
Kozo OKAMOTO
Meteorological Research Institute of Japan Meteorological Agency, Tsukuba, Japan
National Institute of Information and Communication Technology, Koganei, Japan

Toshiyuki ISHIBASHI
Meteorological Research Institute of Japan Meteorological Agency, Tsukuba, Japan

Shoken ISHII, Philippe BARON
National Institute of Information and Communication Technology, Koganei, Japan

Kyoka GAMO
Fujitsu FIP Corporation, Tokyo, Japan

Taichu Y. TANAKA
Meteorological Research Institute of Japan Meteorological Agency, Tsukuba, Japan

Koji YAMASHITA
Numerical Prediction Division of Japan Meteorological Agency, Tokyo, Japan

and

Takuji KUBOTA
Earth Observation Research Center, Japan Aerospace Exploration Agency, Tsukuba, Japan

(Manuscript received 19 April 2017, in final form 13 January 2018)

Abstract

This study evaluated the impact of a future space-borne Doppler wind lidar (DWL) on a super-low-altitude orbit by using an observing system simulation experiment (OSSE) based on a sensitivity observing system experiment (SOSE) approach. Realistic atmospheric data, including wind and temperature, was provided as “pseudo-truth” (PT) to simulate DWL observations. Hourly aerosols and clouds that are consistent with PT winds were also created for the simulation. A full-scale lidar simulator, which is described in detail in the companion paper, simulated realistic line-of-sight wind measurements and observation quality information, such as signal-to-noise ratio (SNR).
1. Introduction

Measurement of the 3D global wind field is crucial for numerical weather prediction, climate studies, and monitoring and prediction of aerosols and other environmental species. The current global observing network contains various wind measurement instruments, such as radiosondes, wind profilers, Doppler radars, ground-based Doppler wind lidars (DWLs), and aircraft-based sensors. These measurements are accurate and provide vertical wind profiles, but observational coverage is limited from a global perspective. Atmospheric motion vectors (AMVs), which are generated by tracking sequential satellite images and ocean surface winds measured by space-borne scatterometers, cover the globe but provide very limited vertical sounding capability. A space-borne DWL is a promising candidate to help complete the measurement requirement with respect to accuracy, resolution, and coverage (World Meteorological Organization; WMO 2009). A DWL transmits laser pulses and receives backscattered signals, from which the Doppler shifts imposed by the atmospheric motion of the distant target are retrieved. Space-borne DWLs can observe wind profiles globally with a high degree of accuracy. Indeed, the European Space Agency (ESA) is preparing to launch a DWL on its Aeolus satellite (ESA 1999; Stoffelen et al. 2005), and scientific communities in the United States are studying the feasibility of space-borne DWLs (Baker et al. 2014).

The Japanese scientific community is also discussing the feasibility of such a mission, and preliminary results were recently published in an overview paper (Ishii et al. 2016). Furthermore, our detailed studies on instrumental development and a lidar simulator are presented in the companion papers of Ishii et al. (2017) and Baron et al. (2017), respectively. The current study aims to investigate the impact of DWL on numerical weather prediction by using an observing system simulation experiment (OSSE) based on a sensitivity observing system experiment (SOSE; Marseille et al. 2008a) approach and discuss trade-offs in DWL configuration and data assimilation system processing. Furthermore, extending the operational assimilation system to use DWL observations ensures that we improve readiness for Aeolus and future space-borne DWL observations.

Many studies have been conducted on OSSE for DWL on Aeolus (Stoffelen et al. 2006; Tan et al. 2007; Marseille et al. 2008a, b, c), OSSE for DWLs discussed in US communities (Atlas 1997; Masutani et al. 2010; Riishøjgaard et al. 2012; Atlas et al. 2015; Ma et al. 2015), and related fundamental topics (Lorenc et al. 1992; Žagar et al. 2004; Riishøjgaard et al. 2004; Žagar et al. 2008; Horányi et al. 2015). These studies accurately simulated DWL observations and assessed their impacts on analysis and forecast skills by using sophisticated data assimilation systems. They agree that positive impacts are obtained with DWL assimilation, and those impacts become greater with an increase in data from more satellites, multiple telescopes, and scanning sampling systems.

Although these studies made significant efforts to create realistic simulations, including wind and clouds,
they seem to use statistical aerosol parameters instead of a day-by-day aerosol field consistent with atmospheric variables. A DWL with a coherent-detection technique is simulated in the current study; this DWL focuses more on the lower troposphere than a direct-detection DWL methods, such as that used by Aeolus (Ishii et al. 2017). Given that aerosols play a critical role in the simulations of DWL backscatter, wind observation distributions and quality, simulations with realistic aerosol variations are essential for coherent-detection DWL OSSE. Thus, we simulated hour-by-hour aerosol fields that are consistent with winds and included them in the lidar simulator together with atmospheric variables and clouds to accurately simulate DWL wind and quality information at each observation location and level. Furthermore, we developed a preprocessing method in the data assimilation system to apply quality control (QC) procedures and assign observation errors to the DWL using quality information. These preprocessings are essential for data assimilation; for example, DWL wind quality was carefully estimated in Marseille and Stoffelen (2003) and Sun et al. (2014). However, the estimated quality information has not been explicitly utilized in previous DWL data assimilation studies. Finally, we conducted data assimilation experiments with simulated DWL in January and August 2010. This allows us to check the seasonal difference of DWL impacts, which has never been published in previous studies as far as we know.

This paper is organized as follows: Section 2 describes the OSSE and data assimilation preprocessing techniques for DWL. The OSSE includes creating a pseudo-truth (PT) atmospheric field, as well as cloud and aerosols, which are used as input for the lidar simulator. The DWL instrument considered in this study and the lidar simulator are also briefly described. More details can be found in the companion papers. Section 3 presents the results of data assimilation experiments for the simulated DWL. Section 4 summarizes the study and shows our future plans. The Appendix lists the acronyms used in this paper.

2. **OSSE configuration and assimilation method**

2.1 **OSSE**

Previous studies have employed three different OSSE approaches for DWL: nature-run (NR)-OSSE (Stoffelen et al. 2006; Masutani et al. 2010; Riishojgaard et al. 2012; Atlas et al. 2015; Ma et al. 2015), SOSE-OSSE (Marseille et al. 2008a), and ensemble-based OSSE (Tan et al. 2007; Megner et al. 2015). Although NR-OSSE has been the most widely used approach, using it for the production of accurate NR and existing observations, including error statistics, is labor intensive (Masutani et al. 2010). Ensemble-based OSSE requires the computing resources of many ensemble members to reduce sampling errors. This study adopted SOSE-OSSE because it can use existing real observations and can be implemented using modest computing resources. SOSE-OSSE defines a synthetic atmospheric state as PT, which represents realistic atmospheric fields and is constructed to reduce forecast error.

SOSE-OSSE is composed of two steps (Fig. 1). The first step is the production of a PT and related fields such as aerosols and clouds, as shown in the upper block of Fig. 1. In this study, we produced hourly PTs, clouds, and aerosol fields in January and August 2010. These were used to simulate DWL observations (see Section 2.3 and Baron et al. 2017). The second step is to assimilate the simulated DWL measurements, together with existing real observations (see Sections 2.4 and 3), which are illustrated in the lower block of Fig. 1. We implemented the first step on an offline basis, that is, we did not interactively update PT and the related fields at every assimilation cycle. An interactive update could produce more accurate PT and related fields but could result in longer forecast lead times (Marseille et al. 2008c). But it would require significant resources because the PT generation process and aerosol model (see later) must be performed at every assimilation cycle step for each experiment.

The way to construct the PT is described in Ishibashi et al. (2014) and is briefly summarized below. The procedure was based on Marseille et al. (2008a): An original initial (analysis) field was corrected using adjoint sensitivity structure and by merging with existing observations. The adjoint sensitivity structure was calculated from the dry total energy norm at 27 h forecast. The adjoint model came from a global data assimilation system of the Japan Meteorological Agency (JMA) (JMA 2007). The data assimilation system was based on a 4D variational (4D-Var) method with a horizontal resolution of TL319 (~60 km) for the outer model and T106 (~120 km) for the inner model and with 60 vertical layers up to 0.1 hPa. It was found that when the global PT field was used as initial condition, forecast errors were reduced throughout a nine-day forecast period and that the error reduction exceeded 6% of a normalized root mean square error (RMSE) verified against the high-resolution operational analysis for temperature and zonal wind at Day 2 (Ishibashi 2014). Furthermore, the assimilation of synthetic wind profile observations made from PT winds on a polar-orbiting satellite with an unrealistic high spatial
sampling showed significant improvements in forecasts from Days 2 to 4 in a single-cycle assimilation experiment (Ishii et al. 2016). These preliminary results convinced us to use PT fields as sufficiently accurate atmospheric fields to simulate DWL wind observations. We will further discuss this through the maximum potential improvement experiment in Section 3b.

PT contains meridional and zonal winds, temperature, surface pressure, and specific humidity. Aerosols and clouds that are compatible with the PT field need to be created for the lidar simulation. Regarding clouds, we employed forecast cloud variables (cloud fraction and cloud mass content) from a global forecast model, as adopted by Marseille and Stoffelen (2003), during the data assimilation cycle of the PT generation process. The evaluation of the cloud forecasted by the JMA global forecast model can be found in Kawai et al. (2014) and Kobayashi et al. (2015). Aerosol distributions were calculated using a global aerosol model called Model of Aerosol Species in the Global Atmosphere (MASINGAR; Tanaka and Chiba 2005; Yukimoto et al. 2012). MASINGAR is a chemical transport model used for research on climate change and data assimilation in the meteorological research institute of JMA and for dust prediction operation at JMA. It treats sea salt, dust, sulfate, black carbon, and organic carbon. The wind field in MASINGAR was hourly adjusted to agree with PT wind by using a nudging scheme based on a Newtonian relaxation forcing. MASINGAR was run with a horizontal resolution of 1.125° and 48 vertical layers up to 0.4 hPa.

2.2 DWL and satellite

We assume a DWL with coherent-detection technique, which has been long developed at the National Institute of Information and Communications Technology of Japan. Two looks of the DWL are pointing toward the ground at 35° off-nadir with azimuth angles of 45° and 135° along the satellite track. This dual-perspective observation capability allows us to construct horizontal wind vectors from the retrieved line-of-sight (LOS) winds. Laser wavelength, pulse energy, and pulse repetition frequency are 2.0 μm, 125 mJ, and 30 Hz, respectively. A scanning sampling function, which would significantly increase observation coverage, is not considered at this stage because of the need to solve a technical difficulty problem in a realistic development timeframe. Target vertical resolution and wind speed accuracy are set to 0.5 km and 1 m s$^{-1}$ at an altitude between 0 and 3 km, 1 km and 3 m s$^{-1}$ between 3 and 8 km, and 2 km and 3 m s$^{-1}$ between 8 and 20 km to meet the WMO observation user requirement (WMO 2009). Horizontal resolution is assumed to be 100 km to obtain enough data with the target vertical resolution and accuracy under the current lidar instrumental specifications. Note that
both resolutions can be varied in a trade-off manner (e.g., higher horizontal resolution is allowed for poorer vertical resolution).

A possible candidate platform for the DWL is a super-low-altitude satellite flying at 220 km or lower. The low orbit allows the DWL to operate with reduced pulse energy and a smaller telescope than a DWL at a normal altitude (e.g., 400–800 km). Satellite orbit can be influential in observation coverage and retrieval quality. Thus, this study discusses the different orbit effectiveness of a dawn–dusk sun-synchronous polar-orbiting satellite passing the equator at 18:00 local time, such as Aeolus, and a non-sun-synchronous, low-inclination satellite focusing on the tropics, such as the Tropical Rainfall Measuring Mission (TRMM) satellite. For this, we compared a sun-synchronous polar-orbiting satellite at 220 km and 96.4° inclination angle (“polar satellite”) and a non-sun-synchronous, low-inclination satellite at 220 km and 35.1° (“tropical satellite”) in this study. More technical details of DWL and satellites are described in Ishii et al. (2017).

2.3 Lidar simulator

The lidar simulator Integrated Satellite Observation Simulator for a Space-Borne Coherent Doppler Lidar (ISOSIM-L) is an end-to-end simulator that calculates backscattered power, background noise power, signal-to-noise ratio (SNR), LOS wind speed, noisy Doppler-shifted signal, retrieved LOS winds, and retrieval quality information. It calculates the scattering and absorption coefficients of cloud and aerosol particles by using Mie theory. The spatial and temporal variability of the clouds is taken into account by randomly generating clouds on the basis of cloud fraction related with PT. Aerosols are sea salt, dust, sulfate, black carbon, and organic carbon, and clouds are stratus and cumulus liquid clouds and cirrus ice clouds. A sunglint effect is not included in the simulation because it is significantly reduced in the coherent DWL system, in which a narrow-bandwidth filter only measures the desired frequencies. Wind speed in the LOS direction from DWL is derived from a Doppler-shifted frequency estimated with a maximum likelihood probability approach with the vertical resolution (0.5, 1.0, or 2.0 km) per laser shot. The final output of ISOSIM-L for OSSE is an LOS wind speed averaged in the horizontal resolution (100 km). It has 19 layers from the surface to approximately 27 km altitude. Baron et al. (2017) shows that LOS wind can be considered a good retrieval if $N^{23}$SNR is over 0.08, where $N$ is the total range-bin number in a resolved volume (for example $N = 2304$ for a 0.5 km × 100 km target resolution).

The details of ISOSIM-L are described in Baron et al. (2017).

The left panel in Fig. 2 shows the sample coverage of DWL on the polar and tropical satellites ISOSIM-L simulated at the 6th layer (11.4 km altitude on the global average) and 14th layer (2.7 km) from 09:00 to 15:00 UTC on August 1, 2010. The dots correspond to the position of the retrievals (100 km horizontal resolution) obtained from both telescopes. The target locations of the two telescopes are very close and cannot be distinguished from the plots. The $N^{23}$SNR varies according to the vertical profile of aerosols and clouds, as shown in Baron et al. (2017). For example, at low latitudes, it becomes larger because of cirrus clouds at high altitude and because of low clouds, dusts and sea salt aerosols at low altitudes.

2.4 Data assimilation system

a. Observation operator

We assimilate individual LOS wind speed independently. Our coherent DWL with two telescopes can retrieve zonal and meridional winds by collocating winds in two different LOS directions under the assumption of the absence of a vertical wind component. However, the matching process slightly reduces data availability and makes it hard to use quality information given for individual retrieval. The observation operator, which maps model wind variables into observation counterparts, for DWL LOS wind speeds is a bilinear interpolation of model-gridded zonal and meridional winds ($u$, $v$) to observation locations and subsequent transformation into LOS wind speed $w$.

$$w_i = -(u \sin \phi_i + v \cos \phi_i) \sin \theta, \quad (1)$$

where $\theta$ and $\phi_i$ are the nadir angle toward the satellite and the angle between the north and LOS direction $i$ (1 or 2). For the DWL, we assume $\theta = 35^\circ$ and $\phi_i = 45^\circ$ or $135^\circ$. Note that Eq. (1) also assumes no vertical wind component.

b. QC

The QC procedure for DWL consists of three steps. First, the anomalous data flagged by ISOSIM-L are rejected. Furthermore, winds in the lowest layers (lower than 100 m) in ISOSIM-L are excluded because of contamination by surface reflection (anomalous QC). Secondly, data with $N^{23}$SNR smaller than 0.1 are rejected (SNR-QC), excluding data retrieved from low signals. This threshold (0.1) is a little larger than the original one (0.08) in Baron et al. (2017; see Section 2.3) and was determined to reduce the risk of low
Fig. 2. DWL distribution simulated at (a, b) the 6th layer (approximately 11.4 km altitude) and (c, d) 14th layer (2.7 km) from the polar-orbiting satellite and at (e, f) the 6th layer and (g, h) 14th layer from the tropical-orbiting satellite between 09:00 and 15:00 UTC on August 1, 2010. The color indicates the logarithm of $N^{1/2}$SNR. DWL availability before (after) the QCs is plotted in the left (right) column.
signal data contamination. Finally, data are removed when observation values are unreasonably departed from first-guess (gross error QC). The thresholds for gross error QC are determined as a function of altitude range on the basis of several trial data assimilation experiments: 0.5 m s\(^{-1}\) below 3 km, 1.0 m s\(^{-1}\) between 3 and 8 km, and 1.5 m s\(^{-1}\) above 8 km. Note that these values are defined for wind speed in the LOS direction and correspond to the horizontal wind speed times sin \(\theta \approx 0.57\). These three QC procedures are applied in the order presented above.

We examined the three QC procedures by comparing observations and PT. Figure 3 shows the probability distribution functions (PDFs) of observation-minus-PT (PT departure) on January 2, 2010, and August 2, 2010. It is obvious that the PT departure decreases (sharper and higher peak of PDF) for QC-passed data, thus indicating that the QC procedure successfully chooses good DWL observations that are close to PT. Interestingly, the PT departure before QC in January is smaller than the one in August, although PT departures after QC are similar. This seasonal difference in DWL quality is perhaps due to seasonal variances in cloud and aerosol distribution. An example of the distribution of DWL winds that passed all QC is shown in the right column of Fig. 2. The comparison with all data in the left column of Fig. 2 suggests that rejection is mostly determined by SNR-QC (DWL winds with small N\(^{1/2}\)SNR disappear). This can be clearly confirmed in Table 1, which shows a monthly averaged number (percentage) of data rejected by each of the three QCs for a 6 h data assimilation window. Although the number varies from one analysis to another because of variable cloud and aerosol distributions, 15 % of DWL winds survive on average, and SNR-QC has the largest contribution, thus removing almost half of the total number. One may think this survival ratio is somewhat too small despite its effectiveness in removing bad quality data. It would be another option to assimilate more data with low quality by assigning large observation errors instead of removing them.

c. Observation error

An observation error in the data assimilation context consists of a measurement error, an observation operator error, and a representativeness error. The measurement error includes instrument noise, wind fluctuation in a target observation volume, and retrieval processing error. Given that ISOSIM-L takes

![Fig. 3. PDF of DWL observation-minus-PT for LOS wind speed (m s\(^{-1}\)) on (a) January 2 and (b) August 2, 2010. Thin gray lines and thick black lines show the PDFs of all samples before QC and samples selected by all the QCs, respectively.](image)

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>reject</th>
<th>use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>reject</td>
<td>Anomalous QC</td>
<td>SNR-QC</td>
</tr>
<tr>
<td>Polar satellite</td>
<td>63,803 (100 %)</td>
<td>6,728 (10.55 %)</td>
<td>39,759 (62.31 %)</td>
</tr>
<tr>
<td>Tropical satellite</td>
<td>63,819 (100 %)</td>
<td>6,735 (10.55 %)</td>
<td>39,629 (62.10 %)</td>
</tr>
</tbody>
</table>
into account these errors, we adopted ISOSIM-L error estimates for measurement error. Figure 4 shows the monthly averaged measurement errors estimated by ISOSIM-L in six latitudinal bands for six layer groups in January and August 2010. The measurement error increases with altitude despite more samples being used for averaging in higher layers. Furthermore, we found that there is a noticeable difference between January and August, thus suggesting a need to assign observation errors in consideration of seasonal variations that are probably due to atmospheric conditions. The observation operator error for DWL is caused by interpolation and variable transformation in Eq. (1). Neglecting the vertical wind component may not be negligible, whereas errors due to interpolation are small. The representativeness error arises when the representative scale is different between the observation and model, such as fine-structured winds that are unresolved by the model. In particular, DWL observation geometry is a 100 km-long line, whereas the inner model of the data assimilation system is representative of a 120 × 120 km² box. This difference can cause nonnegligible representativeness errors, for example, in the presence of horizontal wind shear. As an exact estimation of the observation operator error and representativeness error is difficult to obtain, we assumed that the sum of these two errors is 1.0 m s⁻¹.

Finally, we calculate the observation error $R$ as a summation of these three components:

$$R = C \sqrt{E_m^2 + 1.0^2}.$$  

(2)

$E_m$ is the measurement error estimated by ISOSIM-L, and 1.0 is the sum of the other errors, as described above. $C$ is an observation error inflation factor, which is set to 4.0 in the region equatorward 60° latitude and 6.0 in the other regions in this study. The greater inflation at high latitude has been empirically derived from preliminary cycle experiments with a global constant inflation factor that worsen forecasts in both extra-hemispheres. Given that this degradation seemed to originate from large analysis errors in high-latitude regions, we decided to reduce the weight of DWL in these latitudes. Figure 5 shows a zonal mean of observation error assigned for QC-passed data in August 2010. The observation error is 4–5 m s⁻¹ in most regions equatorward 60°. It gets closer to 7 m s⁻¹ at higher latitude and becomes larger than 7 m s⁻¹ above 300 hPa. This variation corresponds to that of measurement error $E_m$ in Fig. 4. Given that these values are defined in the LOS direction, they are much larger than those for other wind observations. For example, the observation error set for meridional and zonal wind speeds from 1000 to 100 hPa is 2.3–3.1 m s⁻¹ for radiosondes and 4.5–6.8 m s⁻¹ for AMVs. These large observation errors for DWL seem to be counterintuitive in view of the small measurement errors in Fig. 4. Although the inflation factor of 4.0 appears to be too large, trial experiments with smaller values (e.g., 1.0, 1.5, or 3.0) showed clearly negative forecast impacts.

There are some possible reasons for the need to set this large inflation factor. First, it may be related to having no spatio-temporal thinning procedure (selection of one preferred observation in a predefined spatial volume and temporal period) applied for DWL. By contrast, observations other than DWL are thinned in the JMA data assimilation system to counteract neglecting observation error correlation and to reduce the use of computational resources. For example, AMVs are significantly thinned such that one AMV is obtained in a 200 km × 200 km × 100 hPa volume and in a 6 h assimilation window. The lack of thinning procedure for DWL observations may be also related to larger inflation setup in high latitudes in which relatively denser DWL observations are assimilated. Second, inflating observation errors are widely implemented for satellite observations, such as a large amount of space-borne radiance observation, in the JMA data assimilation system. For example, inflation factors for the Advanced Microwave Sounding Unit (AMSU-A) and Microwave Humidity Sounder (MHS) are 1.2 and 4.5, respectively. This treatment aims to avoid the detrimental effect of error correlation and unbalance with other observations and first-guess.
This may also hold for DWL observations, thus suggesting the necessity of thinning and/or error inflation. Another possible reason for inflation is errors in the PT itself. In SOSE-OSSE, PT is not a truth; therefore, simulated observations can have errors. To treat such errors, Marseille et al. (2008a) tuned observation errors by using sensitivity information in their SOSE-OSSE. They inflated observation errors over regions where the analysis adaptation (difference between PT and original analysis) is small, for example, in regions with abundant existing observations. This processing intends to avoid spurious correlation with existing observations and reduce relatively large stochastic error in PT. Our approach, by contrast, is a uniform inflation because of the difficulty in obtaining the analysis adaptation magnitude at each observation location required by Marseille et al. (2008a). Even this simple approach is able to avoid these errors but should underestimate the impact of DWL in sparse regions of existing observations. Restricting the inflation in the observation sparse regions could be a more optimal approach. Further investigations on the reason for large inflation and efforts to assign more reasonable observation errors are still ongoing.

3. Results of data assimilation experiments

3.1 Experiment configuration

We conducted analysis–forecast cycle experiments to assess the impacts of the DWL data simulated by ISOSIM-L in January and August 2010. The data assimilation system was the low-resolution version of the operational global data assimilation system of JMA as of 2010 (JMA 2007). The resolution of the forecast model was TL319L60 (approximately 60 km horizontally and 60 vertical layers), and the model top was at 0.1 hPa. The analysis system was an incremental 4D-Var method with an inner loop of T106L60 (approximately 120 km). We performed 6 h data assimilation cycles from December 20, 2009, to February 7, 2010, (January experiment) and from July 20, 2010, to September 9, 2010 (August experiment). We also ran 168 h forecasts at 12:00 UTC from January 1 to 31, 2010, for the January experiment and from August 1 to 31, 2010, for the August experiment.

The results are shown from three experiments with different observation configurations. One is a reference experiment, denoted as CNTL, where all of the observations used in the operational system were

![Fig. 5. Zonal mean of observation error of LOS wind (m s$^{-1}$) assigned for QC-passed data in the data assimilation system in August 2010.](image-url)
assimilated. The global wind observations in CNTL included AMVs from five geostationary satellites and two polar-orbiting satellites; ocean surface wind vectors from a microwave scatterometer on European Metop polar satellites, buoys, and ships; land surface wind vectors from SYNOP reports; and wind vector profiles from aircraft, wind profilers, and radiosondes. The tropical cyclone (TC) bogus data of wind vector profiles and surface pressure were also assimilated around the center of TCs over the western North Pacific. Other satellite observations in CNTL were the radiances of AMSU-A, AMSU-B, MHS, SSM/I, AMSR-E, TMI, and SSMIS instruments; clear sky radiances at water vapor channels on five geostationary satellites; and atmospheric refractivity from GNSS-RO. The other experiments, denoted as TESTP and TESTT, assimilated DWL LOS winds from the polar and tropical satellite from 00:00 UTC on January 1, 2010, to 18:00 UTC on January 31, 2010, for the January experiment and from 00:00 UTC on August 1, 2010, to 18:00 UTC on August 31, 2010, for the August experiment, respectively, in addition to CNTL observations. TESTP and TESTT are collectively referred to as TESTs. A list of all the experiments discussed in this study is shown in Table 2.

Note that small random perturbations were added to the DWL LOS winds simulated by ISOSIM-L. An additional perturbation has been found in many NR-OSS studies to mimic realistic observation error and avoid using simulated observations that are too close to the truth. In this study, the DWL winds generated by ISOSIM-L already included the measurement error (Section 2.4c). However, because there may be unexpected error sources in real DWL observations, we decided to add a small perturbation of a Gaussian probability distribution with a standard deviation of 0.2 m s\(^{-1}\). Nevertheless, this small perturbation made negligible difference in impacts on forecasts when compared with experiments without the perturbation.

The possible sources of these unexpected errors are height assignment error due to wind shear (Sun et al. 2014) and vertical motion contribution (Marseille et al. 2010), and more accurate estimate is necessary in future works.

3.2 Results of analysis and forecast

a. Analysis

First, we examined the change in analysis error made by assimilating DWL winds to see whether the assimilation system worked properly. An error relative to PT was calculated as the absolute difference between analysis and PT. Figure 6 shows the zonal wind error difference between CNTL and DWL assimilation experiments at 12:00 UTC on August 1, 2010. In this comparison, the DWL assimilation experiments employed the same first-guess as CNTL, instead of TESTs, to clarify the observation impacts in the single assimilation cycle. Figure 6 illustrates that areas with error differences mostly correspond to DWL coverage that passes QC in Fig. 2. Although negative impacts (increasing error) representing the stochastic properties of the data assimilation process occasionally occur (Stoffelen et al. 2006), positive impacts (decreasing error) are more likely to be seen.

Figure 7 shows accumulated DWL numbers over one month for TESTs in 10 hPa pressure bins and 2° latitude bins. The large availability of QC-passed DWLs is related to strong backscattering and weak absorption by clouds and aerosols. For example, there are many DWL observations available in high latitudes and in the upper troposphere in the tropics because of cloud top backscattering. A seasonal change in latitudinal shift associated with the intertropical convergence zone can be seen approximately 15°S for the January experiment (Figs. 7b, e) and approximately 10°N for the August experiment (Figs. 7c, f). A large amount of data available around 30°N for the August experiment (Fig. 7c), despite usually few clouds in the

Table 2. A list of assimilation experiments.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description</th>
<th>Assimilated obs DWL</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNTL</td>
<td>No DWL assimilated</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>TESTP</td>
<td>Assimilate DWL on the polar satellite</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>TESTT</td>
<td>Assimilate DWL on the tropical satellite</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>TESTP2</td>
<td>The same as TESTP except for assigning DWL observation errors in different season</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>TESTtmp</td>
<td>Initialized by PT field to examine the maximum potential impact of SOSE-OSSE</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>TESTqc</td>
<td>Assimilate DWL but deactivate most of the QC</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>TESToe</td>
<td>Assimilate DWL with observation errors of radiosondes</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>
subtropical subsidence region, is explained by a strong
lift of desert dust aerosols depicted in Fig. 2 in Baron
et al. (2017). Furthermore, obvious amount of DWL
below 700 hPa is given by water clouds and aerosols.
Note that the extensive number in polar regions in
TESTP and approximately 35° in TESTT is produced
by the frequent overpass of satellites. The horizontal
stripes of dense data layer can also be seen over most
latitudes especially over ocean. This is caused by the
retrieved wind altitude corresponding to the center
of fixed vertical layers with respect to the surface in
ISOSIM-L. With the vertical layer spacing set to 0.5,
1.0, and 2.0 km according to the altitude in the ISO-
SIM-L retrieval process (see Sections 2.2 and 2.3), the
stripe distribution appears to vary with altitude in the
pressure coordinate.

For comparison, AMVs from five geostationary and
two polar-orbiting satellites are plotted in Figs. 7g and
7h. It is evident that no AMVs are assimilated in the
middle troposphere. This is a result of rigorous QC
procedures based on a blacklisting database, in which
AMVs in this region are identified as poor quality
from observation minus first-guess statistics. More-
over, a large portion of AMVs are excluded in other
regions by thinning and QCs on the basis of the AMV
quality index. This comparison indicates that DWL
can complement the coverage of AMVs, particularly
in the middle troposphere. However, we also need to
keep in mind that there is a possibility that QC and
thinning procedures for actual DWL observations may
remove more data than those for simulated observa-
tions and that AMVs from new-generation geostation-
ary satellites, such as Himawari-8, can produce more accurate and numerous AMVs (Shimoji 2014).

b. Forecast

The impacts of assimilating DWL were assessed with respect to forecast improvement by comparing TESTs and CNTL. The forecast improvement is defined as the relative forecast error reduction, which represents the RMSE difference between the CNTL and TESTs normalized by the CNTL RMSE: \((\text{RMSE}_{\text{CNTL}} - \text{RMSE}_{\text{TEST}}) / \text{RMSE}_{\text{CNTL}} \times 100\). A positive value corresponds to a reduced RMSE or improvement (positive impact) in the TESTs forecast. The RMSE is verified against PT unless otherwise stated.

Fig. 7. Accumulated number of DWL in 10 hPa pressure bins and 2° latitude bins over one month for (top row) TESTP and (middle row) TESTT experiments. (a, d) All DWLs before QC, (b, e) DWLs used in assimilation after QC for the January experiments, (c, f) DWL used for the August experiments. (g, h) Similar to (b, c) but for AMVs used from geostationary and polar-orbiting satellites.
Although 7 d forecasts were run in all experiments, the verification focuses on forecasts up to Day 5 because the evaluation of forecasts beyond Day 5 is less reliable from the examination of the maximum potential improvement experiment using the initial conditions of PT, as discussed later in this section.

Figure 8 shows the relative forecast error reduction (%) of zonal wind as a function of forecasts up to 120 h for (upper) TESTP and (lower) TESTT for the January experiments. It is verified against PT in (a, d) the northern hemisphere (20°–90°N), (b, e) tropics (20°N–20°S), (c, f), and southern hemisphere (20°–90°S). Positive values correspond to forecast improvement in TESTs. The contour lines indicate the statistical significance at the 90 % and 95 % confidence levels.

Fig. 8. Relative forecast error reduction (%) of zonal wind as a function of forecasts up to 120 h for (upper) TESTP and (lower) TESTT for the January experiments. It is verified against PT in (a, d) the northern hemisphere (20°–90°N), (b, e) tropics (20°N–20°S), (c, f), and southern hemisphere (20°–90°S). Positive values correspond to forecast improvement in TESTs. The contour lines indicate the statistical significance at the 90 % and 95 % confidence levels.

Fig. 9. Time sequence of RMSE for zonal wind at (a) 250 and (b) 700 hPa in the tropics at Day 3 for the January experiments. CNTL (black line), TESTP (blue line), and TESTT (red line) are plotted.

Fig. 9. Time sequence of RMSE for zonal wind at (a) 250 and (b) 700 hPa in the tropics at Day 3 for the January experiments. CNTL (black line), TESTP (blue line), and TESTT (red line) are plotted.
specifically in the middle of the month. This finding is encouraging in that DWL improves not only the monthly averaged forecasts but also daily forecasts.

Figure 10 shows that impacts for the August experiment are also generally positive with some exceptions. Positive impacts with statistical significance are seen at Days 2 and 3 in the upper troposphere for TESTP and at Day 1.5 in the upper troposphere of the tropics and between Day 1.5 and 4 in the lower troposphere of the southern hemisphere for TESTT. However, degradation is noticeable in the tropics for both TESTP and TESTT. Moreover, a statistically significant degradation is found in short-range forecasts in the southern hemisphere for TESTP.

A comparison of Figs. 8 and 10 shows that positive impacts are smaller for the August experiment than for the January experiment. This is likely attributable to the lower quality of DWL wind retrievals, as shown in Fig. 3, that shows the variance of PT departure is slightly larger on August 2 even for QC-passed data. This suggests that additional QCs may be necessary to better filter low-quality DWL data. Furthermore, to investigate degradation in some regions for the August experiment, we examined the sensitivity of the improvement or degradation to the observation error assigned. An additional experiment with polar-orbiting DWL for the August experiment, that is, TESTP2, in which January experiment observation errors were assigned, was performed. The relative forecast error reduction for TESTP2, which is displayed in Fig. 11, shows that the degradations is still seen in short-range forecasts in the southern hemisphere, but it is not statistically significant. Despite this favorable change, the impact is negative in the northern hemisphere. The
The result of this additional experiment suggests the high sensitivity of forecast impacts to observation error assignment. Furthermore, it also indicates the possibility of using more appropriate error settings to improve forecasts and the difficulty in achieving global improvement.

Figure 12 shows the relative forecast error reduction in temperature for TESTP and TESTT for the January and August experiments. Degradation is seen in the upper troposphere in the tropics for the August experiment. However, similar to the zonal wind verification, positive impacts are dominant in wide layers and regions and broad forecast ranges. In particular, statistically significant improvements are found in the short-range forecasts in the tropics for all experiments. They are also evident at Days 2 and 3 in the southern hemisphere for the August experiment. This encouraging result means that assimilating DWL wind observations improves forecasts not only for wind fields but also for mass fields. This is consistent with many
previous studies.

To give higher confidence to the results described above, we added verification of the forecast error against the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al. 2011) instead of PT. It is reasonable to use ERA-Interim for this purpose because ERA-Interim is independent from PT and is sufficiently accurate for verification in that it is created using a state-of-the-art data assimilation system and assimilates massive observations at ECMWF. The relative forecast error reduction based on ERA-Interim shows a generally consistent pattern with the one based on PT, although the magnitude of impacts is a little smaller (not shown). For example, the degradation in zonal wind for TESTP in the August experiment is also seen in the ERA-Interim verification, but the statistical significance disappears. Thus, we can have confidence in the results described above, although the detailed quantitative discussion of impacts may require further investigations.

Figure 13 compares the mean position error for up to 18 cases of 5 TCs over the west Pacific Ocean in the August experiment. Adding DWL data slightly reduces position errors for a forecast range beyond 48 h for both TESTP and TESTT. The magnitude of improvements is similar for TESTP and TESTT. However, it is noted that the results are statistically significant only with 60 and 66 h forecasts (not shown) owing to the limited sample number because the forecast is run once a day for a one month period. Furthermore, assimilating TC bogus can be partially responsible for the small difference among all experiments. Figure 14 demonstrates an example of TC track forecast up to 84 h initialized at 12:00 UTC on August 8 for TC Dianmu (2010). A small but constant improvement is found for both TESTP and TESTT.

We compared the impacts between TESTP and TESTT, which is one of the motivations of this study. Figure 15 shows the relative forecast error reduction for zonal wind between TESTT and TESTP. A positive value indicates that TESTT is better than TESTP. Most interesting is that TESTT is superior to TESTP up to Day 2 but is inferior beyond Day 2 in the tropics for both the January and August experiments. One possible explanation for this rapid change of superiority or inferiority is an insufficiently accurate initial state in the tropics due to, for example, difficulty in constructing an appropriate background error covariance in the tropics in a data assimilation system. This can make it more difficult to retain forecast skill for long forecast hours in the tropics than in mid-latitudes (Riishojgaard et al. 2012; Ma et al. 2015). Given that DWL covers the low latitude alone in TESTT, the limitation of forecast skill in the short-range should be more apparent in TESTT than in TESTP. Another possible explanation for the rapid change is the inadequacy of

![Fig. 13. Mean position error (km) of the center of the TCs over the Pacific Ocean in August 2010. The gray line indicates the sample number shown at the right axis.](image1)

![Fig. 14. Track forecasts of TC Dianmu (2010) initialized at 12:00 UTC August 8. The center positions at every 6 h forecast up to 84 h are plotted for CNTL (green), TESTP (blue), and TESTT (red). The black line indicates the best track at the corresponding time.](image2)
PT. Owing to slow error growth in the tropics (Kalnay 2003), PT may not be able to capture errors that grow far beyond when the error norm is computed for the sensitivity structure for a 27 h forecast. Furthermore, the assumption of linear error growth PT may be valid only for short-range forecasts in the tropics. On the other hand, in the extra-tropics, rapidly and linearly growing errors are more likely to occur and can be better captured by PT. These PT characteristics can lead to forecast skills that vary for different periods in TESTT and TESTP. Note that this can depend on the way PT is constructed: Marseille et al. (2008c) demonstrates that a PT that is interactively updated at every assimilation cycle could produce better forecasts than a PT generated on an offline basis, as performed in this study. Further investigations are needed to verify these two explanations (deficiency in data assimilation and inadequacy of PT). In addition to the rapid change in the tropics, TESTT is better in the southern hemisphere throughout the forecast range for the August experiment (Fig. 13f), thus reflecting the degradation in TESTP for the August experiment in Fig. 9. Impacts in the northern hemisphere are unclear. From these mixed results and the fact that there is room to improve preprocessing (QC and observation errors), a firm conclusion on whether TESTT or TESTP are superior cannot be made at this stage of the study.

Finally, we performed three additional experiments (TESTmp, TESTqc, and TESToe) to evaluate the fundamental performance of our SOSE-OSSE setup (Table 2). First, we examined the maximum potential improvement using the hourly PT field of all variables at every grid as initial condition. The intention of this experiment (TESTmp) is to evaluate the maximum limitation of our experimental setup of SOSE-OSSE. Figure 16a shows the global average of the relative forecast error reduction for zonal wind in August experiment. Significant improvements (for example, 15 % at 12 h and 5 % at Day 2) are achieved with the forecasts initialized by the PT field and extends to Day 5 with statistical significance. As Ishibashi (2014) also shows, this confirms that the PT field can substantially reduce forecast errors even at Day 5, thus suggesting that the results in SOSE-OSSE are reliable until at least Day 5. Figure 16b shows much poorer forecast skills in TESTP than TESTmp because DWL assimilation in TESTP only exploits the wind component of PT in the limited spatial and temporal coverage.

Secondly, we implemented two experiments (TESTqc and TESToe) to assess the essential impacts of QC and observation error setup. In TESTqc, we almost deactivated the QC described in Section 2.4b by setting the threshold of 1.e-6 for SNR-QC and 5.0 m s$^{-1}$ irrespective of altitude for gross error QC. These new QC thresholds allowed more data assimilated (on
average 2.2 times more) at the cost of lower quality.
In TESToe, we set the observation errors of DWLs
to those assigned to radiosonde in the LOS direction
without inflation, instead of those defined in Section
2.4c. The radiosonde observation errors varied from
2.3 m s\(^{-1}\) at 1000 hPa to 3.1 m s\(^{-1}\) at 100 hPa for both
the zonal and meridional components but did not have
the seasonal and spatial dependency. The global aver-
age of the relative forecast error reductions in August
experiment for TESTqc (Fig. 16c) and TESToe (Fig.
16d) show significant degradation: the degradation in
TESTqc is evident for the whole forecast range up to
Day 5 in almost all the layers. Although impacts in
TESToe become neutral after Day 3, the degradation
of 2 % to 3 % is found at Days 1 and 2 with statistical
significance. These two experiments again confirm the
crucial importance of adequate QC and observation
error assignment. It is interesting that assimilating
low-quality DWL observations in TESTqc brought
longer degradation than assigning inadequate observa-
tion errors in TESToe, which will be a topic of future
study.

4. Summary and conclusion
This study aims to develop an OSSE system and to
assess the impact of space-borne coherent DWLs. As
an input for a lidar simulation, global hourly PT at-
mospsheric fields were created on the basis of a SOSE
approach, which allowed us to use existing observa-
tions instead of simulating them as in the NR-OSSE
approach. A global, hourly aerosol field was generated
using a global chemical transport model called
MASINGAR that was constrained by PT wind fields.
An extensive lidar simulator ISOSIM-L, calculated re-
alistic LOS wind speeds and their quality information
(N\(^{1/2}\)SNR and measurement errors), which were highly
dependent on aerosol and cloud distributions. A global
data assimilation system with a 4D-Var method opera-
tionally employed as of 2010 was extended to treat
DWL LOS winds. A preprocessor for DWL in the data
assimilation system obtained observations with high
N\(^{1/2}\)SNR and small first-guess departure and assigned
observation errors as a function of the measurement
errors estimated in ISOSIM-L.
Cycle assimilation experiments were conducted
in January and August 2010 to assess the impacts of
The assimilation of DWL LOS winds onboard either polar- or tropical-orbiting satellites showed generally positive impacts on wind and temperature forecasts. This was especially evident for the January experiment. For the track forecasts of TCs, DWL impact was positive but small, partially because of the TC bogus, in both polar and tropical DWL experiments. We were not able to draw a clear conclusion on the relative merits of polar- or tropical-orbiting satellites because of the complicated structure of relative forecast error reduction, which may suggest possible problems in the data assimilation system and SOSE-OSSE in the tropics in our study.

This study developed new preprocessing methods (QC and observation error assignment) that took into account the observation quality information estimated by the lidar simulator. The preprocessing methods successfully brought positive impacts from DWL assimilation, but still need additional consideration to improve them. For example, as suggested in Section 2, it could be better to set smaller observation error inflation where existing observations are sparse. Another interesting finding of this study is the impact difference between January and August. The seasonal difference in impacts is often seen for observation system experiments with actual observations and has plenty of factors, such as the dependence of observation quality and model performance on atmospheric conditions. Thus, the careful development of preprocessing methods is necessary to deal with various atmospheric conditions and avoid overtuning for particular atmospheric cases.

Moreover, there are probably errors that actual DWL observations have but ISOSIM-L does not take into account. For example, Sun et al. (2014) estimated the height assignment errors due to vertical wind shear. Marseille et al. (2010) indicated that the vertical wind component can be generally negligible, as assumed in Eq. (1), but its contribution can be locally significant. Our simple estimate with a global atmospheric field at a certain day suggested that taking the vertical component into account can make at most a 15% difference in retrieving LOS winds for small horizontal wind and relatively large vertical wind. We can probably treat this issue by increasing observation errors when the relatively large vertical motion may be anticipated on the basis of background atmospheric field conditions.

We are planning to implement OSSE with various observation configurations, such as DWL coverage and resolution. For example, it will be interesting to assess the impact of denser DWL observations and compare it with TESTs and TESTmp experiments. Furthermore, to clarify what exactly causes seasonal difference in the quality of DWL simulation and forecast impacts, additional experiments in different years and months are necessary. These experiments will help us distinguish between the seasonal difference and monthly variability. Finally, we emphasize the investigation of the complementarity between DWL and AMV. Although AMVs have a wide coverage and frequent updates, DWL is capable of providing accurate and vertically high-resolution measurements (Ishii et al. 2016). High accuracy in the height assignment, even in the low troposphere day and night, is also an advantage for DWL, which is important for predicting severe weather induced by strong moisture inflow that often happens in Japan. The synergetic use of DWL and AMV will improve the analysis and forecasts of wind fields, for example, by correcting height assignment error in AMVs.

Acknowledgments

We would like to thank Masaki Sato, Toshiki Iwasaki, Tsuyoshi Sekiyama, Takashi Maki, and Riko Oki for the fruitful discussion and support. We are also very grateful to Michiko Masutani for detailed discussion. A part of this research was supported by JSPS KAKENHI Grant Numbers 15K05293 and 15K06129.

Appendix: Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4D-Var</td>
<td>Four-dimensional variational analysis</td>
</tr>
<tr>
<td>AMSR-E</td>
<td>Advanced Microwave Scanning Radiometer for EOS</td>
</tr>
<tr>
<td>AMSU</td>
<td>Advanced Microwave Sounding Unit</td>
</tr>
<tr>
<td>AMV</td>
<td>Atmospheric motion vector</td>
</tr>
<tr>
<td>DWL</td>
<td>Doppler wind lidar</td>
</tr>
<tr>
<td>ECMWF</td>
<td>European Centre for Medium-Range Weather Forecasts</td>
</tr>
<tr>
<td>ERA</td>
<td>ECMWF Reanalysis</td>
</tr>
<tr>
<td>ESA</td>
<td>European Space Agency</td>
</tr>
<tr>
<td>GNSS-RO</td>
<td>Global Navigation Satellite System Radio Occultations</td>
</tr>
<tr>
<td>ISOSIM-L</td>
<td>Integrated Satellite Observation Simulator for a space-borne coherent Doppler lidar</td>
</tr>
<tr>
<td>JMA</td>
<td>Japan Meteorological Agency</td>
</tr>
<tr>
<td>LOS</td>
<td>Line-of-sight</td>
</tr>
<tr>
<td>MASINGAR</td>
<td>Model of Aerosol Species in the Global Atmosphere</td>
</tr>
<tr>
<td>MHS</td>
<td>Microwave Humidity Sounder</td>
</tr>
<tr>
<td>NR</td>
<td>Nature run</td>
</tr>
<tr>
<td>OSSE</td>
<td>Observing system simulation experi-</td>
</tr>
</tbody>
</table>
PT  Pseudo truth
QC  Quality control
RMSE Root mean square error
SNR  Signal-to-noise ratio
SOSE Sensitivity observing system experiment
SSM/I Special Sensor Microwave Imager
SSMIS Special Sensor Microwave Imager/Sounder
SYNOP Report of surface observation from a fixed land station
TC  Tropical cyclone
TMI TRMM Microwave Imager
TRMM Tropical Rainfall Measuring Mission
WMO World Meteorological Organization

References


Marseille, G.-J., and A. Stoffelen, 2003: Simulation of wind


Shimoji, K., 2014: Motion tracking and cloud height assignment methods for Himawari-8 AMV. *Proc. 12th Int. Winds Workshop*, University of Copenhagen, Danish Meteorological Institute, Copenhagen, Denmark, S2–07.


