

Use of an OSSE to evaluate background-error covariances estimated by the NMC method

Ronald M. Errico,^{a,b,*} Nikki C. Privé^{a,b} and Wei Gu^{c,b}

^aGoddard Earth Sciences Technology and Research Center, Morgan State University, Baltimore, MD, USA

^bGlobal Modeling and Assimilation Office, NASA, Greenbelt, MD, USA

^cScience Systems and Applications, Inc., Greenbelt, MD, USA

*Correspondence to: R. M. Errico, Global Modeling and Assimilation Office, NASA/GSFC Code 610.1, Greenbelt, MD 20771, USA.

E-mail: ronald.m.errico@nasa.gov

This article is US Government work and is in the public domain in the USA.

The NMC method has proven utility for prescribing approximate background-error covariances required by variational data assimilation systems. Here, untuned NMC method estimates are compared with explicitly determined error covariances produced within an OSSE context by exploiting availability of the true simulated states. Such a comparison provides insights into what kind of rescaling is required to render the NMC method estimates usable. It is shown that rescaling of variances and directional correlation lengths depends greatly on both pressure and latitude. In particular, some scaling coefficients appropriate in the Tropics are the reciprocal of those in the Extratropics. Also, the degree of dynamic balance is grossly overestimated by the NMC method. These results agree with previous examinations of the NMC method which used ensembles as an alternative for estimating background-error statistics.

Key Words: OSSE; data assimilation; background error; NMC method

Received 27 August 2013; Revised 26 March 2014; Accepted 10 April 2014; Published online in Wiley Online Library

1. Introduction

Modern data assimilation systems (DASs) require estimates of background-error covariances to determine an approximately statistically optimal analysis given imperfect and incomplete observations (Tarantola, 1987; Courtier, 1997). Such covariances cannot be determined from differences between fields of background and truth, since datasets representing truth, especially globally, generally are not much more accurate than a good DAS background. More indirect methods are therefore required for computing the desired covariances. Early methods generally relied on assumptions that statistics estimated for a data-rich region (e.g. as in Hollingsworth and Lönnberg, 1986) or from longer-term forecast error (e.g. as in Parrish and Derber, 1992) provide good starting estimates for the required statistics. Most recently, ensemble techniques (e.g. Bonavita *et al.*, 2012) have been used for this purpose.

The NMC method was introduced at the US National Meteorological Center (now replaced by the National Centers for Environmental Prediction, NCEP) for estimating background-error variances in terms of spherical harmonic coefficients (Parrish and Derber, 1992). It began with global fields of differences between 24 h forecasts and corresponding verifying analyses projected onto the harmonics. Variances of the coefficients determined over a month were then rescaled, by manual tuning, to reflect a 6 h rather than 24 h forecast error. Later, differences between pairs of forecasts valid at the same time (e.g. 48 and 24 h forecasts) were similarly used. One reason

for using such a long lag was to mitigate diurnal signals which otherwise appeared present.

Many reasons can be cited for disclaiming that the statistics determined by the NMC method, prior to tuning, provide a good estimate for what is required. One obvious example is that, if some regions are poorly observed, the differences between lagged forecasts or forecasts and corresponding analysis will tend to be small since such differences vanish in the absence of observations, whereas the background error should tend to be large without observations to constrain them. Other reasons are offered in the following or appear in the literature to be cited. Nevertheless, the background-error covariance models derived using the NMC method have been shown to be useful, and the technique is still used at some operational centres, including the Global Modeling and Assimilation Office (GMAO) and NCEP (Kleist *et al.*, 2009). However, this utility does not imply that other techniques cannot produce much better and much more useful characterizations of the required statistics.

The use of powerful adjoint model tools (Errico, 1997a) has changed our understanding of how forecast errors grow early in a forecast (Langland *et al.*, 1995; Gelaro *et al.*, 2002). In particular, singular vector (SV) analysis has revealed that the structures which grow the greatest as measured by error energy norms (Talagrand, 1981; Errico, 2000a) begin as structures with small vertical and horizontal scale but quickly attain larger scales, becoming almost barotropic throughout the troposphere after 12 h (Farrell, 1989; Palmer *et al.*, 1998). In the Extratropics, they also change in the character of their geostrophy (Errico, 2000b). This suggests that

not only the variances but also correlations may be very different for short-term forecast errors and differences of forecasts at longer time periods. In particular, background errors may be much less barotropic and geostrophic than those inferred from the NMC method.

Revealing theoretical and experimental examinations of the reliability of the NMC method to characterize background error are provided in Berre *et al.* (2006). It shows that there are three mechanisms in the method that can lead to radically different statistics compared to what is desired. One is the use of a longer forecast period than the relevant 6 h, as previously described. Another is the accumulation of several analysis increments used to produce the initial condition for the shorter forecast. Third are the potentially very different natures of analysis increments and analysis errors: the analysis increment may be considered as a response to a low-pass filter whereas the analysis error responds to a high-pass filter. Berre *et al.* (2006) should be consulted for elaboration of these properties. The profound impact of these mechanisms is revealed experimentally in the same reference and in Belo Pereira and Berre (2006) by using an ensemble to make a different, more direct, estimate of background error (but without consideration of model error).

At NASA's GMAO, a framework for carefully validating and conducting an observing system simulation experiment (OSSE) has been developed (Errico *et al.*, 2013; Privé *et al.*, 2013a). Within that framework, background error can be precisely calculated because the truth is known explicitly. Covariances inferred from these errors apply to assimilation for the real atmosphere to the degree that the OSSE validates. Although it is not possible to explicitly validate the OSSE with respect to background-error statistics since these are explicitly unavailable in the real context, the GMAO OSSE has been validated with respect to many other metrics in the two cited references. These demonstrate that the OSSE framework performs realistically in the troposphere except for a small but significant diminution of effects of forecast-model error. Thus, OSSE background error may be reasonably considered as representative of realistic background error, albeit with some caution.

In this article, some variances and correlations are determined for both explicit 6 h background errors and subsequent 48 h minus 24 h forecasts verifying at the same time. The GMAO OSSE is briefly described in section 2. Results are presented in section 3. Their implications are described in section 4.

2. The GMAO OSSE

The general GMAO OSSE development and validation efforts are described in Errico *et al.* (2013) and Privé *et al.* (2013a). Some improvements and the use of the OSSE to estimate analysis error are presented in Errico and Privé (2014). Only a brief summary of the most current OSSE framework is reported here.

The simulation of 'truth' for the OSSE is provided by a nature run (NR) produced by the European Centre for Medium Range Weather Forecasts (ECMWF) using version cy31r1 (ca. 2005) of its forecast model. It is a 'free-running' solution in the sense that its only constraint regarding the real weather specifically occurring during its year-long period of simulation was through its prescribed sea surface temperature, sea ice fraction, and initial conditions. The latter constraint, however, is effectively 'forgotten' due to the chaotic atmospheric dynamics (Lorenz, 1963). The NR fields are specified on 91 atmospheric levels, in addition to the surface, on a reduced Gaussian grid (Hortal and Simmons, 1991) having 512 latitudes and between 18 and 1024 longitudes, with the smaller numbers near the poles.

Although the ECMWF NR for these months was produced using analyzed sea surface temperature and ice fractions for 2005, the observation locations and types considered by the current OSSE are from the corresponding months of 2011, unlike for the original GMAO OSSE described in Errico *et al.* (2013). They are drawn from the NR as described in Errico *et al.*

(2013) and Errico and Privé (2014). Also included are Global Positioning System radio occultation (GPSRO) observations. These are produced by interpolating NR fields to the planes defined by the locations, azimuths, and times of corresponding real observations for the assimilation period and then generating bending angles using the two-dimensional mode of the Radio Occultation Processing Package (ROPP version 6.0; Healy *et al.*, 2007; Culverwell *et al.*, 2013). This observation operator differs from the one-dimensional observation operator used in the GMAO DAS which is described in Cucurull *et al.* (2013). This difference in operators serves as one source of partial observation representativeness error in the OSSE.

Random errors are added to all observations as described in Errico *et al.* (2013) and Errico and Privé (2014) to account for otherwise missing instrument or processing errors and deficient representativeness errors. As such, random vertically correlated errors are added to the GPSRO observations. The error variances are tuned so that innovations for OSSE and real DAS observations at similar elevations have matching variances. The vertical correlation is defined by a Gaussian-shaped function such that the correlation drops to 0.1 at a distance of 500 m.

The current OSSE uses version 5.10.0 of the Goddard Earth Observing System Model and DAS (GEOS-5; Rienecker *et al.*, 2008). The latter is based on the Grid-point Statistical Interpolation (GSI) 3D-Var scheme used at the NCEP (Wu *et al.*, 2002). The incremental analysis update of Bloom *et al.* (1996) is used for enhanced dynamical balance. Assimilations are performed with a 6 h cycle. The background-error statistics it requires are the same ones as derived for the operational GMAO system in 2012 using a tuning of covariances estimated using a previous application of the NMC method. The horizontal grid for the global analysis and forecasts has 361 and 576 equally spaced latitudes and longitudes, respectively. Vertically, there are 72 atmospheric levels in addition to the surface. The vertical coordinate is a hybrid one: a sigma coordinate in the lowest few levels, a pressure coordinate above 150 hPa, and a blending in between. These levels will be designated by a value η that equals the pressure that would be indicated if the surface pressure (p_s) were 1000 hPa everywhere. Background error is determined by differencing backgrounds with corresponding NR 'truth' that has been projected onto the analysis grid as described in Errico and Privé (2014).

The use of different models to produce the NR and to perform the assimilation provides a source of simulated model error. Examinations of the GMAO OSSEs indicate that this simulated model error is a large contributor to analysis and forecast errors. This is especially true in the Tropics as well as above the lower stratosphere (Privé *et al.*, 2013b).

3. Results

Although the NR includes a full year of data, our OSSE validates very poorly for the Northern Hemisphere winter season. Specifically, the OSSE analysis and forecasts are consistently unrealistically accurate in the Northern Hemisphere during the simulated January and February. It may be by chance that this period has simulated weather that is simply more predictable than usual, but with only one such winter season available in the NR dataset, this possibility is difficult to test. For this reason, we currently have restricted all our examinations to the period 1 July to 31 August. All statistics presented here are for averages over 0000 and 1200 UTC during those days.

Attention in this article is focused on the troposphere for reasons presented in Privé *et al.* (2013b). (Essentially, the climates of the NR and GEOS-5 models are so different that observations above $\eta = 5$ hPa cannot rectify their differences.) Fields of background error and 48 minus 24 h forecast differences will be denoted as BKGE and FD24, respectively. In order to limit the amount of diagnostic calculation, some presented correlations were computed only for each of the geographical regions described

Table 1. Geographical regions for computing spatial correlations.

Region	Latitude		Longitude	
	Minimum	Maximum	Minimum	Maximum
NPAC	35°N	65°N	130°E	250°E
TPAC	15°S	15°N	130°E	250°E
SPAC	35°S	65°S	130°E	250°E

in Table 1, representing the South, tropical, and North Pacific Ocean denoted as SPAC, TPAC, and NPAC, respectively.

3.1. Comparison of variances

The FD24 fields are intended to represent forecast-error growth over 24 h. What is desired, however, are error statistics for 6 h forecasts. A rescaling of variances of FD24 fields to match those of BKGE fields is therefore expected.

Ratios of zonal means of temporal variances of BKGE fields with respect to those of FD24 fields are presented in Figure 1 as functions of η -levels and latitudes. The fields are temperature T , specific humidity q , zonal wind u , and meridional wind v . In the extratropical troposphere, except near the surface, all fields show ratios less than 1, with values aloft as small as 0.2. However, in the Tropics ratios are typically larger than 1, with values 2 or larger in some locations. In the Tropics, the FD24 fields underestimate the background errors while in the Extratropics they are an overestimate. So, the required rescaling varies greatly with both pressure and latitude.

In the GMAO OSSE context, model error is implicitly introduced as the difference in formulations of the NR and DAS models. Results in Privé *et al.* (2013b) suggests that this error is the dominant 6 h OSSE forecast-error source in the Tropics. However, the FD24 results instead depend only on this model error in so far as it affects the analysis increments used to create the shorter of the forecasts. Thus, the character of the effects of implicit model error in the FD24 results can be very different from those in the BKGE results.

3.2. Comparison of vertical correlations

SV characterization of forecast-error growth suggests that the dominant errors grow non-modally; i.e. change their shape as well as amplitude with time (Farrell, 1989; Palmer *et al.*, 1998; Errico, 2000b). If so, then spatial correlations are expected to differ for BKGE and FD24 datasets. In the following,

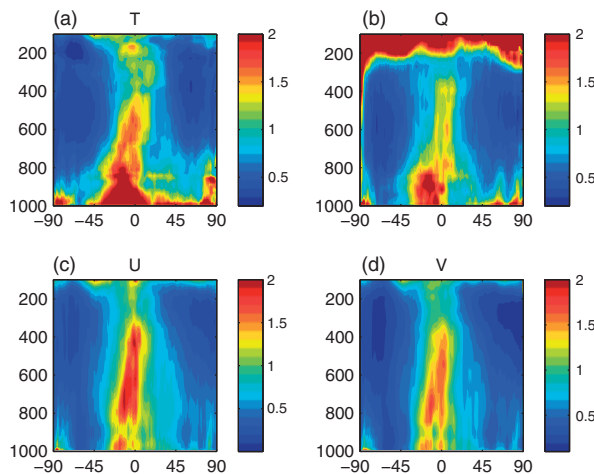


Figure 1. Ratios of zonal mean variances of BKGE data with respect to corresponding variances of FD24 data for (a) temperature T , (b) specific humidity q , (c) zonal wind speed u , and (d) meridional wind speed v , within the range $100 < \eta < 1000$ hPa.

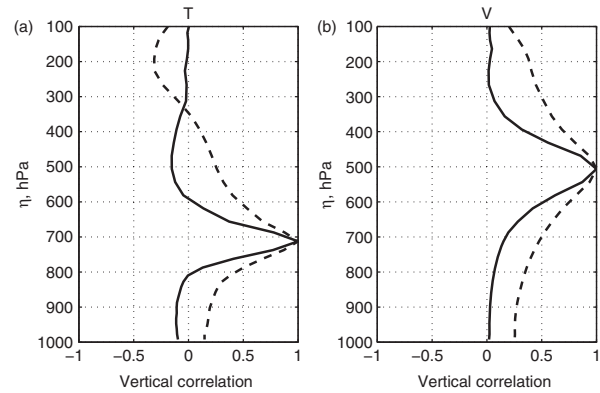


Figure 2. Vertical correlations of (a) T with its value at $\eta = 700$ hPa and (b) v with its value at $\eta = 500$ hPa, computed within the SPAC region for FD24 (dashed) and BKGE (solid) datasets.

geographically dependent statistics are determined by averaging in time and within each of the separate regions indicated in Table 1. Corresponding calculations were also performed for other longitude ranges within the same latitude bands to ensure the robustness of the presented results. This was confirmed for all the qualitative statements here. Quantitatively, results for corresponding wind and q correlations above the boundary layer were negligibly different, but those for T in the Tropics had greater variations.

Typical vertical correlations for the dynamic fields within the extratropical troposphere are presented in Figure 2. Specifically, these are for T at 700 hPa with T at other levels and for v at 500 hPa with v at other levels, computed within the SPAC region. The longer vertical correlation length-scales for FD24 data for both fields are evident. From results such as these, quantitative correlation length-scales L_v are determined as half-widths at half-maxima (i.e. half the distance between locations where the correlation equals 0.5) as functions of the η -level where the correlation is 1, following the procedure described in Errico and Privé (2014).

Values of L_v for $1000 < \eta < 100$ hPa computed for T and v within the SPAC and TPAC regions are presented in Figure 3. For the FD24 SPAC v field, the L_v are almost double their corresponding BKGE values. For SPAC T in the range $\eta < 700$ hPa, FD24 values are roughly 40% larger than corresponding BKGE ones. Results (not shown) for u are almost identical to those for v . Values for q (not shown) are between those for T and v at the same η -level, with values for FD24 also much larger than for BKGE. Within the NPAC region, these differences between the FD24 and BKGE datasets are only slightly less dramatic. However, in the TPAC region corresponding values for FD24 and BKGE are similar to each other for all fields.

Values of L_v were also determined (not shown) for analysis errors computed from differences between each analysis and corresponding truth, as in Errico and Privé (2014). For all fields and regions in the troposphere, these were 5–10% smaller than the corresponding BKGE values. This is expected since the analysis errors are a blending of spatially noisy observation errors with the more spatially correlated background errors.

3.3. Comparison of principal components

Further understanding is aided by consideration of the principal components (PCs) of a matrix constructed from the covariances used to determine the vertical correlations. For this purpose, only levels $100 < \eta < 1000$ hPa are considered. Before PCs are computed from it, each matrix element m_{ij} that considers covariances between η -levels i and j is weighted by the product of the square roots of the fractions of column total mass within each of the atmospheric layers being considered. Specifically,

$$m_{ij} = w_i w_j \{f_i - \langle f_i \rangle \} \{f_j - \langle f_j \rangle \},$$

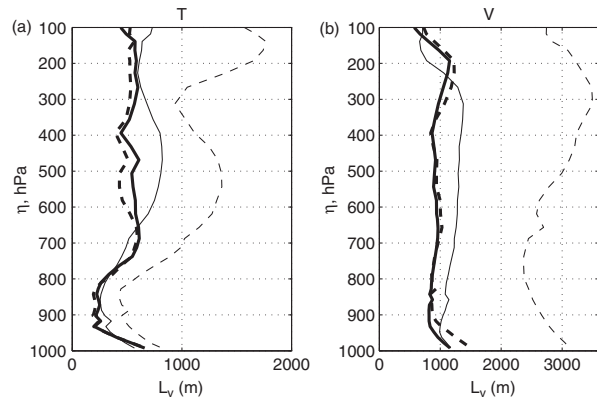


Figure 3. Vertical correlation lengths L_v for (a) T and (b) v for SPAC (thin lines) and TPAC (bold lines) regions for BKGE (solid lines) and FD24 (dashed lines) datasets as functions of η .

where the inner and outer angle brackets respectively denote temporal–mean and regional–mean values, f denotes a generic field, and

$$w_i = \sqrt{\Delta p_i / \sum_i \Delta p_i},$$

with Δp_i the ‘pressure thickness’ of the subscripted layer in units of hPa and the sum is over all layers considered (i.e. 900 hPa).

The pressure thicknesses are computed from the η -levels defining the data’s vertical grid by assuming that the surface pressure is 1000 hPa. This weighting is necessary given the non-uniformity of the η -levels of the GEOS-5 grid. Each PC is then normalized such that the sum of its squared components over the layers through which it is defined is 1. This is equivalent to interpreting its associated eigenvalue as a fractional mass-weighted integral of variance between the surface and 100 hPa for the component of the field having the PC structure considered. However, when a PC structure is presented, it will be unscaled by the weights at each level, so it can be interpreted as the corresponding vertical structure of the field itself.

The variances corresponding to the leading four PCs for the T , v , and q fields in the SPAC and TPAC regions for the FD24 and BKGE data appear in Table 2 along with the percentages of total mass-weighted variance explained by these four PCs. Especially for SPAC v , and apart from T in the Tropics, these explain a large percentage of variance (as large as 89%). In the SPAC region, the ratios of corresponding BKGE versus FD24 values are smaller for more leading PCs, with the smallest ratio 0.16 for PC1 v . Thus, the differences in total variance are accounted for by differences in the variances explained by the leading PCs. However, for T in the Tropics, the ratios for the PCs shown are all between 1.26 and 1.64. For v there, the ratios are even closer to 1.

For all fields, results for NPAC are like those for SPAC, except that the variances associated with corresponding PCs are smaller in the summer (NPAC) hemisphere. In particular, corresponding leading PC structures are very similar. Results for v are very similar to those for u in all regions. Explicit results for u are therefore not shown.

Structures of PC1–4 for T and v in SPAC and TPAC regions appear in Figure 4. Results for BKGE and FD24 appear together. The structures for BKGE and FD24 are similar to each other in most of the panels except for some near 1000 and 100 hPa. So, in what follows, projections of datasets onto one set of leading PCs yield results similar to those if the other set were used. For the SPAC region, each successive leading PC has smaller vertical scale as measured by the number of zero crossings. By the same measure, each PC for T has a smaller vertical scale than that for the corresponding v PC, consistent with the assumption that differences or background errors at large vertical scales are geostrophically balanced. (Errico, 1997b, gives an explanation for this.) For the TPAC region, these changes of scale between successive PCs and relationships between structures for T and v are less apparent. The latter result is consistent with the lack of geostrophy expected there. An explanation of the leading PC structures in the Tropics likely involves effects of convection.

PC results for q are not shown except in Table 2. In all regions, the four leading PCs explain more than 70% of the total q variance. Ratios of corresponding BKGE and FD24 values for q in Table 2 are qualitatively like those for other fields, in agreement with the similarity of panels in Figure 1. Leading PC structures for q have largest magnitudes below $\eta = 500$ hPa due to the approximately exponential decay of q itself as pressure decreases. For PC1 in all regions, the PC structure is of a single sign through the lower half of the troposphere.

3.4. Comparison of horizontal correlations

Horizontal correlations for BKGE and FD24 fields were computed separately in the meridional and zonal directions within each region of Table 1 for each η surface. From these, half-widths at half-maximum were estimated as horizontal correlation lengths L_m and L_z , respectively. Values of L_m are presented for T and v in SPAC and TPAC regions in Figure 5. Values in the NPAC region are not shown since they are very similar to those for SPAC. The same relationships between L_m and L_z values described in Errico and Privé (2014) apply here, so L_z is not discussed further.

For SPAC v at all levels, L_m for BKGE is 60% or smaller than its FD24 counterpart. Values change little with height for $\eta > 400$ hPa. For FD24, unlike for BKGE, there is a rapid increase in L_m as η decreases through the range $300 > \eta > 100$ hPa, such that L_m for BKGE is 30% of its FD24 value at $\eta = 100$ hPa. For TPAC v , L_m for BKGE is approximately 125% of its corresponding FD24 value, except near the surface and top of the range of η presented, where BKGE values are instead smaller. For SPAC T , FD24 values also exceed those for BKGE except near the surface, although the differences are less dramatic than seen for v . For TPAC T at almost all levels, the FD24 values undervalue the BKGE results. Thus the relationships between the horizontal correlation lengths in the FD24 and BKGE datasets are very different, with the nature of that difference reversed in the Tropics compared with the Extratropics.

Values of L_m were also computed for analysis errors (not shown). For all fields, regions, and tropospheric η -levels examined, these were generally 3–7% smaller than corresponding BKGE values. One exception was for SPAC and NPAC T for

Table 2. Variances corresponding to the leading four principal components for the regions, fields, and datasets indicated. Units are K^2 for T , 10^{-7} for q , and m^2s^{-2} for v . Also shown is the percentage of the total variance explained by the leading four PCs.

Region Field Dataset	SPAC						TPAC					
	T		v		q		T		v		q	
	BKGE	FD24	BKGE	FD24	BKGE	FD24	BKGE	FD24	BKGE	FD24	BKGE	FD24
PC1	0.088	0.370	1.110	7.020	0.347	0.829	0.053	0.042	1.050	1.110	5.120	3.380
PC2	0.080	0.165	0.852	2.740	0.181	0.249	0.048	0.032	0.894	0.851	2.480	1.880
PC3	0.061	0.118	0.695	1.240	0.125	0.134	0.041	0.025	0.679	0.579	1.450	1.000
PC4	0.053	0.088	0.406	0.712	0.090	0.084	0.034	0.023	0.574	0.537	1.050	0.714
%Total	47	65	72	89	76	84	37	33	56	58	74	71

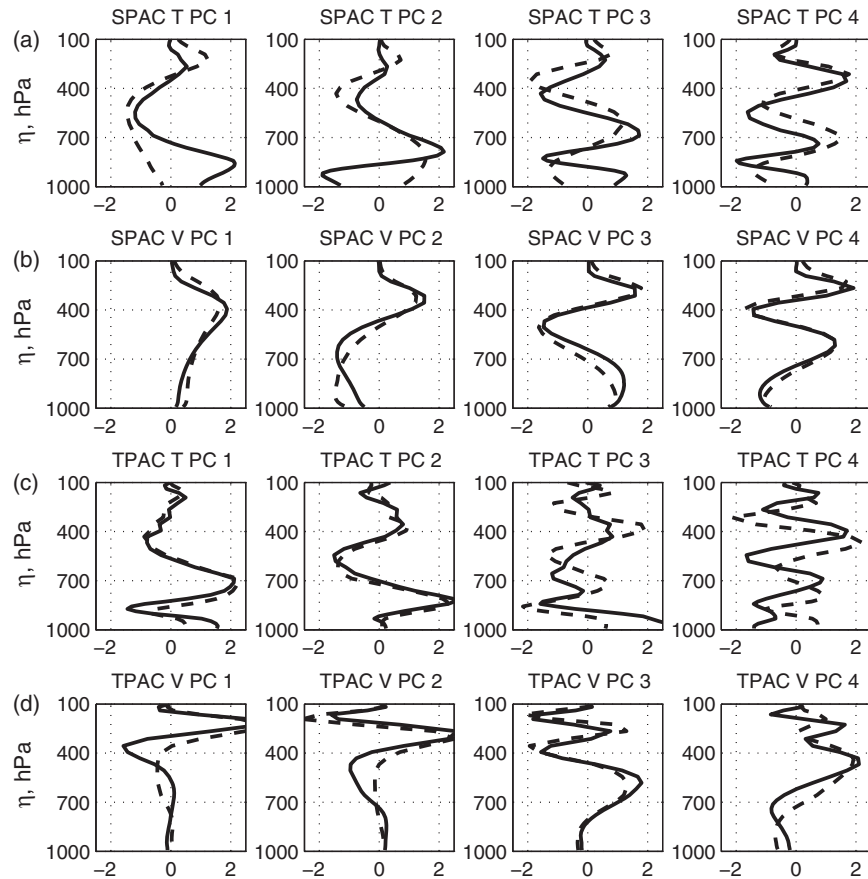


Figure 4. Structures of the four leading PCs (first to fourth columns) computed for temperature (first and third rows) and meridional wind speed (second and fourth rows) within the SPAC (first and second rows) and TPAC (third and fourth rows) for BKGE (solid lines) and FD24 (dashed lines) datasets.

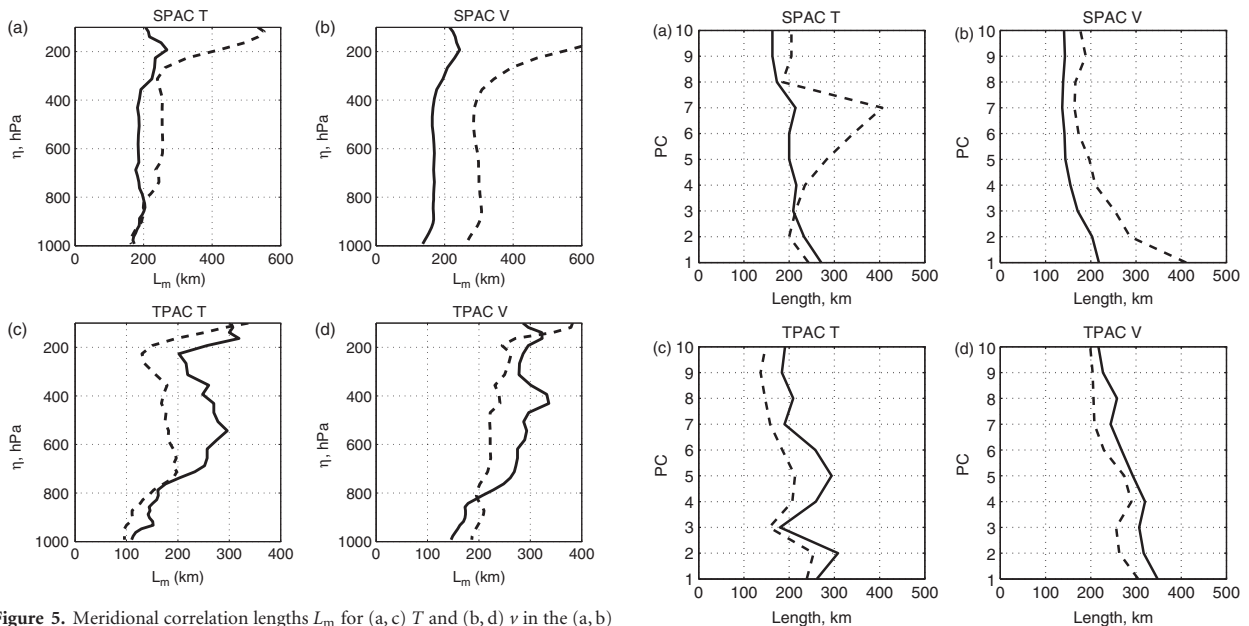


Figure 5. Meridional correlation lengths L_m for (a, c) T and (b, d) v in the (a, b) SPAC and (c, d) TPAC regions for BKGE (solid line) and FD24 (dashed line) datasets as functions of η .

$\eta < 300$ hPa, where L_m values were instead a few percent greater. As is the case for vertical correlations, the typically smaller L_m may be explained by the blending of background errors with less spatially correlated observation errors.

Horizontal correlations were also examined for coefficients of the PCs when the fields are projected onto them, thereby linking vertical and horizontal scales. Values of L_m for PC1–10

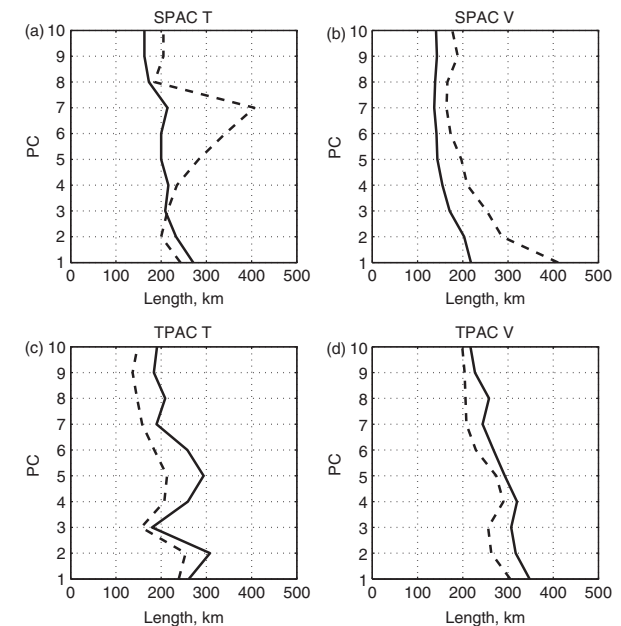


Figure 6. Meridional correlation lengths L_m for (a, c) T and (b, d) v in the (a, b) SPAC and (c, d) TPAC regions for BKGE (solid line) and FD24 (dashed line) datasets for fields of projection coefficients for PC1–10.

are presented in Figure 6 for T and v in the SPAC and TPAC regions for both BKGE and FD24 datasets. For v , they reveal that a difference in L_m between datasets occurs for all PCs, and thus vertical scales, presented, although in the SPAC region, the greatest difference occurs for PC1. Since PC1 also accounts for a large fraction of total variance in either dataset, its difference

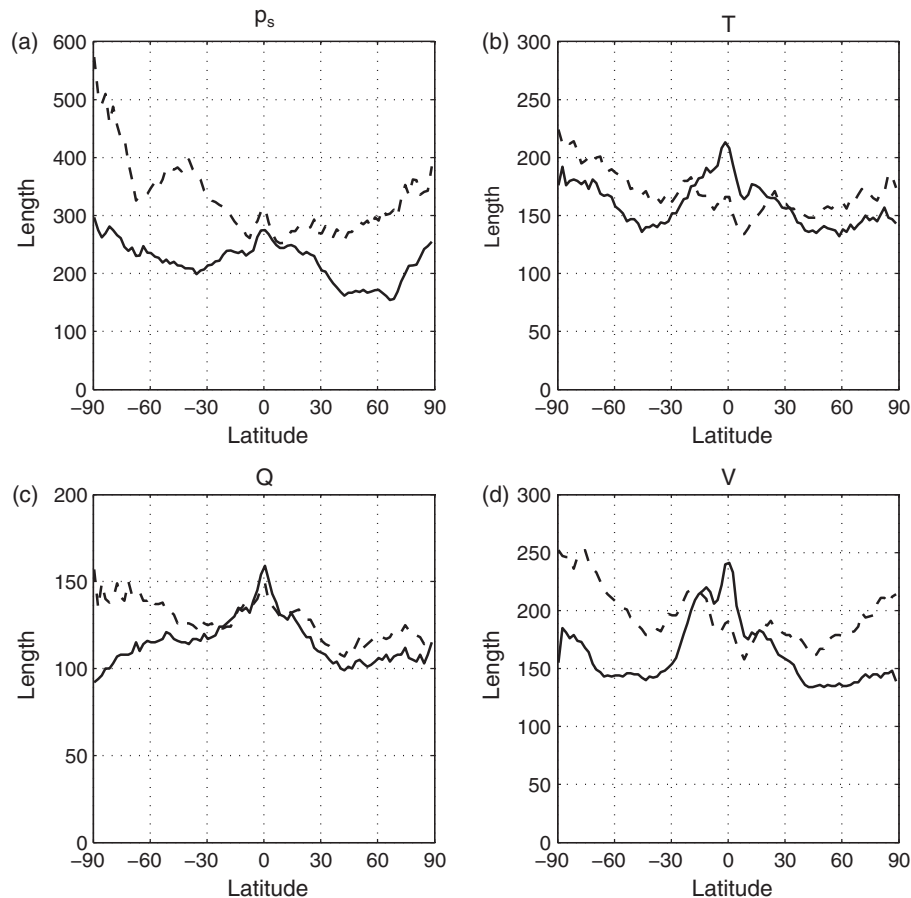


Figure 7. Meridional length-scales (L_c , km) for (a) p_s and for (b) T , (c) q , and (d) v at $\eta = 500$ hPa for BKGE (solid) and FD24 (dashed), computed from the curvature of the correlation function at separation zero.

in L_m values also accounts for much of the difference observed for SPAC v in Figure 5. In contrast, for T in the SPAC region, the differences seen in Figure 5 must be explained by PC n , $n > 3$. In the Tropics all PCs shown reveal the same relationship between BKGE and FD24 datasets as is seen at most η -levels in Figure 5.

A more common way to compute horizontal length-scales is described in Belo Pereira and Berre (2006) with some computational details provided in Appendix A here. It uses local standard deviations of fields and their spatial gradients to efficiently compute curvatures of correlation functions near zero separation distances. This computation therefore does not depend on binning sizes or interpolations between binned values, as required for computing L_m . The length-scales L_c so determined generally differ from the half-widths at half-maxima, with the relationship between the two depending on the particular shape of the correlation function.

Values of L_c in the meridional direction as functions of latitude for p_s and at $\eta = 500$ hPa, for T , q , and v for both BKGE and FD24 appear in Figure 7. Qualitatively, the results for T and v are like those in Figure 5 in the specific sense of which dataset has the larger values in which latitudinal zones. However, the ratios of corresponding BKGE and FD24 values appear less dramatic than those for L_m within the same latitude bands. (This was also confirmed by averaging the local L_c within the regions of Table 1 for direct comparison with Figure 5.) Also, $L_c < L_m$, consistent with the different meanings of the length-scales. For q , corresponding BKGE and FD24 values of L_c are very similar, except at high latitudes in the winter hemisphere, where, as for the other fields, FD24 values are much larger. The fact that the ratios of corresponding FD24 to BKGE values differ for L_m and L_c indicates that the functional shapes of the meridional correlation functions change, not just the length-scales.

Values for L_c appearing for p_s in Belo Pereira and Berre (2006) (Figure 10 there) cannot be compared directly with those in Figure 7 because the former were produced for a lower-resolution DAS and the ensemble they used to estimate BKGE did not consider model error. Both of these differences will tend to create smaller L_c here. This reasoning assumes that the temporally variable part of the model error in the OSSE has very small spatial correlation.

3.5. Comparison of dynamic balances

As a measure of the degree of dynamic balance, the fractions of the total temperature variances contributed by the balanced temperature components as a function of η and latitude are shown in Figure 8 for BKGE and FD24. This fraction is determined using the algorithm described in Appendix B. Almost everywhere in the BKGE troposphere, the fraction of balance temperature variance is less than 0.5. In contrast, values larger than 0.5 are seen in many locations for FD24. It is apparent that the FD24 balance fraction is a gross overestimate of the BKGE value throughout most of the extratropical troposphere.

Balance was also measured for 1 day forecast error using the NR truth. Its values were between those for BKGE and FD24. This indicates that the discrepancy revealed in Figure 8 is not simply because FD24 concerns a 24 h lag rather than forecast error.

4. Conclusions

Estimates of background-error statistics inferred from consideration of forecast differences have always been expected to require significant tuning before they could be successfully employed in a DAS. Here, the specific kinds of tuning required have been

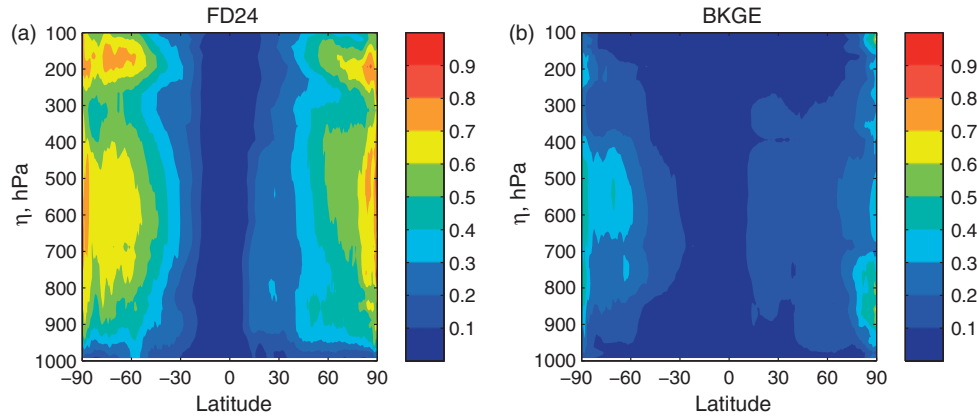


Figure 8. Fractions of temperature variance contributed by dynamical balanced components for (a) FD24 and (b) BKGE.

revealed by examination of explicit background errors determined within an OSSE context.

The ratios of variances produced by OSSE background errors versus 24 h lagged forecast differences vary greatly with both height and latitude. For all fields in the troposphere, ratios range from 0.2 near the extratropical tropopause to more than 2 near the tropical surface. A single global scaling factor should not be applied to variances of the lagged differences when tuning them to approximate background-error variances.

In the Extratropics, corresponding vertical correlation length-scales are much larger for lagged differences than background error. This is as expected from consideration of short-term forecast error growth as revealed by SVs and the dependence of the lagged differences on analysis increments rather than on analysis errors, which can have very different characteristics as revealed in Berre *et al.* (2006). However, in the Tropics there is little apparent change of such length-scales. As for variances, tuning of vertical correlation lengths to modify covariances of the lagged differences to be more like those for background error should be latitude dependent.

The structures of leading PCs of layer mass-weighted vertical covariances in the troposphere are very similar for both lagged differences and background errors. However, the corresponding variances explained by each differ for each dataset. For all fields, greater variance is explained by larger vertical scales.

Horizontal correlations of background error are similar to those for analysis error as revealed in Errico and Privé (2014). For all fields in the Extratropics, correlation lengths are much greater for lagged differences than those for corresponding background errors. This is especially true for wind, where the former is more than 60% larger at all tropospheric levels. However, in the Tropics the relationship is qualitatively reversed. For example, for tropical T for $\eta < 700$ hPa, the meridional length-scale is 40% smaller for the lagged differences.

For wind, larger horizontal length-scales are associated with all vertical PCs for lagged differences in the Extratropics and background errors in the Tropics. In the Extratropics, the leading PC is associated with the greatest difference in horizontal correlation length between datasets. The equivalent relationship for tropical temperature is like that for tropical wind. In the Extratropics, the relationship for temperature is more complex, with the two leading PCs associated with larger horizontal length-scales for the background, and PC4-7 with smaller scales. This mixed behaviour is associated with the varying relationship seen for this field as a function of elevation.

Two kinds of horizontal correlation length-scales were determined: half-widths at half-maxima determined from spatially averaged covariances, and spatially averaged length parameters determined from estimates of the local curvatures of correlation functions at zero separation distances. Although not identical in meaning or the effects of averaging, these yielded qualitatively similar results regarding comparisons of values

for corresponding BKGE and FD24 fields. Some quantitative differences in the results suggest that the shapes of the correlation functions, not just their length-scale parameters, may differ between the BKGE and FD24 datasets.

The fractions of variances of temperature error or forecast differences attributed to dynamical balanced temperature components were determined statistically from covariances between stream functions and temperature at differing vertical levels of the analysis grid. For the FD24 dataset, large fractions of balance are obtained throughout most of the extratropical troposphere. In contrast, the BKGE dataset reveals much less balance. Thus, for this measure also, the pre-tuned NMC method does not yield good estimates for the desired background-error statistics.

In summary, the tuning required to construct appropriate background-error statistics from lagged forecast differences varies much with latitude. For variances in the Extratropics, the appropriate tuning also varies greatly with elevation. Appropriate tuning of vertical and horizontal correlation lengths also differs greatly between the Tropics and Extratropics. In the profound case of horizontal correlations of wind, the length-scale must be made 40% smaller at all η levels in the extratropical troposphere but 25% larger in the Tropics for $800 > \eta > 200$ hPa. Although these results have been produced for an OSSE, validation of the OSSE framework (Errico *et al.*, 2013), consideration of the theory of short-term error growth (e.g. Palmer *et al.*, 1998; Gelaro *et al.*, 2002), and examination of the NMC method (Berre *et al.*, 2006; Belo Pereira and Berre, 2006) suggest they are realistic. In the future we will attempt to use these results to provide a new background-error covariance model to test within the OSSE and for an assimilation of real observations.

Appendix A

Computation of local correlation lengths

The definition of L_c appears in Belo Pereira and Berre (2006):

$$L_c = \sqrt{\frac{\sigma^2(\epsilon)}{\sigma^2\left(\frac{\partial \epsilon}{\partial y}\right) - \left(\frac{\partial \sigma(\epsilon)}{\partial y}\right)^2}},$$

where ϵ is the BKGE or FD24 field in question, y is a meridional distance, and σ^2 is the variance of the indicated variable. The meridional gradients are computed here as centred differences on the GEOS-5 grid that has equidistant separations between latitudes. The variances are computed separately at each latitude and longitude of the grid, excluding the poles, by using temporal averages. The values of L_c discussed here are then determined by simply averaging all the separate grid-point values within either the region or latitude concerned.

For a Gaussian-shaped correlation function,

$$\rho(\Delta y) = \exp\left(-\frac{1}{2} \frac{(\Delta y)^2}{L_c^2}\right),$$

the correlation is 0.5 when the separation distance (Δy) is approximately $1.18L_c$. For a second-order auto-regressive (SOAR) function,

$$\rho(\Delta y) = \left(1 + \frac{\Delta y}{L_c}\right) \exp\left(-\frac{\Delta y}{L_c}\right),$$

the correlation is 0.5 when $\Delta y \approx 1.67L_c$. Note that since ρ generally depends nonlinearly on L_c , determining L_m from spatially averaged covariances as done here and described in Errico and Privé (2014) likely will not produce the same results as spatially averaging locally determined L_c .

Appendix B

Determination of the balance measure

Let \mathbf{T} and ψ denote column vectors whose components are values of T and stream function at successive η levels, defined for each latitude, longitude, and time. The variances of T are determined by the diagonal elements of the matrix

$$\langle (\mathbf{T} - \bar{\mathbf{T}})(\mathbf{T} - \bar{\mathbf{T}})^T \rangle,$$

where T indicates a transpose, the overline denotes a zonal mean and angle brackets here indicate a time and zonal mean. The balanced temperature is defined as

$$\mathbf{T}_b = \mathbf{A}\psi,$$

where

$$\mathbf{A} = \langle (\mathbf{T} - \bar{\mathbf{T}})(\psi - \bar{\psi})^T \rangle \langle (\psi - \bar{\psi})(\psi - \bar{\psi})^T \rangle^{-1}.$$

Variances of \mathbf{T}_b are determined as the diagonal elements of

$$\mathbf{A} \langle (\psi - \bar{\psi})(\psi - \bar{\psi})^T \rangle \mathbf{A}^T.$$

This is the method of computing statistics of dynamic balances to define the background-error covariance models used by GSI at NCEP and the GMAO.

Acknowledgements

The ECMWF nature run was provided by Erik Andersson through arrangements made by Michiko Masutani. The software for simulating GPSRO observations was provided by the Radio Occultation Processing Package with the assistance of Sean Healy at ECMWF. Support for this project was encouraged by Michele Rienecker and provided by GMAO core funding. Attention to Berre *et al.* (2006) and Belo Pereira and Berre (2006) was suggested by two anonymous reviewers, which was very helpful in revising the manuscript.

References

Belo Pereira M, Berre L. 2006. The use of an ensemble approach to study background-error covariances in a global NWP model. *Mon. Weather Rev.* **134**: 2466–2489.

- Berre L, Stefanescu SE, Belo Pereira M. 2006. The representation of the analysis effect in three error assimilation techniques. *Tellus* **58A**: 196–209.
- Bloom S, Takacs L, DaSilva A, Ledvina D. 1996. Data assimilation using incremental analysis updates. *Mon. Weather Rev.* **124**: 1256–1271.
- Bonavita M, Isaksen L, Holm E. 2012. On the use of EDA background-error variances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc.* **138B**: 1540–1559.
- Courtier P. 1997. Variational methods. *J. Meteorol. Soc. Jpn.* **75**: 211–218.
- Cucurull L, Derber JC, Purser RJ. 2013. A bending angle forward operator for Global Positioning System radio occultation measurements. *J. Geophys. Res.* **118**: 14–28, doi: 10.1029/2012JD017782.
- Culverwell I, Offiler D, Lauritsen KB. 2013. *ROM SAF Radio Occultation Processing Package: An Overview*. Ref: SAF/ROM/METO/UG/ROPP/001. EUMETSAT: Darmstadt, Germany. http://www.romsaf.org/romsaf_ropp_overview.pdf (accessed 2 May 2014).
- Errico RM. 1997a. What is an adjoint model? *Bull. Am. Meteorol. Soc.* **78**: 2577–2591.
- Errico RM. 1997b. On the removal of gravitational noise in numerical forecasts. *J. Meteorol. Soc. Jpn.* **75**: 219–227.
- Errico RM. 2000a. Interpretations of the total energy and rotational energy norms applied to determination of singular vectors. *Q. J. R. Meteorol. Soc.* **126A**: 1581–1599.
- Errico RM. 2000b. The dynamical balance of singular vectors in a primitive equation model vectors. *Q. J. R. Meteorol. Soc.* **126A**: 1601–1618.
- Errico RM, Privé NC. 2014. An estimation of some analysis error statistics using the GMAO observing system simulation framework. *Q. J. R. Meteorol. Soc.* **140**: 1005–1012.
- Errico RM, Yang R, Privé NC, Tai KS, Todling R, Sienkiewicz ME, Guo J. 2013. Development and validation of observing system simulation experiments at NASA's Global Modeling and Assimilation Office. *Q. J. R. Meteorol. Soc.* **139**: 1162–1178.
- Farrell BF. 1989. Optimal excitation of baroclinic waves. *J. Atmos. Sci.* **46**: 1193–1206.
- Gelaro R, Reynolds CA, Errico RM. 2002. Transient and asymptotic perturbation growth in a simple model. *Q. J. R. Meteorol. Soc.* **128A**: 205–208.
- Healy SB, Eyre JR, Hamrud M, Thépaut J-N. 2007. Assimilating GPS radio occultation measurements with two-dimensional bending angle observation operators. *Q. J. R. Meteorol. Soc.* **133**: 1213–1227.
- Hollingsworth A, Lönnberg P. 1986. The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus* **38A**: 111–136.
- Hortal M, Simmons AJ. 1991. Use of reduced Gaussian grids in spectral models. *Mon. Weather Rev.* **119**: 1057–1074.
- Kleist DT, Parrish DF, Derber JC, Treadon R, Wu WS, Lord S. 2008. Introduction of the GSI into the NCEP global data assimilation system. *Weather and Forecasting* **24**: 1691–1705.
- Langland RH, Elsberry RL, Errico RM. 1995. Evaluation of physical processes in an idealized extratropical cyclone using adjoint sensitivity. *Q. J. R. Meteorol. Soc.* **121**: 1349–1386.
- Lorenz EN. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**: 130–141.
- Palmer TN, Gelaro R, Barkmeijer J, Buizza R. 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **55**: 633–653.
- Parrish DF, Derber JC. 1992. The National Meteorological Center's spectral statistical–interpolation analysis system. *Mon. Weather Rev.* **120**: 1747–1763.
- Privé NC, Errico RM, Tai KS. 2013a. Validation of the forecast skill of the Global Modeling and Assimilation Office observing system simulation experiment. *Q. J. R. Meteorol. Soc.* **139**: 1354–1363.
- Privé NC, Errico RM, Tai KS. 2013b. The influence of observation errors on analysis error and forecast skill investigated with an observing system simulation experiment. *J. Geophys. Res. Atmos.* **118**: 5332–5346, doi: 10.1002/jgrd.50452.
- Rienecker MM, Suarez MJ, Todling R, Bacmeister J, Takacs L, Liu H-C, Gu W, Sienkiewicz M, Koster RD, Gelaro R, Stajner I, Nielsen JE. 2008. *The GEOS-5 Data Assimilation System—Documentation of Versions 5.0.1, 5.1.0, and 5.2.0*, Suarez M. (ed.), Technical Report Series on Global Modeling and Assimilation, NASA/TM-2008-104606, Vol. 27. NASA: Greenbelt, MD, 118 pp.
- Talagrand O. 1981. A study of the dynamics of four-dimensional data assimilation. *Tellus* **33**: 43–60.
- Tarantola A. 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier Science: Amsterdam.
- Wu W-S, Purser RJ, Parrish DF. 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Weather Rev.* **130**: 2905–2916.