Accounting for Correlated Observation Error in a Dual-Formulation 4D Variational Data Assimilation System

WILLIAM F. CAMPBELL, ELIZABETH A. SATTERFIELD, BENJAMIN RUSTON, AND NANCY L. BAKER
Naval Research Laboratory, Monterey, California

(Manuscript received 24 June 2016, in final form 30 November 2016)

ABSTRACT

Appropriate specification of the error statistics for both observational data and short-term forecasts is necessary to produce an optimal analysis. Observation error stems from instrument error, forward model error, and error of representation. All sources of observation error, particularly error of representation, can lead to nonzero correlations. While correlated forecast error has been accounted for since the early days of atmospheric data assimilation, observation error has typically been treated as uncorrelated until relatively recently. Thinning, averaging, and/or inflation of the assigned observation error variance have been employed to compensate for unaccounted error correlations, especially for high-resolution satellite data.

In this study, the benefits of accounting for nonzero vertical (interchannel) correlation for both the Advanced Technology Microwave Satellite (ATMS) and Infrared Atmospheric Sounding Interferometer (IASI) in the NRL Atmospheric Variational Data Assimilation System-Accelerated Representer (NAVDAS-AR) are assessed. The vertical observation error covariance matrix for the ATMS and IASI instruments was estimated using the Desroziers method. The results suggest lowering the assigned error variance and introducing strong correlations, especially in the moisture-sensitive channels. Strong positive impact on forecast skill (verified against both the ECMWF analyses and high-quality radiosonde data) is shown in both the ATMS and IASI instruments. Additionally, the convergence of the iterative solver in NAVDAS-AR can be improved by small modifications to the observation error covariance matrices, resulting in further reduction in RMS error.

1. Introduction

Errors in observed data are inevitable. Data assimilation (DA) methods based on Kalman filtering, such as the NRL Atmospheric Variational Data Assimilation System-Accelerated Representer (NAVDAS-AR; Xu et al. 2005; Rosmond and Xu 2006), require the specification of error variances for each observation. To weight observations appropriately relative to the background information, such methods require error correlations. If true nonzero error correlations are ignored, the resulting analysis will suffer, as will forecasts initialized from such analyses.

Observation error itself stems from three main sources. The first source is instrument error, which is usually, but not always, uncorrelated. The second source is forward model error [e.g., interpolation, radiative transfer (RT)], which can be correlated due to, for example, errors in RT model spectroscopy. The third source is error of representation, stemming from the difference between the true atmospheric state and its discrete, smoothed representation by a numerical weather prediction (NWP) model state. All of the error variance and correlation information in Kalman filtering–based DA is represented in the observation error covariance matrix (denoted $R$). Typically, DA schemes have assumed that the observation error covariance matrix is diagonal. This assumption is especially problematic for satellite data that are likely to have correlated errors due to their high-resolution (error of representation) or introduced by the radiative transfer model. Data are typically thinned (discarded) or averaged, neither of which adequately addresses error of representation (van Leeuwen 2015), and then assigned an inflated observation error variance to partially compensate for unaccounted correlations.

Recently, several NWP centers have incorporated correlation (off diagonal) terms in $R$. Specifically, the Met Office has been using observation errors with interchannel correlations for the Infrared Atmospheric

DOI: 10.1175/MWR-D-16-0240.1

For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
Sounding Interferometer (IASI) instruments since January 2013 (Weston et al. 2014). Accounting for correlated error for IASI observations has also been explored at the European Centre for Medium-Range Weather Forecasts (ECMWF; Bormann et al. 2016), Environment Canada (S. Heilliette 2016, personal communication), and the National Centers for Environmental Prediction (NCEP; A. Collard 2015, personal communication). Stemming from these efforts, increased emphasis has been placed on methods to estimate such covariance matrices statistically.

In what follows, we assess the benefits of accounting for nonzero vertical correlation for both the Advanced Technology Microwave Satellite (ATMS) and IASI instruments in the Navy’s dual-form four-dimensional variational data assimilation (4DVar) system, NAVDAS-AR. The methodology for estimating and implementing interchannel correlation terms, and subsequent variance adjustment are discussed in section 2. Section 3 describes the model, data assimilation system, and observations. Results verified against ECMWF analysis and radiosonde data are presented in section 4. Discussion and conclusions follow in section 5.

2. Methodology

Error covariance estimation techniques built on partitioning innovation (observation minus background) statistics into observation and background error components were introduced by Rutherford (1972) and Hollingsworth and Lonnberg (1986). A diagnostic method due to Desroziers (Desroziers et al. 2005), which was originally designed as a consistency check, can be used to estimate the full observation error covariance matrix, including correlation terms. Both the Hollingsworth and Lonnberg and Desroziers methods have been used to estimate vertical (interchannel) observation error correlation for radiance observations (Stewart et al. 2009, 2014; Bormann et al. 2003; Bormann and Bauer 2010; Bormann et al. 2010), producing similar results despite the differences in assumptions and approach. The Desroziers method is the easiest to implement, and is used exclusively in this study.

The Desroziers method computes the expected value of the outer product of the analysis residual \( v_a = y - Hx_a \) (observation minus analysis) and the innovation \( v_f = y - Hx_f \) (observation minus background) to recover the observation error covariance matrix:

\[
\langle v_a v_f^T \rangle = R. \tag{1}
\]

If the DA system perfectly specifies both the background and observation error covariances (\( B \) and \( R \), respectively), it can be shown that (1) will recover the original (correct) \( R \). It is well known (Ménard 2016; Waller et al. 2016) that the Desrozier diagnostic will result in suboptimal estimation if the background and observation error covariances are not accurately specified, or if the observation and forecast errors are correlated. Desroziers et al. (2005) suggested an iterative method for converging on the true \( R \); Waller et al. (2016) showed that, although this method is subject to errors associated with prescribed covariances, a useful solution can often be obtained in a single iteration even when iterative techniques cannot be expected to converge. Stewart et al. (2013) demonstrated that it is often better to include approximate correlation structures rather than assume linear independence, even with compensating variance inflation. Based on these studies and the success of Weston et al. (2014), we implemented the Desroziers correlation computed from data using a diagonal \( R \). We also computed a single iteration by rerunning the summer case with updated observation error correlations, leaving the error variances unchanged. The Desroziers estimate from the updated run was not significantly different from the original Desroziers estimate of \( R \) (figure not shown).

The Desroziers estimates for the observation error variances of ATMS and IASI are significantly smaller than those typically used in DA systems, in some cases even smaller than instrument error estimates, which are well characterized. Using them directly leads to poor convergence [verified by testing in the Navy Global Environmental Model (NAVGEM; Hogan et al. 2014)]. There are several potential causes for this low bias, for example, departure-based quality control, and correlations between the forecast error and observation error (introduced through bias correction and/or the forward operator). Investigation of these causes is beyond the scope of the current work. Other centers confirm the need for inflating the Desroziers variances: reconditioning by Weston et al. (2014) resulted in an inflation factor for variances of approximately 1.65, and Bormann et al. (2016) inflated the Desroziers estimate by 1.75. Instead of multiplying the Desroziers variance estimates by a fixed factor, we chose to average the default operational error standard deviations with the corresponding Desroziers estimates for our initial tests. This was an ad hoc choice for our initial experiments; however, because the results were quite good, we never attempted to tune the variances. Having obtained an initial estimate of \( R \) from the Desroziers method, we need to implement it in the context of 4DVar. Note that weak-constraint 4DVar is equivalent to an extended Kalman smoother (Fisher et al. 2005).

Kalman filter–based DA systems can be cast in model space (primal form) or in observation space (dual form). Following Daley (1991), the extended Kalman filter
equations for the analysis increment can be written as follows:
\[ x_a - x_f = K(y - Hx_f), \]  
where the Kalman gain \( K \) is defined as
\[ K = BH^T(HBH^T + R)^{-1}. \]

In the dual formulation, we define a new vector \( z \) as the solution to the following linear equation:
\[ (HBH^T + R)z = (y - Hx_f). \]

Once we have solved (4) for \( z \), we postmultiply by \( BH^T \) to obtain the analysis increment:
\[ x_a - x_f = BH^Tz. \]

To solve (4) for \( z \), denote the diagonal of \( R \) by \( D \), and apply the usual preconditioning by \( D^{-1/2} \) to obtain the following:
\[ D^{-1/2}(HBH^T + R)z = D^{-1/2}(y - Hx_f). \]

Introduce a change of variable, scaling by the diagonal matrix of standard deviations,
\[ w = D^{1/2}z. \]

We can rewrite (6) as
\[ D^{-1/2}(HBH^T + R)D^{-1/2}(D^{1/2}z) = D^{-1/2}(y - Hx_f). \]

We can always write the covariance matrix \( R \) as the product of diagonal standard deviation matrices and a correlation matrix \( C \) as follows:
\[ R = D^{1/2}CD^{1/2}. \]

Combining (8) and (9) yields the following:
\[ (D^{-1/2}HBH^T D^{-1/2} + C)w = D^{-1/2}(y - Hx_f). \]

Equation (10) does not require inverting \( C \), unlike the primal formulation used by most other weather centers:
\[ (I + B^{-1/2}H^TD^{-1/2}C^{-1}D^{-1/2}HB^{-1/2})(B^{1/2}w) \]
\[ = B^{-1/2}H^TD^{-1/2}C^{-1}D^{-1/2}(y - Hx_f). \]

Note the presence of \( C^{-1} \) in (11). When \( C \) is block diagonal with small enough blocks, or has some other easily invertible form, inverting it is a minor issue. If that is not the case (e.g., if an instrument has thousands of correlated channels, or if we try to account for spatial correlations), then the dual formulation (10) has a significant computational advantage. In the next section, we will describe the practical implementation of interchannel error correlations for ATMS and IASI in the NAVGEM system.

3. Description of model and data

a. The NAVGEM model and NAVDAS-AR data assimilation system

The Navy global NWP system consists principally of NAVGEM and the embedded 4DVar NAVDAS-AR. NAVGEM is the Navy’s high-resolution global weather prediction system with an advanced semi-Lagrangian/semi-implicit (SL/SI) dynamical core run, which is run operationally at the Fleet Numerical Meteorological and Oceanographic Center. NAVGEM v1.3, which has 31-km horizontal resolution (spectral triangular truncation at wavenumber 425, i.e., T425) and 60 vertical levels with an effective model top of 0.04 hPa (Hogan et al. 2014), was used for all experiments reported here.

NAVDAS-AR is the observation space (dual form) 4DVar system that is currently the Navy’s operational global DA component of NAVGEM. NAVDAS-AR ingests over 100 million observations for every 6-h DA window; after quality control and data thinning, approximately 3 million observations are used to create the final analysis. Routinely assimilated observations include the following: conventional observations (e.g., radiosondes, dropsondes, buoys, etc.), aircraft observations, feature tracked winds, bending angles from Global Navigation Satellite System Radio Occultation (GNSS-RO), SSMIS and WindSat total precipitable water and wind speeds, scatterometer winds, synthetic observations of tropical storms, and radiances from AMSU-A, MHS, SSMIS, AIRS, IASI, and ATMS. All radiance observations go through a variety of quality control checks, spatial thinning, and variational bias correction. In the present study, we implemented interchannel correlated observation error for both the IASI and ATMS instruments, while keeping a diagonal \( R \) for all other observation types.

b. IASI and ATMS observations and default error specifications

IASI is one of the most important instruments assimilated by global weather forecast models (Hilton et al. 2012). Currently there are two IASI instruments flying on board the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)
MetOp-A and MetOp-B satellites, launched in October 2006 and September 2012, respectively. A third and final first-generation IASI will be launched on MetOp-C in 2018. IASI is a Michelson interferometer designed for high-resolution sounding of the atmosphere, with accuracy requirements of 1 K for temperature and 10% for humidity (Cayla 2001; Chalon et al. 2001). NAVDAS-AR has assimilated IASI data since 2009, and currently assimilates 55 channels in the longwave CO2 band over ocean, 24 of these channels are over land and sea ice. In addition, NAVDAS-AR assimilates 18 humidity channels over ocean. Operationally, only channels assumed above cloud top [as determined by a hyperspectral cloud detection; McNally and Watts (2003)] are assimilated. IASI data is then thinned to approximately 110-km resolution. Prior to considering correlated observation error, the temperature sounding channels were assigned error standard deviations of 0.4–0.6 K, and the moisture channels were assigned significantly higher error standard deviations of 2.5–5.0 K (Fig. 1a). These error standard deviations were adopted from those used successfully with the heritage AMSU-A and MHS instruments.

Both IASI and ATMS consistently show a large positive impact to global NWP in the NAVGEM/NAVDAS-AR system as measured by the 24-h total error norm metric (Langland and Baker 2004). The choice of IASI for this study is made because of the known error correlation that exists between the water vapor channels (Stewart et al. 2009) and because of its prior successful use by Weston et al. (2014). For the ATMS, the impact is similar to that of a combined AMSU-A/MHS sensor suite from the NOAA satellite series. The choice of ATMS for this study is made in part because of the discovery that it has correlated instrument error (Doherty et al. 2012) not seen in the heritage AMSU-A/MHS instruments, but also because of the importance of the instrument to NWP in its own right.

4. Experimental results and discussion

We performed a suite of experiments with the NAVGEM system to test the use of correlated interchannel observation error in assimilating ATMS and IASI data. To obtain an estimate of the observation error covariance matrix for the ATMS and IASI instruments, we applied the Desroziers method to an archive of satellite and NAVGEM model data. The data in the archive ranged from 0000 UTC 1 July to 1200 UTC 30 August 2013 and we considered observations only at 0000 and at 1200 UTC. To apply the Desroziers diagnostic, we used data that had passed the operational

![Fig. 1](http://journals.ametsoc.org/doi/pdf/10.1175/MWR-D-16-0240.1)
quality control and then accepted an observation only if the corresponding full sounding was available. Although we used two months of data for our estimates, previous work (e.g., Weston et al. 2014) has shown that a single day of data is sufficient to obtain robust diagnostics and that there is little daily or seasonal variability. We used the same matrices for both the winter and summer experiments detailed below.

For each experiment, we considered two seasons: summer (1 July–31 August 2013) and winter (1 November 2014–31 January 2015). All runs used the full operational suite of conventional and satellite observations; the winter runs also included the MHS, AMSU-A, and IASI observations from MetOp-B, which were unavailable for the summer runs. The summer runs averaged 2.6 million observations per 6-h DA window; the winter runs had an extra half million observations, most of them due to the inclusion of MetOp-B. The control runs used the operational observation error variances from NAVGEM. We evaluated the results of the experiments with paired comparisons of the time series of zonal and meridional wind, temperature, relative humidity, and geopotential height as a function of region [Northern Hemisphere (20°–80°N), Southern Hemisphere (20°–80°S), and tropics (20°S–20°N)], pressure level, and forecast lead time out to 5 days. Experiments were verified against both ECMWF analyses and conventional observations from a set of 400 high-quality radiosondes. Statistical significance was tested with a predefined set of thresholds and calculated taking into account the serial autocorrelation of consecutive analyses (Wilks 1995, 125–129). (All of the colored boxes in Figs. 4–8 are statistically significant at the 95% level or higher.)

a. Desroziers covariance estimates and adjustments

As discussed in the previous section, the Desroziers diagnostic was computed from two months of NAVGEM output (innovations and analysis residuals), yielding estimates for the interchannel error correlations for the IASI (Fig. 2a) and ATMS (Fig. 2b) channels assimilated by NAVGEM. Note that spatial correlations are not accounted for here, only interchannel (vertical) correlations. The upper-left quadrants of Fig. 2 show the correlations between temperature sensitive channels, from above 100 hPa to the surface, while the lower-right quadrants show the correlations between the moisture sensitive channels from the surface to ~300 hPa. The other quadrants show the correlations between temperature and moisture sensitive channels. The Desroziers diagnostic also yielded estimated error standard deviations for IASI and ATMS, which are shown in Fig. 1 (dashed blue lines). As mentioned in the previous section, these error standard deviations were smaller than estimates of pure instrument noise; therefore, we did not use them directly in any experiments. Instead, we used either the default NAVGEM error standard deviations (Fig. 1, solid red curves) or reduced error standard deviations (Fig. 1, solid green curves), which are a compromise between the default NAVGEM and Desroziers diagnosed error standard deviations. The next section will discuss how the inclusion of correlated error affects the convergence of the conjugate gradient solver in NAVDAS-AR.

b. Condition number, convergence, and observation error covariance reconditioning

The condition number of a symmetric, positive definite matrix is given by the ratio of largest to smallest eigenvalue, and gives an upper bound on the convergence rate of the conjugate gradient solver (Golub and van Loan 1996). [It should be noted that the convergence rate also depends on details of the entire eigenstructure of the given matrix (Ramage 1992)]. In dual space systems such as NAVGEM, the condition number that matters is that of the sum of the scaled representer matrix and the observation error correlation matrix:

\[ D^{-1/2} \mathbf{H} \mathbf{B} \mathbf{H}^T D^{-1/2} + \mathbf{C} \]  

(12)

Note that \( \mathbf{C} \) is primarily the identity matrix, with blocks \( \mathbf{D}^{1/2} \mathbf{H} \mathbf{B} \mathbf{H}^T \mathbf{D}^{-1/2} \) of the norm covariance (12).

We used two methods of covariance matrix reconditioning. After performing an eigenvector decomposition of the original poorly conditioned matrix, the eigenvalue spectrum was adjusted such that the ratio of the largest new eigenvalue to the smallest new eigenvalue was reduced, thus reducing the condition number. The first reduction method changes only the smallest eigenvalue(s), by increasing the value of the smallest eigenvalue until the desired condition number is achieved. To keep the eigenvalues ordered, often other small eigenvalues must also be increased. This procedure approximates a Ky Fan \( p-k \) norm covariance adjustment method (Tanaka and Nakata 2014). The second reduction method adds a positive constant to all of the eigenvalues until the desired condition number is
achieved. This procedure approximates the method of Steinian linear shrinkage (Ledoit and Wolf 2004). In the experiments that follow, we adopt a Ky Fan reconditioning. Additionally, we briefly examine the sensitivity of the results to the choice of reconditioning in section 4d; a more detailed comparison of covariance reconditioning methods and their properties in a dual system is left to a future paper.

Another possible method to speed up convergence of the solver is by preconditioning. Modify (7) to be

\[ w = R^{-1/2} z. \]  

(13)

Then (10) becomes

\[ (R^{-1/2}HBH^T R^{-1/2} + I)w = R^{-1/2} (y - Hx_f), \]  

(14)

which is preconditioned by the inverse square root of the observation error covariance matrix. This formulation is guaranteed to have a minimum eigenvalue of 1, but it may be that the maximum eigenvalue increases by a large amount and yields no improvement in condition number. It has the same drawback of the primal formulation, in that the inverse of \( R \) must be computed at every observation location. For these reasons, and the substantial work required for implementation of this formulation, we chose to leave exploration of this approach to the future. Of course, reconditioning and preconditioning are not mutually exclusive, and combined may yield the best results.

c. Experiments using correlated observation error

All of the experiments discussed below used reconditioned correlation matrices, with condition numbers of 18 for ATMS and 169 for IASI chosen based on convergence experiments for 1 July 2013. We tested the impact of both including correlated observation error and adjusting variances for the ATMS and IASI instruments, versus a control run with default NAVGEM variances and no correlated observation error. We performed two

---

**Fig. 2.** Correlations estimated with the Desroziers method for NAVGEM assimilated channels of (a) IASI and (b) ATMS. Corresponding additive (Weston) reconditioned matrices (c) for IASI and (d) for ATMS used in experiments detailed in section 4d.
experiments: 1) Ky Fan reconditioned correlations with default NAVGEM observation error variances; and 2) Ky Fan reconditioned correlations with reduced observation error variances, formed by averaging the default NAVGEM observation error standard deviations and the (significantly smaller) Desroziers diagnosed error standard deviations.

Figures 4–6 show results for three regions: Northern Hemisphere (NH) (20°–80°N), tropics (20°S–20°N), and Southern Hemisphere (SH) (20°–80°S). In each figure, we used analyses from the ECMWF, taken from the THORPEX Interactive Grand Global Ensemble (TIGGE) archive (Richardson 2005), as verification. These analyses were interpolated (bilinearly) to 1° × 1° resolution to match the NAVGEM model output. Figures 4–6 illustrate the impact of the inclusion of correlations with default variances (left column) and with reduced variances (right column). Each column includes four panels, which display the percentage change of the RMS error of temperature, vector wind, geopotential height, and precipitable water averaged over the two-month experiment. For brevity, we present only results for the summer period and comment on results for the winter period where relevant. Generally, we see less impact for the winter period, likely owing to the increased number of observations due to the inclusion of Metop-B for that run. Another possibility is that the Desroziers estimates were derived from summertime data, but used in winter. Red shading on the plots corresponds to lower RMS error for the experiment than for the control run, while blue shading is the opposite, and white is approximately neutral. All of the colored boxes indicate statistically significant results at the 95% confidence level.

Figure 4 shows the results for the Northern Hemisphere (20°–80°N). For temperature, both experiments showed broad improvement over the control below 200 hPa. Reducing the default variances led to further reduction in RMS error, most notably at the 925–850-hPa layer at early lead times; however, gains can be seen up to 200 hPa and out to a lead time of 5 days. The positive impacts were weaker in the winter run, although still significant at 850 hPa, with some negative impact in the 250–200-hPa layer at a short lead time (not shown). The vector wind results were less compelling for inclusion of correlation alone; however, once the default variances were reduced, broader impacts were seen at early lead times. (The vector wind results for the winter period were largely neutral.) The results for geopotential height showed improvement above the boundary layer out to 5 days, with enhanced error reduction in the reduced variance experiment. Similar results were seen for the winter season. For precipitable water, there was better agreement with analyses in the boundary layer out to 5 days, but in this case, the reduced variance experiment showed a short-term degradation between 500 and 700 hPa.

Figure 5 shows the results for the tropics (20°S–20°N). In the tropics, for temperature, vector wind, and geopotential height, both experiments consistently outperformed the control at virtually every level and lead time (except 200 hPa for temperature, and 50 and 1000 hPa for geopotential height). The reduction of default variances generally led to additional gains in forecast skill, most notably for temperature. Similar results were seen for the winter period. Results for all variables are similar to the NH results but more pronounced, both positive and negative. As in the NH case, reducing the default variances led to further degradation in the midtroposphere for total precipitable water, the causes of which are unknown to us at this time.

In general, we expect satellite data to have a greater impact on the analyses in the tropics and the SH than in the NH, so improvements in the use of that data, such as accounting for correlated error, could potentially make a bigger difference. Comparing Figs. 4, 5, and 6 shows this to be true for both experiments, with the exception of total precipitable water: forecast skill gains are greater at nearly every level and lead time for the tropics and SH. In addition, in the tropics and SH reducing the default variances leads to further gains in nearly every measure, again excepting total precipitable water. Similar results are seen for the winter period, except for some negative impacts for 1000 hPa and in the 250–200-hPa layer.

While the ECMWF analyses provide easily accessible continuous fields for verification, one may question
whether these analyses provide an independent verification, because each center assimilates similar sets of observations and relies on similar physical parameterizations, numerical schemes, etc. One may further question the sensitivity of the results presented here to the choice of verification data. For these reasons, we also present verification versus a set of 400 high-quality global radiosondes (Fig. 7). The disadvantage of using radiosondes as verification is that they are sparse and biased toward the NH over land; however, radiosonde measurements are available at higher levels than TIGGE analyses. Typically, the results seen for radiosonde verification were less compelling. For temperature, there were some slight improvements seen in the
troposphere beyond analysis time, with degradation in the 200–150-hPa layer. (The ECMWF results shown in Fig. 4 also showed such degradation.) Vector winds are in slightly better agreement with radiosondes at upper levels, but in contrast with Fig. 4, reducing default variances provides no benefit. Geopotential heights are in better agreement with radiosondes in the troposphere out through 4 days in both experiments, with mixed results at 5 days. Relative humidity impacts are largely neutral.

To isolate the impact of reducing variances, we repeated the experiment in which the default variances were reduced while excluding correlations (figures not shown). Overall, the results were slightly positive, indicating that the error variances currently used in NAVGEM could be improved. However, we generally saw a much smaller impact than that shown in Figs. 4–7, highlighting the importance of accounting for correlated error.

d. Sensitivity to reconditioning method

To assess the sensitivity of our results to the type of reconditioning, we repeated the second experiment, in which correlations were included and the default variances were reduced, with additive reconditioning replacing Ky Fan reconditioning. The condition number was kept constant by adding 0.218 to the diagonal of the ATMS block, and 0.0694 to the diagonal of the IASI

![Figure 5](http://example.com/fig5.png)

**Fig. 5.** As in Fig. 4, but for the tropics (20°S–20°N).
block. In Fig. 8, the second experiment was treated as the control, so red (blue) shades indicate an improvement (degradation) due to additive reconditioning. The left panel shows results with ECMWF verification; the right panel shows results with global radiosonde verification (as in Fig. 7). Overall, the sensitivity to the choice of reconditioning method was small. Results in temperature were mixed, with Ky Fan (eigenvalue) reconditioning performing better at short lead times in the boundary layer. For vector wind, additive reconditioning showed a benefit at 5 days. Geopotential heights also showed broad improvement at 4 and 5 days for additive reconditioning, with some scattered degradation at early lead times and above the tropopause. However, there was a marked difference in the convergence of the solver, despite identical condition numbers. The control run (no correlated error, default NAVGEM variances) took 56 (±5) iterations to converge. The experiment including correlations with default variances and Ky Fan reconditioning took 78 (±7) iterations; however, when variances were reduced, the convergence was significantly slower, taking 104 (±10) iterations. The experiment with additive reconditioning replacing Ky Fan converged in only 87 (±6) iterations. At least in our system, the computational benefits

Fig. 6. As in Fig. 4, but for the Southern Hemisphere (20°–80°S).
of additive reconditioning over Ky Fan reconditioning outweigh the slightly better forecast performance.

5. Summary and conclusions

We have shown that incorporating correlated observation error for the ATMS and IASI instruments in NAVDAS-AR leads to substantial gains in forecast skill, verified both against independent ECMWF analyses and against high-quality radiosondes. The reductions in RMS error versus the control were statistically significant across many variables, levels, regions, and lead times. Although the benefits were greatest for the summer period, error reductions were still substantial for the winter experiment. Benefits were realized by including correlations without the reduction of variance, although reducing variances typically led to further improvement. An experiment using only reduced variances without correlated error confirmed that the bulk of the benefit came from the correlations rather than the variances. Based on these results, we plan to implement correlated observation error for other instruments assimilated in NAVDAS-AR, most
notably the Cross-Track Infrared Sounder (CrIS), which is similar to IASI, and the AMSU-A/MHS instruments, which are similar to the ATMS.

The Desroziers diagnostic provides an easily computed observation error covariance estimate. Examination of the diagnosed correlation matrices showed that moisture channels in particular have highly correlated error, resulting from the forward model, error of representation, and forecast model bias among other sources. Temperature channel error correlations are much lower, with the exception of some of the ATMS channels, due to correlated instrument error (Bormann et al. 2013). Despite the known shortcomings of the Desroziers technique, small modifications to the computed correlation matrix along with judicious variance inflation yielded large positive impact on forecasts. Our error variances were in line with those used by Weston (2014) and the best values found by Bormann (2016); nevertheless, more effort put into variance tuning

**Fig. 8.** Additive reconditioning (experiment) vs eigenvalue reconditioning (control) verified against (left) NH ECMWF analyses and (right) global radiosondes. Red shading indicates improvement over the control. Colored boxes are all statistically significant at the 95% level or higher.
directly \([\text{matrix–vector multiplication, } \mathcal{O}(n^3)]\) of covariance matrices, which will be needed in the future to account for correlation among thousands of channels instead of hundreds, or to account for horizontal correlation. The condition number of these matrices will continue to be important, and for practical use, some form of reconditioning must be utilized. Even with identical condition numbers, the number of iterations required by the solver can vary substantially; for example, in our studies, additive reconditioning of the ATMS and IASI correlation matrices required only 85% of the number of iterations needed for the Ky Fan reconditioned matrices. Finding the best condition numbers and reconditioning methods for these matrices is a topic for future work.

Several other outstanding issues require further study. Quality control checks that are based on innovations should be altered to consider correlated error. Increasing the forecast model resolution or changing the model physics affects the error of representation, which will require updating \(\mathbf{R}\). Modifying \(\mathbf{R}\) influences many aspects of data assimilation; in particular, it changes the ratio of \(\mathbf{B}\) to \(\mathbf{R}\), which determines how closely the analysis draws to observations. In addition, changes to \(\mathbf{R}\) for ATMS and IASI affect the relative weight of other observations, and could possibly lead to degraded forecasts if the prescribed observation error variances for other observations are suboptimal. Any future changes to the structure of \(\mathbf{B}\) (e.g., going to a hybrid ensemble 4DVar) may reduce the positive impact seen here, and will require recomputing the Desroziers diagnostic, retuning the error variances, reconditioning the correlation matrices, and reassessing forecast results.

Acknowledgments. We thank Dr. Daniel Hodys for helpful discussions on various aspects of this work, especially the error of representation, as well as the anonymous reviewers for their thoughtful comments that helped us greatly improve our manuscript. This research is supported by the Chief of Naval Research through the NRL Base Program, PE 0601153N. The ECMWF forecasts were obtained from the THORPEX Interactive Grand Global Ensemble (TIGGE) data portal at ECMWF.

REFERENCES


Langland, R., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data...


