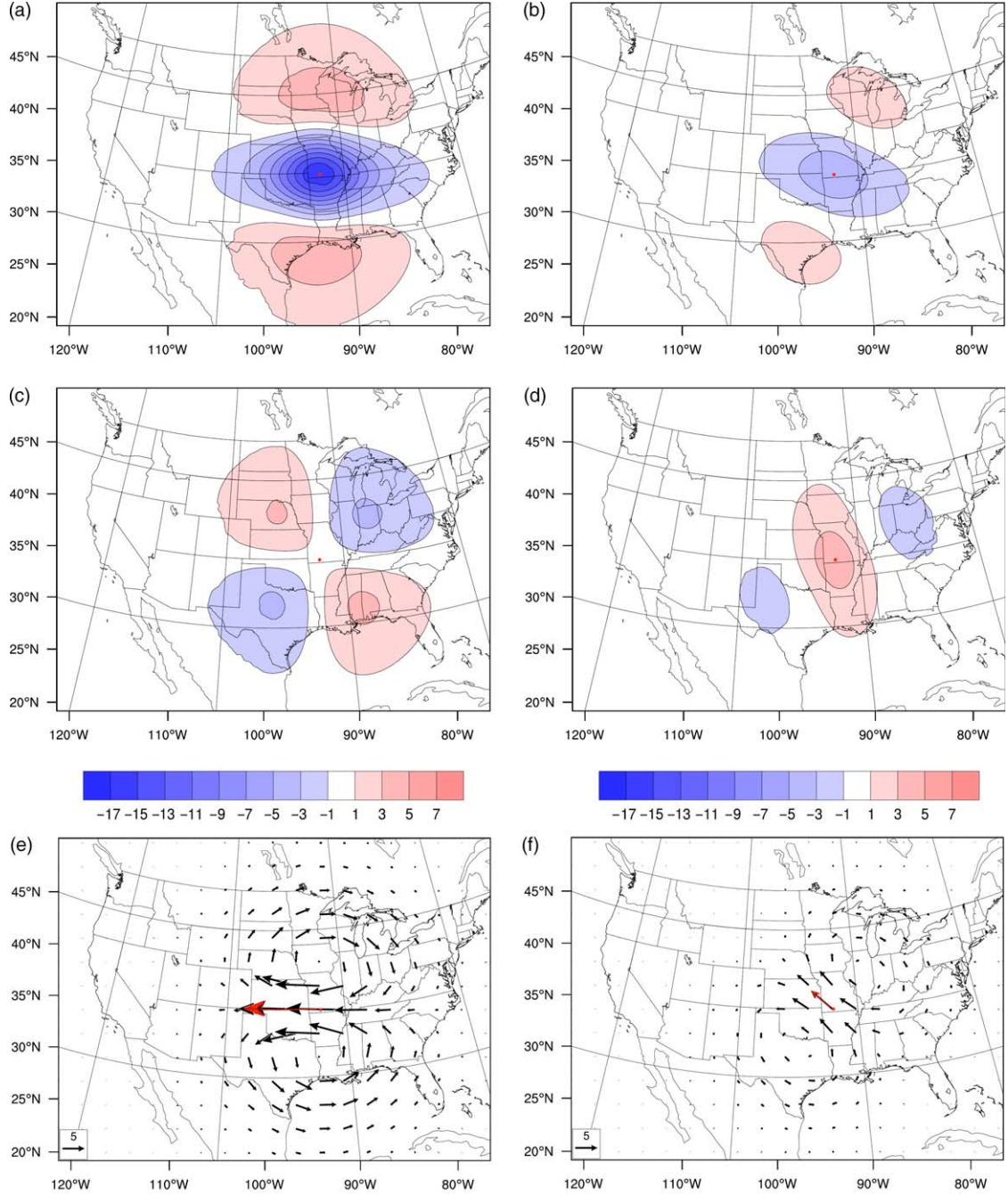


Fig. 5. The wind analysis increments due to a single wind observation pair (-15.0 m s^{-1} , 6.0 m s^{-1}) or (16.2 m s^{-1} , 111.8°) by `asm_uv` (left; a, c, e) and by `asm_sd` (right; b, d, f). The top two panels (a, b) are for u , middle two (c, d) for v and the bottom two (e, f) are for wind vectors (u, v). The observation location is indicated by a red dot. The analysis increment vector at the observation location is highlighted with a red vector.

4.2. Simulated observations

Since we are examining assimilation methods for wind observations, and not observing systems, only wind observations are generated by adding the observational errors to the nature run reference atmosphere every 6 h

over a uniform network with a horizontal resolution of 180 km (i.e. every 30th model grid point, Fig. 7), and across 12 standard pressure levels. From the reference atmosphere, we extract sp and dir at observation locations, and add observation errors to simulate observations, sp^o and dir^o ,

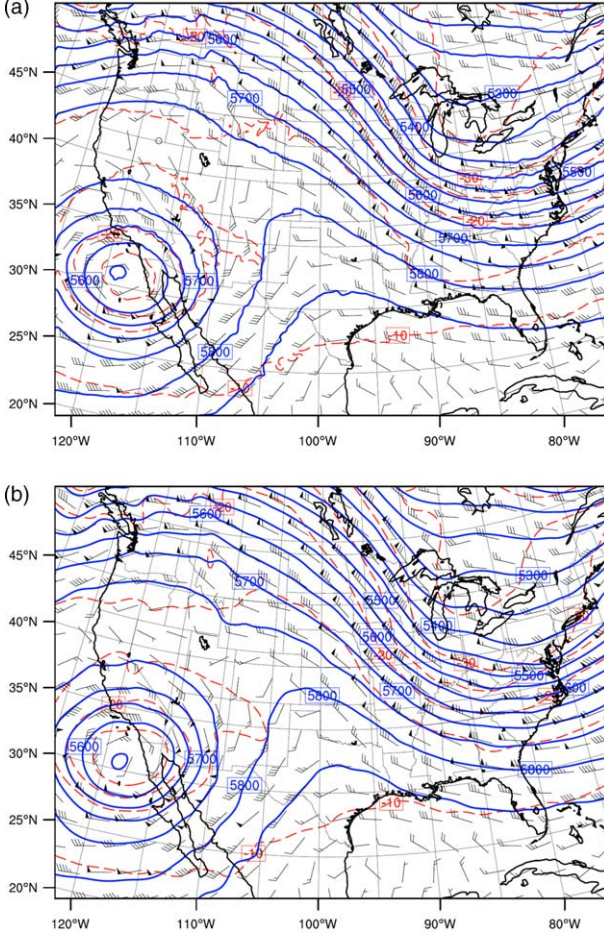


Fig. 6. The 500-hPa geopotential height, winds and temperature of the reference atmosphere (i.e. the nature run) (a) and the FNL analysis (b), valid 0000 UTC 18 December 2011.

over the entire network. Once the simulated observations are obtained, they are converted to u^o and v^o . The observation errors for sp , dir , u and v (denoted as δsp^o , δdir^o , δu^o and δv^o) are then computed by comparing the simulated observations against the direct respective variables within the reference atmosphere.

The sp and dir observation errors are assumed to be uncorrelated and Gaussian with standard deviations estimated for aircraft data by Gao et al. (2012). Previous studies on aircraft and retrieved wind observations (e.g. Leotta and Long, 1989; Gao et al., 2012) have shown that sp error can be characterised by a Gaussian distribution, and dir errors can be characterised with a Gaussian-like distribution. It should be noted that the representativeness error from the observations of sp and dir are not specially considered in this research. In reality, and particularly for near-surface winds, the representativeness error may cause correlation between dir and sp , and may account for a major part of the total

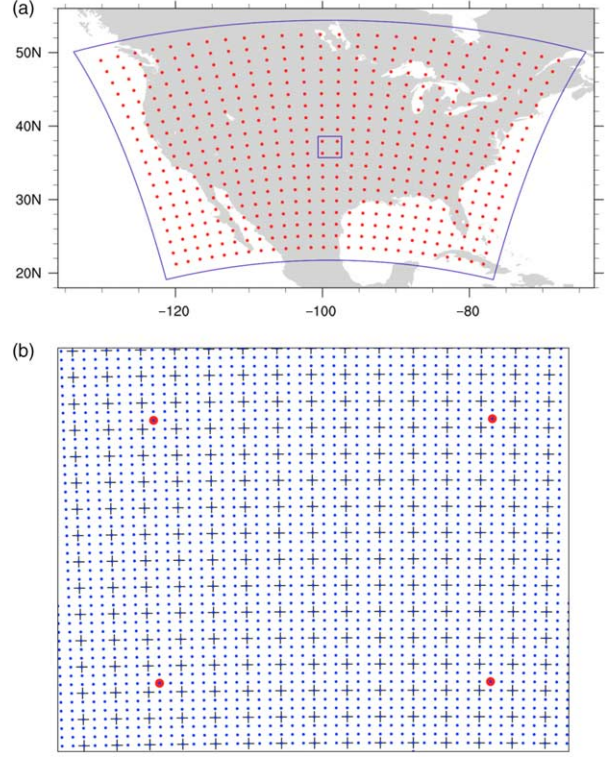


Fig. 7. The model domain with horizontal distribution of simulated observations (dots) in the upper panel and an enlarged view of the small box in the upper panel with 'nature run' model grid points (dots), 'assimilation run' model grid points (crosses) and observation locations (large dots) in the lower panel.

observation error. The error correlation between dir and sp , as well as the distribution of wind dir error, needs further investigation.

The observation error variances are shown in Fig. 8a. The sp and dir observation error variances, $(\delta sp^o)^2$ and $(\delta dir^o)^2$, are close to those used for observation error generation, where $(\delta sp^o)^2$ varies only slightly around the mean value $(2.5)^2 \text{ m}^2 \text{ s}^{-2}$ and $(\delta dir^o)^2$ changes with height with a maximum at the surface and a minimum at 250 hPa. The u and v observation error variances, $(\delta u^o)^2$ and $(\delta v^o)^2$, change with height, but the variations are difficult to directly relate to $(\delta sp^o)^2$ and $(\delta dir^o)^2$. To verify the error estimations, $(\delta u^o)^2$ and $(\delta v^o)^2$ are also computed from $(\delta sp^o)^2$ and $(\delta dir^o)^2$ using the following equations, which can be derived from eq. (A5) in the appendix:

$$(\delta u^o)^2 = \frac{u^2}{u^2 + v^2} (\delta sp^o)^2 + v^2 (\delta dir^o)^2, \quad (16)$$

$$(\delta v^o)^2 = \frac{v^2}{u^2 + v^2} (\delta sp^o)^2 + u^2 (\delta dir^o)^2. \quad (17)$$

The estimated $(\delta u^o)^2$ and $(\delta v^o)^2$ using eq. (16) and eq. (17) are shown in Fig. 8b (labelled as u_linear and v_linear , respectively), and are very close, although not identical, to

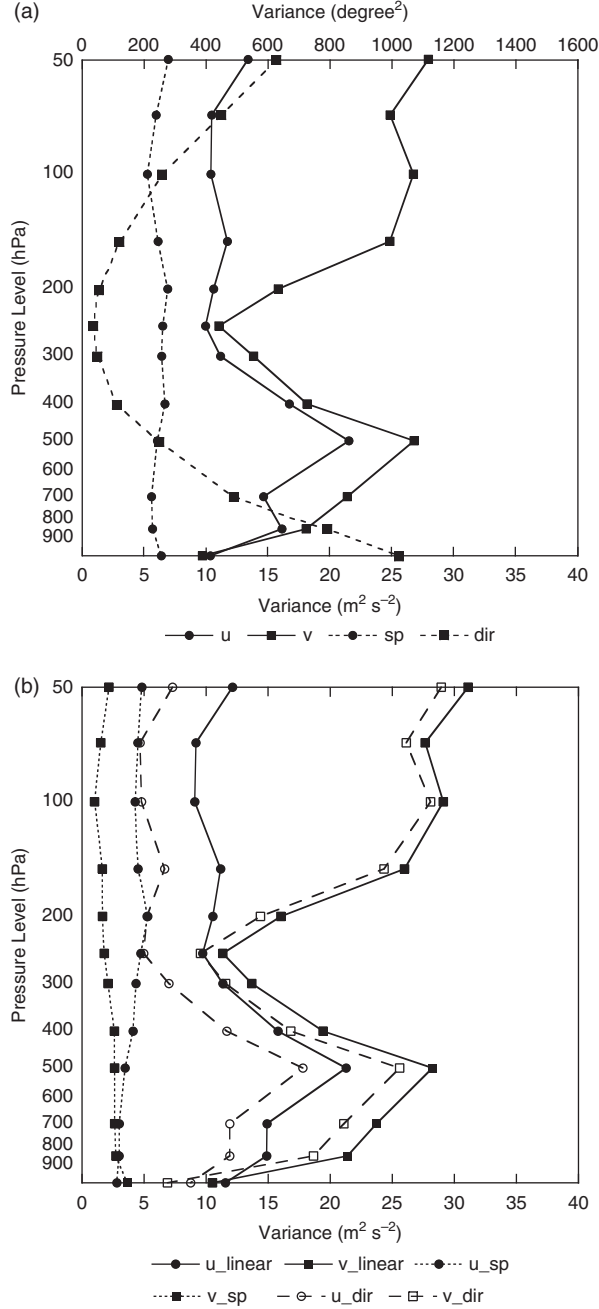


Fig. 8. (a) Observation error variances, $(\delta u^o)^2$, $(\delta v^o)^2$, $(\delta sp^o)^2$ and $(\delta dir^o)^2$, which are computed by comparing the simulated observations against the direct respective variables within the reference atmosphere. (b) Observation error variances, $(\delta u^o)^2$ and $(\delta v^o)^2$, computed from $(\delta sp^o)^2$ and $(\delta dir^o)^2$ using eq. (16) and eq. (17). The contributions from $(\delta sp^o)^2$ only (labelled $_sp$) and from $(\delta dir^o)^2$ only (labelled $_dir$) are also shown.

those shown in Fig. 8a. As shown in eq. (16) and eq. (17), $(\delta u^o)^2$ and $(\delta v^o)^2$ are combinations of $(\delta sp^o)^2$ and $(\delta dir^o)^2$, and are weighted by the background u^2 and v^2 .

Without the contribution from $(\delta dir^o)^2$, both $(\delta u^o)^2$ and $(\delta v^o)^2$ would be smaller than $(\delta sp^o)^2$, and since the upper air mean flow of this experiment configuration is more zonal, $(\delta u^o)^2$ would be larger than $(\delta v^o)^2$ at the higher levels. In Fig. 8b, $(\delta u^o)^2$, the first term of eq. (16), is labelled u_sp , and $(\delta v^o)^2$, the first term of eq. (17) is labelled v_sp . The last term in both eq. (16) and eq. (17) appears as u_dir and v_dir , respectively, in Fig. 8b. The contributions from $(\delta dir^o)^2$ to $(\delta u^o)^2$ and $(\delta v^o)^2$ are larger than those from $(\delta sp^o)^2$ in this experimental configuration. These contributions are proportional to $(\delta dir^o)^2$, as well as u^2 and v^2 .

4.3. Assimilation runs

To ensure that the observations extracted from the nature run would be sufficiently different to the background, a background field with a resolution of 18 km is taken from the 24-h forecast of a no-physics run, which is initialised at 0000 UTC 15 December 2011 on the same domain as the nature run. By following this protocol, we can maximise the difference between the nature run and the background, which will enable a more accurate quantification of the forecast skill improvement between the two assimilation methods.

Beginning with the 24-h forecast, valid 0000 UTC 16 December 2011, WRFDA is run in cycling mode for nine cycles (i.e. 2 d) for the parallel data assimilation experiments. The 6-h forecast from each cycle serves as the background for the analysis of the subsequent cycle. The National Meteorological Center (NMC) method (Parrish and Derber, 1992) is employed to generate the background error covariances using past monthly statistics of differences between WRF 24-h and 12-h daily forecasts. No additional quality control procedures are applied to the observations for either experiment. Finally, a 48-h forecast is produced from each cycle of the parallel runs using identical forecast model configurations. Lateral boundary conditions for the forecasts were obtained from NCEP GFS global forecasts.

4.4. Comparison

For verification, we consider the native 6-km grid of the nature run to be ‘truth’, and the parallel forecasts are compared against this grid at each point. Fig. 9 presents the average RMSE of all analyses and 48-h forecasts over the nine data assimilation cycles. The 0-h RMSE is the verification of the analyses, while the 6-h RMSE is the verification of the background.

The *asm_sd* method produced significantly better verification scores for the *sp* (Fig. 9a) and *dir* (Fig. 9b) analyses and forecasts compared to the *asm_uv* method.

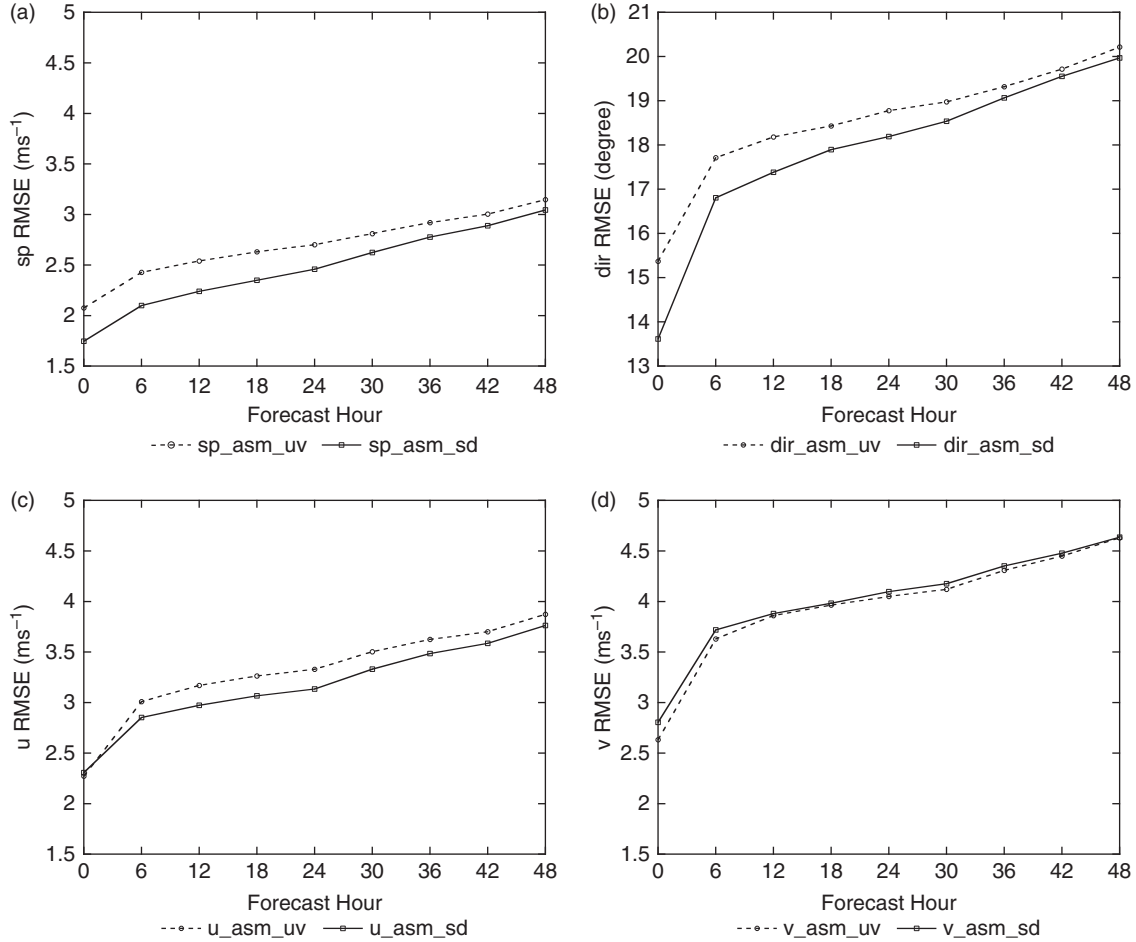


Fig. 9. Time series of the average RMSE of nine cycles of 48-h forecasts, where *sp* and *dir* observation error is generated and then converted to *u* and *v* observation errors. The verification includes the four variables: *sp* (a), *dir* (b), *u* (c) and *v* (d).

The improvements of *sp* and *dir* were 16% and 12%, respectively, at the analysis, and 9% and 4%, respectively, for the mean forecast period. Notable improvements from *asm_sd* over *asm_uv* were also found for *u* forecasts, which are shown in Fig. 9c; however, the dominant zonal flow hinders the ability to highlight differences between *asm_sd* and *asm_uv* when comparing *v* forecasts.

5. Conclusions and outlook

In this study, we present a new method (*asm_sd*) for assimilating wind observations in their observed form of speed and direction. Comparisons between *asm_sd* and the conventional method of assimilating *u* and *v* (i.e. *asm_uv*) are made using simplified arguments and graphic illustrations, followed by single wind observation pair experiments using WRFDA. Finally, idealised experiments using the OSSE framework were employed to compare *asm_sd* to *asm_uv*.

Using a highly simplified configuration (i.e. Figs. 1 and 2), we have shown how *asm_sd* and *asm_uv* can produce very different analyses from the same background and the same observation. Using the WRFDA single observation pair experiments, qualitatively similar results are also produced (Figs. 3 and 4). We have also quantified the difference between *asm_sd* and *asm_uv* in the 3-D analysis increment fields, while the same 3-D background field and the same single observation pair are used (Fig. 5).

The formulation of *asm_sd* ensures that the analysis vector falls between the background and observation in terms of *sp* and *dir*. In the *asm_sd* framework, the analysis can be influenced by the wind direction observation error, which likely plays an important role in most data assimilation systems.

The *asm_sd* and *asm_uv* methods are further compared in a clean and more realistic framework, using simulated wind observations. Because the ‘truth’ was defined by the

nature run, observation errors (Fig. 8) were computed and used for the data assimilation experiments. The results, presented in Fig. 9, show that the *asm_sd* method produces a more accurate analysis, and, as a result, can also lead to a better forecast when compared to the *asm_uv* method. Additional tests are being performed on real observations, and the results will be presented in a subsequent article.

Because the scope of this article focuses on the differences between *asm_sd* and *asm_uv* in the variational formulation using idealised configurations, the quality control aspect of the data assimilation has not been directly addressed. However, for experiments using real observations, the *asm_sd* methodology has the ability to accept or reject observations based on *sp* and *dir* innovations and observation error.

It is well understood that large directional errors in wind observations are correlated to lighter wind speeds for both in-situ observations (e.g. Dobson et al., 1980; Schwartz and Benjamin, 1995; Gao et al., 2012), as well as remotely sensed wind observations (e.g. Plant, 2000; Ebuchi et al., 2002). The observational error for *sp* is positively correlated to *sp*, yet the observational error for *dir* is negatively correlated to *sp*. The new method may be further refined to account for these aspects, as it already deals with *sp* and *dir* directly. For the purposes of this OSSE framework, the *sp* and *dir* observation errors are assumed to be uncorrelated and with a Gaussian distribution. It is of interest to investigate the probability distribution of wind direction error and its relation to the speed error in future research.

6. Acknowledgments

The authors thank Chris Snyder, Fuqing Zhang, Xin Zhang, Xiaoyan Zhang, Dongmei Xu and Peter Childs for fruitful discussions. We are also very grateful for the invaluable comments and suggestions from three anonymous reviewers. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the authors and are not necessarily the views of the National Science Foundation.

7. Appendix A

A.1. Analysis error covariance of *asm_uv* and *asm_sd*

The assumption is made that wind speed and direction observations (sp^o , dir^o) and their errors (δsp^o , δdir^o) are given. The u - and v -wind observations (u^o , v^o), and their observation errors (δu^o , δv^o), are derived from (sp^o , dir^o)

and (δsp^o , δdir^o). For simplicity, we consider one observation pair (sp^o , dir^o) taken at the analysis grid point, and we assume the analysis variables are u and v .

For *asm_sd* (i.e. direct assimilation of sp^o and dir^o), the analysis error covariance $\mathbf{A}$ in the (u, v) space can be expressed as

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}_{sd}^{-1} \mathbf{H}, \quad (\text{A1})$$

where $\mathbf{B}$ is the background error covariance in (u, v) space, $\mathbf{H}$ is the observation operator to convert u and v to sp and dir , respectively, as shown in eq. (8), $\mathbf{H}^T$ is the adjoint of $\mathbf{H}$, and

$$\mathbf{R}_{sd} = \begin{pmatrix} \delta sp^o \\ \delta dir^o \end{pmatrix} \begin{pmatrix} \delta sp^o \\ \delta dir^o \end{pmatrix}^T, \quad (\text{A2})$$

which is the observation error covariance in (sp, dir) space. For *asm_uv* (i.e. assimilation of u^o and v^o), the observation operator is a unit matrix, and the analysis error covariance can be expressed as,

$$\mathbf{A}^{-1} = \mathbf{B}^{-1} + \mathbf{R}_{uv}^{-1}, \quad (\text{A3})$$

where the observation error covariances in (u, v) space are defined as

$$\mathbf{R}_{uv} = \begin{pmatrix} \delta u^o \\ \delta v^o \end{pmatrix} \begin{pmatrix} \delta u^o \\ \delta v^o \end{pmatrix}^T. \quad (\text{A4})$$

It is easily shown that the analysis error covariance in (A1) equals that of (A3) as

$$\begin{aligned} \mathbf{R}_{uv} &= \begin{pmatrix} \delta u^o \\ \delta v^o \end{pmatrix} \begin{pmatrix} \delta u^o \\ \delta v^o \end{pmatrix}^T = \left\{ \mathbf{H}^{-1} \begin{pmatrix} \delta sp^o \\ \delta dir^o \end{pmatrix} \right\} \left\{ \mathbf{H}^{-1} \begin{pmatrix} \delta sp^o \\ \delta dir^o \end{pmatrix} \right\}^T \\ &= \mathbf{H}^{-1} \mathbf{R}_{sd} \mathbf{H}^{-T} \end{aligned} \quad (\text{A5})$$

and

$$\mathbf{R}_{uv}^{-1} = \mathbf{H}^T \mathbf{R}_{sd}^{-1} \mathbf{H}, \quad (\text{A6})$$

where $\mathbf{H}^{-1}$ is the inverse of $\mathbf{H}$,

$$\mathbf{H}^{-1} = \begin{pmatrix} \frac{u}{\sqrt{u^2 + v^2}} & v \\ \frac{v}{\sqrt{u^2 + v^2}} & -u \end{pmatrix}. \quad (\text{A7})$$

In other words, if the u and v errors are linearly calculated from sp and dir errors, respectively, the two methods proposed in this article should produce analyses with the same analysis error covariance, although the analyses can be very different. It should be stressed that in the above derivation, no assumptions are made on the cross error correlation terms. In the WRFDA implementation of *asm_uv* and *asm_sd*, the cross error correlation terms are assumed to be zero, and minor differences between the two WRFDA analysis error covariances do exist.

References

- Andersson, E., Fisher, M., Munro, R. and McNally, A. 2000. Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Quart. J. Roy. Meteor. Soc.* **126**, 1455–1472.
- Barker, D., Huang, X.-Y., Liu, Z., Auligne, T., Zhang, X. and co-authors. 2012. The weather research and forecasting (WRF) model's community variational/ensemble data assimilation system: WRFDA. *Bull. Amer. Meteor. Soc.* **93**, 831–843.
- Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. J. and Xiao, Q. N. 2004. A three-dimensional variational data assimilation system for MM5: implementation and initial results. *Mon. Wea. Rev.* **132**, 897–914.
- Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.* **120**, 1367–1387.
- Daley, R. 1991. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, UK, 457 pp.
- Dobson, F., Hasse, L., and Davis, R. (eds.). 1980. *Air-Sea Interaction: Instruments and Methods*. Chap. 1–4. Springer, New York.
- Ebuchi, N. H., Graber, H. C. and Caruso, M. J. 2002. Evaluation of wind vectors observed by QuikSCAT/SeaWinds using ocean buoy data. *J. Atmos. Oceanic Technol.* **19**, 2049–2062.
- Gao, F., Zhang, X., Jacobs, N. A., Huang, X.-Y., Zhang, X. and co-authors. 2012. Estimation of TAMDAR Observational Error and Assimilation Experiments. *Wea. Forecasting*. **27**, 856–877.
- Huang, X.-Y., Xiao, Q., Barker, D. M., Zhang, X., Michalakes, J. and co-authors. 2009. Four-dimensional variational data assimilation for WRF: formulation and preliminary results. *Mon. Wea. Rev.* **137**, 299–314.
- Kain, J. S. 2004. The Kain–Fritsch convective parameterization: an update. *J. Appl. Meteor.* **43**, 170–181.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D. and co-authors. 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.* **77**, 437–471.
- Leotta, D. F. and Long, D. G. 1989. Probability distribution of wind retrieval error for the NASA scatterometer, July 10–14. In: *Proceedings of IGARSS'89 and 12th Canadian Symposium on Remote Sensing*, Vancouver, Canada, 3, 1466–1469.
- Masutani, M., Schlatter, T. W., Errico, R. M., Stoffelen, A., Andersson, E. and co-authors. 2010. Observing system simulation experiments. In: *Data Assimilation: Making Sense of Observations*. (eds. W. Lahoz, B. Khatatov and R. Menard). Springer, Heidelberg, Germany, ISBN: 978-3-540-74702-4, pp. 647–679.
- van der Molen, M. K., Gash, J. H. C. and Elbers, J. A. 2004. Sonic anemometer (co)sine response and flux measurement. II. The effect of introducing an angle of attack dependent calibration. *Agric. For. Meteorol.* **122**, 95–109.
- Parrish, D. F. and Derber, J. C. 1992. The national meteorological center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.* **120**, 1747–1763.
- Plant, W. J. 2000. Effects of wind variability on scatterometry at low wind speeds. *J. Geophys. Res.* **105**, 16899–16910.
- Schwartz, B. E. and Benjamin, S. G. 1995. A comparison of temperature and wind measurements from ACARS-equipped aircraft and rawinsondes. *Wea. Forecasting*. **10**, 528–544.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M. and co-authors. 2008. *A Description of the Advanced Research WRF Version 3*. NCAR Tec. Note NCAR/TN–475+STR.