Regionally Enhanced Global (REG) 4D-Var

MICHAEL A. HERRERA

National Research Council at Naval Research Laboratory, Washington, D.C.

ISTVAN SZUNYOGH AND ADAM BRAINARD

Texas A&M University, College Station, Texas

DAVID D. KUHL AND KARL HOPPEL

Naval Research Laboratory, Washington, D.C.

CRAIG H. BISHOP,^a TEDDY R. HOLT, AND QINGYUN ZHAO

Naval Research Laboratory, Monterey, California

SABRINA RAINWATER

National Research Council at Naval Research Laboratory, Monterey, California

(Manuscript received 4 August 2017, in final form 17 September 2018)

ABSTRACT

A regionally enhanced global (REG) data assimilation (DA) method is proposed. The technique blends high-resolution model information from a single or multiple limited-area model domains with global model and observational information to create a regionally enhanced analysis of the global atmospheric state. This single analysis provides initial conditions for both the global and limited-area model forecasts. The potential benefits of the approach for operational data assimilation are (i) reduced development cost, (ii) reduced overall computational cost, (iii) improved limited-area forecast performance from the use of global information about the atmospheric flow, and (iv) improved global forecast performance from the use of more accurate model information in the limited-area domains. The method is tested by an implementation on the U.S. Navy's four-dimensional variational global data assimilation system and global and limited-area numerical weather prediction models. The results of the monthlong forecast experiments suggest that the REG DA approach has the potential to deliver the desired benefits.

1. Introduction

Most major numerical weather prediction centers of the world produce both global and limited-area model (LAM) forecasts. These centers typically maintain both a global and a limited-area data assimilation (DA) system, and some of them prepare limited-area analyses and forecasts for multiple limited-area domains. For instance, the Fleet Numerical Meteorology and Oceanography Center (FNMOC) of the U.S. Navy

prepares limited-area weather analyses and forecasts for more than 80 domains, covering a large fraction of the Northern Hemisphere (NH). At these centers, both the development and computational cost of data assimilation could potentially be reduced by preparing the global and limited-area analyses by a single, unified data assimilation system. The savings in development and computational time could be used for the enhancement of the unified data assimilation system.

Streamlining development and production is not the only motivation to search for a unified global–limited-area data assimilation approach, as it is not unreasonable to assume that the quality of both the global and LAM initial conditions may benefit from such an approach. In particular, LAM initial conditions can potentially benefit from the use of global model and observational

^a Current affiliation: School of Earth Sciences, and ARC Centre of Excellence for Climate Extremes, University of Melbourne, Melbourne, Australia.

Corresponding author: Michael A. Herrera, mherr77m@gmail.com

information, because it can aid in the proper interpretation of the contribution of the large-scale component of the atmospheric flow in the limited-area domains. The detrimental effects of the lack of complete information about the large-scale flow on the analyses produced by limited-area data assimilation have been long recognized in operational numerical weather prediction. Some forecast centers, for instance, use an approach called partial cycling of the limited-area data assimilation process: the cycling of the limited-area data assimilation always falls back on a recent global analysis to carry out only a few cycles of limited-area data assimilation (Rogers et al. 2009). Another approach is to modify the cost function minimized by the limited-area data assimilation scheme such that it penalizes departures from the host global analysis that provides the lateral boundary conditions (Guidard and Fischer 2008; Dahlgren and Gustafsson 2012). The global initial conditions may also benefit from the LAM information, because it provides not only an additional estimate of the atmospheric state, but also a higher resolution and presumably more accurate information about the effects about the boundary conditions at the surface (e.g., orography, coastlines, land coverage) and the atmospheric processes (e.g., nonlinear interactions at the smaller scales, cloud physics).

Previous studies with highly idealized Lorenz models of the atmospheric dynamics (Yoon et al. 2012; Kretschmer et al. 2015) demonstrated the potential for improvement of both the global and limited-area forecasts from a unified data assimilation approach. These idealized studies were motivated by the mixed results of an earlier attempt (Merkova et al. 2011) to incorporate higher-resolution limited-area model information into the global data assimilation process. That study used older, reduced-resolution versions of the operational models of the National Centers for Environmental Prediction (NCEP) and the research data assimilation system of Szunyogh et al. (2008) and led to the conclusion that further research was necessary to refine the technical approach. Yoon et al. (2012) explored a *joint-state approach* that prepared an analysis of a *joint state vector* composed of all state vector components of the global and limited-area models. One unexpected difficulty with the implementation of that approach was that in some situations, relatively large differences developed between the global and limited-area analyses in the limited-area region. Fixing this problem required the introduction of an extra term in the cost function of the data assimilation scheme to penalize differences between the global and limited-area analyses in the limited-area regions. While the addition of the penalty term had the desired effect, it made the implementation

of the technique more complicated. Kretschmer et al. (2015) found that similarly good results could be achieved with a simpler *composite-state method*: they estimated the state of a *composite state vector* defined by a linear combination of the global state vector and the limited-area state vectors in the limited-area domains and by the components of the global state vector elsewhere. The components of the global initial conditions in the limited-area domains were obtained by interpolation from the components of the composite state vector. The composite-state method was not only easier to implement than the joint-state approach, but its computational cost was also lower, especially for overlapping limited-area domains, because it estimated a state vector of fewer components.

The present study is the first attempt since Merkova et al. (2011) to test the concept of unified global-limited-area data assimilation on operational models. The technical approach used here is similar, but not completely identical, to that of Kretschmer et al. (2015). To be specific, the equivalent of the composite state vector, called the regionally enhanced global (REG) state vector, is used only to predict the observations in the computation of the innovations, the differences between the observations and their predicted values. This information is then used to obtain the model initial conditions by updating a global background state estimate rather than a background estimate of the REG state vector.

We test our approach by an implementation on the U.S. Navy's operational global and limited-area models and global data assimilation system. The global model is the Navy Global Environment Model (NAVGEN; Hogan et al. 2014), a spectral transform model operationally run at resolution T425. The regional model is the Coupled Ocean–Atmosphere Mesoscale Prediction System (COAMPS; Hodur 1997), which we use with no ocean–atmosphere coupling. COAMPS is implemented operationally on the more than 80 limited-area domains over the world with a varying horizontal resolution from 25 km to 500 m. COAMPS uses partial cycling and a 3D-Var data assimilation scheme (Daley and Barker 2001) to obtain limited-area initial conditions, and the version used for this study does not assimilate satellite radiance observations. The global data assimilation system of NAVGEN, the Naval Research Laboratory Atmospheric Variational Data Assimilation System–Accelerated Representer (NAVDAS-AR), employs a 4D-Var data assimilation scheme (Xu et al. 2005; Rosmond and Xu 2006). A future implementation of REG DA on NAVDAS-AR, a REG 4D-Var, would allow the Navy to unify the DA systems of NAVGEN and COAMPS. Doing so would enable COAMPS initial conditions to be informed by a larger number and variety of satellite observations and would also reduce the overall computation cost of DA

by eliminating it for the many individual limited-area domains.

The paper is organized as follows. [Section 2](#) introduces the general concept of REG DA; [section 3](#) outlines the specific implementation on the Navy's data assimilation system and models; [section 4](#) describes the design of the experiments to evaluate the global and limited-area forecast performance; [section 5](#) presents the results of the experiments; and [section 6](#) offers conclusions and outlines future work.

2. The REG state vector

We expect a more accurate computation of the predicted values of the observations in data assimilation to lead to more accurate analyses. Our hypothesis is that injecting limited-area forecast information into the computation of the predicted values of the observations is an efficient approach to improve their accuracy.

a. Definition

We assume that in addition to a global model, we also have access to a limited-area model that can add resolution and accuracy to the global model forecasts in one or more limited-area domains. In our implementation of REG DA on the Navy's forecast system, which we will describe in [section 3](#), the global and limited-area model states are blended on a high-resolution global grid, and the limited-area domains cover only part of the NH. To simplify notation, in this section, we assume that the limited-area model information is available for the entire globe, and the global and limited-area model state are blended on the grid of the limited-area model.

We define the REG state vector $\mathbf{x}_e$ by

$$\mathbf{x}_e = (1 - \alpha)\mathbf{A}(\mathbf{x}_g) + \alpha\mathbf{x}_\ell, \quad (1)$$

where $\mathbf{A}$ is the interpolation function that maps the discrete global model representation $\mathbf{x}_g$ of the atmospheric state into its limited-area model representation $\mathbf{A}(\mathbf{x}_g)$, $\mathbf{x}_\ell$ is the high-resolution state vector of a limited-area model forecast, and $0 \leq \alpha \leq 1$ is a (possibly location dependent) *blending coefficient*.

b. Implementation

Let $\mathbf{x}_g^b$ be a background estimate of the global state and $\mathbf{x}_\ell^b$ the background estimate of the limited-area state. In modern data assimilation, $\mathbf{x}_g^b$ is a short-term global model forecast, while $\mathbf{x}_\ell^b$ is a short-term limited-area forecast. Both state vectors are considered random vector variables, with the source of the randomness being the forecast errors. The REG background state vector,

$$\mathbf{x}_e^b = (1 - \alpha)\mathbf{A}(\mathbf{x}_g^b) + \alpha\mathbf{x}_\ell^b, \quad (2)$$

is a more accurate estimate of the (true) atmospheric state than either $\mathbf{x}_g^b$ or $\mathbf{x}_\ell^b$, because $\mathbf{x}_e^b$ is a weighted average of the two that filters random errors. If there were no systematic differences between the accuracy of $\mathbf{x}_g^b$ and $\mathbf{x}_\ell^b$ (differences between the means of their errors), the value of α that would minimize the variance of the error of $\mathbf{x}_e^b$ would be determined by the variance of the errors of $\mathbf{x}_g^b$ and $\mathbf{x}_\ell^b$. But we expect $\mathbf{x}_\ell^b$ to provide systematically more accurate information about the atmospheric flow near the surface because of the more detailed representation of the air–surface interactions and atmospheric physical processes in the limited-area model. Hence, we expect $\mathbf{x}_e^b$ to have smaller systematic errors than $\mathbf{x}_g^b$ near the surface.

Let $\mathbf{h}_g(\cdot)$ and $\mathbf{h}_\ell(\cdot)$ be the observation functions operating on $\mathbf{x}_g$ and $\mathbf{x}_\ell$, respectively. Under the assumption that $\mathbf{x}_e^b$ has a smaller systematic error than $\mathbf{x}_g^b$, the vector of innovations,

$$\delta\mathbf{y} = \mathbf{y}^o - \mathbf{h}_e(\mathbf{x}_e^b), \quad (3)$$

more accurately satisfies the assumption of data assimilation that nonzero innovations are the results of random errors in the background and the observations than if it was computed by $\mathbf{h}_g(\mathbf{x}_g^b)$. This motivates using $\mathbf{h}_e(\mathbf{x}_e^b)$ rather than $\mathbf{h}_g(\mathbf{x}_g^b)$ for the computation of the (vector of) innovations.

In the present implementation of REG DA, we use the global data assimilation system to compute the analysis increment $\delta\mathbf{x}_g^a$ from the innovations, which we then add to the global background $\mathbf{x}_g^b$ to obtain a global analysis $\mathbf{x}_g^a$:

$$\mathbf{x}_g^a = \mathbf{x}_g^b + \delta\mathbf{x}_g^a. \quad (4)$$

We generate the limited-area analysis $\mathbf{x}_\ell^a$ from the global analysis $\mathbf{x}_g^a$ by interpolation. An alternative approach would be to obtain an analysis

$$\mathbf{x}_e^a = \mathbf{x}_\ell^b + \mathbf{A}(\delta\mathbf{x}_g^a) \quad (5)$$

of the REG state vector and generate both the global analysis $\mathbf{x}_g^a$ and the limited-area analysis $\mathbf{x}_\ell^a$ by interpolation from $\mathbf{x}_e^a$. We chose the former strategy to implement first, because it is easier to maintain the proper atmospheric balance between the wind and the mass field in the global analysis. As we find the second strategy an intriguing alternative option, at the time of writing, we are working on its implementation.

For $\alpha = 0$, our approach is equivalent to preparing the global analysis the conventional way and obtaining the limited-area analysis by interpolation from the global

analysis. If $\alpha > 0$, however, $\mathbf{x}_g^a$ is different from a conventional global analysis. Most importantly, we expect it to make a smaller correction for systematic errors in $\mathbf{x}_g^b$. This is a desirable behavior in situations where the lower resolution of the global model leads to the type of (possibly flow dependent) systematic error in $\mathbf{x}_g^b$ (e.g., Baek et al. 2006, 2009; Danforth et al. 2007) that can lead to a strong model “spinup” if corrected by the data assimilation. In such situations, $\mathbf{x}_g^a$ could also be a more accurate analysis of the atmospheric state than if it were obtained from a conventional global analysis.

While REG DA can be implemented on any type of DA scheme (e.g., EnKF or 4D-Var), the approximations that have to be made in a practical implementation of the approach depend on the type of the DA scheme. For a discussion of the roots of this dependence, we first recall that the observation function plays a dual role in data assimilation: it is used for the computation of both the innovations and the forecast error covariances in observation space. While the type of the DA scheme has no effect on the computation of the innovations, there is a fundamental difference between the ways EnKF and 4D-Var estimate the forecast error covariances. We expect REG DA to have a positive impact on the analyses and ensuing forecasts primarily because of the more accurate computation of the innovations rather than the more accurate prediction of the error covariances. As the focus of the present paper is on 4D-Var, we only note that an implementation of REG DA on an EnKF would have a limited impact on the estimates of the covariances, because without increasing the number of ensemble members, those estimates would be unlikely to significantly benefit from the availability of higher-resolution information in the limited-area domains. In the case of a 4D-Var, the limiting factor is the resolution of the tangent linear model (TLM) used in the computation of the covariances (see appendix). In particular, in the operational practice of 4D-Var, the TLM and its adjoint are integrated at a reduced resolution, compared to that of the original model. For instance, while the operational resolution of NAVGEM is T425 (1278×639 Gaussian grid), the operational resolution of the TLM in NAVDAS-AR is only T119. At an operational center that currently uses a 4D-Var for limited-area data assimilation, the TLM of the limited-area model is already available, and REG DA could be used to produce higher-resolution predictions of the covariances in the limited-area domains [see Eq. (A14) in the appendix]. Maintaining the TLM (and its adjoint) of the limited-area model, however, would add to the development cost on the long run. Hence, we propose an implementation, in which only the TLM (and its adjoint) of the global model is used. In such an implementation,

the savings in computer wall-clock time from not doing limited-area data assimilation could be used to increase the resolution of the global TLM (and its adjoint). The global analysis would benefit directly, while the limited-area analysis would benefit indirectly from the increased resolution of the TLM (and its adjoint).

3. Implementation on NAVGEM and COAMPS

Our implementation of REG 4D-Var on the Navy’s operational forecast system allows for the use of multiple limited-area domains. We show results for a configuration that has three limited-area domains: a North American, a northeast Pacific, and a European domain (Fig. 1). The NAVGEM global and the COAMPS limited-area model states are blended on a common high-resolution grid rather than the COAMPS grid.

a. Outline of the implementation

The implementation of REG 4D-Var on the Navy’s operational global and limited-area models and global data assimilation system has the following top-level steps (Fig. 2):

- 1) Preparation of the short-term NAVGEM forecast that provides the global background $\mathbf{x}_g^b(t_j)$, $j = 0, \dots, N$, where N is the total number of times at which observations are available in the assimilation window.
- 2) Preparation of the short-term COAMPS forecasts that provide the limited-area background $\mathbf{x}_\ell^b(t_j)$, $j = 0, \dots, N$, for each region.
- 3) Interpolation of both the NAVGEM and the COAMPS forecast fields onto a common high-resolution grid.
- 4) Blending $\mathbf{x}_g^b(t_j)$ and $\mathbf{x}_\ell^b(t_j)$, $j = 0, \dots, N$, in each limited-area region according to Eq. (2).
- 5) Computation of the analysis increment $\delta\mathbf{x}_g^a(t_0)$ by NAVDAS-AR.
- 6) Preparation of the global analysis $\mathbf{x}_g^a(t_0)$ by adding correction to global background.
- 7) Preparation of the limited-area analyses $\mathbf{x}_\ell^a(t_0)$ by interpolating global analysis to regional model’s grid.

Because the observation time window of NAVDAS-AR is a 6-h-long period centered at the analysis time, the data assimilation time window for the background state is from forecast time 3–9h relative to the previous analysis time. NAVGEM and COAMPS both provide hourly outputs, so a composite state is created for each of the 7h of the observation time window, which means that $N = 6$ for steps 1, 2, and 4.

In step 5, we turn on the digital filter initialization option of NAVDAS-AR to ensure proper balance

FIG. 1. The three COAMPS regions used in the experiments.

between the momentum and mass field in the analysis increment. The digital filter of NAVDAS-AR is based on the algorithm of Lynch and Huang (1992) and Lynch et al. (1997) and is applied during the last iteration step of the inner loop of NAVDAS-AR to the TLM solution (T. Rosmond 2017, personal communication). In step 6, the analysis increment is added to the NAVGEM background to create a REG analysis, which also serves as the initial condition of the NAVGEM forecast. The COAMPS initial conditions for the three domains are obtained by the interpolation of the same REG analysis.

b. The interpolation procedures

The common grid used for the blending of the global and limited-area states must have a resolution similar to that of the limited-area model, because otherwise, the data assimilation system could not take advantage of the higher-resolution forecast information in the computation of the innovations. Both the global and limited-area model fields are interpolated onto this common grid. Because NAVGEM is a spectral transform model, for the interpolation of the NAVGEM fields, we first prepare a higher-resolution spectral representation of the fields by zero padding, that is, by adding higher-wavenumber components with zero spectral coefficients to the spectral representation of the fields, then transforming the fields onto a higher-resolution Gaussian grid (the common grid).

Because COAMPS is a finite-difference model that represents the fields on a grid in all phases of the model computations, the interpolation of the COAMPS fields

FIG. 2. Schematic illustration of the implementation of REG 4D-Var on the U.S. Navy's models and DA system. The boxes highlighted in green are excluded for the blend-skip experiment.

requires a grid-space interpolation in both the horizontal and the vertical directions. We use a nearest-neighbor approach for the horizontal interpolation. The vertical interpolation is more involved than the horizontal interpolation, because the two models use different vertical coordinates. In particular, NAVGEM uses a hybrid sigma-pressure coordinate with a pressure-based definition of sigma $\sigma = p/p_s$, where p_s is the surface pressure. The coordinate transitions gradually from pure sigma at Earth's surface to pure pressure at the top of the model atmosphere (Hogan et al. 2014). COAMPS, in contrast, uses a height-based sigma coordinate $\sigma = H(z - z_s)/(H - z_s)$, where z_s is the terrain height, and H is the finite depth of the model atmosphere. The vertical interpolation is performed after completion of the horizontal interpolation of the COAMPS fields on the height-based sigma surfaces of the model. The hybrid sigma-pressure coordinate of each grid point is computed based on the horizontally interpolated COAMPS pressure and surface pressure fields. For this calculation, the horizontally interpolated surface pressure field is corrected for the difference between the orography of NAVGEM and the orography of COAMPS at each horizontal gridpoint location. The procedure follows that of Baek et al. (2009); that is, the corrected surface pressure $\hat{p}_s$ is obtained by

$$\hat{p}_s = p_s e^{g\Delta z/R_d \bar{T}}, \quad (6)$$

where p_s is the original surface pressure, g is standard gravity (9.8 m s^{-2}), Δz is the terrain difference between the two models, R_d is the gas constant for dry air ($287 \text{ J kg}^{-1} \text{ K}^{-1}$), and $\bar{T}$ is the mean temperature of the $|\Delta z|$ deep (hypothetical) atmospheric layer, which we estimate by

$$\bar{T} = \int_0^{\Delta z} T(z) dz = T_s + \frac{1}{2} \gamma \Delta z. \quad (7)$$

FIG. 3. Illustration of the difference between the COAMPS and the NAVGEM surface pressure before and after the correction of the COAMPS surface pressure for the orography differences. The surface pressure values, valid for 0900 UTC 1 Oct 2012, are shown along the latitude segment indicated by a black line in the insert of the top panel. (top) The surface pressure values for NAVGEM and for COAMPS before and after the correction for the orography difference. (bottom) The difference between the NAVGEM surface pressure and the corrected COAMPS surface pressure.

In Eq. (7), γ is the standard atmospheric lapse rate ($6.5 \times 10^{-3} \text{ K m}^{-1}$), and T_s is the surface temperature. Because the temperature is available only at model levels, we estimate the surface temperature by Eq. (A5) of Baek et al. (2009):

$$T_s = T_1 \left(1 - \frac{\gamma R \ln \sigma_1}{g + \frac{1}{2} \gamma R \ln \sigma_1} \right), \quad (8)$$

where σ_1 and T_1 are the values of sigma and the temperature at the lowest NAVGEM level, respectively. Figure 3 illustrates the effect of the correction for the orography difference on the COAMPS surface pressure field. The figure shows that the correction efficiently removes the direct effect of the orography difference. The remaining differences reflect differences in the evolution of the atmospheric flow between the two models, which may include indirect effects of the orography differences that we hope to take into account in our data assimilation approach.

After the computation of the hybrid sigma-pressure coordinate of each grid point of the horizontally interpolated COAMPS fields, the COAMPS fields are interpolated linearly onto the hybrid sigma-pressure surfaces of NAVGEM. A seemingly problematic aspect of this interpolation is that it does not allow for the computation

of the COAMPS fields at locations where the NAVGEM vertical coordinate is associated with a pressure that is higher than the pressure at the lowest COAMPS vertical level. This is the situation at locations where the orography is lower in NAVGEM than in COAMPS. At such locations, our interpolation procedure would introduce discontinuities into the interpolated COAMPS fields, as shown in Fig. 4. These potential discontinuities do not pose a problem in REG 4D-Var, because the interpolated fields are used only for the computation of the innovations, and no observation is expected to fall below the more realistic orography of COAMPS. We did verify that all observations satisfied this condition.

The final step is the computation of the REG background state vector by Eq. (2). In each experiment described in this paper, we use the same value of the blending coefficient α for all COAMPS domains, but near the lateral boundaries (15 grid points from the edge of the domains) and the top of the COAMPS model atmosphere, α is gradually tapered to zero in order to avoid introducing boundary effects into the analysis of the atmospheric fields.

4. Experiment design

Three types of experiments are discussed in this paper (Table 1). While all three are carried out at horizontal

FIG. 4. Lowest COAMPS zonal wind field interpolated onto the lowest hybrid sigma-pressure surface of NAVGEM, valid for 1800 UTC 31 Oct 2012. The dots mark locations where the terrain is higher in COAMPS than NAVGEM.

resolution T119 of NAVGEM (360×180 Gaussian grid) and 32 km of COAMPS and produce a T119-resolution global analysis, they differ in the ways they use NAVDAS-AR. The first type of experiment is a *control experiment*, in which the global analysis is prepared by the conventional configuration of NAVDAS-AR, integrating the TLM of NAVGEM at a reduced resolution of T47. The goal of this experiment is to simulate the feature of the operational DA system that the resolution of the TLM is much lower than that of forecast model. In the second type of experiment, which we call the *blend-skip experiment*, the TLM is integrated at an increased horizontal resolution of T119, and a full implementation of the REG DA code is used with a blending coefficient of $\alpha = 0$. The purpose of this experiment is to test the effects of an increased resolution of the TLM, assess the potential detrimental effects of the extra interpolation steps of REG DA, and provide a clean baseline for the evaluation of the third type of experiments, in which REG DA is used with $\alpha = 0.3$ (30%-blend experiment), $\alpha = 0.5$ (50%-blend experiment), and $\alpha = 1.0$ (100%-blend experiment) blending coefficients.

We chose the particular limited-area domains shown in Fig. 1, because they provide a representative sample of the terrains of the limited-area domains routinely used by the U.S. Navy. The analyses of the experiments are obtained by continuously cycling

NAVDAS-AR for the period 0000 UTC 1 October–0000 UTC 1 November 2012. Analyses are prepared every 6 h by assimilating all observations that were assimilated in real time by the FNMOC of the U.S. Navy.

5. Results

We first present diagnostics of the qualitative behavior of REG 4D-Var, and then we show an example of the analysis increment produced by the technique. Finally, we present diagnostics of the forecast accuracy for NAVGEM for the first 120 forecast hours and for COAMPS for the first 72 forecast hours.

a. Convergence and balance

These diagnostics are prepared to verify that the implementation of REG 4D-Var has no adverse effects on the minimization procedure of NAVDAS-AR. We first examine the speed of the convergence of the minimization algorithm (Fig. 5), as an occasional significant increase of the number of iterations, or a complete lack of convergence of the minimization process, would indicate a failure of REG 4D-Var. The blend-skip and nonzero-blend experiments require, on average, only five additional iterations, compared to the average 50 iterations of the control, which indicates that while the increased resolution of the TLM leads to a slight increase of the necessary iteration steps, the blending of COAMPS information does not lead to an additional increase.

The next set of diagnostics describes the effects of REG 4D-Var on the balance between the mass and the wind field in the analyses (Fig. 6). A standard approach of numerical weather prediction to detect lack of balance in the atmospheric initial conditions is to compute the surface pressure tendency for the first few forecast hours: an initially large, then quickly decaying, surface pressure tendency indicates a rapid adjustment of the mass field, that is, a lack of balance in the initial conditions. The figure shows the evolution of the maximum and minimum (largest magnitude negative) of the surface pressure tendency in the first 6 h for each NAVGEM (global) forecast: each curve is for a different initial

TABLE 1. Summary of the model parameters of the experiments.

	Global resolution	Inner loop resolution	Regional resolution	Intermediate resolution	Blending weight
Control	T119L50	T47	—	—	—
Blend-skip	T119L50	T119	—	T319	—
REG 4D-Var 30%	T119L50	T119	32 km	T319	0.30
REG 4D-Var 50%	T119L50	T119	32 km	T319	0.50
REG 4D-Var 100%	T119L50	T119	32 km	T319	1.00

FIG. 5. The number of iterations of the inner loop required to achieve the prescribed reduction of the cost function. Shown are the number of iterations for the control, the blend-skip, and the 30%-, 50%-, and 100%-blend experiments.

condition (time). The results are very similar for the 50%-blend (top row) and control (bottom row) experiments. The only region in which the 50%-blend experiment shows larger initial tendencies than the control experiment for a few forecasts is the northeast Pacific domain. Because the magnitudes of the tendencies do not indicate serious balance issues even in those exceptional cases, we conclude that blending has no

significant effect on the atmospheric balance in the analyses.

b. An example of the analysis increment

An example of the virtual potential temperature analysis increment produced by our system is shown in Fig. 7 for the model level nearest to the surface. The figure shows that the difference between the two increments is

FIG. 6. Maximum and minimum (largest magnitude negative) surface pressure tendency for the first 9 h of the NAVGEM forecasts for the (top) control and (bottom) 50%-blend experiments. Each pair of top and bottom panels show results for a different region. The different lines of each panel are for the 63 different forecasts verified in the study.

FIG. 7. Virtual potential temperature analysis increment for the (top) blend-skip and (middle) 50%-blend experiments at the model level nearest to the surface. (bottom) The difference between the two increments. The thick black lines mark the boundaries of the three limited-area domains. Valid for 0600 UTC 1 Oct 2012.

confined to the limited-area domains. Interestingly, the increment often has a larger local magnitude and/or the opposite sign for the 50%-blend experiment (middle panel) than for the blend-skip experiment (top panel). The local differences between the two increments (bottom panel) tend to be the largest in regions of high and/or complex orography (e.g., in the European domain). These are locations where the limited-area background typically has more realistic information than the global background about the distance of the pressure levels of the observations above the surface of the Earth. Observe, for instance, the pressure difference between the 900-hPa level and the surface pressure of the two models in Fig. 3. Figure 7 suggests that these are the locations where REG DA (the COAMPS background information) has the largest impact on the analyses.

c. Measures of the forecast errors

We calculate mean-square-error (MSE)-based diagnostics of the forecast performance by using operational analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF) as verification data, $1^\circ \times 1^\circ$ for NAVGEM verification, and $0.5^\circ \times 0.5^\circ$ for COAMPS verification. We calculate the MSE for each state variable and the standard pressure levels in two different ways. We either calculate a single scalar measure ε of the forecast error by

$$\varepsilon = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N (x_{ti}^f - x_{ti}^v)^2, \quad (9)$$

where N is the number of verification locations (grid points), T is the number of forecasts (verification times), x_t^f is the forecast, and x_t^v is the verifying analysis of forecast variable x (e.g., geopotential height) valid at time t ; or we produce a map of the spatial distribution of the errors by displaying

$$\varepsilon_T = \frac{1}{T} \sum_{t=1}^T (x_t^f - x_t^v)^2. \quad (10)$$

For all the results shown using either of these measures, T is equal to 63, the number of forecasts over the entire experiment period that were verified against ECMWF analyses.

We test the statistical significance of the MSE difference between a pair of experiments by using the test statistic

$$z = \left[\frac{1}{T} \sum_{t=1}^T (\text{MSE}_{1,t} - \text{MSE}_{2,t}) \right] / \left(\frac{s_{\text{MSE}_1 - \text{MSE}_2}^2}{T'} \right)^{1/2}, \quad (11)$$

where $\text{MSE}_{1,t}$ and $\text{MSE}_{2,t}$ are the values of the MSE for the two experiments at verification time t , and $s_{\text{MSE}_1 - \text{MSE}_2}^2$ is the sample variance of the difference $\text{MSE}_{1,t} - \text{MSE}_{2,t}$. The effective sample size T' is estimated by

$$T' \approx T \frac{1 - \rho_1}{1 + \rho_1}, \quad (12)$$

where ρ_1 is the lag-1 autocorrelation coefficient for the time series of error differences $\text{MSE}_{1,t} - \text{MSE}_{2,t}$. This test statistic z is then used to find a confidence probability p by using left-tail cumulative probabilities for the standard Gaussian distribution. We mark all MSE differences for which $p < 5\%$ as statistically significant with hashed boxes. We emphasize, however, that p should not be interpreted as the probability that the mean difference between the errors is zero, or that the differences between the errors of the two experiments were produced by random chance alone (Wasserstein and Lazar 2016). The proper interpretation of a statistically significant positive difference is that it is sufficiently large, compared to the sample variance, to warrant further investigations into the forecast benefits of REG 4D-Var.

d. Impact diagnostics

We refer to diagnostics that show the MSE difference between two experiments in percentage format as *impact diagnostics*. For instance, when the results of experiment A are compared to the results of experiment B , the impact diagnostics are computed by $(\varepsilon_B - \varepsilon_A)(\varepsilon_B)^{-1} \times 100$, where ε_A is the MSE ε for a particular forecast variable in experiment A , and ε_B is the MSE ε for the same forecast variable in experiment B . Notice that the difference is taken such that a positive value of the diagnostics indicates that the forecasts of experiment A are more accurate on average than the forecasts of experiment B . While we include results for forecast time 0 h (the analysis time) in the figures, those results should be interpreted with some care, because the errors of the verifying ECMWF analyses are statistically not completely independent of those of our analyses, as they assimilate similar sets of observations.

Figure 8 is the first of a series of impact diagnostics figures. It compares the geopotential height forecast error of the control with those of the blend-skip and 50%-blend experiments. The different panels show the results for different regions: the left panel is for the entire Northern Hemisphere extratropics, while the next three panels are for the three regions (the name of each region is indicated above the top panel). The forecasts of the blend-skip experiment are clearly more accurate than the forecasts of the control, which confirms our

expectation that the quality of the global analyses would benefit from an increase of the resolution of the TLM and its adjoint. It also indicates that the improvements from the higher-resolution TLM and adjoint outweigh the potential degradations from the interpolation errors of our REG DA implementation. Similar to the forecasts of the blend-skip experiment, the forecasts of the 50%-blend experiment are clearly more accurate than those from the control.

In what follows, we investigate the forecast effects of REG DA by computing the impact diagnostics for each experiment that uses a particular $\alpha > 0$ value of the blending coefficient and the blend-skip experiment. The impact diagnostics we show are calculated such that positive values in the figures indicate positive forecast impact for the particular $\alpha > 0$ value of the blending coefficient. Because the TLM is integrated at resolution T119 in both types of the experiments that we compare, the forecast impact shown in these figures is solely due to the use of COAMPS information in the data assimilation step of the particular experiment.

Figures 9–12 each summarize the results for a different state variable. The format of these figures is similar to that of Fig. 8, except they now have six rows of panels. The six rows consist of three pairs of figures for NAVGEM and COAMPS, with each pair showing results for a different $\alpha > 0$ value of the blending coefficient. In addition, the first column for COAMPS shows results for a combination (composite) of the three limited-area domains rather than the NH extratropics. The results of this column represent the overall performance of REG DA. In this column, the diagnostics change more smoothly between the forecast lead times and vertical levels than in the other columns, because they are based on larger spatial samples of the errors.

Figure 9 can be directly compared to Fig. 8, as it also shows results for the geopotential height: blending clearly leads to more mixed results (Fig. 9) than increasing the resolution of the TLM (Fig. 8). But the results are dominantly positive for the 30%- and 50%-blend experiments for both NAVGEM and COAMPS. For these blending coefficients, the number of lead times and vertical levels with statistically significant positive forecast impact is about the same for NAVGEM and COAMPS, but negative forecast impact occurs more rarely for COAMPS. Most importantly, for the composite region of COAMPS, degradations occur only at analysis time and only in the upper troposphere, and none of the degradations is statistically significant. In contrast, for NAVGEM in the NH extratropics, some of the degradations in the upper troposphere at analysis time are statistically significant and also present at 12-h lead time, which suggests that the degradations detected

FIG. 8. Geopotential impact diagnostics comparing the global forecasts of the (top) blend-skip and 50%-blend experiments to those of the control experiment. Red shades indicate forecasts that are more accurate for the blend-skip or 50%-blend experiments, while blue shades indicate forecasts that are more accurate for the control. Statistically significant values ($p < 5\%$) are hatched.

at analysis time may not be pure artifacts of the verification. In general, the forecast improvements propagate from the lower to the upper troposphere as forecast lead time increases, and there is considerable variability in the performance of both NAVGEM and COAMPS between the different limited-area regions.

The performance of the models in the 100%-blend experiment is, in general, poorer than in the 30%- and 50%-blend experiments. The degradation that results from using only the COAMPS background information in the limited-area domains is particularly notable in the higher-tropospheric layers (starting at 700 hPa) in the first 36 h of the NAVGEM forecasts. In the lower layers, at the shortest forecast times, the overall improvements in the 100%-blend experiments are just as large as or even larger than in the 50%-blend experiment. These results suggest that while the use of the higher-resolution COAMPS background information helps the interpretation of the observations near the surface, the poorer COAMPS background information about the larger-scale processes in the free atmosphere leads to a degradation of the analyses in the upper layers. In the northeast Pacific domain, which is dominantly an oceanic region at the exit of the Pacific storm track, the forecasts of the 100%-blend experiment are degraded in the lower layers, which also supports this conclusion.

The forecast impact of REG DA on the temperature forecasts (Fig. 10) is weaker than on the geopotential height forecasts. Otherwise, there are several similarities

between the results for the two state variables: the results are positive for the 30%- and 50%-blend experiments and mixed at best for the 100%-blend experiment, and there are fewer lead times and vertical levels of degradation for COAMPS than NAVGEM. Most notably, the degradations at the lowest two levels in the NAVGEM forecasts at all forecast times are absent in the COAMPS forecasts. There is considerable variability in the forecast performance among the three limited-area domains. Interestingly, degradations and statistically significant improvements are both found more frequently for NAVGEM than COAMPS. For instance, REG DA performs particularly well for NAVGEM in the northeast Pacific domain, which covers a dominantly oceanic region. In addition, the magnitude of the low-level degradations for NAVGEM is the largest in the European domain, the domain that has the most complex orography. The most likely explanation for this result is that it is an artifact of the verification approach. In a mountainous region, the surface pressure is often lower than 1000 or even 925 hPa (Fig. 3). The extrapolated (hypothetical) “below ground” value of the temperature is more sensitive to the presumed lapse rate than the extrapolated value of the pressure. In addition, because the NAVGEM orography is often lower than the orography of the verification dataset, the verification is based on comparing an actual predicted temperature to an extrapolated value at several locations. The possibility that blending truly leads to degradation of the temperature forecasts near

FIG. 9. Impact diagnostics comparing the blend-skip and the (top two rows) 30%-, (middle two rows) 50%-, and (bottom two rows) 100%-blend experiments for the geopotential height. In each pair of rows, the top shows results for NAVGEN, and the bottom shows results for COAMPS. Each column shows results for a particular region. Red shades indicate forecasts that are more accurate for nonzero blending, while blue shades indicate forecasts that are more accurate for the blend-skip experiment. Statistically significant values ($p < 5\%$) are hatched.

FIG. 10. As in Fig. 9, but for temperature.

the surface, however, cannot be ruled out unequivocally. In that case, it is possible that the degradation can be reduced or completely eliminated by the refinement of the interpolation in the blending process.

This is an issue that we hope to further investigate in the future.

The impact diagnostics results are quite positive for the two horizontal components of the wind (Figs. 11 and 12).

FIG. 11. As in Fig. 9, but for the zonal component of the wind.

Similar to the results for the geopotential height and temperature, the forecasts of the 30%- and 50%-blend experiments perform much better than the forecasts of the 100%-blend experiment. Particularly notable is the good performance of REG DA for COAMPS in the

composite domain in the 30%-blend experiment. Similar to the results for the geopotential height, the statistically significant forecast improvements tend to propagate from the lower to the upper troposphere as forecast time increases.

FIG. 12. As in Fig. 9, but for the meridional component of the wind.

In summary, the impact diagnostics suggest that REG DA with a 30% or 50% blending coefficient has an overall positive impact on both the NAVGEN and COAMPS forecasts. The forecast impact for the

100% blending coefficient is more negative. This result suggests that while using some short-term limited-area forecast information in the calculation of the innovations has a positive forecast impact, using

only limited-area forecast information has a mixed impact.

e. Systematic and transient errors

To investigate the origin of the forecast improvements from REG DA, we decompose the MSE ε into an error variance V and a square of the bias b^2 component. That is, $\varepsilon = V + b^2$, where

$$b = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N x_{ti}^f - x_{ti}^v, \quad (13)$$

and

$$V = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \{[(x_{ti}^f - x_{ti}^v) - b]\}^2. \quad (14)$$

The square of the bias b^2 is a measure of the magnitude of the persistent part of the systematic errors, while the error variance describes the magnitude of the transient component of the errors, which includes both the random errors and the flow-dependent part of the systematic errors.

We show results for the geopotential height at the 500-hPa pressure level (Fig. 13). The organization of the rows and columns of this multipanel figure is the same as that of the impact figures (Figs. 9–12). We use a logarithmic scale on the y axis of the panels, because the amplification of the magnitude of the random errors by the chaotic dynamics of the atmosphere leads to an exponential growth of the transient errors in the first few forecast days. The error components for the experiments with $\alpha > 0$ blending coefficients are shown relative to the respective error components for the blend-skip experiment. A positive value indicates lower errors in the experiment with $\alpha > 0$ than in the blend-skip experiment, while a negative value indicates higher errors in the experiment with $\alpha > 0$. While we include results for the 100%-blend experiment in the figures for the sake of completeness, our discussion below is based on the results for the 30%- and 50%-blend experiments.

The results suggest a general picture in which the forecasts of the experiments with $\alpha > 0$ are more accurate than the forecasts of the blend-skip experiment. The advantage of the $\alpha > 0$ experiments grows with the forecast lead time for the entire 72 h of the COAMPS forecasts and the first 96 h of the NAVGEM forecasts. The fact that this advantage largely disappears after 96 h for NAVGEM suggests that the injection of COAMPS information in NAVDAS-AR leads to improvements of the global forecasts primarily at the smaller scales, at which the predictability time limit is shorter.

The impact of REG DA on the variance of the analysis and 12-h lead time forecast errors is positive for

COAMPS and neutral for NAVGEM. This result suggests that the limited-area initial conditions benefit more from the filtering of random errors and the correction of flow-dependent systematic errors. The impact on the variance of the forecast errors becomes increasingly more positive as forecast time increases in the NH extratropics for NAVGEM and the composite domain for COAMPS. The typical trend for the individual domains is similar, but there are some exceptions to that rule at the shorter forecast times. The longer-term (48 h for COAMPS and 96 h for NAVGEM) positive impact of REG DA on the MSE is dominantly due to the positive impact on the error variance.

There is more variability in the impact of REG DA on the bias (mean systematic errors) than the variance among the models, verification regions, and forecast lead times. In particular, while the impact on the bias is negative in the first 48 h for NAVGEM, it becomes positive for COAMPS by the 24-h forecast lead time. Particularly notable is the difference between the behavior of NAVGEM and COAMPS in the European domain: while the impact never becomes positive for NAVGEM, it is not only positive for COAMPS, but also makes a significant contribution to the positive impact on the MSE. This result suggests that REG DA can be particularly beneficial for the COAMPS forecasts in regions of such complex orography as Europe. This is a highly encouraging result, because reducing systematic errors in regions of complex orography is one of the main motivations to use a limited-area data assimilation system for the downscaling of the initial conditions. We also note that the negative impact on the bias component of the NAVGEM forecast errors is not necessarily a negative result for REG DA. In fact, in the presence of model errors whose correction by the data assimilation would lead to a strong model “spinup,” this is an expected tradeoff to reduce the error variance and the MSE (Baek et al. 2006, 2009).

Interestingly, there is a limited-area domain—the North American domain—where REG DA has a positive impact on the bias of the NAVGEM forecasts after 24 h. We use this domain to illustrate the impact of REG DA on the spatial distribution of the error components. We first show the impact on the error components for the geopotential height at several pressure levels (Fig. 14). The bias reduction over the Rocky Mountains and Sierra Madres is confined to the atmospheric layer between 925 and 1000 hPa. Above that layer, the reduction in bias turns into a widespread increase of the bias. We speculate that this increase of the bias is introduced by injecting strong boundary effects from the COAMPS model, which has a top at 4 hPa, into the NAVGEM, which has a top at 0.04 hPa. We hope that

FIG. 13. Evolution of the components of the mean square of the 500-hPa geopotential height forecast error with forecast lead time. Shown are the results for the difference between the blend-skip and the (top two rows) 30%-, (middle two rows) 50%-, and (bottom two rows) 100%-blend experiments. In each pair of rows, the top shows results for NAVGEM, and the bottom shows results for COAMPS. Each column shows results for a particular region.

this bias can be greatly reduced or completely eliminated by a more rapid tapering of the blending coefficient approaching the top of the COAMPS model. The reduction of the error variance for the tropical

cyclones (off the East Coast), however, extends into the upper troposphere.

Figure 15 shows the evolution of the 500-hPa diagnostics from Fig. 14 with forecast lead time. At analysis time, there

FIG. 14. Vertical structure of the decomposition of the MSE difference between the blend-skip and 50%-blend experiments for the geopotential height at 12-h forecast lead time. The different rows show results for different pressure levels, with pressure decreasing from bottom to top. Shown are (left) the MSE difference, (middle) difference in the square bias, and (right) difference in the variance. Red shades indicate improvement, and blue shades indicate degradation due to nonzero blending.

are widespread bias reductions over the mountains, but these improvements disappear by the 12-h lead time. The bias reductions re-emerge at 24-h lead time, and their magnitude and the area they cover continue to grow as forecast time increases. The error variance shows a similar behavior along the tropical cyclone paths off the East Coast.

Diagnostics similar to those shown in Figs. 13–15 for the geopotential height were also prepared for the air temperature and the zonal and meridional wind components. We found that the notable changes in the air temperature forecasts were mainly below 700 hPa, with the largest differences at the surface. As an example, Fig. 16 shows the time evolution of the error components for the temperature at 1000 hPa. It should be noted that the 1000-hPa temperature field is obtained by extrapolation at the many locations where the surface pressure is lower than 1000 hPa. The errors in the prediction of this field, nevertheless, are a good measure of the performance of the model in predicting the air temperature near the surface. One particularly interesting feature of the figure is the large reduction of bias over Louisiana that has a rapidly growing magnitude with increasing forecast time. An inspection of the individual forecasts that contributed to this result revealed that the large bias correction was due to the systematic improvement of the prediction of frontal passages in the region.

6. Conclusions

We described the results of our first attempt at building a global data assimilation system that produces initial conditions for an operational global and limited-area model in multiple domains simultaneously. The motivation to build such a data assimilation system is the hope that it would eventually eliminate the necessity of maintaining a limited-area data assimilation system and doing limited-area data assimilation for multiple limited-area domains. The possibility of eliminating limited-area data assimilation is appealing not only because it would reduce development and computational cost, but also because limited-area data assimilation has fundamental limitations. Most importantly, limited-area data assimilation has no observational or model information about the large-scale atmospheric flow outside the limited-area domain, which makes the proper interpretation of the large-scale component of the flow inside the limited-area domain highly challenging. In practice, this challenge is addressed by partial cycling of the limited-area data assimilation, which is a less-than-ideal solution from both a theoretical and practical point of view. Our hope is that the unified data assimilation system would produce higher-quality limited-area initial conditions than partial cycling. The results of the present paper show that REG DA can produce limited-area initial conditions that lead to more accurate forecasts than the limited-area initial

conditions obtained by the interpolation of a conventional global analysis. Particularly encouraging is the result that REG DA reduced the systematic limited-area forecast errors over complex topography. It is yet to be seen whether or not the magnitudes of these error reductions are comparable to, or even larger than, those from a partially cycled limited-area data assimilation.

The savings in development and computational time from eliminating limited-area data assimilation could be used for the enhancement of the unified data assimilation system, which itself could greatly improve the global initial conditions. As an example, we showed that the performance of the U.S. Navy's global data assimilation system (NAVDAS-AR) could be improved by increasing the resolution of the TLM (and its adjoint) of the global NAVGEM in data assimilation. It is important to point out that the other improvements from REG DA found here are in addition to the improvements from the increased resolution of the TLM. These additional improvements include a small but consistent overall improvement of the global forecasts in addition to the improvements of the limited-area forecasts.

Our technical approach for REG DA was motivated by the composite state method of Kretschmer et al. (2015). Their paper showed for a highly idealized model of the atmosphere that doing data assimilation by estimating a state vector, defined by a linear combination of the state vectors of the global model and the limited-area model in the limited-area domains, produced global and limited-area initial conditions that produced more accurate forecasts than the separate global and limited-area data assimilations. The particular 4D-Var-based implementation of REG DA tested in this paper, which we call REG 4D-Var, computes the innovations based on a model state defined by a linear combination of the state vectors of the global model and the limited-area model in the limited-area domains, but unlike Kretschmer et al. (2015), it updates a background from a global model integration. We find the results with the present implementation sufficiently encouraging to proceed to the testing of other strategies for the generation of the background. The list of other potential refinements to the technical approach includes the adjustment of the prescribed observation error variance in regions where higher-resolution limited-area model information is available (in expectation of reduced representativeness errors) and using blending coefficients that are optimized independently for each domain. In addition, the Navy has recently transitioned to using a hybrid 4D-Var system, which uses an ensemble-based approach to enhance the estimate of the background error covariance matrix. REG 4D-Var could be easily implemented on the hybrid data assimilation system.

FIG. 15. Temporal evolution of the MSE difference between the blend-skip and 50%-blend experiments for the 500-hPa geopotential height. The different rows show results for different lead times. Shown are (left) the MSE difference, (middle) difference in the square bias, and (right) difference in the variance. Red shades indicate improvement, and blue shades indicate degradation due to nonzero blending.

FIG. 16. As in Fig. 15, but for the 1000-hPa temperature.

Acknowledgments. This study was conducted in collaboration with two of the Naval Research Laboratory institutions, NRL Monterey and NRL D.C. Funding was provided from a Texas A&M University Diversity Fellowship and ONR Grants N00014-12-1-0785, N00014-16-1-3011, and N00014-18-1-2509 and NRL Award N00173-16-1-3011.

APPENDIX

REG DA for Incremental 4D-Var

We describe the implementation of REG DA on an incremental 4D-Var scheme for a single iteration of the outer loop.

$$J[\delta \mathbf{x}_g(t_0)] = [\delta \mathbf{x}_g(t_0)]^T \mathbf{P}_b^{-1} \delta \mathbf{x}_g(t_0) + \sum_{j=0}^N [\delta \mathbf{y}^o(t_j) + \mathbf{H}_g(t_j) \mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)]^T \times \mathbf{R}_j^{-1} [\delta \mathbf{y}^o(t_j) + \mathbf{H}_g(t_j) \mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)], \quad (\text{A3})$$

and adding it to the background $\mathbf{x}_g^b$, where $\mathbf{H}_g(t_j)$ is the linearization of $\mathbf{h}_g[\mathbf{x}_g(t_j)]$ about $\mathbf{x}_g^b(t_j)$; $\mathbf{M}_g(t_0, t_j)$ represents the TLM dynamics between times t_0 and t_j for the nonlinear state space trajectory associated with the initial condition $\mathbf{x}_g^b(t_0)$. In Eq. (A3), the expression in the square brackets of the second (observation) term approximates $\mathbf{y}^o(t_j) - \mathbf{h}[\mathbf{x}_g(t_j)]$. The approximate equality of the two expressions,

$$\mathbf{y}^o(t_j) - \mathbf{h}_g[\mathbf{x}_g(t_j)] = \delta \mathbf{y}^o(t_j) + \mathbf{H}_g(t_j) \mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0), \quad (\text{A4})$$

follows from

$$\mathbf{x}_g(t_j) = \mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0), \quad (\text{A5})$$

$$\delta \mathbf{y}^o(t_j) = \mathbf{y}^o(t_j) - \mathbf{h}_g[\mathbf{x}_g^b(t_j)], \quad (\text{A6})$$

and

$$\mathbf{h}_g[\mathbf{x}_g(t_j)] \approx \mathbf{h}_g[\mathbf{x}_g^b(t_j)] + \mathbf{H}_g(t_j) \delta \mathbf{x}_g(t_j). \quad (\text{A7})$$

Equation (A3) describes the cost function minimized by the *inner loop* of the algorithm. When multiple steps of the outer loop are taken, the nonlinear trajectory is recomputed at the beginning of each step of the *outer loop*, using the updated background from the previous step as initial condition. This involves recomputing $\mathbf{H}_g(t_j)$ for the updated values of $\mathbf{x}_g(t_j)$.

a. Incremental 4D-Var

In the incremental formulation of 4D-Var (Courtier et al. 1994), the increment $\delta \mathbf{x}_g$ is defined by

$$\delta \mathbf{x}_g(t_j) = \mathbf{x}_g(t_j) - \mathbf{x}_g^b(t_j), \quad (\text{A1})$$

where $t_j, j = 0, \dots, N$ are the times at which observations are available in the assimilation time window, and $\mathbf{x}_g^b(t_j)$ is the global background state at t_j . The analysis

$$\mathbf{x}_g^a(t_0) = \mathbf{x}_g^b(t_0) + \delta \mathbf{x}_g^a \quad (\text{A2})$$

at time t_0 is obtained by computing the analysis increment $\delta \mathbf{x}_g^a(t_0) = \delta \mathbf{x}_g(t_0)$ that minimizes the cost function (e.g., Szunyogh 2014),

b. The effect of replacing $\mathbf{h}_g(\mathbf{x}_g)$ by $\mathbf{h}_e(\mathbf{x}_e)$

The analog of Eq. (A4) for $\mathbf{h}_e(\mathbf{x}_e^b)$ is

$$\mathbf{y}^o(t_j) - \mathbf{h}_e[\mathbf{x}_e(t_j)] = \delta_e \mathbf{y}^o(t_j) + \mathbf{H}_e(t_j) \mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0), \quad (\text{A8})$$

where

$$\delta_e \mathbf{y}^o(t_j) = \mathbf{y}^o(t_j) - \mathbf{h}_e[\mathbf{x}_e^b(t_j)], \quad (\text{A9})$$

and the linearization $\mathbf{H}_e(t_j)$ of $\mathbf{h}_e[\mathbf{x}_e(t_j)]$ about $\mathbf{x}_e^b(t_j)$ satisfies

$$\mathbf{h}_e[\mathbf{x}_e(t_j)] \approx \mathbf{h}_e[\mathbf{x}_e^b(t_j)] + \mathbf{H}_e(t_j) \delta \mathbf{x}_e(t_j), \quad (\text{A10})$$

where

$$\delta \mathbf{x}_e(t_j) = \mathbf{x}_e(t_j) - \mathbf{x}_e^b(t_j) = (1 - \alpha) \mathbf{A}[\mathbf{x}_g(t_j) - \mathbf{x}_g^b(t_j)] + \alpha [\mathbf{x}_e(t_j) - \mathbf{x}_e^b(t_j)]. \quad (\text{A11})$$

Introducing the notation

$$\delta \mathbf{x}_\ell(t_j) = \mathbf{x}_\ell(t_j) - \mathbf{x}_\ell^b(t_j), \quad (\text{A12})$$

Eq. (A11) can be written as

$$\delta \mathbf{x}_e(t_j) = (1 - \alpha) \mathbf{A}[\delta \mathbf{x}_g(t_j)] + \alpha \delta \mathbf{x}_\ell(t_j). \quad (\text{A13})$$

With the help of $\mathbf{M}_g(t_0, t_j)$ and $\mathbf{M}_\ell(t_0, t_j)$, where the latter represents the linearization of the limited-area dynamics

about the nonlinear trajectory $\mathbf{x}_e(t_j)$, Eq. (A13) can be written as

$$\delta \mathbf{x}_e(t_j) = (1 - \alpha) \mathbf{A}[\mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)] + \alpha \mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0). \quad (\text{A14})$$

Making the assumption that the interpolation function $\mathbf{A}(\cdot)$ can be effectively approximated by a linear mapping,¹ Eq. (A14) can also be written as

$$\delta \mathbf{x}_e(t_j) = \mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0), \quad (\text{A15})$$

where the linear model integration, interpolation, and blending are amalgamated into a single linear mapping $\mathbf{M}_e(t_0, t_j)$. This result shows that REG DA can be implemented by replacing $\mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)$ by Eq. (A14) and $\mathbf{H}_g$ by $\mathbf{H}_e$ in the cost function of Eq. (A3): the new cost function is

$$J[\delta \mathbf{x}_g(t_0)] = [\delta \mathbf{x}_g(t_0)]^T \mathbf{P}_b^{-1} \delta \mathbf{x}_g(t_0) + \sum_{j=0}^N [\delta_e \mathbf{y}^o(t_j) + \mathbf{H}_e(t_j) \mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0)]^T \times \mathbf{R}_t^{-1} [\delta_e \mathbf{y}^o(t_j) + \mathbf{H}_e(t_j) \mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0)]. \quad (\text{A16})$$

The second (observation) term of the cost function depends on the control variable $\delta \mathbf{x}_g(t_0)$, which has the same resolution as in the original cost function [Eq. (A3)], through Eq. (A13).

In the operational practice of 4D-Var, the TLM is integrated at a reduced resolution, compared to the resolution of the original model. That is, the resolution of the TLM solution $\mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)$ is lower than that of $\mathbf{x}_g(t_j)$ in the calculation of $\mathbf{h}[\mathbf{x}_g(t_j)]$. The linearized observation function $\mathbf{H}_g(t_j)$ therefore operates on the reduced-resolution $\mathbf{M}_g(t_0, t_j) \delta \mathbf{x}_g(t_0)$ in the observation term of the cost function [Eq. (A3)]. Likewise, when $\mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0)$ is approximated by an integration of the reduced-resolution global TLM, the linearized observation function $\mathbf{H}_e(t_j)$ operates on a TLM field $\mathbf{M}_e(t_0, t_j) \delta \mathbf{x}_e(t_0)$ that has the resolution of the reduced-resolution global TLM.

REFERENCES

- Baek, S.-J., B. R. Hunt, E. Kalnay, E. Ott, and I. Szunyogh, 2006: Local ensemble Kalman filtering in the presence of model bias. *Tellus*, **58A**, 293–306, <https://doi.org/10.1111/j.1600-0870.2006.00178.x>.
- , I. Szunyogh, B. R. Hunt, and E. Ott, 2009: Correcting for surface pressure background bias in ensemble-based analyses. *Mon. Wea. Rev.*, **137**, 2349–2364, <https://doi.org/10.1175/2008MWR2787.1>.
- Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387, <https://doi.org/10.1002/qj.49712051912>.
- Dahlgren, P., and N. Gustafsson, 2012: Assimilating host model information into a limited area model. *Tellus*, **64A**, 15836, <https://doi.org/10.3402/tellusa.v64i0.15836>.
- Daley, R., and E. Barker, 2001: NAVDAS: Formulation and diagnostics. *Mon. Wea. Rev.*, **129**, 869–883, [https://doi.org/10.1175/1520-0493\(2001\)129<0869:NFAD>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0869:NFAD>2.0.CO;2).
- Danforth, C. M., E. Kalnay, and T. Miyoshi, 2007: Estimating and correcting global weather model error. *Mon. Wea. Rev.*, **135**, 281–299, <https://doi.org/10.1175/MWR3289.1>.
- Guidard, V., and C. Fischer, 2008: Introducing the coupling information in a limited-area variational assimilation. *Quart. J. Roy. Meteor. Soc.*, **134**, 723–735, <https://doi.org/10.1002/qj.215>.
- Hodur, R. M., 1997: The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS). *Mon. Wea. Rev.*, **125**, 1414–1430, [https://doi.org/10.1175/1520-0493\(1997\)125<1414:TNRLSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<1414:TNRLSC>2.0.CO;2).
- Hogan, T. F., and Coauthors, 2014: The Navy Global Environmental Model. *Oceanography*, **27** (3), 116–125, <https://doi.org/10.5670/oceanog.2014.73>.
- Kretschmer, M., B. R. Hunt, E. Ott, C. H. Bishop, S. Rainwater, and I. Szunyogh, 2015: A composite state method for ensemble data assimilation with multiple limited-area models. *Tellus*, **67A**, 26495, <https://doi.org/10.3402/tellusa.v67.26495>.
- Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034, [https://doi.org/10.1175/1520-0493\(1992\)120<1019:IOTHMU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2).
- , D. Giard, and V. Ivanovici, 1997: Improving the efficiency of a digital filtering scheme for diabatic initialization. *Mon. Wea. Rev.*, **125**, 1976–1982, [https://doi.org/10.1175/1520-0493\(1997\)125<1976:ITEOAD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<1976:ITEOAD>2.0.CO;2).
- Merkova, D., I. Szunyogh, and E. Ott, 2011: Strategies for coupling global and limited-area ensemble Kalman filter assimilation. *Nonlinear Processes Geophys.*, **18**, 415–430, <https://doi.org/10.5194/npg-18-415-2011>.
- Rogers, E., and Coauthors, 2009: The NCEP North American Mesoscale modeling system: Recent changes and future plans. *23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction*, Omaha, NE, Amer. Meteor. Soc., 2A4, <http://ams.confex.com/ams/pdfpapers/154114.pdf>.
- Rosmond, T., and L. Xu, 2006: Development of NAVDAS-AR: Non-linear formulation and outer loop tests. *Tellus*, **58A**, 45–58, <https://doi.org/10.1111/j.1600-0870.2006.00148.x>.
- Szunyogh, I., 2014: *Applicable Atmospheric Dynamics: Techniques for the Exploration of Atmospheric Dynamics*. World Scientific, 588 pp.

¹Term $\mathbf{A}(\cdot)$ may not be strictly linear, because the vertical interpolation often involves a nonlinear change of coordinates.

- , E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield, and J. A. Yorke, 2008: A local ensemble transform Kalman filter data assimilation system for the NCEP global model. *Tellus*, **60A**, 113–130, <https://doi.org/10.1111/j.1600-0870.2007.00274.x>.
- Wasserstein, R., and N. Lazar, 2016: The ASA's statement on p-values: Context, process, and purpose. *Amer. Stat.*, **70**, 129–133, <https://doi.org/10.1080/00031305.2016.1154108>.
- Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. *Tellus*, **57A**, 546–559, <https://doi.org/10.3402/tellusa.v57i4.14710>.
- Yoon, Y., B. Hunt, E. Ott, and I. Szunyogh, 2012: Simultaneous global and limited-area ensemble data assimilation using joint states. *Tellus*, **64A**, 18407, <https://doi.org/10.3402/tellusa.v64i0.18407>.