Reviewer comments in black, our responses in red. A list of changes made to the figures is included after the responses. The manuscript showing tracked changes is appended at the end.

Response to RC1 from Alan Geer, 1 Nov 2017

This is an interesting development in a series of studies by the authors investigating the possibility of wind extraction from tracers in the stratosphere. The latest study moves from using a shallow water model to a full operational-quality NWP system, albeit in an OSSE configuration with a restricted set of observations. The study is well presented and interesting, and the OSSE framework seems novel.

Major comments

1) The introduction needs to be improved in order to summarise the reasons why ozone assimilation has not yet been successful in operational NWP systems. A short recap of the points made in Allen et al. (2015) would be useful here, and it is worth restating that one of the big problems seems to have been bias between model and observations, rather than particularly the deficiencies in the data assimilation framework. The introduction also needs to motivate the current study better, for example justifying why, since real MLS ozone observations are available, the framework of an OSSE been chosen. It would be useful to state here what benefit NAVGEM gets from MLS ozone assimilation, and to recap any studies that might have been done.

The introduction has been rewritten to address the reviewer's comments. The modifications include: (1) a brief recap of points made in Allen et al. (2015), (2) statement about impact of bias between model and observations, (3) justification of using simulated rather than real ozone observations, (4) reference to study by Eckermann et al. (2009) on ozone assimilation within NAVGEM, (5) motivation for the use of both real and simulated data, including references to Harnisch et al. (2013) and Tan et al. (2007).

A number of minor issues relate to these issues in the introduction:

- Page 2 Line 27/29: "Theoretical studies" may not be the best description of work like that of Peubey and McNally (2009) which tested a real operational NWP system

Changed "Theoretical studies" to "Various studies" to be more inclusive of all of the referenced studies.

- Page 2 Line 32: "the operational benefit has not yet been obtained" needs to precisely relate to stratospheric tracers, rather than all tracers, as NWP centres routinely benefit from 4D-Var tracing of tropospheric water vapour from IR and MW radiances, as explained by Peubey and McNally (2009)

Changed the discussion to separate the tropospheric studies from the stratospheric studies.

2) Based on shallow-water results, line 10 of the introduction says roughly that "including cross-correlations in B between ozone and other variables provides additional ozone-wind benefit". I would

1

have expected to see experiments in this new study that would have explored whether this remains true in the full NAVGEM framework.

The experiments in Sect. 4.2 with different alpha coefficients was designed to address this issue. The alpha=0.0 result is for standard 4D-Var (neglecting initial cross-correlations), while non-zero alpha includes the initial cross-correlations. We showed that using non-zero alpha improved the winds and temperatures at the upper levels. We note that the sensitivity to alpha was only performed for the case of global ozone assimilation with no imposed errors. The subsequent experiments all used alpha=0.5. We have not yet tested the sensitivity to alpha for the case of realistic ozone with/without radiances.

3) The OSSE design is described and tested in secs. 2.4 and 4.1. The approach of using real observations in the troposphere seems novel, so it deserves more (critical) investigation within this study. If there are any precedents to this design, they should be cited. There is one interesting parallel with the work of Harnisch et al. (2013) which also used a mixture of real and simulated observations, albeit in an EDA framework.

We thank the reviewer for reference to Harnisch et al. (2013). We include in the revised introduction a discussion of the EDA approach and how it relates to our study. Basically, we attempt to reduce error correlations between simulated and real data by separating the data into two regions: troposphere and stratosphere. The true stratosphere is created by a cycling analysis using only tropospheric data, while the simulated data are only in the stratosphere. As discussed in the introduction, this separation is not perfect, since weighting functions and vertical error correlations can extend across the 100 hPa level. In Section 2.3 we now include plots of analysis increments from the truth experiment to examine the separation in more detail.

The statements in sec. 2.4 that the troposphere is "constant, ... essentially independent" and "this gives ... nearly identical tropospheric analyses for each experiment" are imprecise and omit a key idea for understanding this framework. That idea is later supplied in sec. 4.1: "additional observations ... change the numerics of the cost function ... resulting in slight changes to the troposphere".

We deleted the first sentence quoted and changed the wording on the second sentence to avoid confusion. Further discussion on this point is provided below.

I have run some experiments (Geer, 2016) that explore exactly this issue: when assimilating an identical set of observations into an identical data assimilation system, even the slightest numerical perturbation will generate chaotic divergence between different runs of the data assimilation system. As with my experiments, the troposphere in these new OSSE experiments is not fully constrained by observations and hence will exhibit substantial chaotic variation from one run to the next, whenever the slightest numerical difference is introduced. This spread is somewhat smaller than the expected analysis error for reasons explained in Geer (2016) but is still appreciably large. The troposphere in each separate experiment can be seen as being drawn from a potential ensemble of tropospheric analysis states. The really key understanding is that the stratosphere in the truth experiment (TE) must also be just one realisation drawn from a potential ensemble of stratospheric states, this ensemble being constrained by the data being assimilated in the troposphere. Here also the parallel to Harnisch et al. (2013) is clear.

This is very helpful for analyzing our results. We include a discussion of the tropospheric sensitivity to slight perturbations in Sect. 2.4 and enhance the discussion of Sect. 4.1. In Fig. 14 of that section, we now include three perturbation experiments in which we assimilate perfect global, random, and polar-orbiting data with unperturbed initial conditions.

Thinking this way allows some of the more intriguing results of this study to be analysed better. For example the tropospheric errors shown in Fig. 9 will likely be a rough estimate of the spread of this "tropospheric ensemble" - this would help the authors to explain what they are already saying in Sec. 4.1 about the minimum possible errors in the stratosphere.

We now mention that the errors in this figure (now Fig. 14) are a rough estimate of the spread of a hypothetical "tropospheric ensemble" as well as minimum possible errors in the stratosphere.

However it is thus intriguing that by assimilating stratospheric radiance observations simulated from the TE as in Sec. 5.1, Fig. 15, this tropospheric "spread" gets larger, from around 0.2K to 0.5K. This suggests non-optimalities in the NAVGEM data assimilation system or problems in the OSSE framework that need further investigation. Assimilating observations in an optimal system should make errors go down, not up. For me a likely explanation is that the troposphere and the stratosphere are not in reality independent. In the stratospheric radiance assimilation experiment, the troposphere is not identical to that in the truth experiment, and the stratospheric state most compatible with that tropospheric realisation is different to the stratospheric state in the TE. Hence, by trying to make the experiment stratosphere fit that in the TE, it may require the generation of incorrect increments in the troposphere, which then increase the "spread". Sub-optimalities in the NAVGEM system are also a possibility.

We agree that likely the stratosphere and troposphere are not completely independent. In addition to the perturbation that the radiance observations add to the system, some of the radiance observations are influence by the tropospheric observation. This means the observational errors are not independent, even though they are assumed to be in the DAS. This sub-optimality may be the cause of the increased "spread". We add further discussion on this in Sect. 5.1.

4) In section 4.1, line 10, the expectation of "zero analysis error relative to the TE" with "perfect stratospheric observations" is not so obvious, for a number of reasons:

We changed the discussion in this section as follows.

a) The stratospheric analyses are, even in the presence of near-perfect observations, still just realisations from a hypothetical ensemble of possible stratospheric analyses - this is equivalent to what the authors already say in section 4.1

This is addressed in modifications to Section 4.1. We use the phrase "simple four-member null set" rather than "hypothetical ensemble".

3

b) The 0.1 ppmv observation error is not zero, and hence does permit some additional spread within this hypothetical ensemble

We emphasized this allows for variations in the realized state.

c) there may be sub-optimalities of the data assimilation system, such as sampling error in the ensemble-derived part of the background error covariances

We included the possibility of sub-optimalities.

d) It is not clear that the stratospheric state is fully determined by the ozone field.

We attempt to show by the perfect global ozone case the extent to which the stratospheric state can be determined by ozone alone. We do not say "fully determined" in the revision, but rather specify the errors from this experiment.

I agree with the authors that their experiments in section 4.1 define a minimum "limit on the level of errors we can reliably distinguish from the TE" but the reasons are more complex than currently stated.

Hopefully the additional discussion in this section helped to clarify our results.

It would also be good to see the maximum possible level of error in this experiment, i.e. the "errors" (better, spread) between two realisations of the truth experiment, generated for example by starting another TE from perturbed initial conditions. These minimum and maximum errors roughly define the limits of sensitivity of this novel OSSE design.

The baseline experiments were designed to provide the maximum level of error. For the ozone-only experiments, the baseline experiment assimilated the same data as the truth, but included perturbed initial conditions in the stratosphere. The resulting errors are provided, for example, in the dashed lines on revised Fig. 15. For the ozone and radiance tests, we included bounds of the minimum and maximum expected errors by using the noisy and perfect radiance assimilation experiments.

5) An additional limitation to the sensitivity of this experimental design is the statistical significance of differences between experiments. In an ideal world all the relevant figures should have statistical significance bars added. For example, intriguing results like the decrease in analysis errors when ozone observation errors are increased (page 14, lines 11-13) could possibly be explained by a lack of statistical significance. However, the statistical significance would not be easy to estimate in this framework except by using an ensemble of perturbed experiments similar to what I was using in Geer (2016), and analogous to the "spread" between experiments hinted at by Figure 9.

Considerations of sensitivity and statistical significance are important to the conclusion in sec. 5.2 that "in the presence of realistic radiance observations, it is likely that adding ozone assimilation from current ozone retrieval observations .. will have little impact". Again going back to Geer (2016), it is hard to detect the impact of small changes in the observing system in the forecast quality of an

operational-quality NWP system. As shown there, adding a single new instrument in the troposphere will only become statistically significant in an experiment containing around 300 forecast samples. In the work under review, there is only a single sample, so it cannot hope to have the required sensitivity or statistical significance. This means the conclusion is unnecessarily pessimistic. If the OSSE were continued for several months, and statistics computed from that whole period, the benefit of ozone assimilation might be seen to be statistically significant. (Many other results presented in this study are more convincing and would likely be statistically significant, but this one is probably not.)

We agree that our conclusions were premature, based on only one time at the end of the experiment. We decided to include 6 days of averaging (following a 10-day spin-up) in all of the revised error profiles in order to make the results more robust. A plot of the time series of errors for several experiments is provided in Fig. 13 to show that the errors are relatively level during the last 6 days. But even with these 25 analyses, it is unlikely that we can reach a conclusion on statistical significance. See further discussion on our altered approach in the response to the second review.

6) Page 16, line 11-14: For similar reasons to those explained in comment 5, the conclusion that current NWP systems can't benefit from ozone-wind interaction in a 4D-var system is probably incorrect. It is more a matter of quantifying the size of that benefit, which is something this current OSSE presentation does not have the sensitivity to explore.

Good point. We altered this discussion accordingly.

Minor comments

Page 3, Lines 12-13, from "will perturb..." are difficult to follow and need rewriting for clarity.

We modified this sentence.

Page 4, line 27: it is confusing to refer to X' first as an "ensemble state" and then as a "perturbation from the ensemble mean".

We clarify the wording. X is ensemble state, X' is perturbation from the ensemble mean.

Page 5, line 8: "the standard suite" of observations would nowadays include satellite radiances, so this would be better described as a "baseline" or "conventional only" suite of observations.

Changed "standard suite" to "conventional".

Page 6, line 26: Figure 3 caption is missing this piece of information: the level is 10.5hPa

Included the level in the figure caption (now Figure 6).

Page 6, line 31-32: That the ozone initial conditions are biased with respect the ozone scheme in the experiments seems a major flaw and needs more explanation or investigation - probably it is no real

problem, as implied by section 4.1, that the initial conditions don't matter to the results on 1 December.

We do not think this is a major problem, given the relative insensitivity to initial conditions. We decided to remove the comment about initializing with ozone based on another scheme to avoid confusion.

Page 7, line 10: "By using a different stratospheric analyses" would be clearer if it was written "by replacing the initial conditions with a different stratospheric analysis"

Good suggestion. We change the wording as suggested.

Page 9, line 8: The high ensemble spread in the vortex is non-intuitive and deserves some explanation. How is this being generated? It should be fairly clear if, for example, the central location of the vortex varies quite a lot across the ensemble.

While individual maps of ensemble members show some variability in the location, orientation, and shape of the vortex, the ozone shows even larger variability. We think this is due to slight variations in the vortex evolution in each ensemble member that result in differences in ozone advection, which accumulate with time due to the long photochemical lifetime of ozone in the NH winter polar region. This causes the initially small spread to increase with time over the experiment. The discussion was modified accordingly.

Page 10, line 10: "It is likely that radiances are the limiting factor..." - please explain this better - I don't understand it.

We removed this sentence, since it was not helpful in the discussion.

Citations - these citations are now all included in the manuscript

Allen, D. R., Hoppel, K. W., and Kuhl, D. D.: Wind extraction potential from ensemble Kalman filter assimilation of stratospheric ozone using a global shallow water model, Atmos. Chem. Phys., 15, 5835-5850, https://doi.org/10.5194/acp-15-5835-2015, 2015.

Geer, Alan J. "Significance of changes in medium-range forecast scores." Tellus A: Dynamic Meteorology and Oceanography 68, no. 1 (2016): 30229, http://dx.doi.org/10.3402/tellusa.v68.30229

Harnisch, F., S.B. Healy, P. Bauer, and S.J. English, 2013: Scaling of GNSS Radio Occultation Impact with Observation Number Using an Ensemble of Data Assimilations. Mon. Wea. Rev., 141, 4395–4413, https://doi.org/10.1175/MWR-D-13-00098.1

Peubey, C., and A. P. McNally. "Characterization of the impact of geostationary clearâ˜A ˘ Rsky radiances on wind analyses in a 4Dâ˜A ˘RVar context." Quarterly Journal of the Royal Meteorological Society 135, no. 644 (2009): 1863-1876.

Response to RC2 from Thomas Milewski, 2 Nov 2017

This article addresses the potential of improving wind and temperature analyses in the stratosphere and mesosphere through the assimilation of ozone observations in a reduced-resolution NWP model. It follows up on previous studies that investigated this potential for a variety of data assimilation systems in a simpler model (e.g. global shallow water model), which pointed towards the quality of the Hybrid (covariances) 4D-VAR for this particular purpose. This study is a significant step forward in that it continues to investigate this outstanding question in a more realistic, closer to operational NWP DA setting.

Specifically, this study is an OSSE that focuses on the assimilation of stratospheric ozone observations and its potential added-value over more traditional radiance assimilation. The overall qualities of the study are the well-prepared experimental setup, with a clear progression between experiments, the tests in sensitivity to different parameters and the insights given about the impact of ozone assimilation on the other analyzed variables. However, in the reviewer's opinion, some aspects need to be improved for the article to be ready for final publishing.

Major comments:

The authors are making negative conclusions on the potential benefit of ozone assimilation from the diagnostics of a single case (Dec 1, the final date of a 14-day experiment). There is generally high quality in the experimental setup and the angle of analysis in this study, but it is difficult to objectively distinguish between the random noise in the results and an actual robust signal, in order to draw general conclusions. If the authors intend this article to be a case study, it needs to be firmly stated in the abstract/introduction/conclusions, a more detailed analysis of the current conditions and error patterns, and more caution in making conclusions are needed. Otherwise, the authors need to be more convincing on how this case is representative of more general conditions or, even better, extend the length of the experiments and provide time-averaged results, with statistical significance tests.

We agree that the general conclusions that we drew from a single case was premature, and therefore are editing the text accordingly (see also response to the first review). We also changed the error profiles to be time-averaged results over the last 6 days of the 16-day experiment, allowing for 10 days of spin-up. The justification for spin-up time is discussed further below. The time-averaging decreases noise from what we had in a single time, but the limited extent of the experiment does not permit testing of statistical significance (see also comments and responses from the first reviewer). We therefore intend this to be a case study, and added further results (see next paragraph) and text to indicate this.

We also included additional discussion of the current conditions and error patterns. Maps of ozone and height in the lower stratosphere (new Fig. 5) and upper stratosphere (new Fig. 7) were added. In addition, we included time series of the globally- and vertically-averaged errors (new Fig. 13), zonal mean plots of errors for the ozone/radiance experiments (new Fig. 22 and 25), and further diagnostics including the ozone tendency (Fig. 23) and the mean increment (Fig. 3). These enhance the paper as a case study of a particular event, providing guidance for future work that would examine the statistical significance of ozone assimilation on the winds and temperatures.

Minor comments:

Section 2.1, line 31: "low resolution of T47", please compare it to the operational resolution. This is important considering that this study is addressing the potential benefit of assimilating ozone in NWP systems.

Operational resolution is T425L60 (0.28° Gaussian grid spacing at the equator) for the outer-loop and T119L60 (1.0°) for the inner-loop, which is the resolution of the data assimilation. T47 is 2.5° spacing at the Equator. We include this comparison in the revision.

Also, How does the reduced resolution of the model might affect the results of radiance assimilation versus ozone assimilation? In other words, could a higher resolution in the ensembles and/or the background fields help favor assimilation of ozone profiles versus assimilation of radiances?

It is difficult to address this without actually performing the experiments, and we want to avoid too much speculation. However, we added a brief discussion of the potential impact of higher resolution on ozone assimilation in Sect. 6.

Section 2.2, line 18, "60 vertical levels": maybe specify the number of vertical levels and approximate vertical resolution in the stratosphere and mesosphere.

In the stratosphere (lower mesosphere), there are 18 (7) levels. The vertical spacing ranges from ~1.5 km in the lower stratosphere to ~2.5 km at the stratopause to ~5 km at the model top. This is now indicate in the revision in Section 2.1.

Section 2.2, line 25, "with a prescribed estimate of the analysis error variance": how is it estimated in this context ?

Modified text to include the following description of the analysis error variance. The ET scheme transforms the previous 6-h ensemble perturbations into a new set of initial perturbations such that the initial ensemble covariance is consistent with a prescribed climatological 3D-Var based estimate of

8

the analysis error variance. The climatological variances are averaged from 10 June 2015 to 10 August 2015, and are the same as those used in the operational NAVGEM system.

Section 2.5, line 10, "the perturbation is performed ... different stratospheric analyses": but presumably valid at the same date ? Please specify.

Yes, these are valid at the same time (15 November 2014, 00Z). We indicate this in the revision.

Section 3.1: Did you look at the temporal evolution of sigma_ens, to make sure that the ensemble system has finished its spinup phase ?

We included analysis of the ensemble spin-up in the revision (see revised Fig. 10). The ensemble spin-up took about 10 days for T and Z, and about 8 days for U and V. We updated several of our results to be time averages over the last 6 days of the experiments, following the spin-up period.

Section 3.2: What motivated this choice of latitude and height ? You state that the PV charge analogy is particularly valid in regions "where strong ozone gradients and geostrophic balance occurs", but 28.6S is not a typical region for these two criteria.

We revised this sentence to avoid confusion. The horizontal ozone gradients are actually relatively large at this latitude for this time period. We added zonal mean maps of ozone to the revised Figure 4 in order to illustrate this (see white dotted line on Fig. 4c). The geostrophic balance at ~30S would be limited somewhat by the Coriolis parameter being around sin(30) = 0.5, but the circulation still resembles geostrophy at this latitude.

Section 4.1: In comparing the ozone-assimilation experiments with perturbed and unperturbed initial conditions, you are also perturbing the troposphere, which can roughly be considered as a lower boundary condition in your experimental setup. The title of the section "dependence on initial conditions" might be a bit limited or ambiguous. The baseline experiment RMS errors are more representative of the dependence on initial conditions only.

We changed the title of this subsection to "Sensitivity of the analysis to perturbations in the DAS," and we added more relevant discussion, also answering some questions from the first reviewer.

P17, line 9: please correct "The mechanisms through which".

We corrected this phrase (changed "though" to "through").

List of Changes to the Figures

New	Original	Changes from Original
Figure 1	Figure 1	None
Figure 2	Figure 2	None
Figure 3	----------	New figure: Analysis increments
Figure 4	Figure 5	Included zonal mean ozone plots
Figure 5	----------	New figure: Ozone maps at 77 hPa
Figure 6	Figure 3	None
Figure 7	----------	New figure: Ozone maps at 1.1 hPa
Figure 8	Figure 4	None
Figure 9	Figure 6	None
Figure 10	----------	New figure: Ensemble standard deviation time series plots
Figure 11	Figure 7	Included continents on panels (a) and (g)
Figure 12	Figure 8	None
Figure 13	----------	New figure: Time series of globally and vertically averaged errors
Figure 14	Figure 9	See 1 below, now show perfect global, random, and polar cases
Figure 15	Figure 10	See 1 below
Figure 16	Figure 11	Changed 10^6 to 10^{-6}
Figure 17	Figure 12	Changed 10^6 to 10^{-6}
Figure 18	Figure 13	See 1, 2 below, changed dotted line from green to black
Figure 19	Figure 14	See 1, 2 below, removed dotted lines, changed x-axes
Figure 20	Figure 15	See 1 below, colors changed
Figure 21	Figure 16	See 1 below
Figure 22	Figure 17	See 1 below, replaced line plots with zonal mean differences
Figure 23	----------	New figure: Hourly ozone tendency
Figure 24	Figure 18	See 1, 2 below
Figure 25	Figure 19	See 1, 2 below, replaced line plots with zonal mean differences

1. Data shown are now averaged from 25 Nov - 1 Dec 2014, rather than only at 1 Dec 2014.

2. For the experiments that assimilated "noisy" ozone data, we decided that in the specification of the observation error standard deviation that we limit the minimum value to 0.1 ppmv, the same number used for the "perfect" ozone (see Table 1), rather than use the same percent value. This was to avoid having the system pull too tightly to very precise data in the lower stratosphere.

Extraction of wind and temperature information from hybrid 4D-Var assimilation of stratospheric ozone using NAVGEM

Douglas R. Allen, Karl W. Hoppel, David D. Kuhl

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

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

Abstract.

Extraction of wind and temperature information from stratospheric ozone assimilation is examined within the context of the Navy Global Environmental Model (NAVGEM) hybrid 4D-Var data assimilation (DA) system. Ozone can improve the wind~~s~~ and temperature~~s~~ through ~~the~~ two different DA mechanisms~~:~~. (1)~~First,~~ through the "flow-of-the-day" ensemble background error covariance~~s~~ that ~~is~~are blended together with the static background error covariance and (2)~~. Second,~~ via the ozone continuity equation in the tangent linear model and adjoint used for minimizing the cost function. All experiments assimilate actual conventional data ~~and satellite-derived wind vectors~~ in order to maintain ~~nearly the same~~a similar realistic troposphere. In the stratosphere, the experiments assimilate simulated ozone and/or radiance observations in various combinations. The simulated observations are ~~taken~~constructed f~~r~~or~~rom~~ a case study based on a 16-day cycling truth experiment (TE), which is an analysis with no stratospheric observations. The impact of ozone on the analysis is evaluated by comparing the experiments to the TE for the last 6 days, allowing for a 10-day spin-up. Ozone assimilation ~~is found to~~ benefit~~s~~ the wind~~s~~ and temperature~~s~~ when data are of sufficient quality and frequency. For example, assimilation of perfect (no applied error) global hourly ozone data ~~c~~ ~~with no error~~ constrains the stratospheric wind~~s~~ and temperature to within ~2 ms^{-1} and ~1 K~~, respectively~~. This demonstrates that there is dynamical information in the ozone distribution that can potentially be used to improve the stratosphere. This is particularly important for the tropics, where radiance observations have difficulty constraining wind~~s~~ due to ~~their broad weighting functions and~~ breakdown of geostrophic balance. Global ozone assimilation provides the largest benefit when the hybrid blending coefficient is an intermediate value (0.5 was used in this study), rather than 0.0 (no ensemble background error covariance~~s~~) or 1.0 (no static background error covariance~~s~~), which is consistent with other hybrid DA studies. When perfect global ozone is assimilated in addition to radiance observations, wind and temperature error decreases of up to ~3 ms^{-1} and ~1 K occur in the tropical upper stratosphere. Assimilation of noisy global ozone (2% errors applied) results in error reductions of ~1 ms^{-1} and ~0.5 K in the tropics and slightly increased temperature errors in the Northern Hemisphere polar region. Reduction of the ozone sampling frequency also~~, addition of observational noise, or inclusion of radiance observations all~~ reduce~~s~~ the benefit of ozone throughout the stratosphere, with noisy polar-orbiting data having only minor impacts on wind and temperature when assimilated with radiances. ~~For example, a single polar orbiting ozone measurement set with realistic errors has no significant impact on the wind analysis when a full suite of radiance observations is also assimilated.~~ An examination of ensemble cross-correlations between ozone and other variables shows that a single ozone observation behaves like a potential vorticity (PV) "charge", or a monopole of PV, with rotation about a vertical axis and vertically oriented temperature dipole. Further understanding of this relationship may help in designing observation systems that would optimize the impact of ozone on the dynamics.

1 Introduction

The spatial-temporal variability of long-lived tracers such as stratospheric ozone contains dynamical information that can potentially be exploited to improve analyses of wind~~s~~ and temperature in the stratosphere and mesosphere, where direct wind

2

observations are largely absent. ~~Theoretical~~ Various studies have examined tracer-wind interactions within a variety of data assimilation (DA) systems including extended Kalman Filter (EKF) (Daley, 1995, 1996), 4D Variational assimilation (4D-Var) (Andersson et al., 1994; Riishøjgaard, 1996; Peuch et al., 2000; Andersson et al., 2007; Peubey and McNally, 2009; Semane et al., 2009; Han and McNally, 2010; Dragani and McNally, 2013; Allen et al., 2013, 2014), Ensemble Kalman Filter

5 (EnKF) (Milewski and Bourqui, 2011, Allen et al., 2015), and hybrid 4D-Var (Allen et al., 2016). Initial 1D and 2D investigations by Daley (1995, 1996) and Riishøjgaard (1996) and 3D investigations by Peuch et al. (2000), Semane et al. (2009), and Allen et al. (2013) showed that coupling the tracer continuity equation with the dynamical equations could allow wind information to be extracted from tracer observations in either 4D-Var or EKF. These studies illustrated the potential of tracer assimilation to influence winds, but also highlighted limitations on this process from observation quality and sampling,

10 inadequate tracer modeling, and geophysical variability. Further theoretical studies by Allen et al. (2014, 2015, 2016) examined the tracer-wind mechanisms in the shallow water model framework using a hierarchy of DA systems: 4D-Var, EnKF, and hybrid 4D-Var. Additional work by Milewski and Bourqui (2011) examined the assimilation of ozone in the EnKF framework using a 3D model, highlighting the propagation of information from ozone to wind via background error covariances.

15 As seen from these studies, t~~T~~here are two primary ways that ozone can influence winds in hybrid 4D-Var DA. The first way is via the ensemble cross-correlations between ozone and other variables that are blended into the initial background error covariance. These so-called "errors of the day" allow ozone to influence dynamical variables directly. Allen et al. (~~2014, 2015~~2016~~), in shallow water model studies,~~) showed that including these cross-correlations provide~~s~~s additional ozone-wind benefit over conventional 4D-Var that excludes initial cross-covariances between ozone and other variables. Second, if ozone

20 is included in the cost function, increments to the dynamical fields at the beginning of the analysis window (i.e., strong-constraint 4D-Var), will ~~perturb~~ be adjusted to minimize the differences between the tangent linear ~~linearized~~ ozone forecast and the ozone observations distributed throughout the analysis time window. This linear approximation and adjoint of the tracer continuity equation propagate the ozone sensitivities over the analysis time window (see Allen et al. (2013) for a 1-D heuristic analytical solution to this problem to illustrate ozone influence in 4D-Var). Thereby ozone observations can influence

25 the winds indirectly as the system attempts to reduce the ozone innovations via both wind and ozone increments. Note that two additional ~~a third~~ ways that ozone assimilation could potentially benefit winds are~~is~~ that improved ozone fields could result in improved radiative calculations in the forecast model (e.g., Cariolle and Morcrette, 2006) as well as improved representation of ozone in forward modeling of ozone-sensitive radiation channels (e.g., Dragani and McNally, 2013); we do not attempt to address the~~is~~se two mechanisms in the current study.

30

While the potential for tracers to influence winds in DA systems has been well established, the ultimate goal is to obtain operational benefit from this process. Andersson et al. (2007) and Peubey and McNally (2009) showed a tropospheric benefit when infrared and microwave humidity channels from geostationary and polar-orbiting satellites were assimilated in the

European Centre for Medium-Range Weather Forecasts (ECMWF) 4D-Var system. As demonstrated by Peubey and McNally (2009), the dominant factor involves the adjustment of the winds to match observed humidity features (the so-called "tracer advection effect"). Semane et al. (2009) found a slight reduction (< 0.1 ms^{-1}) in the global wind bias (relative to radiosondes) in the lower stratosphere when assimilating Microwave Limb Sounding (MLS) ozone with the Météo-France 4D-Var system coupled to an offline chemistry transport model. However, other attempts to assimilate stratospheric ozone using 4D-Var algorithms and the resultant dynamical coupling have resulted in problems in operational numerical weather prediction (NWP). For example, Han and McNally (2010) state that biases between ozone observations and model background led to erroneous wind and temperature increments in the stratospheric analyses in the ECMWF system. These biases could potentially be alleviated by including a bias correction scheme for ozone data (e.g., Dethof and Hólm (2004)), but this (as far as we are aware) has yet to be accomplished. Other potential problems could include breakdown of the tangent linear model in the presence of large ozone gradients (Riishøjgaard, 1996), errors in parameterized ozone photochemistry, and improper characterization of observational and/or forecast model errors. ~~While these studies show that strong potential exists for wind extraction from tracer assimilation, the operational benefit has yet to be obtained.~~ The dynamical benefit of stratospheric ozone assimilation in an operational framework is also challenged by the huge number of additional competing observations from microwave and infrared sounders (on the order of millions of observations per cycle). Whether ozone assimilation can add significant value to operational analyses and forecasts is yet to be determined. ~~This study attempts to move one step further toward determining whether ozone assimilation can benefit winds and temperatures in operational numerical weather prediction (NWP) models, focusing on ozone-dynamical (i.e., wind and temperature) interactions within a hybrid 4D-Var system, the Navy Global Environmental Model (NAVGEM), at a reduced operational model resolution but with conventional (non-radiance) operational observational data along with simulated radiance and ozone observations based on a truth experiment.~~

~~There are two primary ways that ozone can influence winds in hybrid 4D-Var DA. The first way is via the ensemble cross-correlations between ozone and other variables that are blended into the initial background error covariance. These so-called "errors of the day" allow ozone to influence dynamical variables directly. Allen et al. (2014, 2015), in shallow water model studies, showed that including these cross-correlations provides additional ozone-wind benefit over conventional 4D-Var that excludes initial cross-covariances between ozone and other variables. Second, if ozone is included in the cost function, increments to the dynamical fields at the beginning of the analysis window (i.e., strong-constraint 4D-Var), will perturb the difference between the linearized ozone forecast and the ozone observations distributed throughout the analysis time window. This linear approximation and adjoint of the tracer continuity equation propagate the ozone sensitivities over the analysis time window (see Allen et al. (2013) for a 1-D heuristic analytical solution to this problem to illustrate ozone influence in 4D-Var). Thereby ozone observations can influence the winds indirectly as the system attempts to reduce the ozone innovations via both wind and ozone increments. Note that a third way that ozone assimilation could potentially benefit winds is that improved~~

~~ozone fields could result in improved radiative calculations in the forecast model; we do not attempt to address this mechanism in the current study.~~

This study attempts to move one step further toward determining whether stratospheric ozone assimilation can benefit ~~winds and temperatures~~ analyses in operational numerical weather prediction (NWP) models, focusing on ozone-dynamical (i.e., wind and temperature) interactions within a hybrid 4D-Var system, the Navy Global Environmental Model (NAVGEM). Ozone observations have been assimilated in research versions of NAVGEM for some time (e.g., Eckermann et al., 2009). These have produced reliable ozone fields, but the impact of ozone on the dynamics has not been determined, except for a single update cycle demonstration by Allen et al. (2013). An observing system experiment (OSE) could be performed with NAVGEM (as in Semane et al., 2009 with the Météo-France NWP suite and MLS data). We have opted, however, ~~, at a reduced operational model resolution but with conventional (non-radiance) operational observational data along with simulated radiance and ozone observations based on a truth experiment.~~ to use simulated ozone observations in this study in order to have controlled experimental conditions. One reason for using simulated observations is to eliminate the impact of biases between observations and model that have caused problems in earlier studies. Another reason is to probe the potential dynamical information content available from stratospheric ozone assimilation by using global sampling patterns with unrealistic temporal and spatial coverage, in addition to the more realistic sampling provided by MLS. The overall goal is to build understanding of how ozone assimilation influences the dynamics that can help in design and interpretation of future OSEs.

Data impact experiments with simulated observations are generally performed using the observing system simulation experiment (OSSE) approach. In the traditional OSSE, all data are produced from a model forecast truth (or "Nature Run"). After randomly perturbing the observations in a manner consistent with assumed error statistics, these data are assimilated into a data assimilation system (DAS) in varying combinations. The OSSE approach is used for testing future observations, where data are not yet available. Since the truth is known, the analyses from these experiments can be directly compared to the truth to obtain absolute observation impact. The difficulty is that OSSEs are computationally expensive, requiring the simulation of large numbers of observations of various types. An alternative approach, used by Harnisch et al. (2013), combined real and simulated data in an ensemble of data assimilations (EDA) in order to determine relative impact of future data systems. The real observations included the operational global observing system (GOS), while the simulated Global Navigation Satellite System (GNSS) radio occultation (RO) measurements were simulated from ECMWF analyses. Harnisch et al. (2013) explain that data simulated in this way are not independent of the real observations, even if a different DAS is used, since there is a common database of observations in the GOS. Therefore, assimilating the real and simulated data together poses some challenges if the observation errors are assumed to be uncorrelated. The analysis of relative changes in ensemble spread is useful for determining the information content of the new observations, but one cannot assess the impact of the new observations on the absolute skill of the EDA mean state with this approach. A similar EDA approach was taken by Tan et al. (2007) to determine the impact of future Atmospheric Dynamics Mission Aeolus wind-profiling LIDAR data.

In this paper, we take a novel approach that combines real and simulated data sets, but in a way that attempts to reduce the error correlations between the two sets. This is performed by spatially separating the data into two regions, nominally defined as troposphere (pressures greater than 100 hPa) and stratosphere (pressures less than 100 hPa). The truth is created from a cycling analysis that assimilates real conventional data in the troposphere, while simulated observations are all located in the stratosphere. This separation is not complete, since vertical background error covariances can extend upward into the stratosphere, and observation weighting functions can extend downward into the troposphere. In addition, separating the troposphere and stratosphere in this way does not perfectly decouple the model fields, since the stratosphere is affected by forcing from below and vice-versa. However, this dynamical coupling is expected to weaken with increased altitude. This is illustrated in the paper by widely divergent stratospheres resulting from simulations using the same tropospheric observations, but using unperturbed and perturbed stratospheric initial conditions. As far as the decoupling assumptions being valid, this approach allows us to calculate absolute impact, since the truth is known, but avoids having to perform a complete OSSE.

Simulated stratospheric observations (ozone profiles and microwave and infrared radiance measurements) are created using the truth. In addition to simulating ozone observations from a typical polar-orbiting limb sounder, we also examine the impact of globally distributed ozone observations using two different sampling patterns in order to explore the information content available in the stratospheric ozone fields.

The ozone data are assimilated with and without simulated satellite radiance measurements in order to determine the impact on the stratosphere with and without a realistic GOS. The impact of ozone on the wind and temperature analyses is examined for a case study in the Northern Hemisphere late fall, 15 November - 1 December 2014. Examining this case study allows us to lay the framework for future work to assess the statistical impact of ozone assimilation over longer periods.

The paper is outlined as follows. The NAVGEM ~~The overall goal of this work is to evaluate how ozone interacts with winds and temperatures in a full~~ NWP system is ~~(~~described in Sect. 2~~)~~. T~~We first examine t~~he characteristics of the ensemble cross-covariances are examined in Sect. 3 in order to understand how ozone and other variables relate ~~(Sect. 3)~~. Section 4 explores how ~~We next look at how~~ well ozone by itself can constrain the stratospheric dynamics. ~~(Sect. 4). The last part of the paper~~ An~~(Sect. 5)~~ examination of~~es~~ ozone assimilation in the presence of simulated radiance observations is presented in Sect. 5, and~~in order to determine whether ozone can provide added value.~~ Sect. 6 provides a summary and conclusions.

2. Model description

2.1 Forecast model

This study uses a reduced resolution version of the operational NAVGEM described in Hogan et al. (2014). The NAVGEM global forecast model uses a semi-Lagrangian/semi-implicit integration of the hydrostatic equation, the first law of

6

thermodynamics, and conservation of moisture and ozone. This study uses a 60 level hybrid sigma-pressure coordinate (top at 0.05 hPa) as described in Eckermann et al. (2009). There are 18 levels in the stratosphere and 7 in the lower mesosphere (defined in this study as pressures ranging from 1.0 to 0.05 hPa), and the vertical spacing ranges from ~1.5 km in the lower stratosphere to ~2.5 km at the stratopause to ~5 km at the model top. The model is run at a relatively low resolution of T47

5 (144 longitudes × 72 latitudes, for a Gaussian grid spacing of ~2.5° at the Equator). The model time step is 1800 s. The same forecast configuration and resolution are used for the control (outer loop) and the ensemble forecasts and DA (inner loop). The current (early 2018) operational resolution uses the same 60 vertical levels and a horizontal resolution of T425 (Gaussian grid spacing of ~0.28° at the Equator) for the outer loop and T119 (Gaussian grid spacing of ~1.0° at the Equator) for the innner loop. ~~The model time step is 1800 s. The same forecast configuration and resolution are used for the control (outer loop) and~~

10 ~~the ensemble forecasts (inner loop).~~

2.2 Hybrid 4D-Var data assimilation system

NAVGEM employs a hybrid 4D-Var method, which is becoming increasingly popular at operational NWP centers (e.g., Buehner et al., 2010; Bonavita et al., 2012; Clayton et al., 2013; Kuhl et al., 2013; Kleist and Ide, 2015). NAVGEM minimizes a quadratic cost function using the accelerated representer approach as described in Xu et al. (2005) and Rosmond and Xu

15 (2006). The conventional initial background error covariance $\mathbf{B}_0^{con}$ is calculated using an analytic formulation that employs the hydrostatic relationship in the vertical between geopotential and temperature, and wind-geopotential correlations based on approximate geostrophic balance on an *f*-plane, i.e., constant Coriolis parameter with latitude (Daley, 1991; Daley and Barker, 2001; Kuhl et al., 2013). There is no coupling between ozone and dynamical variables in $\mathbf{B}_0^{con}$, but coupling between these variables does develop implicitly over the 4D-Var time window. The only difference between the hybrid 4D-Var used in this

20 study and Kuhl et al. (2013) is the incorporation of ozone observations in the analysis and ozone in the ensemble forecasts and forecast error covariance.

The tangent linear model (TLM) currently used in NAVGEM is based on linearization of the Navy Operational Global Atmospheric Prediction System (NOGAPS) global spectral forecast model (Hogan et al., 1991), which was the forerunner of

25 NAVGEM. Relevant details of the TLM and adjoint (ADJ) used in this study are provided in Rosmond (1997). The TLM and ADJ are also run at T47 resolution with 60 vertical levels, as is the nonlinear forecast model, but with a reduced time step of 900 s. The TLM has parameterizations for surface flux and vertical mixing based on Louis (1979), but does not include ~~any~~ other physical parameterizations such as radiation, ozone chemistry, and gravity wave drag, which are in the nonlinear model. The 4D-Var system runs with a 6-hour analysis window, and the analysis at the middle of one window is used to initialize a

30 9-h~~our~~ forecast that serves as the background for the next update cycle. Each analysis and resulting 1-5 h forecast are saved for use in creating the simulated observations.

The ensemble consists of 80 members, which are updated each cycle using the ensemble transform (ET) approach described by McLay et al. (2008, 2010) and Kuhl et al. (2013). The ET scheme transforms the previous 6-h ensemble perturbations into a new set of initial perturbations such that the initial ensemble covariance is consistent with a prescribed climatological 3D-Var based estimate of the analysis error variance. The climatological variances are averaged from 10 June 2015 to 10 August 2015, and are the same as those used in the operational NAVGEM system. The ensemble covariance,

$$\mathbf{B}_0^{\text{ens}} = \mathbf{X}'\mathbf{X}'^{\text{T}} / (N_{\text{ens}} - 1),$$ is calculated at the start of each 6-hour window using the ensemble states $\mathbf{X}$; the prime indicates perturbation from the ensemble mean, the superscript T indicates transpose, and N_{ens} is the ensemble size. The ensemble covariance is then blended together with $\mathbf{B}_0^{\text{con}}$ using $\mathbf{B}_0^{\text{hybrid}} = (1-\alpha)\mathbf{B}_0^{\text{con}} + \alpha\mathbf{S} \circ \mathbf{B}_0^{\text{ens}}$, where α is a blending coefficient between 0zero and 1one, $\mathbf{S}$ is the localization function, and the open circle indicates the Schur product. The horizontal localization is based on a second-order autoregressive function and the vertical localization employs a Gaussian log-sigma correlation as described in Daley and Barker (2001) and Kuhl et al. (2013). The ensemble was initialized using NAVGEM analyses from a separate experiment that included conventional data along with radiances and Microwave Limb SounderMLS ozone and temperature profiles. This separate experiment was run at T119 with 74 levels and was downscaled to the T47L60 resolution used for this study. The data were sampled at 18 Z on 80 consecutive days, starting 2 May 2014. EInitial ensemble standard deviations are examinedpresented in Sect. 3.1.

2.3 Observations

The experiments in this study assimilate both actual observations (in the troposphere) and simulated stratospheric observations computed from the truth experiment (described in Sect. 2.4). The tropospheric observations include a standard suiteconventional of operational measurements, such asincluding surface observations from ships, buoys, and land surface stations, upper air observations from radiosondes and aircraft, and satellite-derived winds. (Global Positioning System (GPS) radio occultation observations are not included for this particular study). The standard NAVGEM operational data quality control and thinning algorithms are used, and the resulting tropospheric observation counts range from ~750,000 to 1,000,000 observations for each 6-h cycle. The actual observations are limited to pressures greater than 100 hPa, which mainly affects the radiosonde profiles.

For the simulated ozone observations, three different sampling patterns are used: global, polar-orbiting, and random (See Fig. 1). The global observations (Fig. 1a) are provided on an approximately equal-area sampling, generated by subdividing an icosahedral base into a triangular grid with ~300 km spacing (3840 elements). To avoid both horizontal and vertical interpolation by the DA, the ozone observation locations were moved to the nearest model Gaussian grid points, and the ozone was sampled vertically on the model levels. Seventeen vertical levels in the stratosphere are used, ranging from 778 to 1.12

8

hPa (~20 to 50 km altitude). Note that for the 60-level NAVGEM configuration, these stratospheric levels are all constant pressure levels. The temporal sampling for the global observations is 1 h, matching the forecast output sampling. ~~For the assimilation of these observations, the observation error covariance is uncorrelated with a specified standard deviation σ_{ob}. We will examine cases with perfect observations (i.e., no random error added) and cases with the applied random error. Note~~

5 ~~that for the perfect observations, we do not set σ_{ob} exactly to zero, but to a reasonably small value of 0.1 ppmv. Setting σ_{ob} to zero causes the cost function to become singular and prevents the solution from converging.~~ The second set of ozone observations simulates a polar-orbiting limb sounder, and was created by sampling the TE at the observation locations of the Aura ~~Microwave Limb Sounder (MLS)~~MLS instrument. There are approximately 3500 observations per day for MLS~~;, along~~ a sample flight track is shown in Fig. 1b. For the polar-orbiting ~~data~~data, we also move the locations to the nearest model grid

10 point and sample vertically on the same model levels. The third set of ozone observations (Fig. 1c), which we call "random," sub-samples the global observations randomly in space and time at a frequency of 3500 per day, which is ~~a~~similar ~~frequency~~ to that of the polar-orbiting observations. This tests~~provides a test of~~ whether spreading the information from polar-orbiting sampling would make the data more useful for extracting wind and temperature information.

For the assimilation of these ozone observations, the observation error covariance is uncorrelated with a specified standard

15 deviation σ_{ob}. We will examine cases with "perfect" (i.e., no random error added) and "noisy" (with applied random error) observations. Note that for the perfect observations, we do not set σ_{ob} exactly to zero, but to a reasonably small value of 0.1 ppmv. Setting σ_{ob} to exactly zero causes the cost function to become singular and prevents the solution from converging. Calling these observations "perfect," while having non-zero observation error in the DAS, follows the naming convention of Peuch et al. (2000). For the "noisy" observations, we apply 2%, 5%, or 10% random error to the perfect observations, and we

20 set σ_{ob} to the same percent value, except that we limit σ_{ob} to a minimum value 0.1 ppmv. The lower limit is to prevent the DAS from constraining too tightly to highly precise ozone observations, which could result in spurious wind increments.

Some of the experiments assimilate simulated stratospheric radiance observations, in order to assess the value of ozone in combination with a typical set of global~~radiance~~ observations. NAVGEM routinely assimilates microwave and infrared radiances from a number of sounders. In this study, simulated radiances are created using actual data sampling from Advanced-

25 Microwave Sounding Unit (AMSU-A), Atmospheric Infrared Sounder (AIRS), Advanced Technology Microwave Limb Sounder (ATMS), and Infrared Atmospheric Sounding Interferometer (IASI) (see Fig. 2 for an example of radiance observations for one update cycle). During the creation of the truth experiment, actual radiance observations are processed to the point where background radiance values are calculated, but are then omitted from the DA solver. The background radiance values then become "perfect" simulated radiances. Because the simulated radiances are created and then assimilated using

same radiative transfer model, they are unbiased "perfect model" data. For the radiance and ozone assimilation experiments (Sect. 5), we add Gaussian random noise to the perfect radiance data, matching the observation error values used in NAVGEM for the actual instruments. These provides realistic "noisy" radiance observations. The variational radiance bias correction scheme in the DA is disabled for the simulated radiances. This is a best-case scenario for the radiances, since the DA of true

5   radiances always includes biases. Addressing the impact of bias correction in the context of ozone assimilation is beyond the scope of this paper.

As explained in the Sect. 1, the separation of the troposphere and stratosphere is not perfect, since tropospheric observations, as well as specified forecast errors, can have vertical error correlations that extend upward into the stratosphere. For example,

10   tropospheric temperature increments can raise the geopotential height at altitudes in the stratosphere, and wind/height balances in the conventional and ensemble covariances will result in stratospheric circulation increments. These, in turn, can influence the stratospheric ozone field via advection. In addition, the adjoint and tangent linear integrations can propagate information vertically to and from the observation locations. To illustrate the combined effects of these non-local processes, Fig. 3a,b shows the time mean and zonal mean of the absolute value of the u and T increments over the 6-day period, 25 November - 1

15   December 2014 from the truth experiment (described in the next section). While all observation point locations for the truth experiment are located below 100 hPa, small increments extend above 100 hPa. The increments in the stratosphere are limited to less than around 1.0 ms^{-1} for zonal wind and ~0.4 K for temperature. The stratospheric increments to u and T are larger in the NH than in the SH, likely due to more radiosondes over the continental regions. The largest wind increments are in the tropical upper troposphere, where we also see the largest ozone increments (plotted in Fig. 3c as percentage of the local

20   analyzed ozone). The ozone increments are small in the region of simulated ozone observations (region between the white dashed lines on Fig. 3), generally less than 1%, suggesting that the coupling will not adversely affect our ozone assimilation experiments. The simulated radiances will have a weak correlation with the tropospheric observations from the temperature increments in the stratosphere, as well as from radiance weighting functions that extend into the troposphere. The wind increments will not directly influence the simulated radiance and ozone observations, but may affect them indirectly via

25   advection over the assimilation window.

2.4 Truth experiment and meteorological conditions

The experimental design (described in more detail in Sect. 2.5) is based on a truth experiment (TE) that is used to simulate observations that are assimilated back into the system and alsoand to evaluate all other experiments. Our goal was to examine stratospheric assimilation and analysis with a constant, realistic troposphere that is essentially independent of our experiments.

30   This TE could be caccomplished by created bying a free-running nature forecast, which would be used and then simulateing a full set of conventional tropospheric observations, as in conventional OSSEs. As explained in Sect. 1, weInstead, we took a simpler approach in which our TE is a normal cycling analysis in which only tropospheric data (pressures >greater than 100 hPa) are assimilated. The stratospheric analysis of the TE can be viewed as a mechanistic forecast forced by a realistic

10

troposphere, which is then used to create simulated stratospheric observations. All subsequent experiments, with differing stratospheric observations, assimilate the same set of tropospheric observations. This gives ~~nearly identical~~similar (but not exactly the same, as discussed below) tropospheric analyses for each experiment, but the ~~differing~~ stratospheric analys~~is~~es can vary widely. This approach provides a realistic evaluation of how differing stratospheric observations impact a typical global

5 analysis. We note here (and will show later) that when stratospheric observations are included, the analyzed tropospheric state will be different from the TE. As explained by Geer (2016), even a slight numerical perturbation will generate chaotic divergence between two analyses. In Sect. 4.1 we will examine the extent to which the troposphere is sensitive to "perturbations" in the data set being assimilated. The stratospheric truth should therefore be considered as one realization drawn from a potential ensemble of stratospheric states, given the potential variation in the troposphere caused by these slight

10 perturbations.

We initialized the TE on 15 November 2014 (00Z) and ran through 1 December 2014 (00Z), for a total of 16 days with 64 cycles. The initial zonal mean latitude/pressure cross-sections for zonal wind, temperature, and ozone are plotted on the top row of Fig. 4. The Arctic winter stratospheric vortex is seen in the Northern Hemisphere (NH), extending to the top of the

15 model with peak winds of ~70 ms^{-1}. In the polar region of the Southern Hemisphere (SH), westerlies occur in the lower stratosphere and easterlies occur in the upper stratosphere and lower mesosphere. The westerlies are the remnant of the Antarctic winter stratospheric vortex, which is in the process of breaking down. In the tropics, a complicated pattern of alternating easterlies and westerlies is observed, with a large region of easterlies in the lower stratosphere. The zonal mean temperature shows typical solstice conditions, with warm summer and cold winter stratosphere, a cold tropical tropopause,

20 and a warm troposphere. The ozone mixing ratio maximizes in the tropical middle stratosphere, with the peak shifted towards the SH. Low ozone mixing ratios occur in the troposphere, mesosphere, and SH polar lower stratosphere (i.e., ozone hole conditions).

25

The meteorological conditions for the lower, middle, and upper stratosphere are illustrated in Figs. ~~3~~5, 6, and 7, which show~~s~~ the ozone and geopotential height at 77 hPa (~18 km), 10.5 hPa (~32 km), and 1.1 hPa (~48 km), respectively, from 15-30 Nov 2014 (all at 00Z) at 3-day intervals. At 77 hPa (Fig. 5), t~~The Northern Hemisphere (NH)~~ NH stratospheric polar vortex can be identified by closed height contours, with the center displaced off the pole towards Asia. The NH ozone has generally

30 higher mixing ratio in the vortex and lower in the tropics at this level. Over the course of the next two weeks, the vortex shape is modulated by ridges forming on both sides of the vortex, resulting in a dumbbell-shape on 30 November. This dynamical activity is accompanied by ozone advection eastward and northward from the tropics (for example, see the tongue of low ozone air forming at the bottom of Fig. 5e). High ozone occurs along the edge of the vortex, and the mixing ratio increases with time over late November. In the SH, ozone depletion is evident within the Antarctic vortex, while higher ozone occurs in the

extratropics. The Antarctic vortex shifts over the course of late November, being drawn into an oval shape by the end of the month and displaced well off the pole. The low ozone contours follow the shape of the vortex over this period. ,

At 10.5 hPa (Fig. 6) the NH polar vortex, indicated by low ozone mixing ratio and enclosed height contours, is seen on 15 November centered slightly off the pole. A tongue of low-latitude air moves northward and eastward over the next few days as an Aleutian high starts to spin up, pushing the vortex off the pole. The vortex elongates and the minimum ozone mixing ratio increases, indicating some mixing of the vortex air outward. The Aleutian high is still strong on 30 November. Note that the ozone mixing ratio in the tropics decreases over the course of the TE; this is due to the initial conditions being drawn from an experiment that used different parameters in the ozone photochemical scheme.

In the Sthe Southern Hemisphere (SH), the large-scale circulation in the middle stratosphere is generally becoming easterly in November, with accompanying high pressure and lower ozone mixing ratio. On 15 November, a low pressure cyclone between South America and Antarctica disrupts the normal otherwise easterly flow. Low latitude air with high ozone is pulled clockwise around the cyclone on 18 and 21 November. The cyclone diminishes in strength from 21-30 November, and the prevailing anticyclonic flow center moves back towards the pole.

At 1.1 hPa (Fig. 7), The Arctic vortex is initially centered close to the pole, with low ozone inside. The ozone mixing ratio in the vortex increases sharply by 18 November. This is likely due to parameterized photochemistry drawing the ozone towards a climatological state that was different from the simulation used for the initial conditions. The upward extension of the growing Aleutian high is seen at this level as well, forcing the vortex off the pole and stretching it into a "comma" shape by 27 November. In the SH, the height contours are nearly zonal throughout this period, with steady easterly circulation, and the ozone becomes nearly zonal as well. This suggests that the wave activity observed at 10.5 hPa becomes trapped before it reaches the upper stratosphere, due to the presence of the zero wind line (see Fig. 4a), which serves as a critical line for stationary planetary waves. The overall meteorological situation for this period is characterized by a decaying Antarctic vortex in the lower stratosphere and quiescent SH easterlies in the upper stratosphere/lower mesosphere, while in the NH, the Arctic vortex is being influenced by moderate wave activity that is causing the vortex to be pushed off the pole and stretched. This provides a range of dynamical conditions for testing the impact of ozone assimilation on the winds and temperature.

2.5 Experimental design

We perform two types of stratospheric assimilation experiments: ozone-only assimilation and ozone/radiance assimilation. To illustrate the differences between these experiments, Fig. 4 8 provides a schematic diagrams. Both types of experiment use the same TE and observation database (note that all experiments assimilate the same tropospheric data). Except for the TE and a few test experiments, most of the all experiments use unperturbed initial conditions in the troposphere and perturbed initial

12

conditions in the stratosphere. The perturbation is performed ~~by using a different stratospheric analyses~~by replacing the initial conditions with a different stratospheric analysis (at pressures less than 100 hPa), valid at the same time (15 November 2014, 00Z), but based on NAVGEM experiments that differ in terms of model resolution and data assimilated, resulting in slightly different dynamical fields. The ozone fields, however, are initially identical. Figure ~~5~~4 (second row) shows the zonal mean cross-sections of zonal wind, temperature, and ozone for the perturbed state, and Fig. 4 (third row) shows the differences between perturbed and unperturbed initial conditions ~~for zonal wind and temperature~~. Large dynamical differences occur in the tropical upper stratosphere and throughout the lower mesosphere.

For the ozone-only assimilation (Fig. 8, top row, results presented in Sect. 4), a baseline experiment (BE) is performed by running the system from the perturbed initial conditions and assimilating only the tropospheric conventional observations. As will be shown below, the stratospheric winds and temperature~~s~~ in this BE deviate significantly from the TE (after 16 days, zonal mean differences of up to ~80 ms^{-1} occur for vector wind and ~25 K for temperature). When experimenting with the blending coefficient ($\alpha = 0.0$, 0.5, and 1.0), we must run separate TE and BE cases for each value. This is because the blending coefficient affects the tropospheric assimilation, and hence changes the reference TE for each case. Changing α examines the sensitivity of the amount of ozone-wind correlation being used from the ensemble covariances. Ozone-only assimilation experiments are next performed, which examine the limit to which ozone could potentially constrain the winds without any other data present in the stratosphere. Global data are assimilated for all three values of α (Sect. 4.2), while random and polar-orbiting data are only assimilated for $\alpha = 0.5$ (Sect. 4.3). In addition, to examine sensitivity to data quality, experiments are performed for $\alpha = 0.5$ with assimilation of ~~global~~ ozone data with imposed observational errors (Sect. 4.4).

In the ozone/radiance assimilation experiments (Fig. 8, bottom row, results presented in Sect. 5), we test the extent to which assimilating ozone data can reduce the ~~initial~~ errors relative to ~~errors in~~ a system constrained by realistic radiance observations. First, we create a BE by assimilating noisy radiances created from the TE. Then, experiments are performed in which either global, random, or polar-orbiting ozone data are assimilated in addition to the radiances using either perfect ozone ~~and~~or ozone with imposed random errors. These experiments are all run for 16 days, which allows for a 10-d spin-up of the ensemble and errors (discussed below) and evaluation of experimental errors over 6 days. This single case study does not include enough data to adequately assess statistical significance (e.g., Geer, 2016), but will provide a framework for analyzing and guidance in designing future long-range experiments.

3. Discussion of background errors

The background error covariance is a critical component of the hybrid 4D-Var system. Hybrid 4D-Var combines a conventional error covariance with a localized ensemble covariance in order to take advantage of both the high-rank properties of the conventional and flow-of-the-day properties of the ensemble components. In this section, we first examine the latitude/pressure

cross-sections ~~of the conventional and ensemble errors (Sect. 3.1) and the ensemble spin-up (Sect. 3.2). We next examine~~and horizontal maps of the ensemble background error standard deviations (i.e., the square root of the diagonal terms of the covariance matrix) in Sect. 3.3. ~~We denote the background error standard deviations as~~ σ_{con} ~~,~~ σ_{ens} ~~, and~~ σ_{hyb} ~~for the conventional, ensemble, and hybrid, respectively.~~ ~~Finally~~Next, in Section 3.4, we examine the cross-correlation terms, which

5 indicate how errors are correlated with other variables and spatial locations. Our particular interest is the patterns that describe how ozone correlates with other variables. For simplification, w~~We~~ denote the background error standard deviations as σ_{con}

~~,~~ σ_{ens} , and σ_{hyb} for the conventional, ensemble, and hybrid, respectively.~~.~~

3.1 Comparing conventional and ensemble error standard deviation

Figure ~~6~~9 shows latitude/~~-~~pressure cross-sections of σ_{con} and σ_{ens} for zonal wind (u), temperature (T), and ozone (O$_3$). The

10 σ_{ens} has been zonally averaged, while σ_{con} is formulated as a zonal mean model. The conventional errors are shown for 15 November 2014 (00Z) only, while the ensemble errors are provided for 15 November and 1 December 2014 (both at 00Z). The σ_{con} for zonal wind (Fig. 9a) increases with altitude from ~2 m-s^{-1} in the troposphere to ~8 m-s^{-1} at 0.1 hPa. The σ_{ens} for zonal wind on 15 November (Fig. 9d) shows more structure, with higher values in the upper troposphere and lower mesosphere, and low~~er~~ values in the extratropical stratosphere. As the ensemble evolves over the next 16 days, the zonal wind

15 σ_{ens} generally increases, particularly in the lower mesosphere, as seen in Fig. 9g, peaking at over 15 ms^{-1}.

The temperature σ_{con} (Fig. 9b)~~is relatively small, ranging~~ranges from ~0.5 to ~2 K, with lower values in the tropics. The initial σ_{ens} for temperature (Fig. 9e) has a similar geographic structure as zonal wind, with elevated values in the upper troposphere and mesosphere. The temperature σ_{ens} generally increases ~~over the next two weeks~~with time, with ~~very~~ large

20 values (>5 K) occurring above ~1 hPa on 1 December (Fig. 9h). For ozone, the σ_{con} is prescribed as a constant value of 0.3 ppmv, except for elevated values in the tropical troposphere (Fig. 9c). The initial ozone σ_{ens} (Fig. 9f) is elevated in the tropical middle stratosphere and SH lower and upper stratosphere. The ozone errors evolve by 1 December to have three regions of enhanced σ_{ens} located in the middle stratosphere in the tropics and extratropics of each hemisphere, with relative minima in the subtropics (Fig. 9i). Lower values are seen ~~both in the troposphere and upper~~in both the troposphere and upper

25 stratosphere/mesosphere, which largely reflect the lower ozone mixing ratios in these levels.

3.2 Ensemble spin-up

To determine when the ensemble has finished its spin-up phase, we calculated the globally averaged σ_{ens} for u, T, geopotential height (Z), and ozone at each model level and then vertically averaged over all levels. For the vertical average, we weighted the profile by a layer thickness in km. The thickness was calculated by first choosing a nominal pressure value for each NAVGEM level, based on a model pressure profile with surface value of 1000 hPa, then calculating the log-pressure height for each level using a constant scale height of 7 km, and finally differencing adjacent layers to get thicknesses. The resulting time-series plots (daily values at 00Z from 15 November - 1 December 2014) in Figure 10 show that the first decrease in spread for u (similar results were obtained for meridional wind, v) occurs on 25 November (after 10 days of cycling), while for T and Z the first decrease occurs on day 23 November (8 days). We therefore consider the ensemble to be spun up by 25 November. The ozone σ_{ens} increases monotonically from 16 November - 1 December, and therefore we cannot assign an objective spin-up time using this approach. However the rate of increase becomes quite small by 25 November (less than 1.5% per day), so we consider the ensemble dependence on the initial ozone ensemble to be small after 10 days of cycling. Further examination of the evolution of ozone σ_{ens} will be presented in Sect. 3.3. Neglecting the 10-day spin-up, error results presented in Sections 4 and 5 are generally averages over the last 6 days (25 November, 00Z - 1 December 2014, 00Z), which includes 25 separate analyses.

3.~~2~~ 3 Horizontal maps of ozone σ_{ens}

The ozone σ_{ens} shows strong geographic patterns that are related to the flow-of-the-day. Figure ~~7~~ 11 shows horizontal maps of the ozone σ_{ens} for the same level (10.5 hPa) and dates as in Fig. ~~3~~6. Geopotential contours are overlaid on the σ_{ens} to facilitate comparison with the flow. On 15 November, the initial σ_{ens} is not aligned with the flow, since the initial ensemble was constructed with analyses on consecutive days from an offline experiment; after 3 days, however, flow-like patterns start to emerge. On 18 November, the σ_{ens} in the NH is larger within the polar vortex ~~and in the tropics~~, while smaller values occur outside of the vortex. This pattern strengthens over the next several days so that by 24 November the vortex/extra vortex distinction is prominent. The high σ_{ens} in the vortex is in a location where the ozone mixing ratio is actually low (see Fig. ~~3~~6). While individual maps of ensemble members (not shown) indicate some variability in the location, orientation, and shape of the vortex, the ozone maps exhibit even larger variability. We think this is due to slight variations in the vortex evolution over time in each ensemble member that result in differences in ozone advection that accumulate with time due to the long photochemical lifetime of ozone in the NH winter polar region. This process causes the initially small ozone spread to increase over the experiment (as seen in Fig. 10d). Further work is necessary to elucidate the exact mechanisms that force changes in

15

ozone ensemble spread. Long streamers of high σ_{ens} are visible in the NH throughout this period, circling around the outer edges of the polar vortex and Aleutian high, where the ozone gradients are large. These patterns are significant for data assimilation, since they will affect the weight that is given to the observations. For example, in the polar vortex the σ_{ens} is large, so ozone observations in this region would be expected to ~~will~~ have a larger impact than those ~~outside of the vortex~~in

5   regions of low σ_{ens} .

The ozone σ_{ens} in the SH also shows a rapid spin-up from an initial state that is approximately constant in the zonal direction. Flow-dependent patterns~~Swirls~~ are seen on 18 November in the cyclonic region between South America and Antarctica~~flow~~, with a low σ_{ens} "tongue" surrounded by high σ_{ens} . From 18 to ~~-~~27 November, the σ_{ens} in the anticyclonic closed height

10   contours increases. A tongue of low σ_{ens} occurs between the two flow regimes on 24 and 27 November, apparently advected by the nearly cross-polar flow. By 30 November, the σ_{ens} pattern shows generally large values at high latitudes, small values in the tropics, and complicated structure in the extratropics. While a complete analysis of the causes of these features is beyond the scope of this paper, it is clear that the σ_{ens} is strongly flow-dependent and may (at least in the experiments with α = 1.0) result in large differences in weighting of ozone observations. ~~;~~ ~~N~~note that the errors have a large dynamic range from ~0.04

15   to 0.99 ppmv at 10.5 hPa for 30 November.

3.~~3~~ 4 Ensemble ozone-wind cross-covariances

As discussed in Sect. 1, ozone can influence wind~~s~~ and temperature~~s~~ via the ensemble background error cross-correlations. Here we show an example of these cross-correlations. Figure ~~8~~12 provides a composite view of the impact of a single ozone observation at pressure of 10.5 hPa and latitude of 28.6° S. The composite was created by separately calculating the spatial

20   correlations of ozone with all other points and variables at 36 longitude~~points~~ (0°, 10°, 20°, …, 350°) ~~with all other points and variables~~. The correlations were then shifted to a common longitude of 180° E and averaged to reduce spurious noise. The top row of Fig. ~~8~~12 shows the horizontal correlations. ~~-~~The ozone-ozone correlation (Fig. ~~8a~~12a) has a maximum of 1.0 at the observation point, and then decreases gradually in each direction, with a larger decorrelation length in the zonal direction. The ozone correlates strongly with vorticity (Fig. ~~8b~~12b), with the ozone-vorticity correlation having a similar ~~zonally-~~

25   ~~oriented~~zonally oriented shape. The ozone-height correlation (Fig. ~~8c~~12c) is more isotropic and represents an anticyclonic circulation, which is counter clockwise in the SH, as seen in the correlations with zonal (Fig. ~~8d~~12d) and meridional (Fig. ~~8e~~12e) wind. The ozone-temperature correlation (Fig. ~~8f~~12f) is weak at the level of the observation, but vertical cross-sections in longitude (Fig. 12~~8~~l) and latitude (Fig. ~~8r~~12r) reveal a strong dipole pattern with cold (warm) temperature~~s~~ above (below) the observation. Vorticity (Fig. ~~8h~~12h, n) and height (Fig. ~~8l~~12l, o) correlations are vertically oriented similar to the ozone-

16

ozone correlation, with slight westward and southward tilting with height; the wind cross-sections (Fig. ~~8j~~12j, k, p, q) show that the anticyclonic circulation extends above and below the observation.

The temperature and circulation patterns revealed in the correlations of Fig. ~~8~~ 12 are similar to those associated with the potential vorticity (PV) "charge" concept developed by Bishop and Thorpe (1994). In this analogy to electrostatics, an elementary PV charge is associated with a field that produces a circulation about the vertical axis and a vertically oriented temperature dipole (see also Fig. 14 of Allen et al., 1997). That a single ozone observation would produce the same circulation patterns as a monopole of PV makes sense, since PV and ozone are both quasi-conserved quantities and will therefore have strong correlations. The pattern is also seen at other latitudes, although its strength varies due to differing ozone gradients and geostrophic coupling. In the NH, the ozone-vorticity correlation is negative and the circulation is clockwise (anticyclonic in the NH), and the temperature dipole is similar with cold (warm) temperature~~s~~ above (below) the observations. These results indicate that ozone observations ~~can~~ may be considered as pseudo-PV observations, at least in the regions ~~of~~where strong horizontal ozone gradients ~~and geostrophic balance occur; this would likely be true of other long-lived trace gases as well~~. As seen in the zonal mean ozone (Fig. 4c), the ozone contours are approximately vertically oriented at this latitude (indicated by white dashed line), so the horizontal ozone gradients are relatively strong.

4. Ozone-only assimilation

~~Here~~ In this section, we evaluate the influence of ozone-only assimilation on the wind and temperature ~~(T)~~ analyses. There are several factors that will affect the ozone-wind/temperature~~T~~ relationships in the system. The experiments focus on the sensitivity to ~~the initial conditions~~perturbations in the DAS, blending coefficient, sampling pattern, and observation error (see Table 1 for a complete list of experiments). To quantify the ozone impact, we calculate the root mean square (RMS) error profiles for the background and Ozone Assimilation Experiment~~s~~ (OAE) for vector wind and T ~~on 1 December 2014 (00Z)~~ in three latitude bands (NH, ~~30~~90°N - ~~90~~30°N; TR, 30°~~S~~N - 30°~~N~~S; SH, 30°~~S~~ - 90°S). These were calculated by first computing the RMS error (RMSE) for ~~zonal wind (u)~~, ~~meridional wind (v,)~~ and ~~temperature (T)~~ using the following formula (shown below for u, but similar for v and T).

$$u_{RMSE}^2(k) = \frac{\sum\limits_{j=j\min}^{j\max}\left[\dfrac{1}{nlon}\sum\limits_{i=1}^{nlon}\left(u(i,j,k)-u_{TE}(i,j,k)\right)^2\cos(lat(j))\right]}{\sum\limits_{j=j\min}^{j\max}\cos(lat(j))} \tag{1}$$

Here i, j, and k are indices for longitude (lon), latitude (lat), and vertical level, while nlon indicates the number of longitudes and jmin and jmax indicate the latitude indices corresponding to the bounding latitudes for each region (NH, TR, or SH). To ~~get~~ calculate the vector wind error, we combine the u and v errors as follows.

17

$$V^2_{RMSE}(k) = \sqrt{u^2_{RMSE}(k) + v^2_{RMSE}(k)} \qquad (2)$$

To reduce random noise, the errors are averaged over the last 6 days of the experiment (25 November 2014, 00Z - 1 December 2014, 00Z), allowing for the ensemble spin-up as well as reduction of initial errors. Time series of the vertically averaged vector wind and temperature errors (using the same approach as in Sect. 3.2, but limited to the pressure range of the ozone observations, 77 - 1.1 hPa) are provided in Fig. 13. These show results for the assimilation of ozone and radiances in different combinations, which will be examined in more detail below. Here we simply point out that the errors in each experiment level out around 25 November. We can therefore consider that the sensitivity to initial conditions has been generally lost by 25 November, and the 6-day average is reasonable. Due to limited independent analyses in this study, we do not actually test for statistical significance; Geer et al. (2016) suggest that large numbers (on the order of several hundred) of independent tests may be necessary to determine statistical significance to changes in a NWP system.

4.1 ~~Dependence on initial conditions~~Sensitivity of the analysis to perturbations in the DAS

Before comparing the results from the various ozone assimilation experiments, we first examine the dependence ~~on the initial conditions~~of the analysis on perturbations to the system. As discussed by Geer (2016), a perturbation to the observational dataset or the numerics of the system will generate chaotic divergence in a given DAS. By running the perturbed and unperturbed system, we can estimate the analysis variance, which is referred to as the "null hypothesis" by Geer (2016). We generate a simple four-member null set as the TE and experiments that ~~If we start with zero initial error ("unperturbed"), and~~ assimila~~te~~te perfect global, random, or polar-orbiting stratospheric ozone observations, starting from unperturbed initial conditions. ~~(of any type),~~ For an ideal DA system starting from perfect initial conditions, adding perfect observations would produce zero innovations, which would have no effect on the analysis. However, ~~we might expect to have zero analysis error relative to the TE. However,~~ the inclusion of additional observations, ~~even if they are perfect,~~ relative to the TE changes the numerics of the cost function minimization, resulting in slightly different analysis increments. Also, the specified ozone σ_{ob} standard deviation is not zero (0.1 ppmv), allowing some variation in the realized state. In addition, sub-optimalities in the data assimilation system, such as sampling error in the ensemble background error covariances, may also result in changes to the analyses. These behaviors of the system create a limit on the level of errors we can reliably distinguish from the TE. We also note that small changes in the troposphere will lead to differences in gravity waves that grow exponentially with increasing altitude (see also Sect. 4.3) ~~(which includes all the tropospheric observations).~~ For a high-altitude analysis, such as the one used here, this should cause the chaotic divergence, or variance, to be largest in the lower mesosphere (Liu et al, 2008). ~~This results in slight changes to the troposphere, which thereby forces the stratosphere away from the TE. This behavior of the system creates a limit on the level of errors we can reliably distinguish from the TE.~~

18

The results is areis illustrated in Fig. ~~9~~14, which shows ~~errors~~ error profiles ~~for perfect ozone assimilation experiments~~ averaged over the last 6 days of the experiments. ~~started from both perturbed and unperturbed initial conditions (first 3 experiments in Table 1).~~ The ~~resulting~~ vector wind and *T* errors are very similar below about 300 hPa, suggesting that the lower tropospheric response is independent of the choice of stratospheric perturbations. These tropospheric errors could be considered a rough estimate of the spread in a hypothetical tropospheric ensemble created by a large number of analyses with slightly different perturbations. As expected, the differences relative to the TE increase with altitude in the stratosphere and lower mesosphere. In the stratosphere, there are larger variations in errors among the three experiments, with wind errors ranging from ~0.5 - 2 ms^{-1} and *T* errors ranging from ~0.1 - 1 K. The global experiment gives the largest differences from the TE in the lower stratosphere. This might be expected, since it has the largest number of observations. However, the global experiment has generally smaller errors in the lower mesosphere. This may due to the dense spatial and temporal coverage allowing the best restraint of the error growth from gravity waves. The sparse coverage from polar ozone assimilation, on the other hand, shows the greatest differences from the TE at high altitudes, since it likely has the least ability to limit forecast divergence. Further analysis (beyond the scope of this paper) would be necessary to evaluate this mechanism in detail. Due to the limited number of independent analyses in this study, we do not use this null set to determine statistical significance. However, t~~after 16 days show relatively small errors in the troposphere and stratosphere (wind errors less than ~2 ms^{-1} in the tropics, while *T* errors are generally less than ~1 K. The results with unperturbed initial conditions (black lines) show slightly smaller errors than with perturbed initial conditions (red lines), but the general profile shapes are the same. This suggests that after sufficient time, the experiments are not very sensitive to the size of the initial error.~~hese results do provide a preliminary estimate of the sensitivity of the analyses to perturbations in the DA system, and the maximum stratospheric errors in Fig. 14 could be considered a rough estimate of the minimum possible errors for our subsequent ozone assimilation experiments. ~~The results with unperturbed initial conditions (black lines) show slightly smaller errors than with perturbed initial conditions (red lines), but the general profile shapes are the same. This suggests that after sufficient time, the experiments are not very sensitive to the size of the initial error.~~

4.2 Dependence on blending coefficient

The next set of experiments assimilate perfect global ozone data in the stratosphere, using blending coefficients of $\alpha = 0.0, 0.5,$ and 1.0. For each choice of blending coefficient, three experiments are completed for the TE, BE (no ozone), and OAE (ozone); see nine experiments listed for this subsection in Table 1. ~~Separate~~As discussed above, separate TE and BE are necessary for each case, since the blending coefficient affects not only the stratosphere, but also the tropospheric analysis ~~(see nine experiments listed for this subsection in Table 1)~~.

Figure ~~10~~15 shows RMS~~E~~E ~~errors~~ for the BE (dotted) and OAE (solid) for the three blending coefficients. The BE wind errors increase with altitude throughout the stratosphere, ranging from ~1-2 ms^{-1} at 100 hPa to ~~70~~60 ~~80~~70 ms^{-1} at 1 hPa in the NH,

~2520-30-25 ms^{-1} in the tropics, and ~43-5 ms^{-1} in the SH. The differences in wind errors in different latitude bands reflect different sensitivities to perturbations in the initial conditions. Due to the low sensitivity in the SH, our discussion will focus mainly on the NH and TR. The BE errors for different blending coefficients show slight differences, indicating the sensitivity of the stratosphere to changes in the DA system used for tropospheric analysis. The BE T errors are also largest in the NH and

5   TR, with generally increasing errors with height in the stratosphere. BE T errors in the stratosphere reach up to ~20-25-30 K in the NH, ~5-5-6 K in the TR and ~1.5-2 K in the SH.

For the OAEs, there are generally large reductions in vector wind errors throughout the stratosphere and mesosphere relative to the BE. In the NH and TR, the results with non-zero blending coefficients are better than with $\alpha = 0.0$ above about 10 hPa. This indicates that the ensemble correlations are playing a large role at higher altitudes. This is expected, since the conventional

10   balance approximations were designed to simulate tropospheric balance conditions, and they do not take into account the influence of resolved unbalanced modes such as gravity waves. The $\alpha = 0.5$ results are slightly better than the $\alpha = 1.0$ results, suggesting that combining covariances is helpful for the system, a well-documented result in the stratosphere (e.g., Kuhl et al., 2013). The T errors for the OAEs also show reductions in the NH relative to the BE, with the $\alpha = 0.5$ producing the consistently smallest errors throughout the stratosphere and mesosphere, followed by $\alpha = 1.0$ and $\alpha = 0.0$. In the TR and SH, the $\alpha = 1.0$

15   results are worse than $\alpha = 0.0$ in the stratosphere.

The larger errors in $\alpha = 1.0$ may be related to spurious resolved gravity waves being generated in the system. To identify gravity wavesGW, Fig. 1116 shows global divergence patterns maps on 1 December 2014 (00Z) at 10.5 hPa for the OAE with three blending coefficients, along with the zonal standard deviation of the divergence as a function of latitude. The divergence

20   at this level clearly increases over the globe with more ensemble information added to the system. The globally averagedglobally averaged divergence profiles are provided in Fig. 1217. These show increasing divergence with altitude, as expected for upward propagating gravity waves. Also, , showing that the he enhanced divergence is enhanced occursat all vertical levels with larger value of α. This suggests that local imbalance due to the use of localized ensemble covariance may be causing gravity waves that are propagating upward into the stratosphere and mesosphere (see Keypert, 2009 and Allen et

25   al., 2015 for discussions of imbalance in the framework of the shallow water model and EnKF). Although more work is necessary to sort out the details, using $\alpha = 0.5$ likely provides the best results by combining reduced spurious imbalance relative to $\alpha = 1.0$ as well as enhanced flow-of-the-day information relative to $\alpha = 0.0$. We will use $\alpha = 0.5$ as the blending coefficient for the following sensitivity tests as well as for the combined ozone/radiance assimilation experiments in Sect. 5.

4.3 Dependence on sampling pattern

30   The previous results show that with global hourly coverage, ozone observations are able to constrain the stratospheric to winds to error of less than about 2 ms^{-1} and T errors less than about 1 K. These are approximately the limits indicated by the bounds of the null hypothesis set examined in Sect. 4.1, indicating that the results are near the discernable limit for this system. The

globalThis sampling is, of course, unrealistic in both horizontal and temporal coverage. Here we examine sensitivity to sampling by repeating the $\alpha = 0.5$ experiments with polar-orbiting and random sampling (see Fig. 1b,c). The polar-orbiting sampling would be similar, for example, to the MLS or Ozone Mapping and Profiler Suite (OMPS). The random sampling is not realistic, but provides a hypothetical test of what would happen if random observations occurred with the same frequency

5   as the polar orbiter. In each case, we assume perfect ozone observations and σ_{ob} of 0.1 ppmv. The RMS error profiles, averaged over the last 6 days of the experiments, are provided in Fig. 1318.

The OAE error profiles for vector wind show that assimilation of both polar-orbiting (bluegreen) and random (red) perfect ozone observations reduce the errors relative to the BE (dotted line). Particularly in the lower stratosphere, from about 100 to

10   10 hPa, the wind errors remain are relatively small, less than about 4 ms⁻¹. In the upper stratosphere (above 10 hPa), the errors for the polar-orbiting observations increase sharply with altitude to ~50 ms⁻¹ at 1.0 hPa. The wind errors for random sampling (red) are consistently lower than for polar-orbiting, even though both contain approximately the same number of observations. While there may be some redundancy in the polar-orbiting observations due to closely spaced along-track profiles, it is also likely that the large gaps between orbit tracks (see Fig. 1b) make it difficult for the polar-orbiting observations to completely

15   constrain the winds. While the random sampling does better than polar-orbiting, there are still rather large wind errors in the random sampling, up to ~20 25 m/s in the NH and ~10 m/s in the tropicsTR. We note that the error reductions occur even in the mesosphere, where there are no observations, suggesting suggesting that that improving the stratospheric analyses will also also improves the mesosphere.

20   The OAE error profiles for T also show improvements relative to the BE when polar-orbiting observations are assimilated, with the smallest errors occurring in the lower stratosphere. However, in the NH upper stratosphere the polar-orbiting observations only constrain T to about ~15 K. The experiment with random observations has smaller T errors, which are similar to the global errors in the NH and tropics TR up to about 10 hPa. In the SH lower stratosphere, the polar-orbiting and random cases actually have smaller T errors than the global case, and the global case has errors larger than the BE. However, the

25   magnitude of these errors are near the error limit discussed in Sect. 4.1. Overall, we see that the ozone-dynamical influence is strongly sensitive to the sampling pattern, but wind and T improvements are possible even with a realistic polar-orbiting satellite.

4.4 Dependence on observation error

Next, we examine the sensitivity of the analysis to the ozone observation error. First, we assimilate polar-orbiting data with

30   2% error (blue green solid lines on Fig. 1318). This is a realistic error value for the middle stratosphere; for example, Aura MLS V4.2 precision specifications are rated at 2% at 22, 10, and 5 hPa and greater than 2% elsewhere (Livesy et al., 2016). The results show slightly larger vector wind and T errors in the NH for 2% error than when perfect data were assimilated, but

the errors are less than the background throughout the stratosphere and mesosphere, suggesting value added by these observations. In the tropics and SH, the 2% case is also very similar to the 0% case for both vector wind and T. These results suggest that assimilating actual profile measurements with realistic errors can potentially benefit the analyses.

5   We now add random noise to the global observations using Gaussian errors of 2%, 5%, and 10% to further examine sensitivity to errors. For each of these three cases, the specified σ_{ob} is also set to the same percent value, but with a lower limit of 0.1 ppmv (see Table 1, Sect. 4.4). The results, in Fig. ~~14~~19, show that adding noise increases the vector wind errors over the perfect observations. Below about ~~5~~10 hPa, the wind errors are similar for 2%, 5%, and 10% cases, while above ~~5~~10 hPa, there is generally increased error with increased observational noise. The stratospheric wind errors are still relatively small (less than
10   ~10 ms^{-1}, even with 10% error, suggesting that the dynamic variability of ozone is large enough to allow wind information on this error level. ~~In the SH lower stratosphere, adding ozone errors does cause errors larger than the BE case. This may be related to small variability of the ozone fields in this region.~~

The OAE ~~T~~ T errors (Fig. 19d,e,f) ~~show reduction relative to the BE in the NH and TR, and errors~~ are also generally larger
15   with more observational noise. ~~With~~At 10% applied ozone errors, the stratospheric T errors are constrained to within ~~~6~~4 K in the NH, ~2K and TR, and ~1~~.5~~ K in the SH. In the SH lower stratosphere, there is a reversal of the T errors, with the 10% case showing smaller errors than the 5% or 2% cases. The cause of this reversal is uncertain, but it may be that using higher σ_{ob} in the 10% case reduces the weight of the observations and therefore results in reduced spurious errors relative to the TE. ~~This may become important in regions of weak tracer tendency. Allen et al. (2013) discussed this possibility in both idealized 1-D~~
20   ~~advection simulations and NAVGEM simulations of a single cycle of ozone assimilation.~~ Overall, we conclude here that noisy observations will generally reduce the amount of wind information that can be derived from ozone~~, but if the σ_{ob} is specified consistent with the actual errors, then the result is generally an improvement over the no-ozone baseline case~~. As a caveat, we remind the reader that we are only simulating random error and not biases, which could be a significant source of additional error. ~~In the next section, we examine how ozone impacts winds and T in experiments that include realistic stratospheric~~
25   ~~radiance observations.~~

5. Ozone and radiance assimilation

5.1 Baseline experiment for radiance assimilation

~~In Sect. 4, we showed that ozone assimilation can benefit the winds and T in the stratosphere and mesosphere. Here, we~~We next examine the impact of ozone when the stratosphere is already constrained by radiance observations. As described in Sect.
30   2.3, we simulated infrared and microwave radiance observations for AMSU-A, AIRS, ATMS, and IASI for the α = 0.5 TE

case, and then added random noise. ~~As with the ozone (Sect. 4.1), we ran experiments assimilating radiance data with both perturbed or unperturbed initial conditions. The profiles after 16 days of assimilation (not shown) were again very similar, suggesting that all the results we show are not too sensitive to initial conditions.~~ The vector wind and T error profiles for radiance-only experiments are provided in Fig. ~~15~~20. The ~~black~~grey lines shows the results for assimilation of noisy radiances. In the stratosphere, wind errors range from around 2 to 4 ms^{-1}, while T errors range from around 0.~~4~~5 to 1.5 K. These are relatively small errors, ~~such that~~making ~~reducing the~~ error reductions ~~further~~ via ~~by adding~~ ozone assimilation ~~is~~ more challenging ~~from the perspective of ozone dynamical interactions~~than when only ozone was assimilated in the stratosphere. Comparing the noisy radiance results in Fig. 20 with the noisy global ozone (2% error) from Fig. 19, we see that ozone assimilation has smaller wind errors throughout most of the stratosphere in all three latitude bands. For T, however, the radiance assimilation has generally smaller errors in the extratropics, while T errors are similar for ozone or radiances in the tropics.

Before ~~combining~~adding ozone ~~with radiances~~into the system, we also performed an experiment in which "perfect" radiances were assimilated with unperturbed initial errors, ~~a~~. ~~As~~ we did with ~~the~~ozone in ~~(~~Sect. 4.1.~~),~~ ~~w~~Wind and T error profiles from this case are shown ~~with~~in the ~~black~~red lines on Fig. ~~15~~20. ~~As with the assimilation of perfect ozone with unperturbed initial conditions,~~ ~~T~~the experiment with perfect radiances results in errors ~~both in the troposphere and stratosphere~~in both the troposphere and stratosphere that are slightly larger than the case of perfect ozone with unperturbed initial conditions (Fig. 14). Potential reasons for the difference may include the interaction of deep vertical weighting functions of radiances observations with the tropospheric state or the much larger number of radiance observations compared to ozone observations. ~~W~~As explained in Sect. 4.1, in our experimental design, inclusion of additional observations (relative to the TE), even if they are perfect, results in slight changes to the troposphere, which thereby forces the stratosphere away from the TE. Comparing Fig. 15 with Fig. 9, we see that the error profiles for perfect radiances are slightly larger than for perfect global ozone. Therefore, it is likely that radiances are the limiting factor when combining ozone and radiance data together. For this reason, we~~ expect the errors for the combined ozone/radiance experiments to lie within ~~black and red~~the grey and black lines of Fig. ~~15~~ 20~~(discussed further below)~~.

5.2 Ozone and radiance assimilation experiments

In the next set of experiments, ozone data (global, random, and polar-orbiting) are ~~added to~~assimilated along with the noisy radiance observations~~assimilation~~. Figure ~~16~~21 shows vertical profiles of the resulting errors for perfect ozone observations for the three sampling patterns. We also include in Fig. ~~16~~21 the error profiles (black and ~~gray~~grey lines) from the radiance assimilation experiments shown in Fig. ~~15~~20 as a comparison. In the TR, the impact of ozone assimilation is positive for all three sampling patterns, with generally increasing impact with altitude throughout the stratosphere. Global observations reduce tropical vector wind errors by up to ~2 ms^{-1} at the stratopause, while random and polar-orbiting data reduce tropical wind errors by about 1.0 and 0.4 ms^{-1}, respectively. In the NH and SH, global observations benefit winds throughout the stratosphere, but at a reduced amount compared with the tropics. The impact of random observations in the NH and SH is positive throughout

the stratosphere, but at smaller levels (~0.1 to 0.3 ms^{-1}) than for global data. For polar-orbiting observations, the impact on NH and SH winds is even smaller, but still generally positive. We note that the error profiles generally lie within the perfect and noisy radiance profiles, and the vector wind errors for global ozone is close to the perfect radiance profile. This suggests that the ozone is reducing wind errors to near the minimum possible values, identified by the perfect radiance case.

5

Temperature error reductions show a similar pattern to the wind errors, with largest impact from the global observations in the tropical upper stratosphere, where reductions of ~0.7 K occur. The random and polar-orbiting observations also impact the tropical upper stratosphere, with T error reductions of ~0.3 and ~0.1 K, respectively. In the extratropics, the impact of the random and polar-orbiting observations on T is small, generally less than ~0.1 K. We also note that impact of ozone
10 observations on both winds and T is generally positive in the mesosphere, above the highest observation level. Global ozone assimilation results in temperature errors that are slightly lower (higher) than the perfect radiance case in the tropics (NH and SH). In terms of vertically-averaged errors (over the range of ozone observations, 77 - 1 hPa), the overall impact on both winds and T is positive throughout the time period of consideration (see Fig. 13a,b). While six days is likely too small for statistical significance tests, the pattern of error reductions is consistent, with polar data providing a slight improvement, followed by
15 random data. The errors for global data are similar to the perfect radiance error case.

For further quantification of ozone impact on the dynamics, ~~Because the resulting OAE errors are very close to the radiance only results,~~ we also ~~plot~~ calculated the error difference (OAE errors minus the ~~black line~~ noisy radiance errors~~s from Fig. 15~~), where~~.~~ negative ~~Positive~~ values of this difference indicate value added due to ozone assimilation. ~~Figure 17 shows vertical
20 profiles of these differences for perfect ozone observations.~~

~~In the TR, the impact of ozone assimilation is positive for all three sampling patterns, with generally increasing impact with altitude throughout the stratosphere. Global observations reduce vector wind errors by up to ~1.5 ms^{-1} at the stratopause, while random and polar-orbiting data reduce tropical wind errors by about 0.5 and 0.3 ms^{-1}, respectively. In the NH and SH, global
25 observations benefit winds throughout the stratosphere, but at a reduced amount compared with the tropics. The fact that ozone has the largest benefit in the tropics is consistent with there being less wind constraint from radiance observations. The impact of random observations in the NH and SH is positive throughout the stratosphere, but at much smaller levels (~0.1 - 0.3 ms^{-1}) than for global data. For polar-orbiting observations, the impact on NH and SH winds is even smaller, but still generally positive. We note that the error profiles generally lie within the perfect and noisy radiance profiles of Fig. 15, and the vector
30 wind errors for global ozone is very close to the perfect radiance profile. This suggests that the ozone is reducing errors to near the minimum possible values, identified by the perfect radiance case.~~

~~Temperature error reductions show a similar pattern to the wind errors, with largest impact from the global observations in the tropical upper stratosphere, where reductions of ~0.7 K occur. The random and polar-orbiting observations also impact the~~

24

tropical upper stratosphere, with *T* error reductions of about 0.3 and 0.2 K, respectively. In the extratropics, the impact of the random and polar-orbiting observations on *T* is small, generally less than about 0.1 K, but generally positive. We also note that impact of ozone observations on both winds and *T* is generally positive in the mesosphere, above the highest observation level. Figure 22 plots these quantities as zonal mean cross-sections. For global ozone, the vector winds and *T* are improved throughout

5   much of the stratosphere and lower mesosphere, with peak reductions of ~3 ms^{-1} and ~1 K in the tropical upper stratosphere. The random and polar sampling patterns also improved the tropical winds and temperature, but to a lesser extent than the global data. One reason for the maximum impact is in the tropics is due to the dynamics being less constrained by radiance measurements due to lack of geostrophy. As demonstrated by Allen et al. (2015), using a shallow water model with simulated ozone and height observations, ozone assimilation particularly improved the tropical winds. In addition, as discussed by Daley

10   (1996), the impact on winds from tracer assimilation also depends on tracer time tendency. When this tendency is small, wind recovery will occur slowly or not at all. Small tendency can occur when the tracer gradient is small, when the wind speed is low, or when the tracer and streamfunction are highly correlated. To test this, we averaged the absolute value of the 1-h time tendency over the last 6 days of the TE at each grid point and then calculated the zonal mean. The results (Fig. 23) show the strongest tendencies in the tropical lower and upper stratospheres, in the NH polar upper troposphere/lower stratosphere, and

15   in the SH lower stratosphere. The tropical maxima roughly coincide with regions of large wind and *T* error reductions in Fig. 22, but there are no corresponding error reductions in the extratropical lower stratospheres. This may be due to the radiances having a large influence there with many observations and strong geostrophic coupling, making it difficult for ozone observations to add value.

20   Next, in Figs. 18 24 and 1925, we repeat the above comparison with noisy ozone data (2% applied error)ozone error. The wind error profiles (Fig. 24) show much less variation from the noisy radiance profile. Global ozone benefits the winds over a large altitude range and also benefits *T*, particularly in the tropical upper stratosphere. It is difficult to discern impacts of random and polar ozone from these plots, except in the lower mesosphere, where there is a slight benefit. This is also seen in the vertically integrated error time series (Fig. 13c,d), where the global ozone is consistently smaller than the noisy radiance

25   alone. The random and polar errors are not easily distinguishable from noisy radiance. The global ozone observations tend to have the largest positive impact, in the tropical middle stratosphere. However, the magnitude of the impact, for both wind errors and *T* errors, is much smaller than for perfect ozone. The impact of random and polar ozone observations is also smaller throughout the stratosphere. It is not surprising that the radiances overwhelm these noisy polar and random ozone observations since the total number of ozone observations (3500/4 profiles x 17 observations/profile = 14,875) is only ~1% of the number

30   of radiance observations (e.g., 1,768,409 observations shown in Fig. 2) for a given 6-h cycle.
. The very small benefit from global ozone with reasonable errors suggests that any benefit that current NWP systems gain from adding stratospheric ozone assimilation is likely not due to the dynamical ozone-wind interaction in hybrid 4D-Var. [DRA1]

The zonal mean error differences for the 2% ozone error case are provided in Fig. 25. Wind and *T* error reductions are seen for the global ozone in the tropical upper stratosphere and lower mesosphere, while the random and polar only show strong influence in the tropical lower mesosphere. Slight increases in temperature errors also occurs for global ozone in the NH polar region. Whether these error impacts are statistically significant cannot be determined from the limited number of samples. However, a general pattern of improvement from ozone assimilation in the tropical stratosphere and mesosphere emerges from these experiments, providing encouragement for future work with more realistic observing systems, model resolution, and extended analyses to allow significance testing.

These results provide a rather sobering conclusion that in the presence of realistic radiance observations, it is likely that adding ozone assimilation, from current ozone retrieval observations, will have little to no impact on the winds through the ozone-wind interactions investigated in this study. However, we add the caveat that the ozone, radiances, and model used in the study are unbiased. We did not test the effect that ozone might have when the radiances or model have bias. We also did not investigate the degradation that ozone assimilation might cause if the ozone observations are biased. Semane et al. (2009) showed a slight reduction in wind bias in the lower stratosphere when assimilating of MLS ozone data, suggesting that the ozone might help provide further benefit in the presence of model or observation biases. Further work is necessary to examine ozone impact on mean wind and temperature errors in more detail. [DRA2]

6. Summary and Conclusions

This study examined the potential impact of stratospheric ozone assimilation on the wind and temperature analyses in the stratosphere and lower mesosphere for a case study in late November/early December 2014. We used unbiased simulated measurements and a perfect model to test the windozone-dynamics interaction in hybrid 4D-Var DA, which arises from background ensemble covariances and the tracer advection in the linear/adjoint model. The structures of the ensemble cross-correlations for ozone with other variables were illustrated with a composite single ozone observation increment, formed by averaging the spatial cross-correlations for 36 points around a latitude circle. Clear patterns emerged that included rotation around a vertical axis and a vertical temperature dipole. These patterns resembled the potential vorticity "charge" concept, discussed by Bishop and Thorpe (1994). This suggests that an ozone observation, at least in the presence of sufficient spatial gradients and geostrophic balance, acts like an observation of potential vorticity. This is likely due to both quantities being quasi-conserved in the stratosphere and therefore forming compact relationships. Further work on the understanding of these relationships may provide insight into designing ozone observingozone-observing systems that would optimize the ozone-wind relationship.

Experiments were then conducted in which simulated stratospheric observations were assimilated in a cycling hybrid 4D-Var system. The resulting analyses were compared with a truth experiment that was used for simulating the observations and

verifying the analyses. All experiments included a suite of conventional ~~tropospheric~~ observations ~~and satellite derived winds~~ to constrain the troposphere. This approach allowed a controlled method for determining ozone impact on the stratospheric dynamics, while maintaining a realistic troposphere. Experiments assimilated various combinations of stratospheric ozone and radiances. The mechanisms through which ozone can impact the winds in hybrid 4D-Var include both the application of cross-covariances of ozone with other fields in the initial blended background error covariance and the use of the ozone continuity equation in the tangent linear model/adjoint. We showed that using a blending coefficient of 0.5 provided better results than either 0.0 or 1.0. This is likely due the combined positive effects of the ensemble flow-of-the-day information with the negative aspects of spurious unbalanced modes spawned by the localized ensemble covariance. These aspects were discussed in the shallow water model context in Allen et al. (2016), where it was shown that the optimal blending coefficient also depends both on the data being assimilated and on the ensemble size.

Ozone assimilation ~~clearly~~ can benefit the winds and temperature~~s~~ if sufficient high-quality observations are available. For example, global hourly ozone data with no error constrained the stratosphere to within a few ms^{-1} for the winds an~~d~~ ~1 K for temperature, which was better than noisy radiance assimilation, but worse than perfect radiance assimilation. When ozone is assimilated with radiances, wind improvement is mostly found in the tropics, where the lack of geostrophic balance renders radiances less effective. For example,. ~~This demonstrates that there is dynamical information embedded in the ozone field that could potentially be "mined" to obtain stratospheric wind and temperature information. This is particularly important for the tropics, where radiances alone have difficulty due to their large weighting functions and breakdown of geostrophic balance.~~ assimilating realistic radiance data and perfect global ozone data resulted in additional tropical wind and temperature error decreases in the upper stratosphere of ~3 ms^{-1} and ~1 K relative to the noisy radiance data. However, when a realistic 2% error was added to the global ozone data, the tropical error decreases were reduced to ~1 ms^{-1} and ~0.5 K, and slight error increases occurred in the NH polar region. Reduction of the sampling frequency also reduced the~~and/or addition of~~ ~~observational noise~~ ~~reduced the~~ benefit of ozone.. ~~Without radiance assimilation, ozone improved winds and temperatures, even for a single polar orbiting measurement with realistic error. When added to simulated radiance assimilation, ozone had a very small benefit, to the extent~~ that realistic 2% ozone error resulted in insignificant wind and temperature changes using our methodology. [DRA3]

We are unable to establish the statistical significance of the additional error reductions for the cases of noisy ozone assimilation, since much longer assimilation tests are required to establish significance for changes of less than a percent (Geer, 2016). When looking for very small improvements from ozone assimilation, several other factors are likely to become important. We ~~also~~ limited the study to unbiased ozone and radiance observations. Further work is necessary to determine the impact of ozone assimilation in a system with model and/or observation biases~~,~~. Also, our experiments simulated ozone measurements on the model vertical grid. For lower vertical resolution ozone measurements such as Solar Backscatter Ultraviolet or OMPS Nadir Profiler (Flynn et al., 2009, 2014), or ozone sensitive channels of infrared sounders, the observation operator must include a vertical weighting function. Further studies are required to determine the ozone vertical resolution requirements for achieving

27

wind improvements. Similarly, unlike this study, the horizontal resolution of current NWP systems is higher than the resolution of limb sounding ozone measurements such as MLS (Livesy et al., 2015) and OMPS Limb Profiler (Jaross et al., 2014). This requires the use of horizontal weighting functions in the observation operator for optimal assimilation. All these issues will be important for achieving wind improvements from ozone assimilation with current NWP systems and ozone measurement technology.

In this study, we only simulated ozone vertical profile measurements, since we expect that vertical resolution is essential for the ozone-wind relationship to be robust. However, vTotal column observations ertical sounders suchfrom as Ozone Monitoring Instrument (OMI) or OMPS could provide supplementary information to constrain the winds, particularly in the lower stratosphere (see, for example, the study by Peuch et al., 2000). Certain radiance channels also have ozone sensitivity that could potentially be exploited (Dragani and McNally, 2013). We also limited the study to unbiased ozone and radiance observations. Further work is necessary to determine the impact of ozone assimilation in a system with model and/or observation biases. Other approaches to the ozone-dynamical impact could include assimilation of radiances channels that are sensitive to ozone as well as assimilation of ozone radiances directly into the system rather than retrieved profiles. In addition, the impact of assimilation of other tracers could be tested in a similar framework. For example, Andersson et al. (1994, 2007) have shown dynamical impacts from assimilation of radiance channels sensitive to water vapor, and Allen et al. (2014), in a shallow water model study, showed that nitrous oxide and water vapor could additionallyso potentially benefit improve winds in 4D-Var DA.. Given the potential benefit of tracer assimilation on the dynamics in NWP (referred to by Daley (1995) as a "tantalizing possibility"), it is our hope that this study will motivate future work in this area and eventually result in improved operational analyses and forecast skill.

Acknowledgments

References

Allen, D. R., Stanford, J. L., Elson, L. S., Fishbein, E. F., Froidevaux, L., and Waters, J. W.: The 4-day wave as observed from the Upper Atmosphere Research Satellite Microwave Limb Sounder, J. Atmos. Sci., 54, 420-434, 1997,

Allen, D. R., Hoppel, K. W., Nedoluha, G. E., Kuhl, D. D., Baker, N. L., Xu, L., and Rosmond, T. E.: Limitations of wind extraction from 4D-Var assimilation of ozone, Atmos. Chem. Phys., 13, 3501-3515, doi:10.5194/acp-13-3501-2013, 2013.

Allen, D. R., Hoppel, K. W., and Kuhl, D. D.: Wind extraction potential from 4D-Var assimilation of stratospheric O_3, N_2O, and H_2O using a global shallow water model, Atmos. Chem. Phys., 14, 3347-3360, doi:10.5194/acp-14-3347-2014, 2014.

Allen, D. R., Hoppel, K. W., and Kuhl, D. D.: Wind extraction potential from ensemble Kalman filter assimilation of stratospheric ozone using a global shallow water model, Atmos. Chem. Phys., 15, 5835-5850, doi:10.5194/acp-15-5835-2015, 2015.

Allen, D. R., Hoppel, K. W., and Kuhl, D. D.: Hybrid ensemble 4DVar assimilation of stratospheric ozone using a global shallow water model, Atmos. Chem. Phys., 16, 8193-8204, doi:10.5194/acp-16-8193-2016, 2016.

Andersson, E., Palleux, J., Thépaut, J.-N., Eyre, J. R., McNally, A. P., Kelly, G. A., and Courtier, P.: Use of cloud-cleared radiances in three/four-dimensional variational data assimilation, Q. J. Roy. Meteor. Soc., 120, 627-653, 1994.

Andersson, E., E. Hólm, P. Bauer, A. Beljaars, G. A. Kelly, A. P. McNally, A. J. Simmons, J.-N. Thépaut, and Tompins, A. M.: Analysis and forecast impact of the main humidity observing systems, Q. J. Roy. Meteor. Soc., 133, 1473-1485, doi:10.1002/qj.112, 2007.

Bishop, C. H., and Thorpe, A. J.: Potential vorticity and the electrostatics analogy: Quasi-geostrophic theory, Q. J. Roy. Meteorol. Soc., 120, 713-731, 1994.

Bonavita, M., Isaksen, L., and Hólm, E., 'On the use of EDA background error variances in the ECMWF 4D-Var,' ECMWF Tech Memo 664. Available at https://www.ecmwf.int/en/elibrary/8272-use-eda-background-error-variances-ecmwf-4d-var http://old.ecmwf.int/publications/library/do/references/show?id=90381 (last access: 21 December 20171 April 2015), 2012.

Bishop, C. H., and Thorpe, A. J.: Potential vorticity and the electrostatics analogy: Quasi-geostrophic theory, Q. J. Roy. Meteorol. Soc., 120, 713-731, 1994.

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B.: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: One-month experiments with real observations, Mon. Wea. Rev., 138, 1567-1586, doi:10.1175/2009MWR3158.1, 2010.

Cariolle, D., and Morcrette, J.-J.: A linearized approach to the radiative budget of the stratosphere: influence of the ozone distribution, Geophys. Res. Lett., 33, L05806, doi:10.1029/2005GL025597, 2006.

Clayton, A. M., Lorenc, A. C., and Barker, D. M.: Operational implementation of a hybrid ensemble/4D-Var global data assimilation system at the Met Office, Q. J. R. Meteorol. Soc, 139, 1445–1461. doi:10.1002/qj.2054, 2013.

Daley, R., Atmospheric Data Analysis, Cambridge University Press, Cambridge, pp 150-185, 1991.

Daley, R.: Estimating the wind-field from chemical-constituent observations - experiments with a one-dimensional extended Kalman filter, Mon. Wea. Rev., 123, 181-198, doi: 10.1175/1520-0493(1995)123<0181:ETWFFC>2.0.CO;2, 1995.

Daley, R.: Recovery of the one and two dimensional windfields from chemical constituent observations using the constituent transport equation and an extended Kalman filter, Meteorol. Atmos. Phys., 60, 119-136, 1996.

Daley, R., and Barker, E.: NAVDAS Source Book 2001, NRL Publication NRL/PU/7530—01-441, available at http://www.dtic.mil/docs/citations/ADA396883, last access: ~~16~~ 21 ~~August~~ December 2017, Naval Research Laboratory,
5   Moneterey, California, USA, 163pp, 2001.

Dethof, A., and Hólm, E.: Ozone assimilation in the ERA-40 reanalysis project, Q. J. Roy. Meteorol. Soc., 130, 285102872, doi:10.1256/qj.03.196, 2004.

Dragani, R., and McNally, A. P.: Operational assimilation of ozone-sensitive infrared radiances at ECMWF, Q. J. Roy. Meteor. Soc., 139, 2068-2080, doi:10.1002/qj.2106, 2013.

10   Eckermann, S. D., Hoppel, K. W., Coy, L., McCormack, J. P., Siskind, D. E., Nielsen, K., Kochenash, A., Stevens, M. H., Englert, C. R., Singer, W., and Hervig, M.: High-altitude data assimilation system for the northern summer mesosphere season, J. Atm. Sol.-Terr. Phys., 71, doi: 10.1016/j.jastp.2008.09.036, 531-551, 2009

Flynn, L. E., McNamara, D., Beck, C. T., Petropavlovskikh, I., Beach, E., Pachepksy, Y., Li, Y. P., Deland, M., Huang, L.-K., Long, C. S., Tiruchirapalli, R., and Taylor, S.: Mearurements and products from the Solar Backscatter Ultraviolet (SBUV/2)
15   and Ozone Mapping and Profiler Suite (OMPS) instruments, International Journal of Remove Sensing, 30:15, 4259-4272, doi:10.1080/01431160902825040, 2009.

Flynn, L., Long, C., Wu, X., Evans, R., Beck, C. T., Petropavlovskikh, I., McConville, G., Yu, W., Zhang, Z., Niu, J., Beach, E., Hao, Y., Pan, C., Sen, B., Novicki, M., Zhou, S., and Seftor, C.: Performance of the Ozone Mapping and Profiler Suite (OMPS) products, J. Geophys. Res. Atmos., 119, 6181-6195, doi:10.1002/2013JD020467, 2014.

20   Geer, Alan J.: Significance of changes in medium-range forecast scores, Tellus A: Dynamic Meteorology and Oceanography, 68, 30229, http://dx.doi.org/10.3402/tellusa.v68.30229, 2016.

Harnisch, F., Healy, S. B., Bauer, P., and English, S. J.: Scaling of GNSS Radio Occultation Impact with Observation Number Using an Ensemble of Data Assimilations. Mon. Wea. Rev., 141, 4395–4413, https://doi.org/10.1175/MWR-D-13-00098.1, 2013.

25   Han, W., and McNally, A. P.: The 4D-Var assimilation of ozone-sensitive infrared radiances measured by IASI, Q. J. Roy. Meteor. Soc., 136, 2025-2037, doi:10.1002/qj.708, 2010.

Hogan, T. F., and Rosmond, T. E.: The description of the Navy Global Operational Prediction System's spectral forecast model, Mon. Wea. Rev., 199, 1786-1815, 1991.

Hogan, T., Liu, M., Ridout, J., Peng, M., Whitcomb, T., Ruston, B., Reynolds, C., Eckermann, S., Moskaitis, J., Baker, N., McCormack, J., Viner, K., McLay, J., Flatau, M., Xu, L., Chen, C., Chang, S.: The Navy Global Environmental Model. Oceanography 27, 116–125, 2014.

5   Jaross, G., Bhartia, P. K., Chen, G., Kowitt, M., Haken, M., Chen, Zhong, C., Xu, P., Warner, J., and Kelly, T.: OMPS Limb Profiler instrument performance assessment, J. Geophys. Res. Atmos., 119, 4399-4412, doi:10.1002/2013JD020482, 2014.

Kepert, J. D.: Covariance localisation and balance in an ensemble Kalman filter, Q. J. Roy. Meteor. Soc., 135, 1157-1176, doi:10.1002/qj.443, 2009.

Kleist, D. T., and Ide, K.: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part II: 4DEnVar and hybrid variants, Mon. Wea. Rev., 143, 452-470, doi:10.1175/MWR-D-13-00350.1, 2015.

10 Kuhl, D. D., Rosmond, T. E., Bishop, C. H., McLay, J., and Baker, N. L.: Comparison of Hybrid Ensemble/4DVar and 4DVar within the NAVDAS-AR data assimilation framework, Mon. Wea. Rev., 141, 2740-2758, doi:10.1175/MWR-D-12-00182.1, 2013.

Liu, H.-L., Sassi, F., and Garcia, R. R.: Error growth in a whole atmosphere climate model, J. Atmos. Sci., 66, 173-186, doi:10.1175/2008JAS2825.1, 2009.

15 Livesy, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Millan Valle, L. F., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Martinez, E.: Version 4.2x Level 2 data quality and description document, Jet Propulsion Laboratory, California Institute of Technology, JPL D-33509 Rev B, 2015.

Louis, J. F.: A parametric model of vertical eddy fluxes in the atmosphere, Boundary Layer Meteorol., 17, 187-202, 1979.

McLay, J. G., Biship, C. H., Reynolds, C. A.: Evaluation of the ensemble transform analysis perturbation scheme at NRL,
20 Mon. Wea. Rev., 136, 1093-1108, doi:10.1175/2007MWR2010.1, 2008,

McLay, J. G., Bishop, C. H., and Reynolds, C. A.: A local formulation of the ensemble transform (ET) analysis perturbation scheme, Wea. Forecasting, 25, 985-993, doi:10.1175/2010WAF2222359.1, 2010.

Milewski, T., and Bourqui, M. S.: Assimilation of stratospheric temperature and ozone with an ensemble Kalman filter in a chemistry-climate model, Mon. Wea. Rev., 139, 3389-3404, doi:10.1175/2011mwr3540.1, 2011.

25 Peubey, C., and McNally, A. P.: Characterization of the impact of geostationary clear-sky radiances on wind analyses in a 4D-Var context, Q. J. Roy. Meteor. Soc., 135, 1863-1876, doi: doi:10.1002/qj.500, 2009.

Peuch, A., Thepaut, J. N., and Pailleux, J.: Dynamical impact of total-ozone observations in a four-dimensional variational assimilation, Q. J. Roy. Meteor. Soc., 126, 1641-1659, doi: 10.1002/qj.49712656605, 2000.

Riishøjgaard, L. P.: On four-dimensional variational assimilation of ozone data in weather-prediction models, Q. J. Roy. Meteor. Soc., 122, 1545-1571, doi:10.1002/qj.49712253505, 1996.

Rosmond, T. E.: A technical description of the NRL adjoint modeling system, NRL Publication NRL/MR/7532-97-7230, available at: http://www.dtic.mil/docs/citations/ADA330960, last access: ~~16 August~~21 December 2017, Naval Research Labortory, Monterey, California, USA, 57 pp., 1997.

Rosmond, T., and Xu, L.: Development of NAVDAS-AR: non-linear formulation and outer loop tests, Tellus, 58A, 45-58, 2006.

Semane, N., Peuch, V. H., Pradier, S., Desroziers, G., El Amraoui, L., Brousseau, P., Massart, S., Chapnik, B., and Peuch, A.: On the extraction of wind information from the assimilation of ozone profiles in Meteo-France 4-D-Var operational NWP suite, Atmos. Chem. Phys., 9, 4855-4867, doi:10.5194/acp-9-4855-2009, 2009.

Tan, D. G. H., Andersson, E., Fisher, M., and Isaksen, L.: Observing-system impact assessment using a data assimilation ensemble technique: application to the ADM-Aeolus wind profiling mission, Q. J. R. Meteorol. Soc., 133, 381-390, doi:10.1002/qj.43, 2007.

Xu, L., Rosmond, T., and Daley, R.: Development of NAVDAS-AR: formulation and initial tests of the linear problem, Tellus, 57A, 546-559, 2005.

Experiment Type	α	Ozone Observation Sampling	Ozone Error [%]	σ_{ob} [constant or %/minimum]	Radiance Observations	Radiance Error	Initial Conditions
Sect. 4.1							
Truth	0.5						Unperturbed
Ozone Only	0.5	global	0	0.1 ppmv			Unperturbed
Ozone Only	0.5	random	0	0.1 ppmv			Unperturbed
Ozone Only	0.5	polar-orbiting	0	0.1 ppmv			Unperturbed
Ozone Only	0.5	global	0	0.1 ppmv			Perturbed
Sect. 4.2							
Truth	0.0						Unperturbed
Truth	0.5						Unperturbed
Truth	1.0						Unperturbed
Baseline	0.0	global	0	0.1 ppmv			Perturbed
Baseline	0.5	global	0	0.1 ppmv			Perturbed
Baseline	1.0	global	0	0.1 ppmv			Perturbed
Ozone Only	0.0	global	0	0.1 ppmv			Perturbed
Ozone Only	0.5	global	0	0.1 ppmv			Perturbed
Ozone Only	1.0	global	0	0.1 ppmv			Perturbed
Sect. 4.3							
Ozone Only	0.0	global	0	0.1 ppmv			Perturbed
Ozone Only	0.5	random	0	0.1 ppmv			Perturbed
Ozone Only	0.5	polar-orbiting	0	0.1 ppmv			Perturbed
Sect. 4.4							
Ozone Only	0.5	global	0	0.1 ppmv			Perturbed
Ozone Only	0.5	polar-orbiting	2%	2%/0.1 ppmv			Perturbed
Ozone Only	0.5	global	2%	2%/0.1 ppmv			Perturbed
Ozone Only	0.5	global	5%	5%/0.1 ppmv			Perturbed
Ozone Only	0.5	global	10%	2010%/0.1			Perturbed
Sect. 5.1							
Baseline	0.5				All	Noisy	Perturbed
Baseline	0.5				All	Perfect	Unperturbed
Sect. 5.2							
Radiance & Ozone	0.5	global	0	0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	random	0	0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	polar-orbiting	0	0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	global	2%	2%/0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	random	2%	2%/0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	polar-orbiting	2%	2%/0.1 ppmv	All	Noisy	Perturbed
Radiance & Ozone	0.5	global	2	2%/0.1 ppmv	All	Noisy	Perturbed

33

Table 1. Experiment descriptions used in each subsection of this study. Columns indicate (1) experiment type, (2) covariance blending value, (3) ozone observation sampling, (4) ozone observation error, (5) background ozone error standard deviation (given as constant mixing ratio in ppmv or % value with minimum threshold in ppmv), (6) radiance observations,
5 (7) radiance error, and (8) initial conditions.

(a) Global: 92160 Observations (b) Polar: 3460 Observations

(c) Random: 3500 Observations
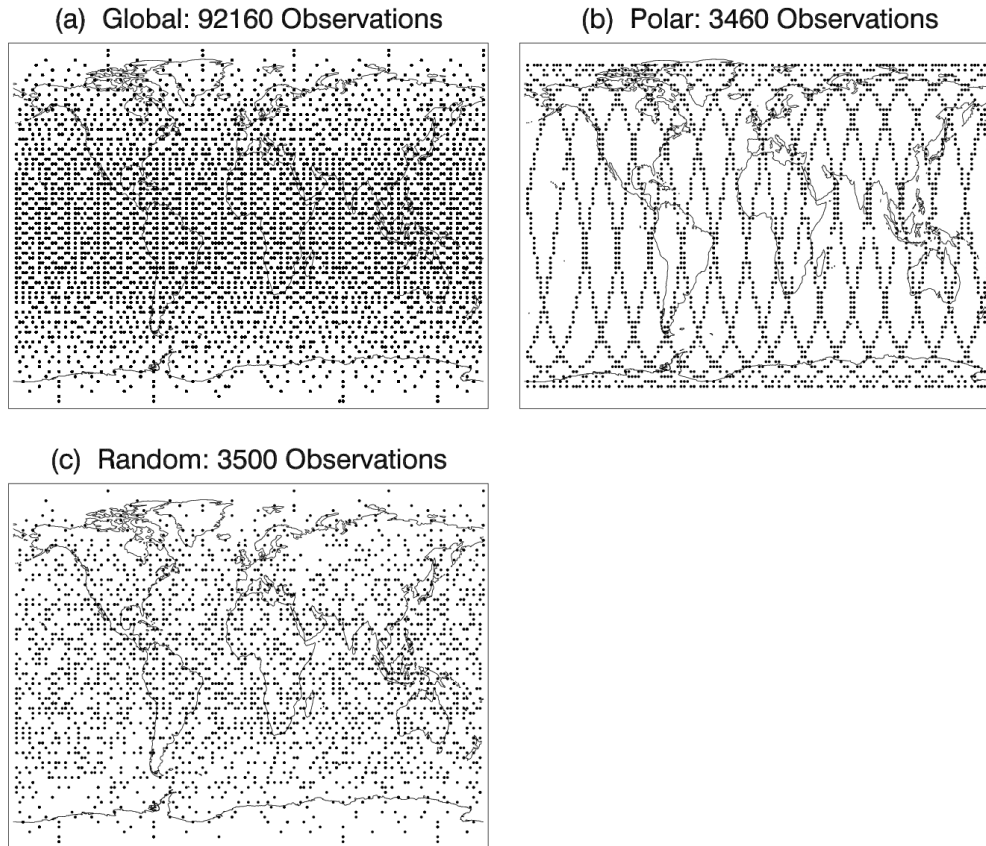
10

Fig. 1. Locations of simulated ozone observations for (a) global, (b) randompolar-orbiting, and (c) randompolar-orbiting data. For global data, these represent hourly coverage, while for polar and random data these are all the observations over a 24-hour period.
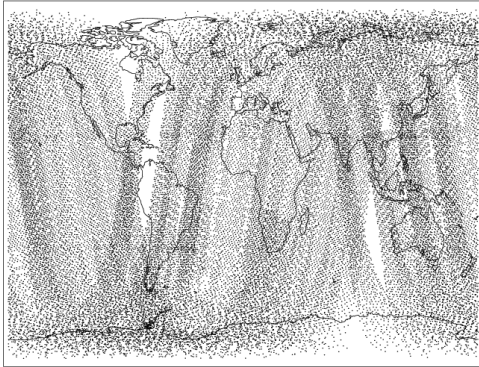
34
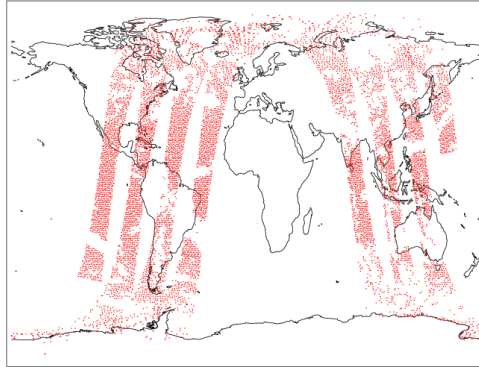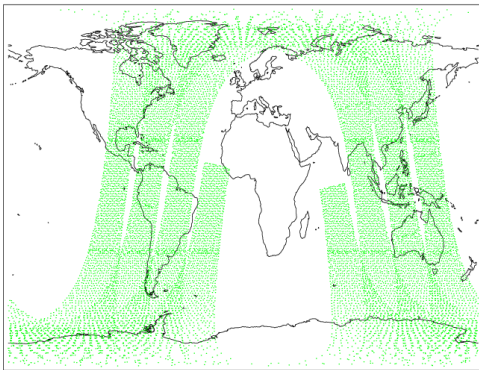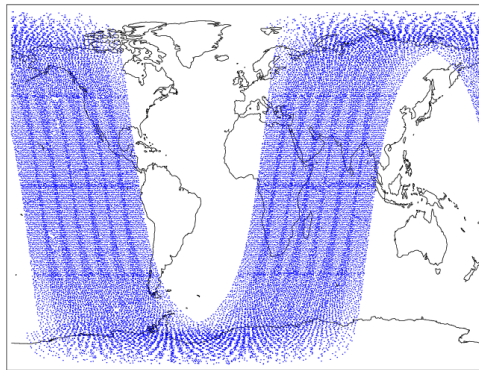
Fig. 2. Locations of simulated radiance observations for one 6-h update cycle centered ~~at 06Z~~ on 15 November 2014 (06Z). Panels show observation locations for (a) AMSU-A, (b) AIRS, (c) ATMS, and (d) IASI.

5
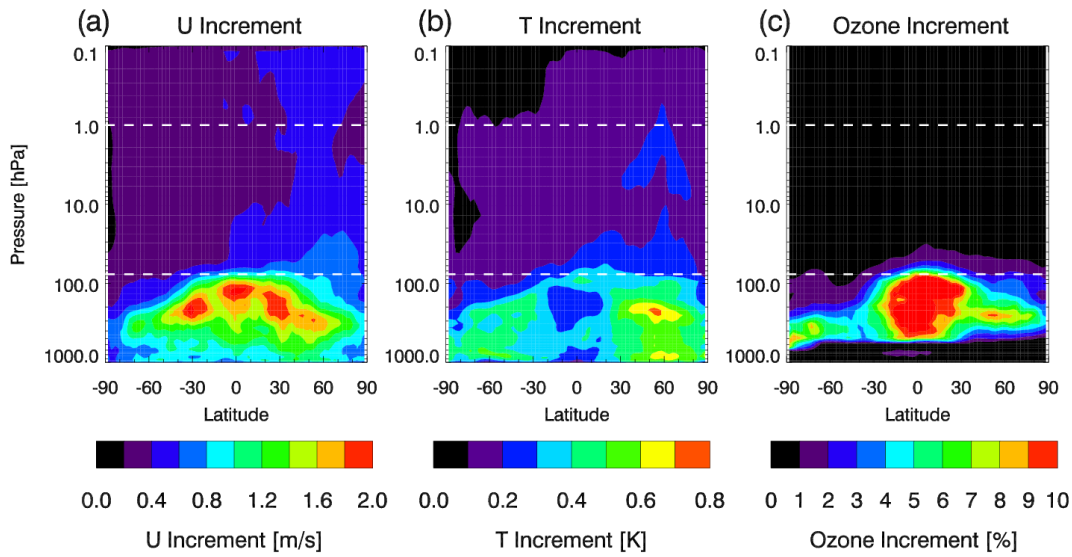
Fig. 3. Zonal mean (pressure vs. latitude) plots of (a) the absolute value of the u increment [in ms^{-1}], (b) the absolute value of the T increment [in K], and (c) the absolute value of the ozone increment divided by the analyzed ozone [in %] for the truth experiment, averaged over all cycles from 25 November - 1 December 2014. Red (blue) indicating high (low) values. Region between the white dashed lines indicates vertical range of assimilated stratospheric ozone observations.
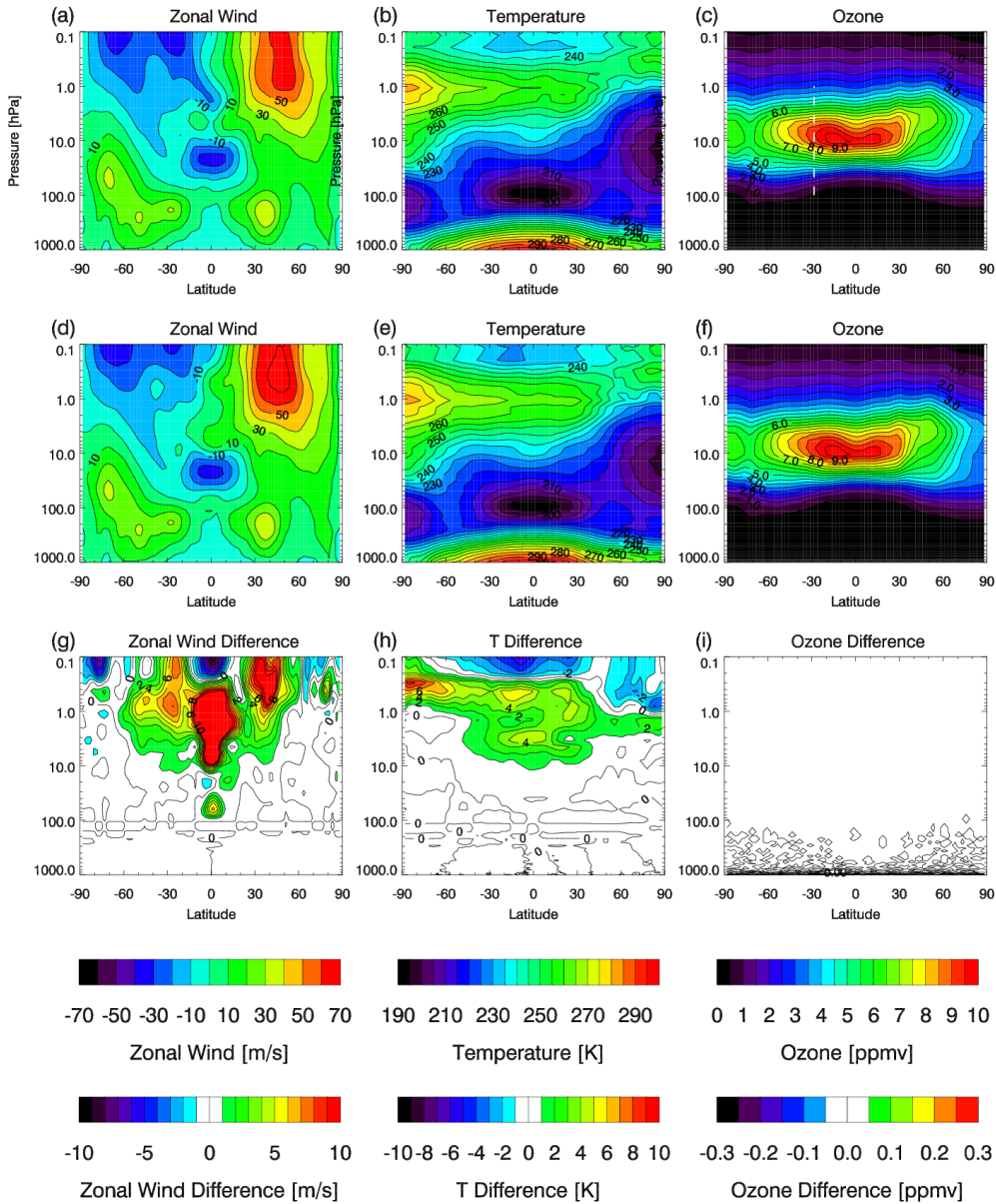
Fig. 54. Analyses for 15 November 2014 (00Z). Zonal mean zonal wind [ms⁻¹] for the (a) truth experiment and (bd) baseline experiments and (eg) baseline-truth zonal wind difference. Zonal mean temperature [K] for the (db) truth experiment and (e) baseline experiment and (fh) baseline-truth temperature difference. Zonal mean ozone [ppmv] for the (c) truth experiment and (f) baseline experiment and (i) baseline-truth ozone difference. Color bars are provided for zonal mean quantities and differences. Red (blue) indicates high (low) values for each quantity. White dashed line in panel (c) indicates latitude of 28.6°S, which is the central latitude of the cross-correlations examined in Sect. 3.4
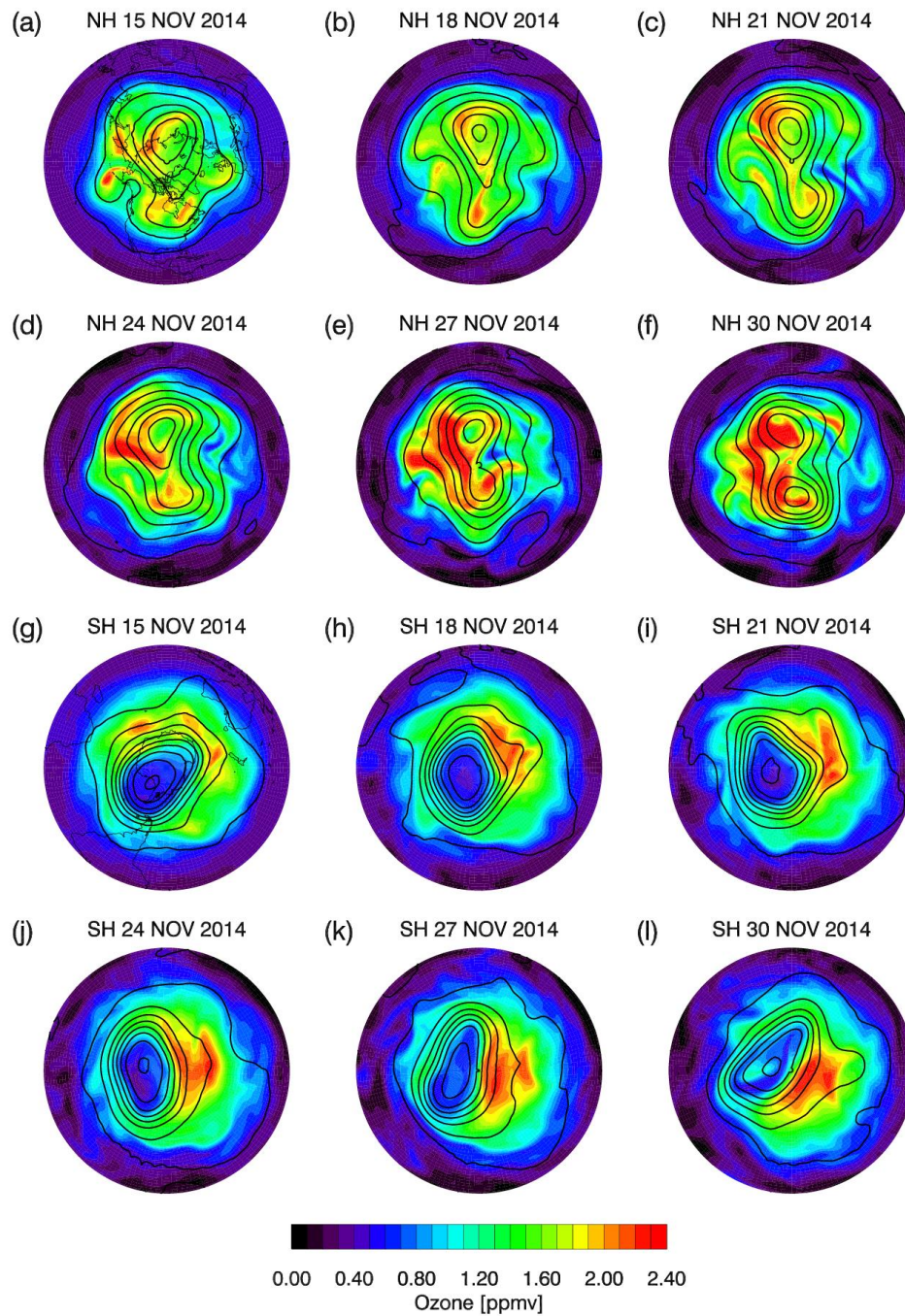
Fig. 5. Maps of ozone [ppmv] (colors) at 77 hPa (~18 km) overlaid with geopotential height (black lines) at 200 m intervals for 15, 18, 21, 24, 27, and 30 November 2014 (all at 00Z). (a-f) are NH and (g-l) are SH. Red (blue) contours indicated high (low) ozone values. Continent lines are placed on the maps for 15 November.
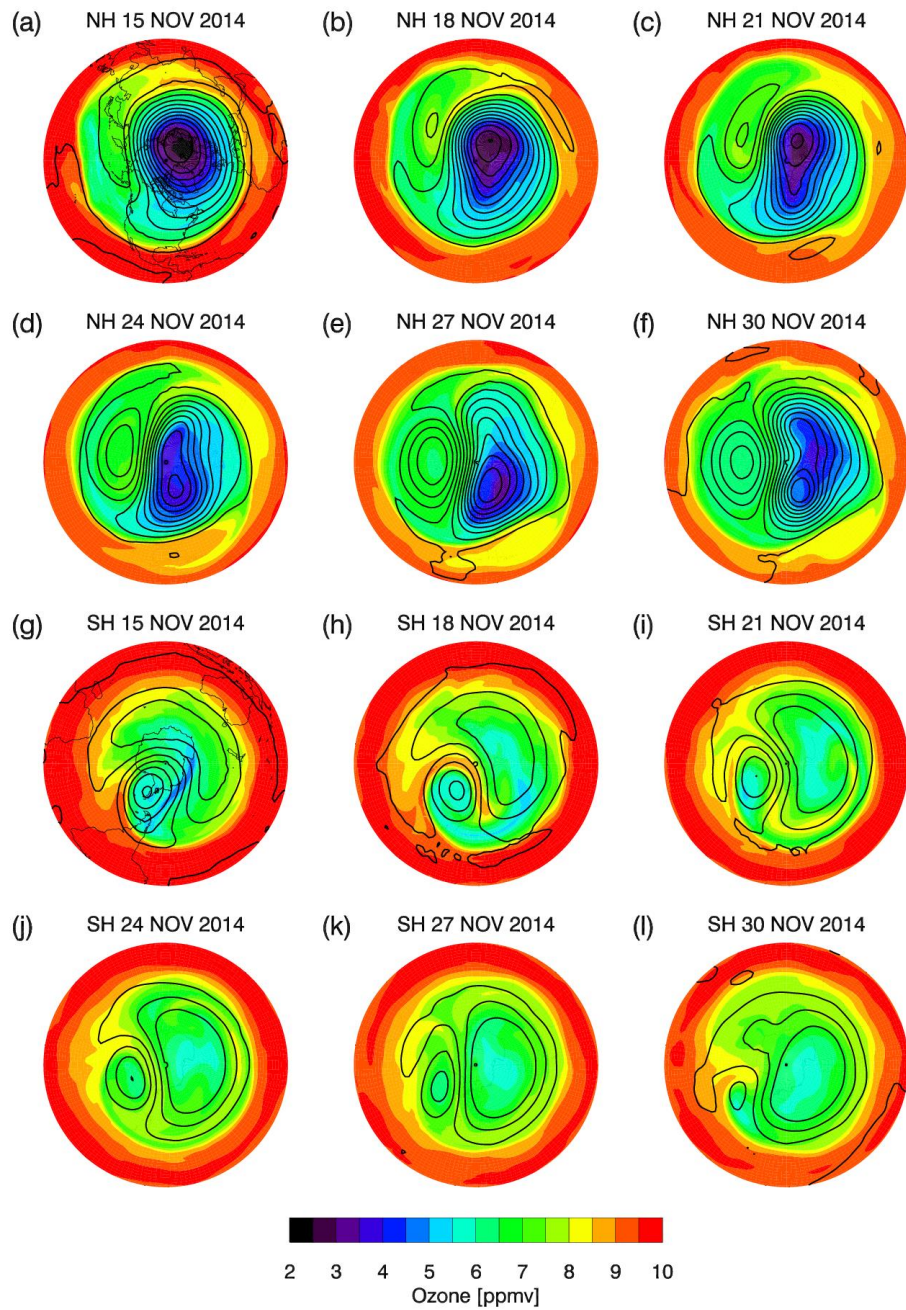
Fig. 36. Maps of ozone [ppmv] (colors) at 10.5 hPa (~32 km) overlaid with geopotential height (black lines) at 200 m intervals for 15, 18, 21, 24, 27, and 30 November 2014 (all at 00Z). (a-f) are NH and (g-l) are SH. Red (blue) contours indicated high (low) ozone values. Continent lines are placed on the maps for 15 November.

39
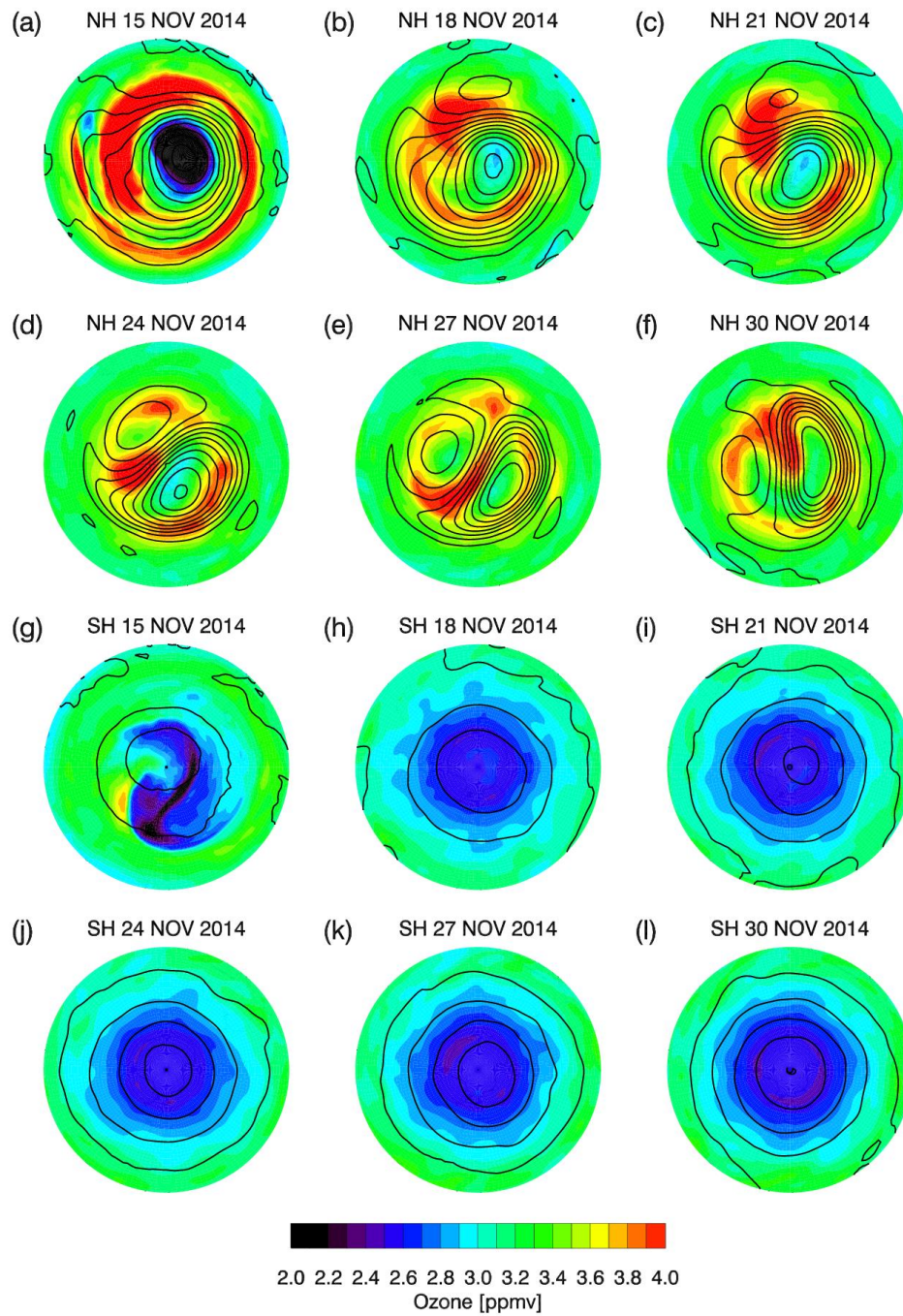
Fig. 7. Maps of ozone [ppmv] (colors) at 1.1 hPa (~48 km) overlaid with geopotential height (black lines) at 500 m intervals for 15, 18, 21, 24, 27, and 30 November 2014 (all at 00Z). (a-f) are NH and (g-l) are SH. Red (blue) contours indicated high (low) ozone values. Continent lines are placed on the maps for 15 November.

40
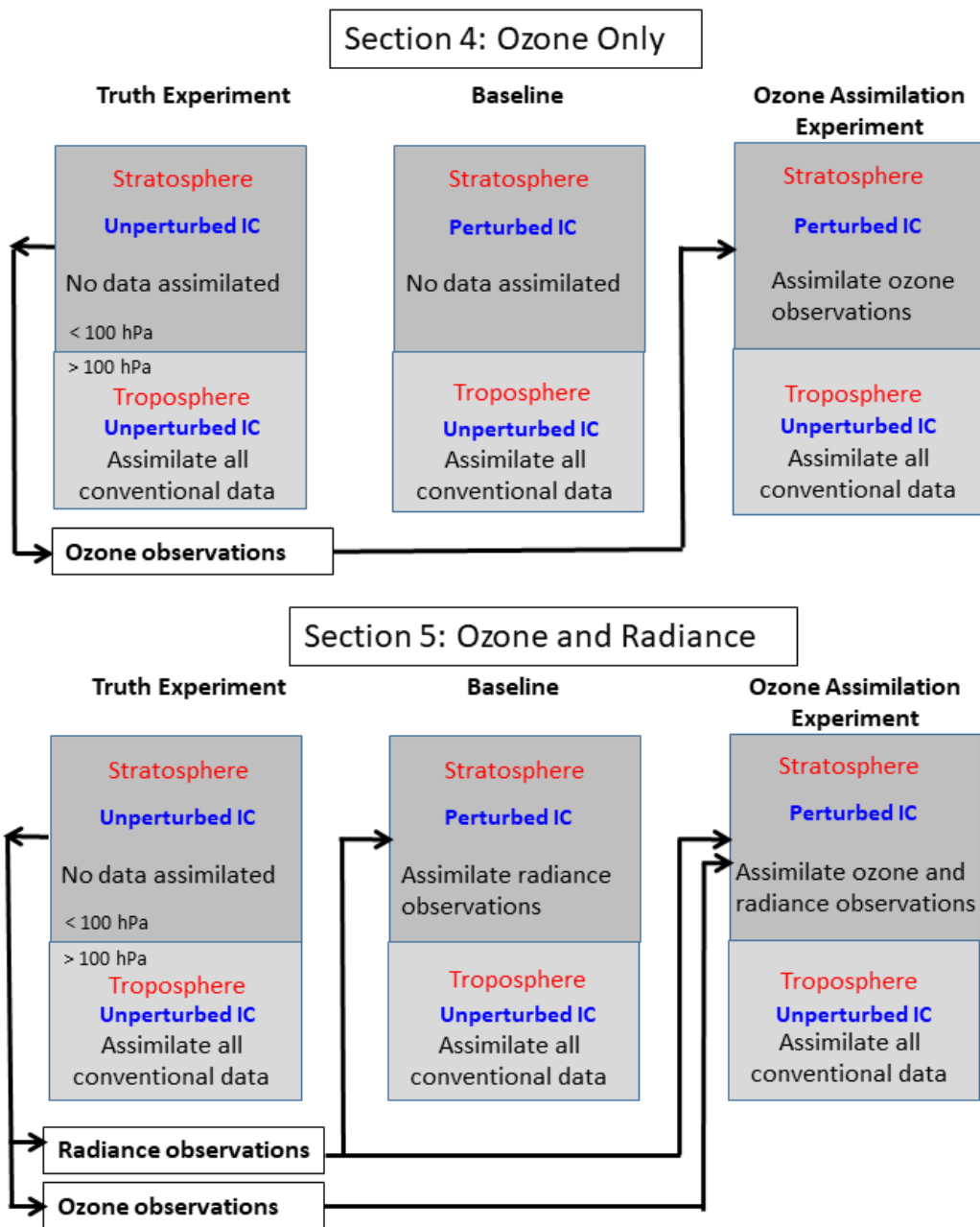
Fig. 48. Schematic diagram to illustrate the design for the experiments analyzed in Sect. 4 and Sect. 5. The truth experiment (TE) in both cases assimilates all conventional data in the troposphere (pressures ~~<~~greater than 100 hPa) and no data in the stratosphere (pressures less than~~<~~ 100 hPa). In Sect. 4, the baseline experiment is the same as the TE, except for perturbed initial conditions (IC) in the stratosphere. The ozone assimilation experiment for Sect. 4 assimilates ozone observations created by the TE. In Sect. 5, the baseline includes assimilation of noisy radiance observations created from the TE. The

41

ozone assimilation experiment in Sect. 5 includes both radiance and ozone observations from the TE.
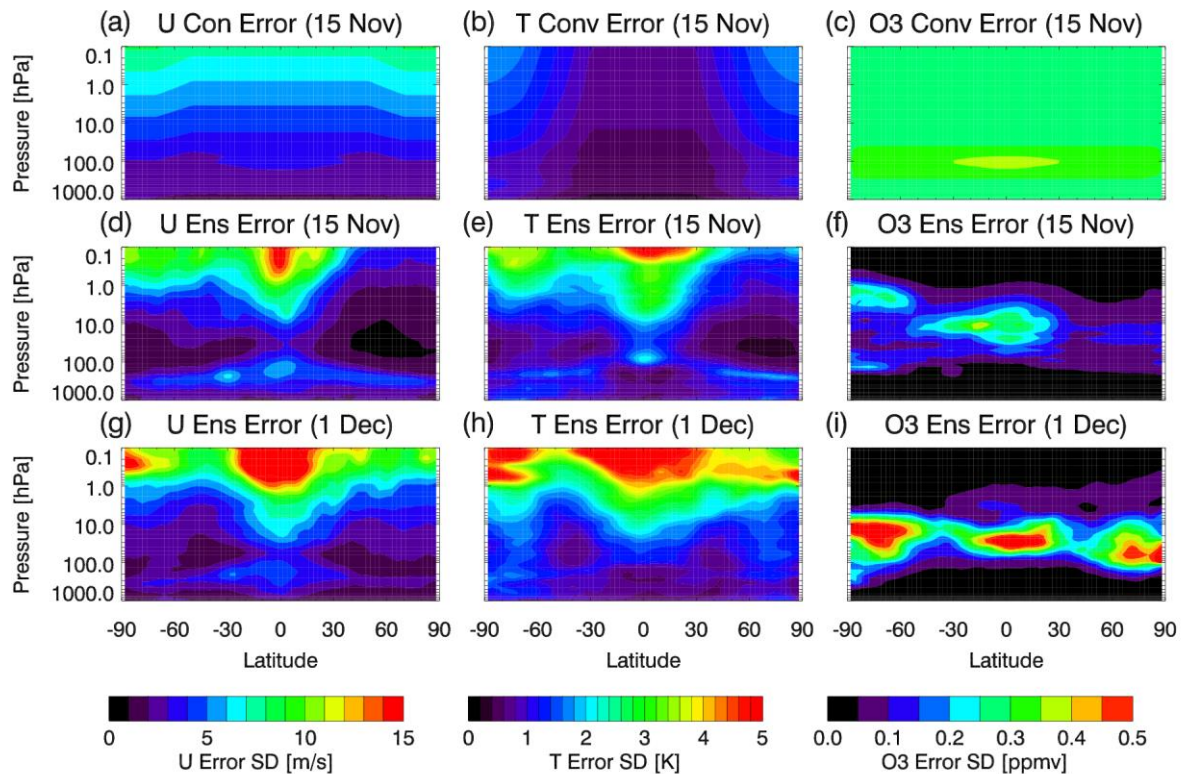
Fig. 6̶9. Conventional background error standard deviations for (a) zonal wind [ms⁻¹], (b) temperature [K], and (c) ozone
[ppmv] for 15 November 2014 (00Z). Ensemble background error standard deviations for (d) zonal wind, (e) temperature, and
(f) ozone for 15 November 2014 (00Z). Ensemble background error standard deviations for (g) zonal wind, (h) temperature,
and (i) ozone for 1 December 2014 (00Z). Color bars are provided for zonal wind, temperature, and ozone, with red (blue)
indicating high (low) values.
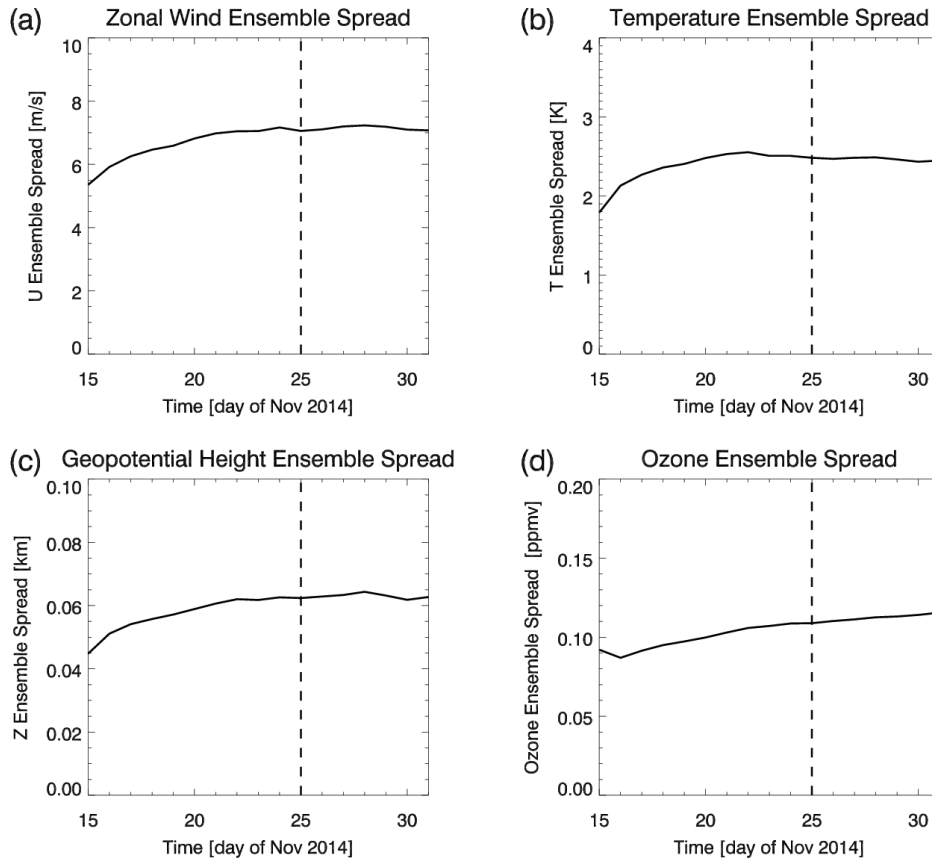
Fig. 10. Globally and vertically (from suface to model top) averaged (surface to model top) ensemble background error standard deviation (or "spread") for 15 November - 1 December 2014, evaluated at 00Z for each day. Panels are for (a) zonal wind [ms^{-1}], (b) temperature [K], (c) geopotential height [km], and (d) ozone [ppmv]. The dashed line on 25 November indicates the end of the spin-up period.

43
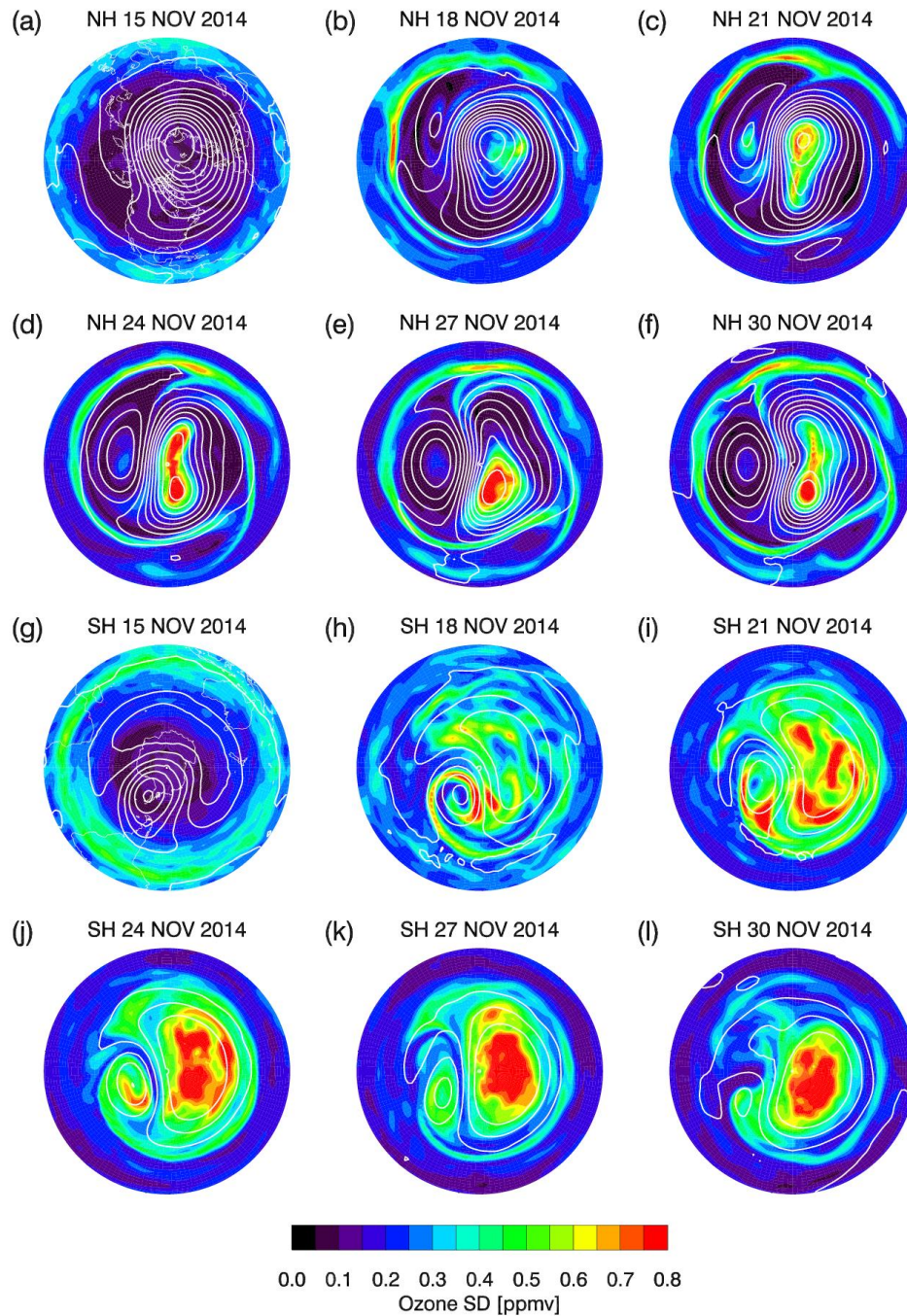
Fig. 11. Maps of ozone σ_{ens} [ppmv] (colors) at 10.5 hPa (~32 km) overlaid with geopotential height (white lines) at 200 m intervals for 15, 18, 21, 24, 27, and 30 November 2014 (all at 00Z). (a-f) are NH and (g-l) are SH. Red (blue) contours indicated high (low) ozone values. Continent lines are placed on the maps for 15 November.

44
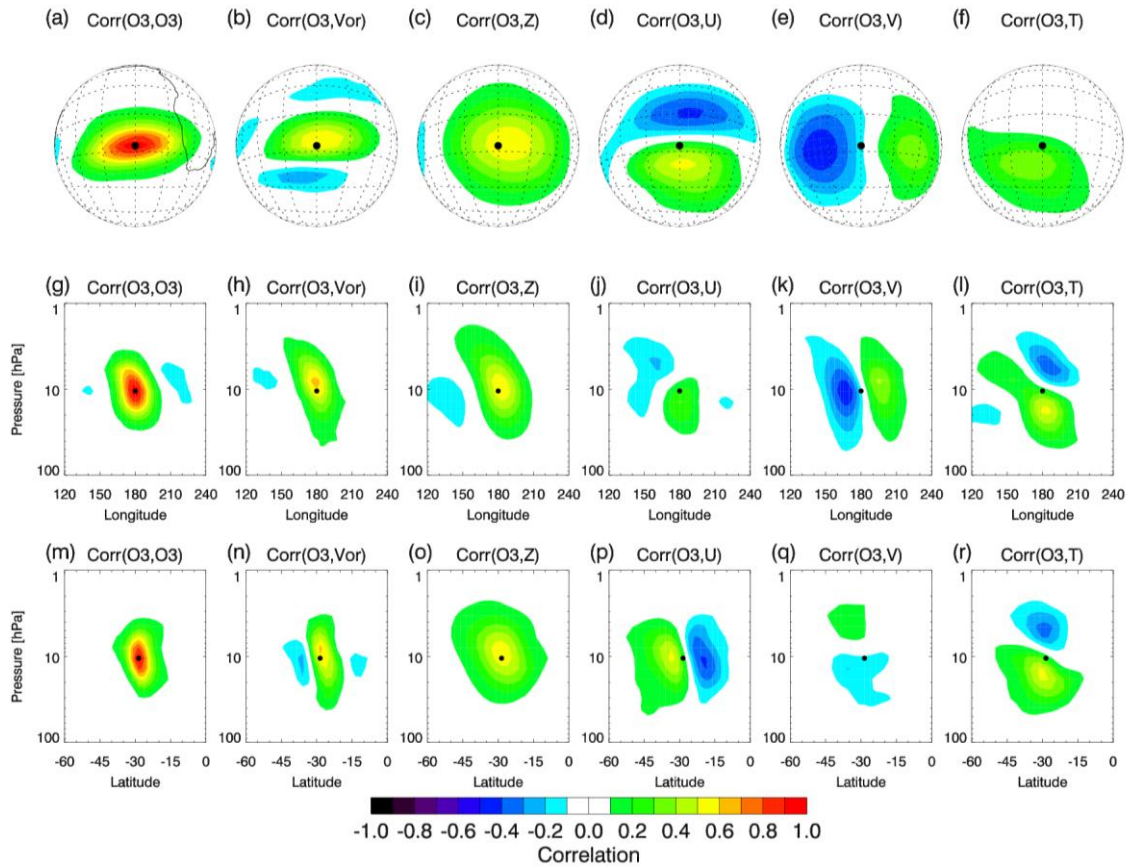
Fig. 812. Composite analysis of the ensemble cross-correlations ("Corr") between ozone and other variables on 30 November 2014 (00Z). Calculation is an average at 10.5 hPa and latitude of 28.6°S, and the observation (black dot) is centered at 180° longitude (see text for details). Top row is the horizontal correlation using a satellite projection with (grid lines at 10° spacing in longitude and latitude, and continental outline is showneen for southern Africa on panel (a). M, middle row is the longitude-pressure cross-section, and bottom row is the latitude-pressure cross-section. Columns indicate correlation between ozone and (1) ozone, (2) vorticity, (3) geopotential height, (4) zonal wind, (5) meridional wind, and (6) temperature. Colors are correlation with red (blue) indicating high (low) values.
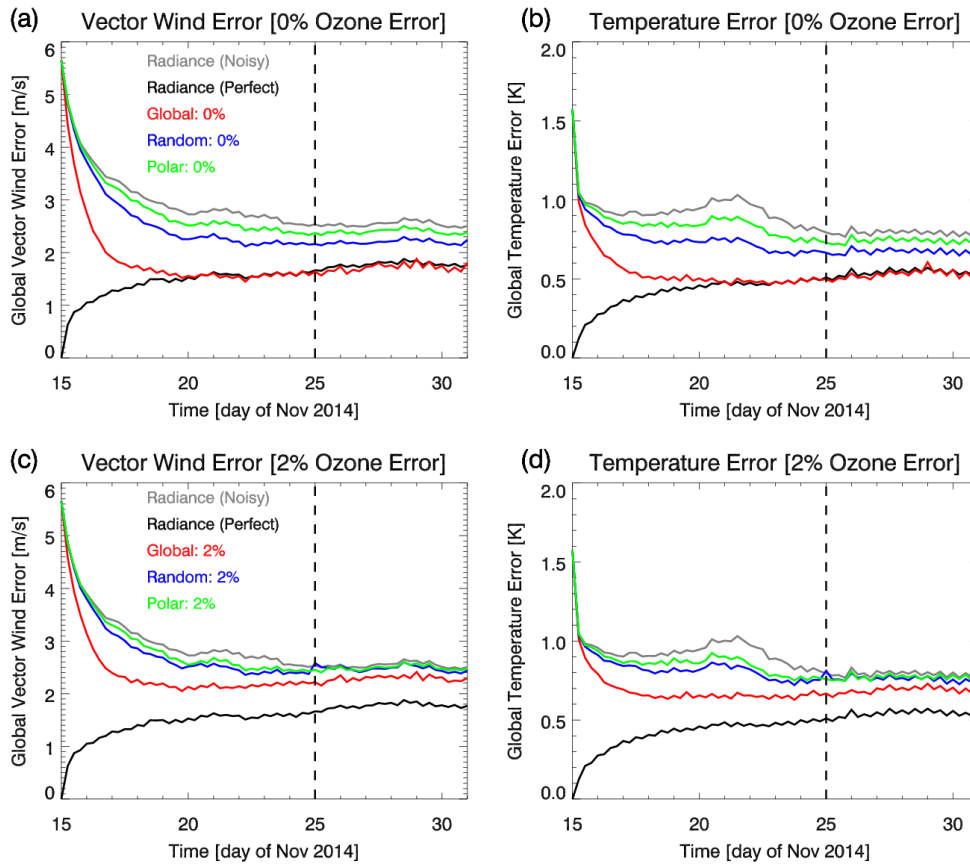
Fig. 13. Globally and vertically (77 hPa to 1.1 hPa) averaged RMS errors for five different cycling experiments from 15 November - 1 December 2014, calculated every 6 hours. Black (grey) lines indicate experiments that assimilate perfect (noisy) radiance data only. Colored lines are for experiments that assimilate noisy radiance data along with global (red), random (blue), or polar-orbiting ozone data (green). Panels (a) and (c) show vector wind errors [ms^{-1}] for cases with perfect ozone (0% added noise) and noisy ozone (2% added noise), respectively. Panels (b) and (d) show temperature errors [K] for cases with perfect ozone and noisy ozone, respectively. Dashed lines indicate the end of the experimental spin-up phase. See Table 1, Sect. 5.1 and Sect. 5.2.
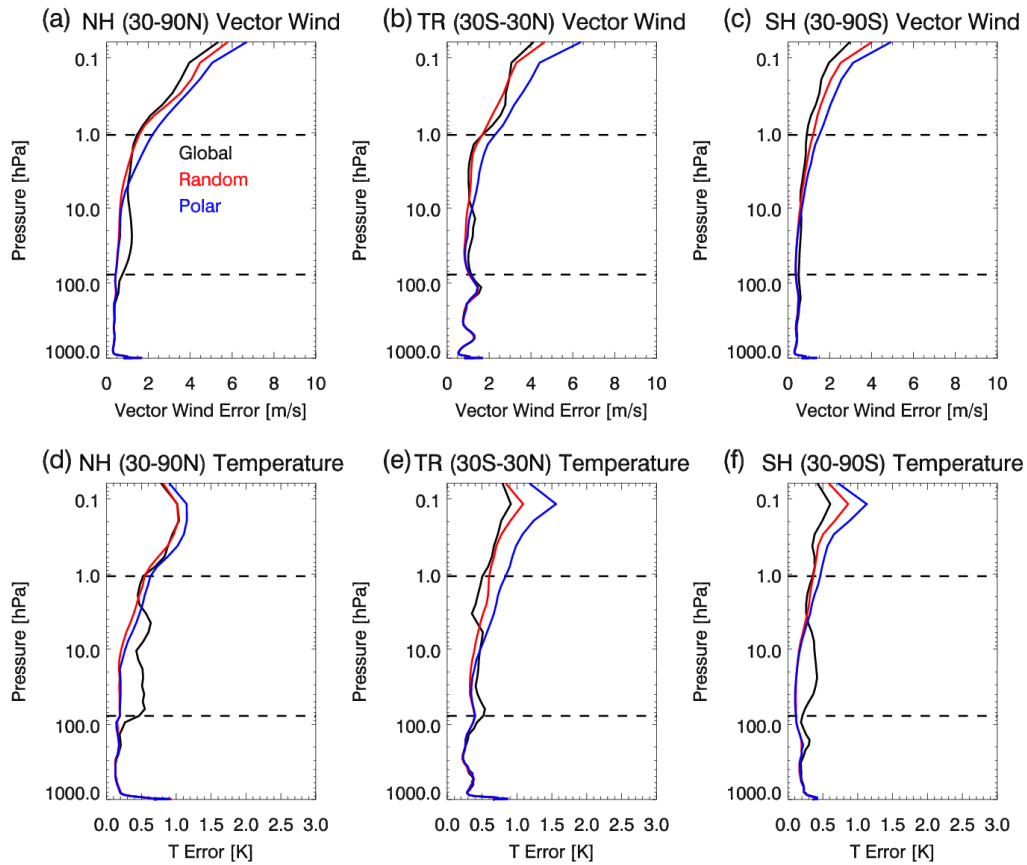
5   **Fig. 9̶14.** RMS vector wind errors [ms^-1] o̶n̶ ̶1̶ ̶D̶e̶c̶e̶m̶b̶e̶r̶ ̶2̶0̶1̶4̶ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K]

for the (d) NH, (e) TR, and (f) SH averaged from 25 November - 1 December 2014.̶ ̶B̶l̶a̶c̶k̶ ̶(̶r̶e̶d̶)̶ ̶l̶i̶n̶e̶s̶ ̶a̶r̶e̶ for experiment̶s̶ that

assimilated perfect global (black), random (red), and polar-orbiting (blue) ozone data with unperturbed ̶(̶p̶e̶r̶t̶u̶r̶b̶e̶d̶)̶ initial

conditions. Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. See Table 1, Sect.

4.1.
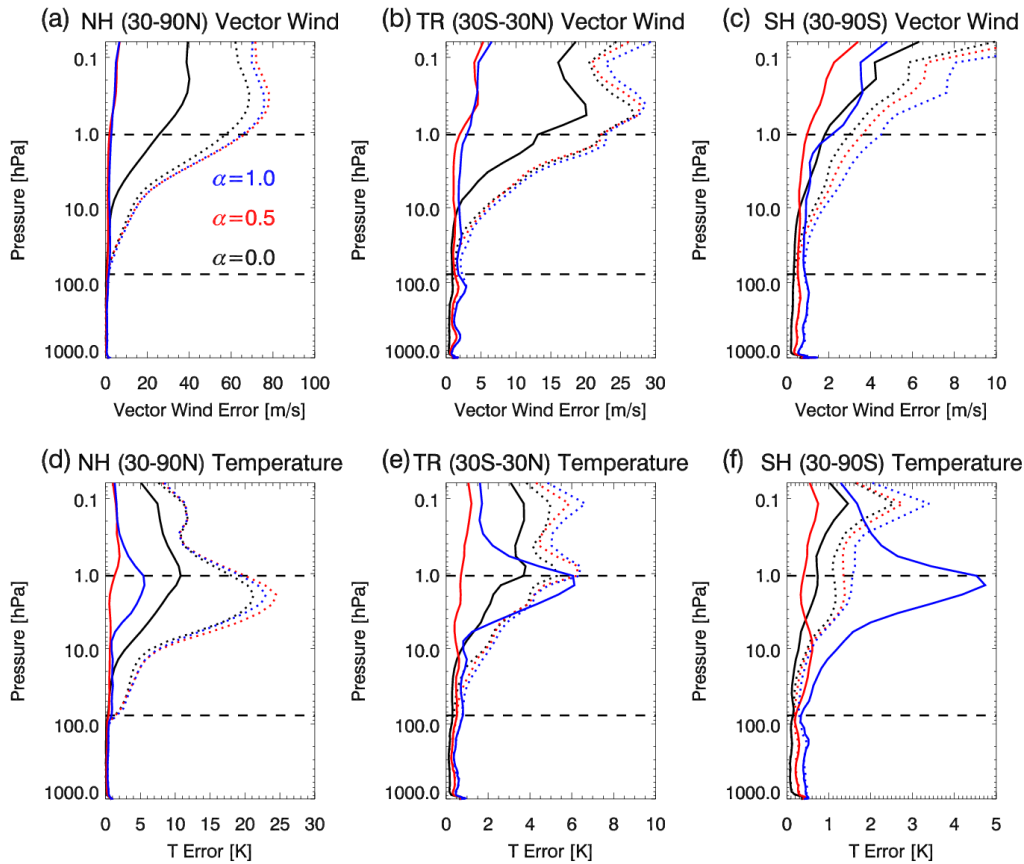
Fig. ~~10~~15. RMS vector wind errors [ms⁻¹] ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH averaged from 25 November - 1 December 2014. Solid (dotted) lines are for ozone assimilation experiments (baselines) with α = 0.0 (black), 0.5 (red), and 1.0 (blue). Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum baseline errors. See Table 1, Sect. 4.2.

48
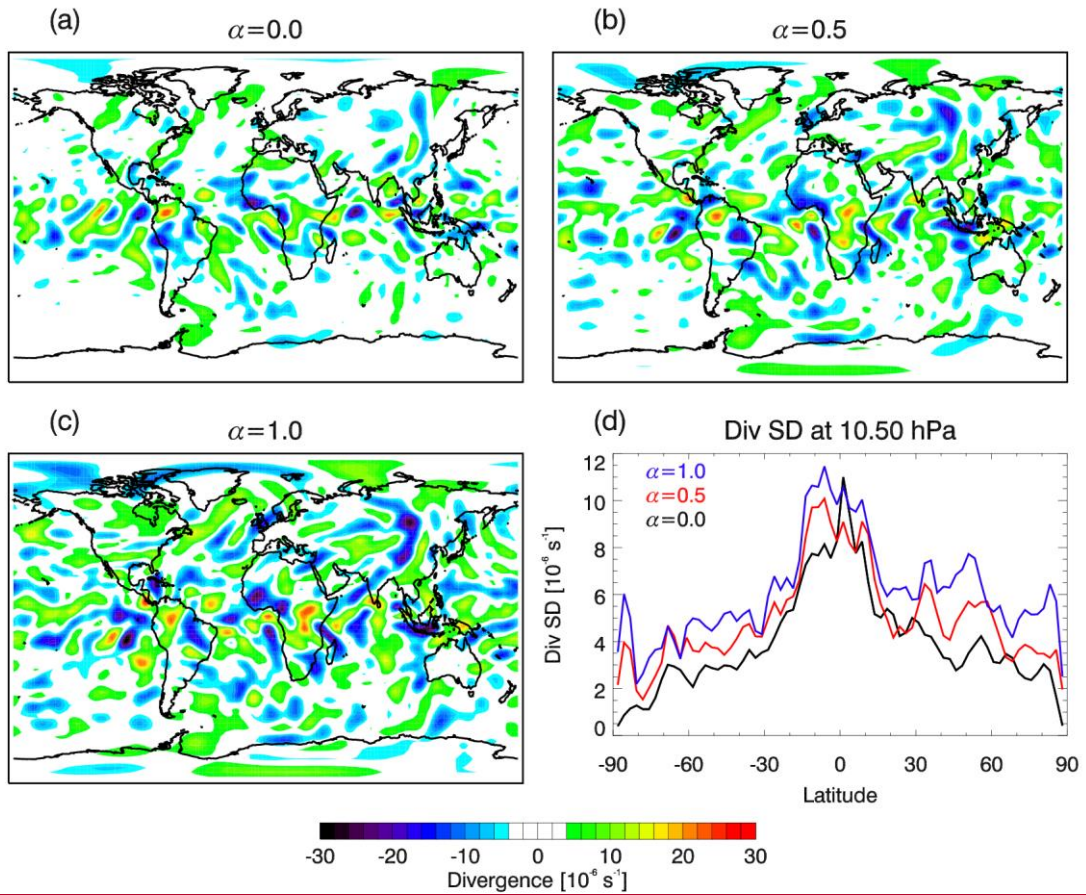
Fig. 1116. Horizontal maps at 10.5 hPa of the divergence [s⁻¹] on 1 December 2014 (00Z) for ozone assimilation experiments with (a) α = 0.0, (b) α = 0.5, and (c) α = 1.0. Colors show divergence with red (blue) indicating high (low) values. (d) The standard deviation (SD) of the divergence as a function of latitude at 10.5 hPa for 1 December 2014 (00Z). Lines are α = 0.0 (black), α = 0.5 (red), and α = 1.0 (blue). See Table 1, Sect. 4.2.

Fig. ~~12~~17. Vertical profiles of the global standard deviation of the divergence [s⁻¹] for ozone assimilation experiments on 1 December 2014 (00Z). Lines are α = 0.0 (black), α = 0.5 (red), and α = 1.0 (blue). See Table 1, Sect. 4.2.

5

Fig. ~~13~~18. RMS vector wind errors [ms⁻¹] averaged from 25 November - 1 December 2014 ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH. Solid ~~(dotted)~~ lines are for ozone assimilation experiments ~~(baselines)~~ for global: 0% error (black), random: 0% error (red), polar: 0% error (blue), and polar: 2% error (green). Dotted line is the background for the ozone assimilation experiments. Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum baseline errors. See Table 1, Sect. 4.3 and 4.4.
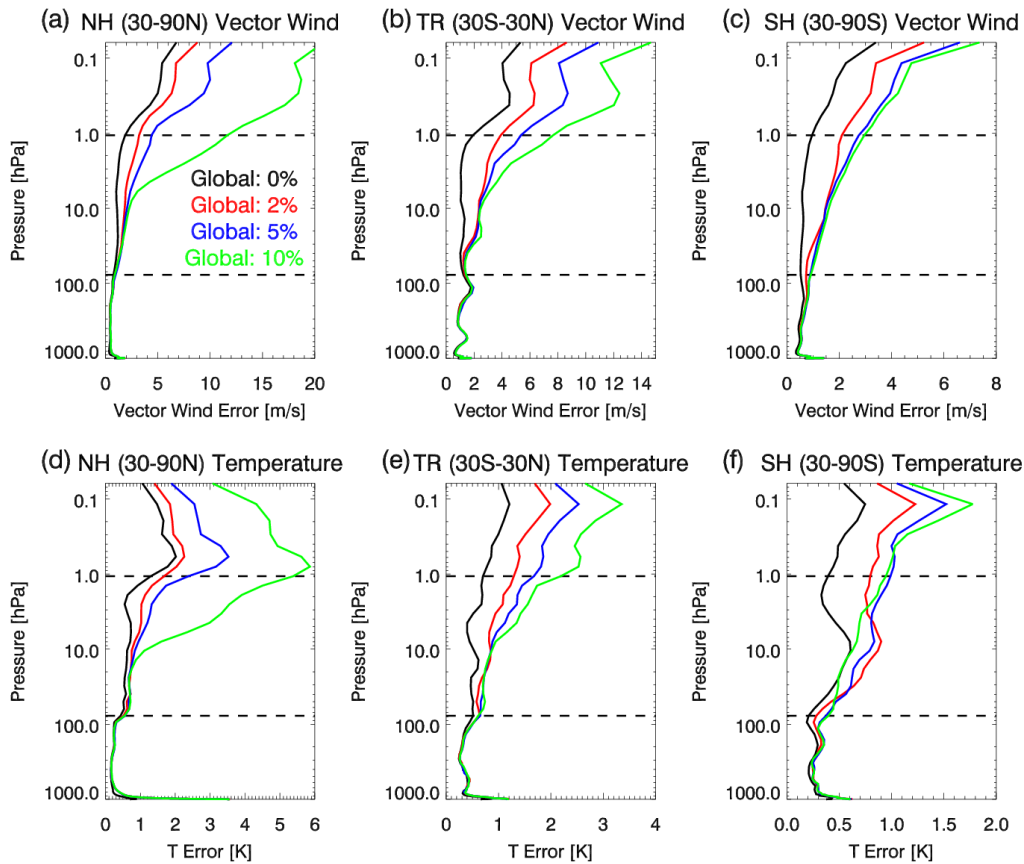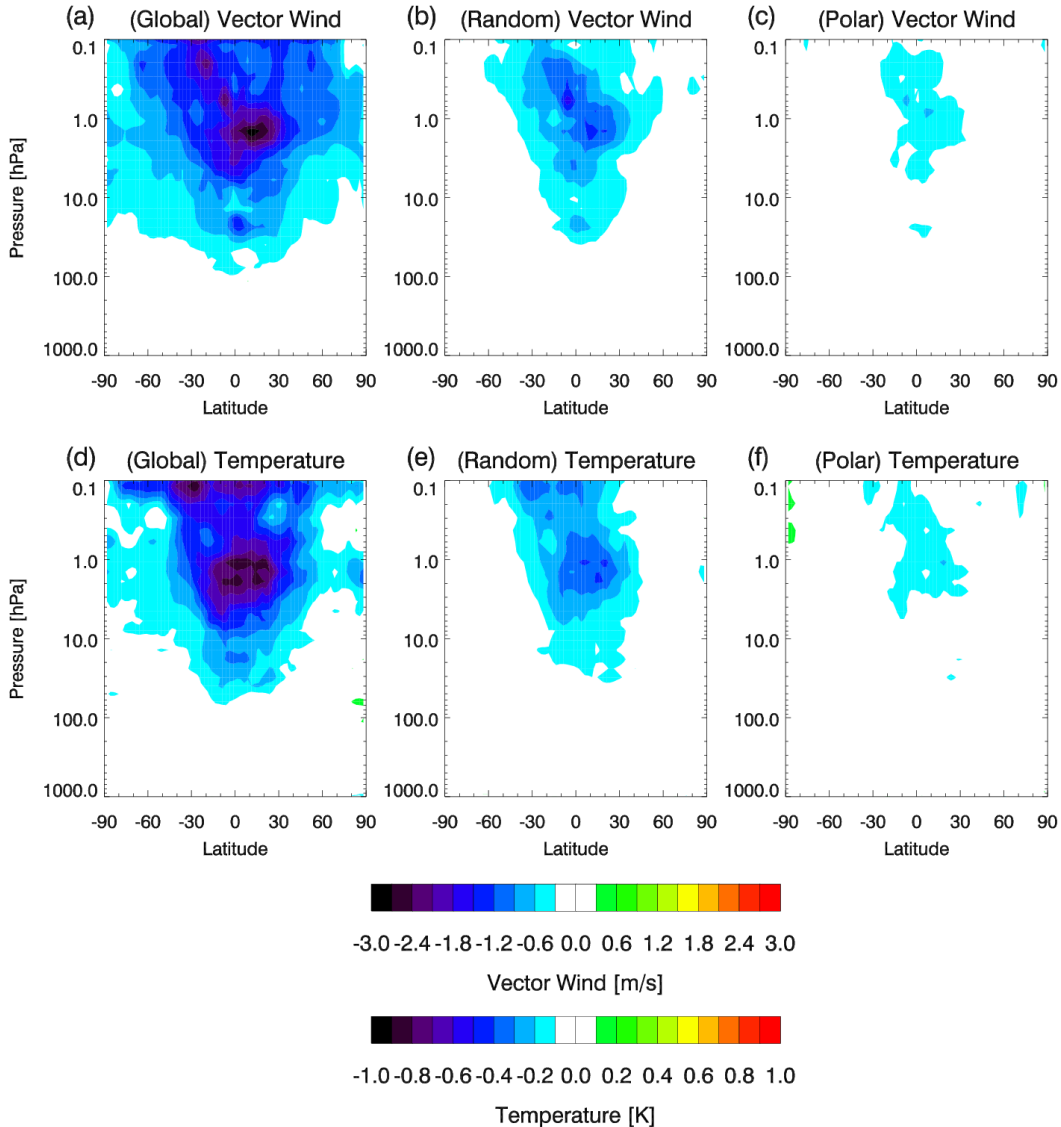
Fig. ~~14~~19. RMS vector wind errors [ms^{-1}] averaged from 25 November - 1 December 2014 ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH. ~~Solid (dotted)~~ Llines are for ozone assimilation experiments for global ozone with ~~(baselines) for~~ observation errors of 0% (black), 2% (red), 5% (blue), and 10% (green). Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum ~~baseline~~ errors. See Table 1, Sect. 4.4.
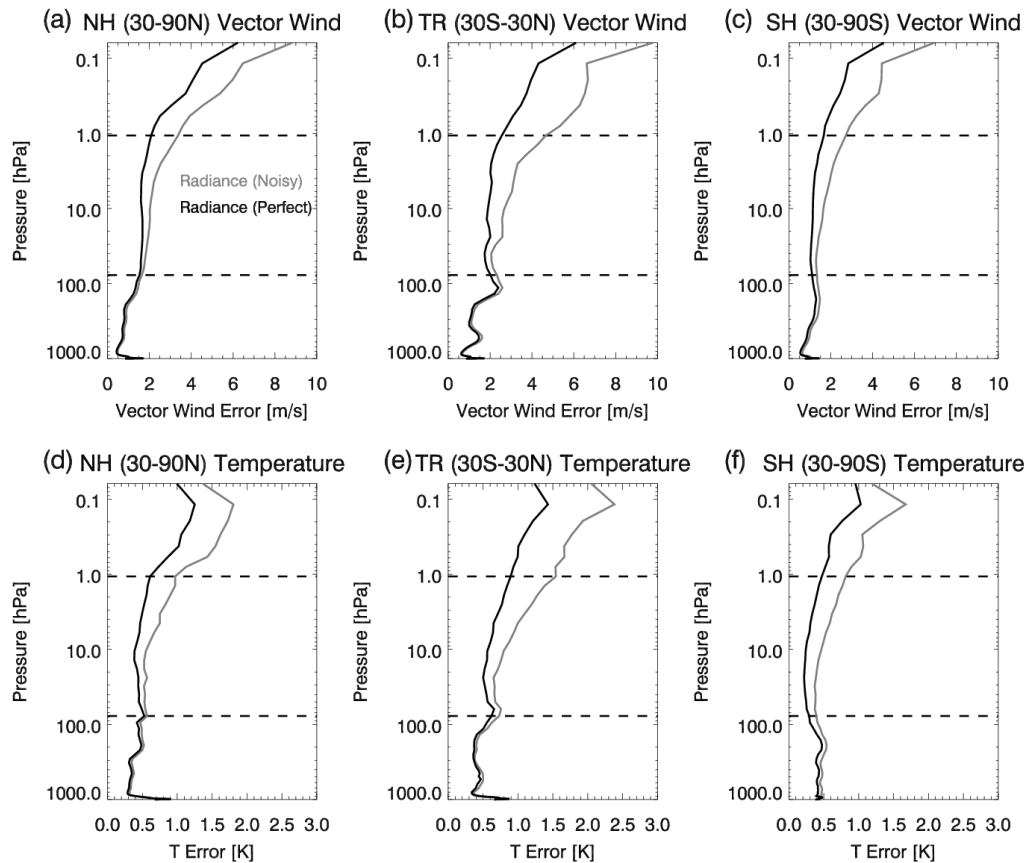
Fig. ~~15~~20. RMS vector wind errors [ms⁻¹] averaged from 25 November - 1 December 2014 ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH. Black (grey~~red~~) lines are for assimilation of perfect~~noisy~~ (noisy~~perfect~~) radiance data. ~~-~~Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum errors. See Table 1, Sect. 5.1.
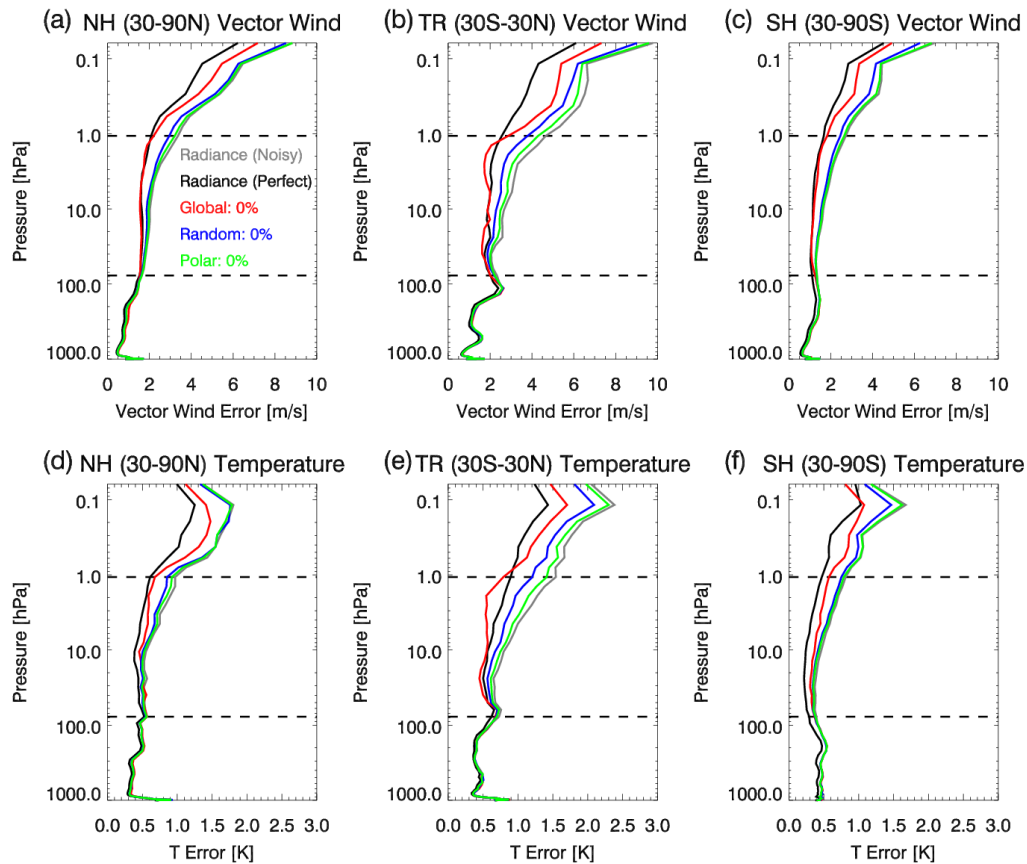
Fig. ~~16~~21. RMS vector wind errors [ms⁻¹] averaged from 25 November - 1 December 2014 ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH. Colored lines are for ozone assimilation experiments with noisy radiance and perfect ozone data using sampling pattern of global (red), random (blue), and polar (green). Black (g~~G~~rey) lines are for radiance only assimilation for ~~noisy and~~ perfect (noisy) data (same as in Fig. 20~~15~~). Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum errors. See Table 1, Sect. 5.1 and 5.2.
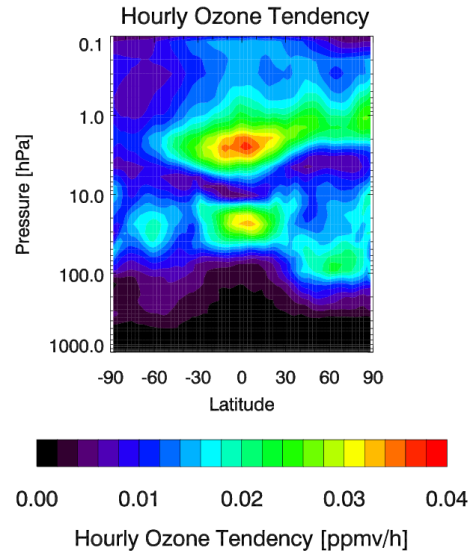
Fig. 22. Zonal mean (pressure vs. latitude) plots of the difference in RMS error between assimilation experiments with ozone and radiance data versus those with radiance alone. The results are averaged from 25 November - 1 December 2014. Wind

55

vector results [ms^{-1}] are given for perfect ozone (0% imposed error) for (a) global, (b) random, and (c) polar sampling. Temperature results [K] are given for perfect ozone (0% imposed error) for (d) global, (e) random, and (f) polar sampling. See Table 1, Sect. 5.2.

5

Fig. 23. Zonal mean (pressure vs. latitude) plots of the ozone tendency [ppmv/h] for the truth experiment, averaged over forecasts from 0-1 h, 1-2 h, 2-3 h, 3-4 h, 4-5 h, and 5-6 h for 25 November - 1 December 2014. Colors show tendency with red (blue) indicating high (low) values.
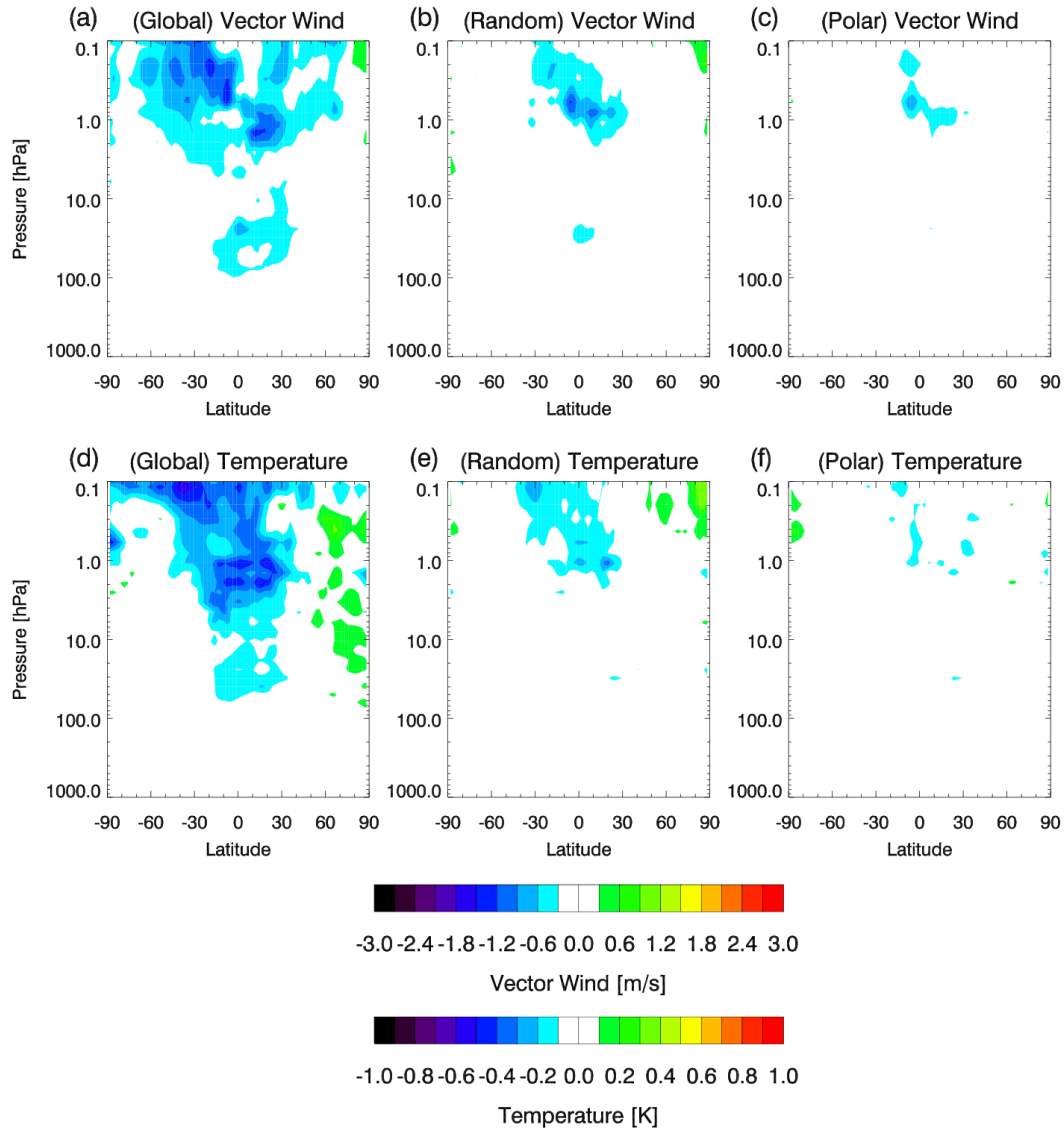
10
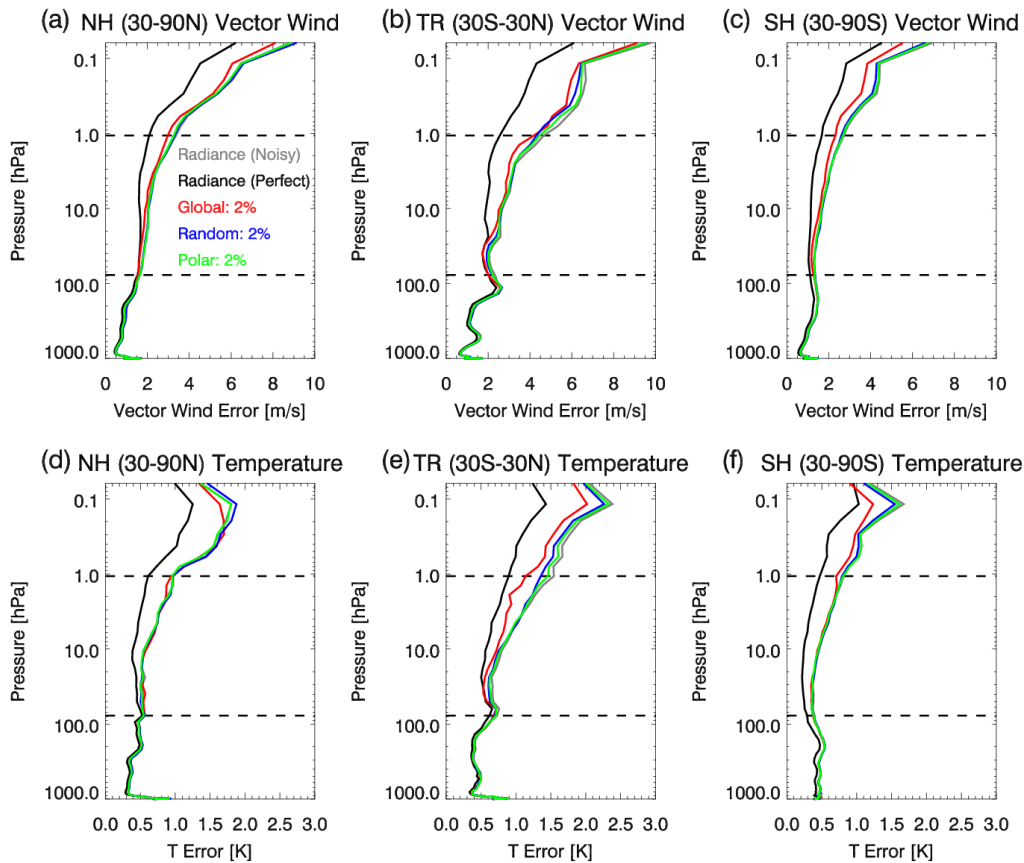
Fig. 2418. RMS vector wind errors [ms⁻¹] averaged from 25 November - 1 December 2014 ~~on 1 December 2014~~ for (a) NH, (b) TR, and (c) SH and RMS temperature errors [K] for the (d) NH, (e) TR, and (f) SH. Colored lines are for ozone assimilation experiments with noisy radiance and ozone data with 2% error using sampling pattern of global (red), random (blue), and polar (green). Grey lines are for radiance only assimilation for noisy and perfect data (same as in Fig. 2015). ~~Note that the ranges of the horizontal axes for each panel varies based on maximum errors.~~ Horizontal dashed lines indicate vertical range of assimilated stratospheric ozone observations. Note that the ranges of the horizontal axes for each panel varies based on maximum errors. See Table 1, Sect. 5.1 and 5.2.

~~**Fig. 19.** RMS vector wind error differences (with respect to assimilation of radiance only) on 1 December 2014 for (a) NH, (b) TR, and (c) SH and RMS temperature error differences for the (d) NH, (e) TR, and (f) SH. Lines are for ozone assimilation experiments with ozone data that has imposed errors (2%) and sampling pattern of global (red), random (blue),~~

Fig. 25. Zonal mean (pressure vs. latitude) plots of the difference in RMS error between assimilation experiments with ozone and radiance data versus those with radiance alone. The results are averaged from 25 November - 1 December 2014. Wind vector results [ms^{-1}] are given for noisy ozone (2% imposed error) for (a) global, (b) random, and (c) polar sampling. Temperature results [K] are given for noisy ozone (2% imposed error) for (d) global, (e) random, and (f) polar sampling. See Table 1, Sect. 5.2.