Wind extraction potential from 4D-Var assimilation of O$_3$, N$_2$O, and H$_2$O using a global shallow water model

D. R. Allen, K. W. Hoppel, and D. D. Kuhl

Remote Sensing Division, Naval Research Laboratory, Washington, DC, USA

Received: 3 September 2013 – Accepted: 18 September 2013 – Published: 1 October 2013

Correspondence to: D. R. Allen (douglas.allen@nrl.navy.mil)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The wind extraction due to assimilation of trace gas (tracer) data is examined using a 4D-Var data assimilation system based on the shallow water equations coupled to the tracer continuity equation. The procedure is outlined as follows. First, a Nature Run is created, simulating middle stratospheric winter conditions. Second, ozone (O₃), nitrous oxide (N₂O), and water vapor (H₂O) (treated in this study as passive tracers) are initialized using Microwave Limb Sounder (MLS) mixing ratios at 850 K potential temperature and advected by the Nature Run winds. Third, the initial dynamical conditions are perturbed by using a 6 h offset. Fourth, observations based on the simulated tracer data are then assimilated with a 4D-Var system in which the tracer and winds are coupled via the adjoint of the tracer continuity equation. Finally, the wind extraction potential (WEP) is calculated as the reduction of the Root Mean Square (RMS) vector wind error due to tracer assimilation relative to the total possible reduction from the initial perturbed conditions. For a single 6 h assimilation cycle of “perfect” tracer (unbiased and no imposed random errors), WEP values are 70 % for O₃, 49 % for N₂O and 16 % for H₂O. O₃ and N₂O provide more wind information than H₂O due to stronger background gradients relative to the tracer precisions. 10 day multi-cycle simulations with “perfect” tracer result in WEP of 98 % for O₃, 97 % for N₂O, and 90 % for H₂O. There is therefore sufficient information in these fields to nearly completely specify the dynamics, even without assimilation of dynamical information. For assimilation of tracer observations with realistic random noise (based on MLS precision at 10 hPa), the WEP after 10 days decreases to 57 % for O₃, 42 % for N₂O, and 28 % for H₂O. The root-mean-square (RMS) wind errors level out at ~ 1–2 ms⁻¹ for these runs, suggesting a limit to which realistic tracers could constrain the winds, given complete global coverage. With higher observation noise levels, the WEP values decrease further, with negative WEP occurring in cases of very large errors for H₂O, indicating that assimilation of very noisy observations may worsen the wind fields.
1 Introduction

This study examines the problem of extraction of global wind information from the assimilation of atmospheric trace gas (tracer) observations. This inverse problem for the 3-D atmosphere poses many challenges that have heretofore prevented the dynamical coupling of trace gas observations and winds in operational numerical weather prediction (NWP) systems (Han and McNally, 2010; Dragani and McNally, 2013). Among these challenges are paucity of observations, insufficient data quality, insufficient modeling of tracers in the forecast model, as well as phenomenological limitations due to tracer variations with respect to the wind vector. The problem of tracer-wind extraction has been examined using a series of increasingly sophisticated models ranging from 1-D (Daley, 1995; Allen et al., 2013), 2-D (Daley, 1996; Riishøjgaard et al., 1996), and 3-D (Peuch et al., 2000; Semane et al., 2009; Allen et al., 2013). A review of the results from these studies is provided in Allen et al. (2013). The basic conclusion is that with complete data coverage, sufficient precision, a high-quality background tracer field, and a 4D-Var or Extended Kalman Filter algorithm, wind extraction is possible across the globe and throughout the stratosphere and upper troposphere. However, the ability to extract wind is greatly hindered by limitations in any of these parameters, making the current value of operational tracer-wind coupling in NWP systems questionable.

Given the observational and modeling challenges, one may ask whether the continued pursuit of tracer-wind extraction is worthwhile. One major motivation is the current deficiency of adequate global wind profile information (World Meteorological Organization, 2000). Upper air wind observations from pilot reports, radiosondes, and cloud and water vapor feature-tracking leave large gaps, particularly in the tropics, Southern Ocean, and in most of the stratosphere and mesosphere. In the extratropics, traditional mass-wind balance provides some constraint on the wind from temperature information based on radiance observations, but this balance fails in the tropics and for sub-synoptic scales (less than ~1000 km) in the extratropics. While spaceborne Doppler Wind Lidar (DWL) has the potential to provide wind profile measurements in
the troposphere and lower stratosphere, it is extremely expensive (National Research Council, 2007), and the only currently planned DWL mission, Atmospheric Dynamics Mission (ADM-Aeolus) (Stoffelen et al., 2005), will be limited to a single line-of-sight wind component, altitudes below $\sim 26$ km, and simple along-track sampling.

Assimilation of tracer observations for wind information may help to fill this gap. In the studies to date, the main trace gas that has been considered is ozone ($O_3$), due to its relatively long lifetime, the roles it plays in the radiative heating/cooling in NWP systems and infrared radiances assimilation, and the availability of high quality satellite observations (e.g., from the Suomi National Polar-Orbiting Partnership (NPP) Ozone Mapping and Profiler Suite (OMPS); Flynn, 2009). There are other long-lived tracers, however, that could be considered as potentially useful for wind extraction. These include nitrous oxide ($N_2O$), water vapor ($H_2O$), methane ($CH_4$), chlorofluorocarbons, etc. Modeling each of these species in a complete assimilation system poses many challenges, but in principle these tracers contain dynamical information that could be “harvested” to benefit the winds, particularly in the stratosphere. This paper examines the relative impact on the winds due to assimilation of three of these tracers: ozone, nitrous oxide, and water vapor.

The experiments are performed using a 4D-Var data assimilation system based on the shallow water model (SWM) equations. Application of the SWM equations to the variational assimilation problem has proved useful for algorithm development and basic understanding of the problem (e.g., Courtier and Talagrand, 1990). The SWM system provides sufficient complexity that realistic flow situations can be simulated, while maintaining greater control of the system parameters than in the full NWP system. This study builds on the 2-D tracer assimilation studies by Daley (1996) and Riishøjgaard et al. (1996). In the stratosphere, which is the region we are focusing on, the tracer-wind extraction should be to first order a 2-D quasi-isentropic problem, since the horizontal transport timescales are smaller than the vertical timescales (Plumb, 2002). Near the tropopause, strong vertical ozone gradients and vertical motion make the full 3-D modeling necessary (Riishøjgaard et al., 1996). The experiments in this study
use an observation grid that provides complete global coverage of the tracer in order to maximize the wind extraction. The goal is to determine the relative wind extraction potential due to the geophysical variability of each of the tracers, rather than to determine the impact of the observing grid. This idealized approach will provide guidance as to which tracers theoretically provide the most dynamical information. Future work is planned to examine the impact of various observation systems as well as simultaneous assimilation of tracer and dynamical fields.

Section 2 describes each aspect of SWM data assimilation system, which includes the forecast model, the tangent linear and adjoint models, and the cost function definition and minimization algorithm. Section 3 describes the experimental design for wind extraction from tracer assimilation, including development of a realistic Nature Run, construction of simulated observations, and explanation of perturbed initial conditions. Section 4 presents the results of both single-cycle and multiple-cycle experiments, focusing on the relative impact of the different tracers O_3, N_2O, and H_2O. Sections 5 and 6 provide a discussion and conclusions, respectively.

2 Model description

2.1 Forecast model

The forecast model is a spectral model based on the vorticity-divergence form of the SWM equations as formulated in Section 2a of Ritchie et al. (1988), with the inclusion of fourth-order diffusion applied to the vorticity, height, and geopotential following the semi-implicit formulation. In addition, an advection equation is included, which solves for the mixing ratio of a passive tracer as a function of time with the same fourth-order diffusion operator. The system is run at triangular truncation T42 with a timestep of 1200 s for both the forward model and for the tangent linear model (TLM) and adjoint integrations discussed in Sect. 2.2. The diffusion coefficient is set to $5.0 \times 10^{15} \text{ m}^4 \text{ s}^{-1}$, as in Jakob-Chien et al. (1995). The model fields are saved on the Gaussian grid.
with 128 longitudes × 64 latitudes, for a grid resolution of ∼2.8° at the equator. The discretization uses a leap-frog time integration and a semi-implicit approximation for terms that produce gravity waves (Ritchie, 1988). A constant geopotential height of 10 km is specified for the free upper surface of the fluid. The algorithm is very similar to the “STSWM” code described by Jakob-Chien et al. (1995). The forecast model was validated using tests developed by Williamson et al. (1992) and Galewsky et al. (2004), with excellent results.

2.2 Data assimilation system

We employ a 4D-Var data assimilation system (DAS) using the “accelerated representer” approach (Xu et al., 2005; Rosmond and Xu, 2006), which is based on the solution to a standard 4D-Var cost function defined as follows (note: the “perfect model” assumption is used).

\[
J = J_b^0 + J^r
\]

\[
J_b^0 = \frac{1}{2} \left[ \mathbf{x}_b^0 - \mathbf{x}_0 \right]^T \mathbf{P}_b^0^{-1} \left[ \mathbf{x}_b^0 - \mathbf{x}_0 \right]
\]

\[
J^r = \frac{1}{2} \sum_{n=1}^{N} \sum_{n' = 1}^{N} \left[ \mathbf{y}_n - H(\mathbf{x}_n) \right]^T \mathbf{R}^{-1} \left[ \mathbf{y}_{n'} - H(\mathbf{x}_{n'}) \right]
\]

Here \(J_b^0\) is the cost function for deviations of the initial analyzed state vector \(\mathbf{x}_0\) from the initial background state \(\mathbf{x}_b^0\). In the DAS algorithm, the horizontal flow is converted from the forecast model variables (vorticity and divergence) to zonal and meridional wind (\(u\) and \(v\)). \(\mathbf{P}_b^0\) is the initial background error covariance matrix (BECM). \(J^r\) is the cost function for the observations, where \(\mathbf{y}_n\) is the vector of observations at time \(t_n\), \(H(\mathbf{x}_n)\) is the observation operator that maps the model state vector onto the observations, and \(\mathbf{R}\) is the observation error covariance matrix (OECM). The cost function is defined over the analysis window of length 6 h with a total of \(N = 18\) time steps. The representer
method is used to obtain the solution, incorporating the linearized forecast model $M$ and the linearized observation operator $H$.

$$x^a = x^b + P^b H^T [H P^b H^T + R]^{-1} [y - H x^b]$$

where

$$P^b = M \left[ M_0 P_0 M_0^T \right] M^T$$

The background state is calculated by integrating the non-linear forecast model $M$, subject to the initial conditions $x^b_0$:

$$x^b_n = M \left( x^b_{n-1} \right) \text{ for } 1 \leq n \leq N$$

The analysis $x^a$ is obtained with the solver and post-multiplier approach described in Xu et al. (2005) using a standard conjugate gradient descent algorithm. In Eq. (2) both the TLM, $M$, and its transpose (or adjoint), $M^T$, are integrated over the analysis window during each iteration of the descent algorithm in order to “sweep through” all observations and to calculate the time-evolving BECM. In 4D-Var the coupling of tracer and dynamical fields (wind, temperature, and geopotential) occurs implicitly via the adjoint of the TLM, which propagates sensitivities of the cost function with respect to the tracer observations backwards in time to obtain sensitivities of the cost function to the initial model state (see Allen et al., 2013, for further discussion). The solution is iterated until the gradient of an intermediate cost function (see Daley and Barker, 2001 for details) drops by a factor of 20 (ratio to initial gradient of 0.05). If the solution does not reach this threshold, the solver is terminated after 100 iterations. The TLM is constructed from the full nonlinear model by expanding each forecast model state variable into time-evolving background and perturbation terms and removing terms that are quadratic in perturbations. The quality of the TLM, as discussed in Sect. 3, is quite good over the 6 h assimilation window for the initial conditions used in this paper. The adjoint model development was performed using a line-by-line approach, following the
adjoint algorithms of Giering and Kaminski (1998) and guidance from Rosmond (1997). Each piece of code was validated by the “gradient test” to ensure correct coding.

The initial BECM ($P_0^b$) is defined following the approach described in Daley (1991) and Daley and Barker (2001). The wind-wind correlations are calculated from the streamfunction correlations using the Helmholtz equations, while the streamfunction correlations are isotropic and homogeneous with a spatial correlation based on the single-order auto regressive (SOAR) function. Various values of the horizontal correlation lengths and background error standard deviations were tested to maximize tracer-wind extraction. We decided to use a length that varied smoothly with latitude from 1500 km in the tropics to 1000 km in the extratropics (see Fig. 1a). This length scale works well for the simulations done here at T42, but may need to be altered for simulations at higher resolution. The wind-geopotential correlations are based on approximate geostrophic balance, with a coupling between the geopotential and the streamfunction that varies with latitude as illustrated in Fig. 1b. The ratio of the divergent to total horizontal kinetic energy in the analysis increments was defined as 0.05. Background error standard deviations are based on a wind value of 4 m s$^{-1}$. The background error standard deviation for geopotential height is calculated from the wind value and the scaling defined in Daley (1991). Figure 1c and d show the background wind and geopotential height error standard deviations for these settings. The tracer background error covariance is specified to be diagonal with a standard deviation equal to a constant value (0.1 parts per million by volume (ppmv) for O$_3$, 0.01 ppmv for N$_2$O, and 0.3 ppmv for H$_2$O); therefore no tracer information is passed to the dynamical variables through the initial BECM. There is also no spatial correlation imposed on the tracer increments, allowing tracer increments of very fine scale.

The observation operator $H$ is defined by interpolating from tracer data on the Gaussian grid, which is regular in longitude, but irregular in latitude. The longitudes were wrapped around the zero meridian and the latitudes were wrapped across each pole to form an intermediate extended array. This array was then used to bi-linearly interpolate the background tracer to the locations of the observations. Since the operator is de-
fined to be linear, we have $H = H$. The hourly observations are specified at an integral number of time steps, so no time interpolation is necessary. The adjoint of the observation operator was also derived using the adjoint algorithms of Giering and Kaminski (1998). The OECM was defined as diagonal for these experiments, with no spatial or temporal correlations between observations. Since only tracer observations are assimilated, there is no information passed to the dynamical variables through the OECM. The observation error standard deviation $\sigma_{ob}$ is a constant value in terms of the tracer mixing ratio, as specified below.

3 Experimental design

The wind extraction experiments are designed using an approach similar to Allen et al. (2013) in which a Nature Run is first performed, which is considered to be the truth, both for the tracer and dynamical fields (winds and geopotential). This run is used to create “perfect” observations on the full T42 Gaussian grid. The observations may or may not include imposed errors, as specified for each experiment. An initial, imperfect, 6 h background forecast (or “background trajectory”) is then created by adding perturbations to the dynamical fields of the Nature Run. A 6 h data assimilation cycle is performed to create a 6 h “analysis trajectory.” Comparing the background and the analysis with the Nature Run allows determination of the impact of tracer assimilation on the dynamical fields. Below we describe each component of these wind extraction experiments in more detail. We note here that since the same forecast model is used for the Nature Run and for the assimilation, the results are likely over optimistic. However, the intent of this study is not to determine the impact of any realistic observing system, but rather to examine the theoretical limitations of the wind information contained within the tracer field.

The Nature Run was designed to simulate Northern Hemisphere (NH) winter conditions in the middle stratosphere. This was based on previous simulations by Juckes (1989), Norton (1994), and Thuburn and Lagneau (1999), in which an SWM system
is forced by the bottom topography being raised and lowered in order to simulate planetary-scale waves. The model was initialized with a zonally-symmetric zonal wind that varies with latitude ($\phi$) as $u(\phi) = u_{\text{max}} \sin(2\phi)$ in the NH and is zero in the Southern Hemisphere (SH). The jet strength is set by $u_{\text{max}} = 60 \text{ m s}^{-1}$. The initial geopotential height $h = \Phi/g$, where $\Phi$ is geopotential and $g$ is the Earth’s gravitational acceleration, is specified using the gradient wind balance and given a mean height of 10 km. The topographic forcing is produced following the approach of Norton (1994). A mountain of height 1250 m is created with a 20 day cycle (4 days ramping up, 12 days constant, and 4 days ramping down). The mountain is a zonal wave 1 feature that peaks at 45° N. The topography is turned off after 20 days. Since the assimilation period starts at day 20 of this Nature Run, there is therefore no surface topography during the assimilation.

In the Nature Run the tracer is initialized using Aura Microwave Limb Sounder (MLS) trace gas data (Waters et al., 1999; Livesey et al., 2011). The data are selected for a period with weak planetary wave activity (1–15 March 2011) and are interpolated to the 850 K isentropic level (approximately 32 km altitude or 10 hPa), representative of middle stratosphere conditions. The zonal mean and time mean mixing ratio as a function of latitude was calculated for this period and interpolated to the Gaussian grid. Simulations were performed with O$_3$, N$_2$O, and H$_2$O. The tracer in each case is treated as passive (i.e., no chemical source/sink) and there is no radiative interaction between tracer and dynamics. N$_2$O and H$_2$O have relatively long photochemical lifetimes in the middle stratosphere, as does O$_3$ in the winter hemisphere. However, the photochemical lifetime of O$_3$ becomes shorter (about 10 days) in the tropics and summer hemisphere at 10 hPa (e.g., McCormack et al., 2006, Fig. 2), and therefore photochemical affects will become more important. The impacts of chemistry and radiation on the wind extraction process are reserved for future work.

Figure 2 plots the potential vorticity ($Q = \zeta_a/h$, where $\zeta_a$ is the absolute vorticity) in 5 day intervals starting on the 15th day during the Nature Run along with the three tracers. The wave forcing produces an “Aleutian High” (centered at 30° longitude) type wave response in the NH with the polar vortex (centered near 180° longitude) pushed
off the pole. Low $O_3$ and $N_2O$ (high $H_2O$ and PV) are visible in the polar vortex and high $O_3$ and $N_2O$ (low $H_2O$ and PV) are entrained into the “Aleutian High”. The sharpening of the PV gradients with time along the vortex edge is seen, as expected for these types of wave-breaking simulations. The three tracer fields have very similar contour shapes, as expected for long-lived tracers in a 2-D flow (Rhines and Young, 1983). However, the dynamic range of the tracers is considerably different, resulting in different horizontal gradients. The minimum and maximum values (in ppmv) at initialization (Day 0) are as follows: $O_3$ [4.1, 11.0], $N_2O$ [0.013, 0.23], $H_2O$ [4.6, 6.4]. It is clear that $N_2O$ allows the largest dynamic range, followed by $O_3$ and then $H_2O$.

To start the assimilation experiments, the background forecast is initialized from the Nature Run using the true tracer distribution but with dynamical fields sampled at a six hour offset. This method allows for perturbed initial dynamical fields that are approximately in balance. Although the tracer is perfect at the first time step, tracer errors develop over the six hour background forecast by advection with imperfect dynamical fields. The first assimilation cycle represents the best case scenario for wind extraction because all the tracer error, and hence the deviation relative to simulated observations, is caused by wind error. On further cycling of the assimilation, the background tracer field will retain error from the previous cycle in addition to new error from imperfect advection. The assimilation is initialized 20 days into the Nature Run ($T = 20 \text{ d}, 0 \text{ h}$), and the initial dynamical fields are then taken from 20 days +6 h, while the tracer field is taken from 20 days +0 h. The background forecasts always use the full forecast model with the same settings as the Nature Run. In the multi-cycle simulations the background for each successive cycle is produced from an integration starting with the previous 6 h analysis.

Observations are simulated at each hour and at each Gaussian grid point by sampling the Nature Run and adding varying amounts of random error. Table 1 lists the random error added for each experiment, and the observation error standard deviation, $\sigma_{ob}$ specified in the 4D-Var solution. Experiment 1 is a special “perfect” observation case where no random error is added the observations. In Experiment 2, we specify ob-
observation errors that are consistent with MLS retrieval error at 10 hPa (~30 km), with values of 0.1 ppmv for O₃, 0.01 ppmv for N₂O, and 0.3 ppmv for H₂O (see MLS Version 3.3 level 2 data quality and description document, Livesey et al., 2011). Note that errors are constant values in mixing ratio units, and the relative error will vary depending on the tracer distribution. The MLS values are used here to represent typical state-of-the-art measurements, even though the simulated observation density and frequency are better than any foreseen measurement system. In Experiments 3 and 4, the random error is increased from Experiment 2 by factors of 5 and 10, respectively, to examine the sensitivity of wind extraction to varying observation error.

4 Results

We first examine the results of the first, or single-cycle, assimilation. Examining the first cycle for each experiment allows us to compare wind extraction potential for different tracers and observation errors using the same background wind error and perfect initial background tracer fields. We then examine results from cycling experiments to determine the extent to which tracer assimilation alone is able to constrain the model dynamics. In a full NWP system, tracer assimilation will normally be combined with a full suite of meteorological observations, rather than tracer observations alone. However, the goal of this study is to determine the theoretical limit of the information that can be extracted solely from the tracer field.

First, we examine the O₃ assimilation from Experiment 1. To quantify the results, we show the root-mean-square (RMS) of the background and analysis errors, calculated relative to the Nature Run and averaged in longitude. The 4D-Var analysis should yield an analysis trajectory that is optimal over the entire six hour time window. In Fig. 3 we verify this by showing the errors for the first assimilation cycle at the beginning, middle, and end of the analysis window, along with global maps of the analysis increments (analysis – background) and ideal analysis increments (truth – background). Results are shown for the zonal wind, meridional wind, and height. For all three fields, the anal-
ysis errors are less than the background errors across the globe. The most dramatic reduction occurs for the meridional wind, where errors drop from $\sim 2-4\text{ m s}^{-1}$ down to $\sim 0.5-1.5\text{ m s}^{-1}$. For zonal wind, the reduction also occurs globally. For geopotential, the reduction is greatest in the extratropics of both hemispheres, with smaller reductions observed in the tropics. This is consistent with expectations. The tracer most directly affects the winds via the adjoint of the tracer continuity equation. The influence on the geopotential is secondary via advection of the geopotential and via the wind-geopotential cross-covariances in $P_0$. Since the latter is less constrained by geostrophy in the tropics, it is expected that it will be more difficult to influence the tropical height fields by tracer assimilation. The quality of the assimilated wind fields does not change very much with time over the analysis window. The height field, however, shows more sensitivity with time, with smaller RMS errors in the extratropical SH and the subtropical NH toward the end of the analysis window. In some places (e.g., near $50^\circ$ N) the height errors actually increase with time over the analysis window.

The ability to extract wind is partially dependent on the quality of the TLM integration. To estimate of the TLM error, we also plot in Fig. 3 the RMS difference between the Nature Run and the sum of the background simulation and a TLM integration that was initialized with the ideal increments. For a perfect TLM integration these errors should be zero. Non-zero errors indicate the effects of neglecting the non-linear terms in the SWM equations. The TLM errors are everywhere smaller than the analysis errors, and they show some similar features. For example, the peak in the analyzed zonal wind errors at $40^\circ$ N is co-located with a similar peak in the TLM errors, suggesting that non-linear terms are playing a role there. The TLM errors give an estimate of the absolute limit of wind extraction in the 4D-Var approach. Zonally-calculated RMS wind errors are less than $\sim 0.5\text{ m s}^{-1}$ across the globe and height errors are less than $\sim 5\text{ m}$. The residual difference is then due to other factors such as limited tracer gradients, inexact definitions of the BECM, incomplete observations, etc. Comparing the increments in Fig. 3, we see that the analysis increments (middle row) match most of the large-scale features in the ideal wind and height increments (bottom row). There are some small-
scale structures in the ideal wind increments that are smoothed out in analysis due to the horizontal correlations specified in the $P^b_0$, but these appear to be a second-order effect.

We next compare Experiment 1 single-cycle results for all three tracers. Figure 4 shows the background and analysis errors for $O_3$, $N_2O$, and $H_2O$. For all three tracers, wind and height errors decrease after tracer assimilation. The greatest error reductions are due to $O_3$ assimilation. $N_2O$ provides less wind error reduction than $O_3$, while $H_2O$ has clearly smallest impact on the winds. We further quantify these results by introducing the Wind Extraction Potential (WEP), which we define as the ratio of the reduction in global mean RMS vector wind error to the total possible reduction, multiplied by 100 (see Allen et al., 2013). A WEP value of 100% indicates the analysis equals the truth. Table 1 lists the WEP values for all experiments in the study. For Experiment 1, the WEP after one cycle is 70% for $O_3$, 49% for $N_2O$ and 16% for $H_2O$. These results can be qualitatively explained by relating the global dynamic tracer range (maximum – minimum mixing ratio) to the specified precision of each tracer. This ratio (at the beginning of this first assimilation cycle) is $\sim 73$ for $O_3$, $\sim 23$ for $N_2O$, and $\sim 6$ for $H_2O$. Although the relative precision of $N_2O$ is larger than that of $H_2O$, the dynamic range of $N_2O$ is greater, causing it to be a better tracer for wind extraction. This is a simple rule-of-thumb relationship. More precisely, the wind extraction is influenced by magnitude and orientation of the local tracer gradients relative to the wind errors (see Allen et al., 2013, for further discussion).

These results represent only one level (850 K or $\sim 10$ hPa) and one set of dynamical conditions. Analysis at other levels and conditions will result in varying wind-extraction potentials. For example, we ran similar tests using a Nature Run that was initialized with MLS data at 460 K ($\sim 70$ hPa) and used precision values representative of that level, but used the same dynamical fields. The resulting WEP values after one cycle (48% for $O_3$, 18% for $N_2O$, and 2.6% for $H_2O$) were smaller than at 850 K. This can be qualitatively explained by smaller range-to-precision ratios at 460 K ($\sim 37$ for $O_3$, $\sim 12$ for $N_2O$, and $\sim 4$ for $H_2O$) relative to those at 850 K. While the precision value of
O\textsubscript{3} is smaller at 460 K (0.04 ppmv), the dynamic range is much smaller due to lower mixing ratios. For N\textsubscript{2}O, the precision value is larger at 460 K (0.02 ppmv), while the range is slightly smaller. For H\textsubscript{2}O, the precision is the same as at 850 K (0.3 ppmv), but the range is smaller. These results are consistent with Allen et al. (2013), who showed greater wind extraction in the middle and upper stratosphere than in the lower stratosphere for a 4D-Var ozone assimilation experiment using a full NWP system.

Next we examine Experiments 2–4, which impose increasing random error on the observations. As shown in Fig. 5, increasing error reduces wind extraction. For O\textsubscript{3} and N\textsubscript{2}O there is still noticeable reduction in wind and height errors even with an error 10 times larger than MLS. This is consistent with the results from Allen et al. (2013), who performed single-cycle O\textsubscript{3} assimilation with a full 3-D system. For H\textsubscript{2}O the increased errors result in almost no wind extraction, and in some cases there is a slight increase in both wind and height errors (see, e.g., the height error at 60° S for imposed errors of 1.5 and 3.0 ppmv). Care must obviously be taken when assimilating very noisy data into such a system. As shown in Table 1, WEP values decrease for each tracer as the data errors increase, with values becoming negative (i.e., worsening winds) for Experiment 4 for H\textsubscript{2}O (WEP = −0.14). The single-cycle results show that the amount of tracer-wind extraction depends on the chosen tracer as well as on imposed errors. A single update cycle starting from a perfect background tracer field is far from realistic, however.

We next performed experiments that are based on 10 days of assimilation, corresponding to days 20 to 30 of the Nature Run. Although the initial background tracer field is perfect for the first cycle, background tracer errors develop over subsequent cycles due to imperfect wind fields. For these runs, we continue to use observation coverage on the full Gaussian grid at hourly intervals. The WEP results after 10 days of assimilation are summarized in the last column of Table 1.

Figure 6 shows the globally-averaged RMS wind, height, and tracer errors and WEP as a function of time for Experiment 1. The zonal and meridional wind errors drop from initial values of ~ 3 m s\textsuperscript{-1} to ~ 0.06 m s\textsuperscript{-1} for O\textsubscript{3}, ~ 0.1 m s\textsuperscript{-1} for N\textsubscript{2}O, and ~ 0.3 m s\textsuperscript{-1} for H\textsubscript{2}O after 10 days. The height errors drop from ~ 40 m to ~ 3 m for O\textsubscript{3}, ~ 4 m for N\textsubscript{2}O,
and $\sim 5$ m for $\text{H}_2\text{O}$. The WEP values after 10 days are 98 % for $\text{O}_3$, 97 % for $\text{N}_2\text{O}$, and $\sim 90$ % for $\text{H}_2\text{O}$. It is clear that assimilation of perfect hourly global tracer data is sufficient to globally constrain the dynamical fields as well as the tracer field. The tracer errors, scaled by the specified observation error standard deviation for each experiment (listed in Table 1), are also shown in Fig. 6. Even though tracer errors increase over the first couple cycles from initially perfect conditions, tracer errors become quite small (less than 1 %) after 10 days. Figure 7 shows the variation with latitude of the errors after 10 days for Experiment 1. For $\text{H}_2\text{O}$, the wind errors are considerably larger in the tropics than near the poles. Some latitude dependence may be expected due to the sampling of tracer observations on the Gaussian grid, which results in more observations near the poles. The cost function minimization would therefore, in principle, preferentially reduce tracer errors near the poles, as is seen in the $\text{H}_2\text{O}$ error plot on Fig. 7. In addition, since $\text{H}_2\text{O}$ mixing ratio is smaller in the tropics (see Fig. 2), the fractional observation error will be larger than in the extratropics, thereby providing less constraint to the winds. For $\text{O}_3$ and $\text{N}_2\text{O}$, the wind errors are relatively constant across the globe, with a slight increase at the North Pole.

In Fig. 8, the impact of imposed random tracer observation errors based on MLS precision values (Experiment 2) is examined for the multi-cycle experiments. Winds are far less constrained by the noisy tracer observations than in the case of perfect data. The global RMS zonal wind errors drop initially, but then level out with time to values of $\sim 1.3$ m s$^{-1}$ for $\text{O}_3$, $\sim 1.8$ m s$^{-1}$ for $\text{N}_2\text{O}$, and $\sim 2.2$ m s$^{-1}$ for $\text{H}_2\text{O}$ (errors are slightly less for meridional wind than for zonal wind). The height errors also drop initially, but then increase slowly with time before leveling out at values of $\sim 25$ m for $\text{O}_3$, and $\sim 30$ m for $\text{N}_2\text{O}$ and $\text{H}_2\text{O}$ after 10 days. The tracer errors also level out after a few cycles, showing that the simultaneous analysis of winds and tracer is possible, even with realistic data errors. Final RMS tracer errors are $\sim 25$ % of the specified random observation error standard deviations for each tracer. Overall, the relative impact on winds from the three tracers is consistent with previous results, with WEP for after 10 days of 57 % for $\text{O}_3$, 42 % for $\text{N}_2\text{O}$, and 28 % for $\text{H}_2\text{O}$. Ozone consistently provides the
greatest benefit to winds, followed by N$_2$O and then H$_2$O. Figure 9 shows the latitude dependence of the errors after 10 days for Experiment 2. Wind and height errors are relatively constant with latitude for each tracer. The tracer errors, however, show a very strong dependence with latitude, maximizing in the tropics and minimizing at the poles. The multiple cycle runs with increased error (5 times and 10 times the MLS value, Experiments 3 and 4) exhibit a corresponding decrease in the WEP after 10 days (see Table 1). For O$_3$ and N$_2$O the WEP is still positive, but for H$_2$O it decreases to $-49\%$, indicating a significant worsening from the initial conditions when assimilating very noisy data.

5 Discussion

The use of trace gas data to benefit winds in 4D-Var NWP systems is an attractive prospect, given the paucity of direct wind observations in the stratosphere and over the oceans in the troposphere. The actual implementation, however, is challenging due to a number of factors such as data quality, temporal sampling, spatial coverage, etc. Although there has been some promising work in the area (e.g., Semane et al., 2009), operational tracer-wind coupling has not (to our knowledge) yet been implemented. Han and McNally (2010), for example, reported that Solar-Backscatter Ultra Violet (SBUV) ozone assimilation could degrade the operational European Centre for Medium-Range Weather Forecasts (ECMWF) 4D-Var wind analyses. In order to prevent erroneous wind increments, they stated that “the observation operator that links wind adjustments to changes in ozone concentration has been artificially cut.”

The main tracer that has been used in previous studies is O$_3$, for which observations are readily available, and O$_3$ is also generally included as a prognostic variable in NWP forecast models. H$_2$O is also included in NWP systems, so it provides an additional tracer for wind extraction. However, the results of this study show that the impact of H$_2$O on the dynamics in the middle stratosphere via the 4D-Var mechanisms is far less than for O$_3$, when the observation errors are constrained by MLS precisions. N$_2$O
provides a third interesting possibility, which for these experiments rivals the potential benefits from O₃. One advantage of N₂O over O₃ is that the photochemistry of N₂O is less important, given a long photochemical lifetime in the stratosphere. In this study, the tracers were assumed to be passive. Inclusion of O₃ and N₂O errors in photochemical parameterizations could tip the scales towards N₂O as being the more useful tracer for wind extraction at certain levels. However, it is likely that the impact of N₂O would decrease with altitude in the upper stratosphere due to smaller mixing ratio and shorter photochemical lifetime.

This study modeled the wind-tracer coupling using a 4D-Var system, which has certain inherent limitations. In 4D-Var the coupling of tracer and dynamical fields (wind, temperature, and geopotential) occurs implicitly via the adjoint of the TLM, which propagates sensitivities of the cost function with respect to the tracer observations backwards in time to obtain sensitivities of the cost function to the model state at the beginning of the analysis window. The optimization problem involves perturbing the initial model state such that the resulting forecast minimizes the “distance” of the analysis with the background and with the observations, subject to error (and possibly other) constraints. In general, tracer-wind coupling is not specified explicitly in the BECM. Realistic tracer-wind coupling is more difficult to impose with a static initial BECM than the wind-temperature coupling. The latter takes advantage of dynamical balance approximations (Daley, 2001). No such balances are available between tracer and winds, although correlations between certain tracers and potential vorticity may prove useful (Li et al., 1998; Elbern et al., 2010).

One way to obtain more realistic tracer-wind correlations in the BECM is to use ensemble statistics to propagate information from one observed variable to the other ones. This approach avoids using the TLM and adjoint completely, relying on an ensemble of forecasts based on the full non-linear forecast model. Milewski and Bourqui (2011) used an Ensemble Kalman Filter (EnKF) with a chemistry-climate model in a system that assimilated realistic stratospheric temperature and ozone observations. It was shown that assimilation of ozone profiles was able to significantly improve the
wind analysis due to the cross-covariances. This suggests that a well-specified BECM may be used to constrain the wind field when assimilating tracer observations. Further work by Massart et al. (2012) examined using ensemble-based ozone BECM specification within a 4D-Var assimilation suite in order to improve ozone analyses. This “hybrid” approach is gaining momentum in the NWP community (e.g., Kuhl et al., 2013), but is relatively unexplored in the context of chemical data assimilation. While Massart et al. (2012) showed that ozone analyses benefitted from improved flow-dependent ozone BECM, their system was univariate (based on a chemistry and transport model driven by offline winds), and so the ozone assimilation did not feedback on the model dynamics. A full hybrid ensemble approach to the complete chemical and dynamical system may produce further improvements to tracer-wind extraction. Although the EnKF and hybrid-EnKF approaches for tracer-wind extraction are very promising, there are still many challenges that must be overcome, not least of which is the computational cost. Therefore, many operational NWP centers (including the US Navy’s operational Fleet Numerical Meteorology and Oceanography Center) are still currently using 4D-Var assimilation techniques.

6 Conclusions

A 4D-Var data assimilation system based on a global shallow water model coupled to a tracer advection equation was used to examine the tracer-wind coupling problem. While a number of additional tracers (methane, chlorofluorocarbons, etc.) could have been tested, we focused on \( \text{O}_3 \), \( \text{N}_2\text{O} \), and \( \text{H}_2\text{O} \), which are readily available from instruments such as MLS. Using realistic tracer data based on a Nature Run, we assimilated hourly observations on the full Gaussian grid. In both idealized single-cycle as well as more realistic 10 day experiments, \( \text{O}_3 \) provided the most dynamical information, closely followed by \( \text{N}_2\text{O} \). Due to weaker gradients relative to the precision, \( \text{H}_2\text{O} \) provided the least impact on the winds. Wind extraction is more difficult in the presence of data errors, although the overall impact on the winds is generally positive for a single cycle
even with large errors, as long as the observation error covariance matrix is specified consistently with the imposed errors. Assimilation of perfect tracer data for 10 day multiple cycle experiments shows that the observations are able to sufficiently constrain the model winds, height, and tracer, suggesting that there is, in principle, enough information in the tracer field to constrain complete system. Assimilation of multiple cycles of data with imposed random errors also results in stable conditions, but with larger final analysis errors and reduced wind extraction potential. However, a slow increase of height errors with time in the case of noisy data may indicate the need for some additional dynamical observations to constrain the system.

Acknowledgements. This work was funded by the US Office of Naval Research.

References


Wind extraction from 4D-Var trace gas assimilation

D. R. Allen et al.


Table 1. Specifications of assimilation experiments used in this paper. Wind extraction potential (WEP) is percentage of potential global vector wind reduction (see text for details). For the WEP after 10 days (last column), the value shown is the average value over the past four cycles. Note that ppmv refers to parts per million by volume.

<table>
<thead>
<tr>
<th>Tracer</th>
<th>Experiment</th>
<th>$\sigma_{ob}$ (ppmv)</th>
<th>Imposed Obs Error (ppmv)</th>
<th>WEP (%) (one cycle)</th>
<th>WEP (%) (after 10 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_3$</td>
<td>1</td>
<td>0.10</td>
<td>0.0</td>
<td>70</td>
<td>98</td>
</tr>
<tr>
<td>$N_2O$</td>
<td>1</td>
<td>0.01</td>
<td>0.0</td>
<td>49</td>
<td>97</td>
</tr>
<tr>
<td>$H_2O$</td>
<td>1</td>
<td>0.30</td>
<td>0.0</td>
<td>16</td>
<td>90</td>
</tr>
<tr>
<td>$O_3$</td>
<td>2</td>
<td>0.10</td>
<td>0.10</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>$N_2O$</td>
<td>2</td>
<td>0.01</td>
<td>0.01</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>$H_2O$</td>
<td>2</td>
<td>0.30</td>
<td>0.30</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>$O_3$</td>
<td>3</td>
<td>0.50</td>
<td>0.50</td>
<td>34</td>
<td>33</td>
</tr>
<tr>
<td>$N_2O$</td>
<td>3</td>
<td>0.05</td>
<td>0.05</td>
<td>18</td>
<td>31</td>
</tr>
<tr>
<td>$H_2O$</td>
<td>3</td>
<td>1.50</td>
<td>1.50</td>
<td>1.2</td>
<td>6.0</td>
</tr>
<tr>
<td>$O_3$</td>
<td>4</td>
<td>1.00</td>
<td>1.00</td>
<td>23</td>
<td>15</td>
</tr>
<tr>
<td>$N_2O$</td>
<td>4</td>
<td>0.10</td>
<td>0.10</td>
<td>13</td>
<td>4.5</td>
</tr>
<tr>
<td>$H_2O$</td>
<td>4</td>
<td>3.00</td>
<td>3.00</td>
<td>−0.14</td>
<td>−49</td>
</tr>
</tbody>
</table>
Fig. 1. (a) Specified background error correlation length scale (in km) as a function of latitude. (b) Geostrophic coupling parameter (unitless) as a function of latitude. (c) Wind background error standard deviation (in m s\(^{-1}\)). (d) Geopotential height background error standard deviation (in m).
**Fig. 2.** Global maps of PV, O$_3$, N$_2$O, and H$_2$O (columns 1, 2, 3, and 4, respectively) for days 15, 20, 25, and 30 (rows 1, 2, 3, and 4, respectively) during the Nature Run. For all maps red (blue) contours indicate high (low) values. Units are $10^{-9}$ m$^{-1}$ s$^{-1}$ for PV and parts per million by volume (ppmv) for O$_3$, N$_2$O, and H$_2$O.
Fig. 3. (Top row) RMS zonal wind, meridional wind, and height errors as a function of latitude for the O$_3$ assimilation from Experiment 1. Black lines show the background errors at $t = 0$ h. Blue lines show analysis errors at $t = 0$ h (solid), 3 h (dotted), and 6 h (dashed). Green lines show the TLM errors after 6 h of integration. (Middle row) The analysis increments as a function of longitude and latitude at $t = 0$ h. (Bottom row) The ideal increments as a function of longitude and latitude at $t = 0$ h. For the increment maps red (blue) contours indicate high (low) values. Units are m s$^{-1}$ for zonal and meridional wind and m for height.
Fig. 4. RMS zonal wind, meridional wind, and height errors as a function of latitude for Experiment 1. Background errors are black, while analysis errors for Experiment 1 at \( t = 0 \) h are shown for \( \text{O}_3 \) (red), \( \text{N}_2\text{O} \) (green), and \( \text{H}_2\text{O} \) (blue). Units are m s\(^{-1}\) for zonal and meridional wind and m for height.
Fig. 5. Background (black) and analysis (colored) errors for zonal wind (top row), meridional wind (middle row), and height (bottom row) for Experiments 2 (red), 3 (green), and 4 (blue) with O$_3$, N$_2$O, and H$_2$O that include imposed random errors with standard deviations as specified in the top row. The errors are calculated at the beginning of the trajectory ($t = 0$ h). Units are m s$^{-1}$ for zonal and meridional wind and m for height.
Fig. 6. Global RMS errors for zonal wind, meridional wind, height, and tracer along with WEP as a function of time for Experiment 1: O₃ (red), N₂O (green), and H₂O (blue). Units are m s⁻¹ for zonal and meridional wind, m for height, and percentage for tracer and wind extraction.
Fig. 7. RMS errors for zonal wind, meridional wind, height, and tracer as a function of latitude Experiment 1: O₃ (red), N₂O (green), and H₂O (blue). Results after 10 days of cycling are shown. Units are m s⁻¹ for zonal and meridional wind, m for height, and percentage for tracer.
Fig. 8. Global RMS errors for zonal wind, meridional wind, height, and tracer along with WEP as a function of time Experiment 2: O$_3$ (red), N$_2$O (green), and H$_2$O (blue). Units are m s$^{-1}$ for zonal and meridional wind, m for height, and percentage for tracer and wind extraction.
Figure 9. RMS errors for zonal wind, meridional wind, height, and tracer as a function of latitude for Experiment 2: O₃ (red), N₂O (green), and H₂O (blue). Results after 10 days of cycling are shown. Units are m s⁻¹ for zonal and meridional wind, m for height, and percentage for tracer.