Quadratic Polynomial Regression Using Serial Observation Processing: Implementation within DART

DANIEL HODYSS

Naval Research Laboratory, Monterey, California

JEFFREY L. ANDERSON AND NANCY COLLINS

Data Assimilation Research Section, National Center for Atmospheric Research, Boulder, Colorado

WILLIAM F. CAMPBELL AND PATRICK A. REINECKE

Naval Research Laboratory, Monterey, California

(Manuscript received 3 April 2017, in final form 16 August 2017)

ABSTRACT

It is well known that the ensemble-based variants of the Kalman filter may be thought of as producing a state estimate that is consistent with linear regression. Here, it is shown how quadratic polynomial regression can be performed within a serial data assimilation framework. The addition of quadratic polynomial regression to the Data Assimilation Research Testbed (DART) is also discussed and its performance is illustrated using a hierarchy of models from simple scalar systems to a GCM.

1. Introduction

Many ensemble-based Kalman filtering (EBKF) algorithms process the observations serially (e.g., Houtekamer and Mitchell 2001; Whitaker and Hamill 2002; Anderson 2003; Snyder and Zhang 2003). Serial observation processing views the data assimilation process as an iterative sequence of scalar update equations. What is useful about this data assimilation algorithm is that it has very low memory requirements and does not need complex methods to perform the typical high-dimensional inverse calculations of many other algorithms.

Recently, the push has been toward the prediction, and therefore the assimilation of observations, for regions and phenomena for which high resolution is required and/or highly nonlinear physical processes are operating. For these situations, a basic hypothesis is that the use of the EBKF is suboptimal and performance gains could be achieved by accounting for aspects of the non-Gaussianty. To this end, we develop here a new component of the Data Assimilation Research Testbed (DART; Anderson et al. 2009) to allow for a wide variety of users to test this

hypothesis. This new version of DART allows one to run several variants of the EBKF, as well as several variants of the quadratic polynomial filter (Hodyss 2011), while employing the exact same forecast model and observations. Differences between the results of the two systems will then highlight the degree of non-Gaussianity in the system being examined.

This new version of DART will employ the polynomial filtering theory of Hodyss (2011, 2012), which is extended here to serial observation processing data assimilation systems. The key useful feature of polynomial filtering is that its structure is homomorphic to the EBKF. This ensures that changes to already constructed codes are minor and fit within the overall flow. The method essentially takes an augmented state vector approach in which one augments both the states and the observations with new states and synthetic observations corresponding to their squares. As we show below, this process of state augmentation corresponds with quadratic polynomial regression.

Section 2 will review both polynomial regression and serial observation processing, and then will subsequently show how to merge the two methods. Section 3 illustrates the basic ideas in a simple scalar system and shows under what regimes we believe that polynomial

---

Corresponding author: Daniel Hodyss, daniel.hodyss@nrlmry.navy.mil

TABLE 1. Definitions of variables.

First-step variables	
N_o	Total No. of observations
y_i^o	ith observation
y_{ki}^p	kth prior member in observation space for the ith observation
$\bar{y}_i^p$	Prior mean in observation space for the ith observation
$\bar{y}_i^u$	Posterior mean in observation space after the ith observation
σ_i^p	Prior ensemble variance for the ith observation and after assimilation of the previous $(i-1)$ observation
σ_i^0	Same as σ_i^p, but before any observations have been assimilated
σ_i^u	Posterior ensemble variance after the ith observation
T_i^p	Prior third moment for the ith observation
F_i^p	Prior fourth moment for the ith observation
r_i	Observation error variance for the ith observation
r_{3i}	Third moment of the observation likelihood for the ith observation
r_{4i}	Fourth moment of the observation likelihood for the ith observation
Second-step variables	
N_x	Total No. of state variables to be updated
$\bar{x}_j^p$	Prior ensemble mean for the jth variable
$\bar{x}_{ji}^u$	Posterior mean for the jth variable after the ith observation
x_{kj}^p	kth prior member for the jth variable
x_{kji}^u	kth posterior member for the jth variable after the ith observation
σ_{ji}^p	kth prior covariance for the jth variable and the ith observation
T_{ji}^p	Prior third comoment for the jth variable and the ith pseudo-observation
F_{ji}^p	Prior fourth comoment for the jth squared variable and the ith pseudo-observation

filtering can be significantly better than the EBKF. Section 4 shows the performance of the new algorithm within the DART framework with several different modeling tests. A summary of the most important results and our conclusions from them are provided in section 5.

2. Polynomial regression using serial observation processing

In this section we will first review the previous work on extending ensemble methods to perform polynomial regression and then we will show how to connect this previous body of work to serial observation processing. Because our intended application is the DART infrastructure, we will attempt to remain consistent with the symbology of Anderson and Collins (2007). Table 1 provides a synopsis of the symbols that will be used throughout this manuscript.

a. Polynomial regression

Polynomial regression is useful when the true posterior mean is a nonlinear function of the observations. When this is the case, the linear fit of the EBKF will lead to larger error for large innovations than polynomial regression because nonlinear functions are inherently curved [see Hodyss (2011) for further discussion]. To account for this curvature of the posterior mean, we create a state estimate with curvature in the form of powers of the observations. The way we do this is to first

note that the posterior mean can be represented as a polynomial (Hodyss 2011); namely,

$$\bar{x}_{ji}^u(y_i^o) = \bar{x}_j^p + \left(c_{ji} - b_{ji}\frac{T_i^0}{\sigma_i^0 + r_i}\right)d_i + b_{ji}[d_i^2 - (\sigma_i^0 + r_i)] + \cdots,$$

(1)

where $d_i = y_i^o - \bar{y}_i^p$ is the innovation and

$$c_{ji} = \frac{\sigma_{ji}^0}{\sigma_i^0 + r_i},$$

(2)

$$b_{ji} = \frac{T_{ji}^0}{\Pi_i} - c_{ji}\frac{T_i^0}{\Pi_i},$$

(3)

$$\Pi_i = F_i - (\sigma_i^0)^2 - \frac{(T_i^0)^2}{\sigma_i^0 + r_i} + r_{4i} - r_i^2 + 4\sigma_i^0 r_i.$$

(4)

Please see Table 1 for variable definitions.

Equation (1) reveals three important features of the impact of curvature on state estimation. First, the magnitude of two different third moments (the third moments in observation space and the third moments between the observation space squared and the state space to be updated) needs to be significant for the new terms to contribute. If they vanish, an expansion truncated at these terms reduces to the EBKF. Second, the new information from the third moment does not just provide a new term that is quadratic in the observations but also provides for a correction to the slope of the term linear in the observations. Finally, when the observation

and the prior mean happen to be equal, such that the innovation vanishes, a correction is still made to the prior mean when the prior is skewed. This occurs because when the observations happen by chance to be equal to the prior mean when the prior is skewed, then the implied curvature of the posterior is such that its center bends away from the prior mean (e.g., Hodyss 2012).

In this manuscript we truncate (1) to the quadratic term. If we do this, we may write our quadratic polynomial estimate of the posterior mean as

$$\overline{x}_j^u(y_i^o) = \overline{x}_j^p + \mathbf{G}_{ji}\hat{\mathbf{d}}_i, \tag{5}$$

where

$$\mathbf{G}_{ji} = \left[c_{ji} - b_{ji}\frac{T_i^0}{\sigma_i^0 + r_i} \quad b_{ji} \right], \tag{6}$$

$$\hat{\mathbf{d}}_i = \left[\begin{array}{c} d_i \\ d_i^2 - (\sigma_i^0 + r_i) \end{array} \right], \tag{7}$$

which has now been transformed into a form reminiscent of the EBKF. The difference of course is that we have doubled the number of observations and doubled the number of elements in the gain matrix.

Hodyss (2012) showed that (5) could also be written with a gain that is in a more familiar form:

$$\hat{\mathbf{G}}_{ji} = \hat{\mathbf{B}}_{ji}[\hat{\mathbf{A}}_{ii} + \hat{\mathbf{R}}]^{-1}, \tag{8}$$

where

$$\hat{\mathbf{B}}_{ji} = \left[\begin{array}{cc} \sigma_{ji}^0 & T_{ji}^0 \\ T_{(N_x+j)i}^0 & F_{(N_x+j)i}^0 - (\sigma_{(N_x+j)i}^0)^2 \end{array} \right], \tag{9}$$

$$\hat{\mathbf{A}}_{ii} = \left[\begin{array}{cc} \sigma_i^0 & T_i^0 \\ T_i^0 & F_i^0 - (\sigma_i^0)^2 \end{array} \right], \tag{10}$$

$$\hat{\mathbf{R}}_{ji} = \left[\begin{array}{cc} r_i & 0 \\ 0 & r_{4i} - r_i^2 + 4r_i\sigma_i^0 \end{array} \right], \tag{11}$$

and the third moments are calculated as

$$T_{ji}^0 = \langle (x_j^0 - \overline{x}_j^0)(y_i^0 - \overline{y}_i^0)^2 \rangle, \tag{12}$$

$$T_{(N_x+j)i}^0 = \langle (x_{(N_x+j)}^0 - \overline{x}_{(N_x+j)}^0)^2(y_i^0 - \overline{y}_i^0) \rangle, \tag{13}$$

$$T_i^0 = \langle (y_i^0 - \overline{y}_i^0)^3 \rangle, \tag{14}$$

where angle brackets denote an expectation with respect to the prior. In analogy with the variance and the covariance, we will throughout this manuscript refer to the "third moment" as (14) and the "third comoment" as (12). The fourth moments follow the same convention as

the third moments defined above and the notation on the subscript (i.e., $N_x + j$) is a prelude to the state augmentation to be discussed in the next section.

The reader may readily verify, using the known formula for the inverse of a 2×2 matrix, that the first row in (8) is the same as (6) and therefore consistent with (1). The key feature of the gain in (8) is that the calculation of third and fourth moments has been rendered identical to the calculation of a covariance. More specifically, the calculation of the third moment is simply the "covariance" between the jth state variable and the ith "squared" observation. Thinking of the calculation of the moments higher than the second as simply the covariance between various powers of perturbations allows for the convenient calculation of these higher moments using the algorithmic tools of the EBKF.

Further, because the structure of (5) with gain (8) is precisely of the form of the EBKF, we know that we may serially process the observation and the new pseudo-observation. To do this, we associate the new pseudo-observation with a portion of the new pseudoinnovation, which is the second element in (7). In the pseudoinnovation, we associate $d_i^2 - r_i$ with the pseudo-observation and the remaining portion of the pseudoinnovation, σ_i^0, is associated with the "prior" estimate of the pseudo-observation. Note that the expected value of the pseudo-observation is σ_i^0, which is consistent with our designation of the prior estimate being the prior variance.

The reader may have noticed that our decomposition of the pseudo-observation as $d_i^2 - r_i$ was not the only way to partition the pseudoinnovation. In particular, we could have defined the pseudo-observation as d_i^2 and the prior estimate as $r_i + \sigma_i^0$, but this would have required carrying the observation error variance r_i around whenever we were to calculate σ_i^0 for the prior estimate of the pseudo-observation. By defining the pseudo-observation as $d_i^2 - r_i$, which can be calculated efficiently before assimilating any observations, we do not need any reference to the observation error variance while assimilating any of the observations or even the pseudo-observations

b. Serial observation processing within DART

1) LINEAR REGRESSION

We use the ensemble adjustment Kalman filter (EAKF; Anderson (2001) to illustrate this section. The modifications required to implement a stochastic EBKF (Burgers et al. 1998; Evensen 2003) are illustrated later. The state vector will have N_x elements with each element denoted by j. The observation vector will have N_o elements with each element denoted by i. Please see Table 1 for a complete list of symbols and their definitions.

The DART algorithm performs data assimilation in a two-step process following Anderson (2003). We imagine that we are now ready to assimilate the ith observation. The first step of the DART algorithm is to calculate the observation increment, Δy_{ki}, for each ensemble member, $k = 1, 2, \ldots, N_e$:

$$\Delta y_{ki} = a_i(y_{ki}^p - \bar{y}_i^p) + \bar{y}_i^u - y_{ki}^p, \tag{15}$$

where

$$\frac{\sigma_i^u}{\sigma_i^p} = \frac{r_i}{\sigma_i^p + r_i}, \tag{16}$$

$$\bar{y}_i^u = a_i^2\left(\bar{y}_i^p + \frac{\sigma_i^p}{r_i}y_i^o\right), \tag{17}$$

$$a_i = \sqrt{\frac{\sigma_i^u}{\sigma_i^p}}. \tag{18}$$

At this point we have now created an ensemble of observation increments.

The second step is to project each increment onto the jth state variable in the following way:

$$x_{kji}^u = x_{kj}^p + \beta_{ji}\Delta y_{ki}, \tag{19}$$

where

$$\beta_{ji} = \frac{\sigma_{ji}^p}{\sigma_i^p}. \tag{20}$$

Localization is performed by simply noting the spatial distance between the ith observation and the jth state variable and multiplying the regression coefficient in (20) by the distance-dependent function of the users choice (typically Gaspari–Cohn; Gaspari and Cohn 1999).

2) QUADRATIC POLYNOMIAL REGRESSION

We emphasize here that we will make the approximation of Hodyss (2012) and neglect the quadratic cross terms between observations. This approximation reduces the computational cost and memory requirements considerably at the expense of discarding information about higher-order relationships between variables.

To convert the linear regression algorithm of the previous section to quadratic polynomial regression, one needs to perform the two modifications discussed in section 2a: 1) augment the innovation with the new pseudo-observations and 2) augment the state vector with pseudosquared states in order to produce a covariance matrix of the form (9). We discuss the form of these modifications next.

The first step is to increase the number of observations by creating pseudo-observations of the form

$$y_{(N_o+i)}^o = (y_i^o - \bar{y}_i^0)^2 - r_i, \tag{21}$$

along with their corresponding observation error variances,

$$r_{(N_o+i)} = 2r_i^2 + 4r_i\sigma_i^0, \tag{22}$$

where $i = 1, 2, \ldots, N_o$ and the subscript $(N_o + i)$ refers to the state augmentation procedure whereby these new pseudo-observations and observation error variances fill the bottom half of an extended state vector that is $2N_o$ long.

As discussed by Hodyss (2012), the form of (22) is due to the Gaussian assumption on the observation likelihood. If the observation likelihood is symmetric but non-Gaussian, one needs to recalculate (22) using the second and fourth moments:

$$r_{(N_o+i)} = r_{4i} - r_i^2 + 4r_i\sigma_i^0. \tag{23}$$

If the observation likelihood is not symmetric, then this generally implies a correlation between the observations and the pseudo-observations because nonsymmetric distributions will typically have a nonzero third moment. If the observation likelihood has a nonzero third moment, then the observations and their corresponding pseudo-observations are correlated and therefore the basic assumption of serial observation processing is violated.

Finally, we append to the state vector new state variables that we refer to as "pseudosquared states," which are calculated as

$$x_{k(N_x+j)}^p = (x_{kj}^p - \bar{x}_j^p)^2, \tag{24}$$

where $j = 1, 2, \ldots, N_x$. We emphasize that the state vector is now $2N_x$ long and the subscript in (24) makes note of this. We note that the mean of (24) with respect to the index k is $(K-1)/K$ times the variance of the jth regular state variable. Therefore, when we subtract the mean of (24), and subsequently calculate the variance, we obtain the element in the bottom-right corner of (9). Furthermore, the concatenation of the regular state variables with these pseudosquared states produces a prior covariance matrix also of the form (9), which of course was the intended goal.

Augmenting the state vector with these pseudosquared states of the form (24) results in new regression coefficients that correspond to the update of a regular state variable by a pseudo-observation,

$$\beta_{ji} = \frac{T_{ji}^p}{F_i^p - (\sigma_i^p)^2}; \qquad (25)$$

as well as a new regression coefficient corresponding to the update of a pseudosquared state by the pseudo-observation,

$$\beta_{ji} = \frac{F_{ji}^p - \sigma_i^p \sigma_{ji}^p}{F_i^p - (\sigma_i^p)^2}; \qquad (26)$$

and finally the regular observation also updates the pseudosquared states, and its regression coefficient is of the form

$$\beta_{ji} = \frac{T_{(N_x + j)i}^p}{F_i^p - (\sigma_i^p)^2}. \qquad (27)$$

Note that this implies that there are now four different regression coefficients and therefore the cost of quadratic polynomial filtering is approximately 4 times that of the EBKF.

Because we are doing data assimilation for the pseudosquared states in (24), this begs the question: What does the state update of a pseudosquared state mean? The posterior mean estimate for a pseudosquared state is an approximation to the following integral:

$$P_j^*(y_i^o) = \int_{-\infty}^{\infty} (x_j^p - \overline{x}_j^p)^2 p(x_j^p \,|\, y_i^o)\, dx_j^p = P_j^u(y_i^o) + (\overline{x}_j^u - \overline{x}_j^p)^2,$$
$$(28)$$

where $P_j^u(y_i^o)$ is the posterior variance about the posterior mean. In other words, the state estimate of the pseudosquared state is an estimate of the true posterior variance but calculated around the prior mean. For now, the DART code does not make use of this state estimate and simply discards it after the assimilation. Using this estimate of the true posterior variance in the ensemble generation step may be beneficial but has been left for future work.

3) Issues with ensemble generation

We will begin this section by discussing issues that we had with developing the quadratic polynomial version of the stochastic EBKF [Burgers et al. (1998); Evensen (2003); sometimes erroneously referred to as "perturbed" observations; see Hodyss (2011), Hodyss and Campbell (2013), and Posselt et al. (2014) for further discussion] and then follow this up with a brief discussion of the differences between stochastic and deterministic ensemble generation methods.

It is important to realize that when the observation likelihood is Gaussian that the likelihood for the pseudo-observations is not. Hence, when generating an ensemble using the framework of the stochastic EBKF, it is technically incorrect to use Gaussian noise for the pseudo-observations. Nevertheless, the present version of DART associates the pseudo-observations with Gaussian noise because of the computational complexity of using the correct noise. The main goal of this section is to illustrate and explain this issue and how we chose to tackle it.

This issue is identical to the corresponding issue in the traditional stochastic EBKF when applied to skewed observation likelihoods [see Hodyss (2011, 2012) for further discussion]. When the observation likelihood is skewed, one must draw the noise for the observations from the correct skewed distribution and must also change the form of the update equation. While the change to the update equation is minor, the change in the way one draws the noise for the quadratic polynomial filter is quite complicated.

To illustrate the issue, we construct a simple scalar non-Gaussian example and compare the correct ensemble generated by a stochastic ensemble generation method and that from the approximation we use in DART. We assume the prior is χ^2 with one degree of freedom and the observation likelihood is Gaussian with variance $r = 1$. The experiments below will employ 10^8 trials in which the truth and observational errors are resampled for each trial. We emphasize that because the prior is χ^2 we know all of the required moments without sampling error, and we will use this knowledge to calculate the true quadratic polynomial state estimate. However, the DART algorithm described above is fundamentally an ensemble method and therefore we will use an ensemble with a size of 10^8 in its calculations.

As discussed in several places recently (see Hodyss 2011; Hodyss and Campbell 2013; Posselt et al. 2014; Hodyss et al. 2016), the intended result of the stochastic EBKF is not the true posterior moments but in fact their expected values with respect to the marginal distribution of observations. The expected posterior moments of a state estimation method may be generated by calculating the state estimate for each observation of the set of trials, then subtracting this state estimate from the truth, raising this error to the power of the moment to be estimated, and finally averaging these quantities over all trials. These true expected posterior moment values for the quadratic polynomial filter are reported upon in Table 2.

These true posterior moments are then compared to two serial observation processing algorithms employing a stochastic generation method. The first method (referred to as stochastic QF) correctly accounts for the non-Gaussian nature of the pseudo-observations by using

TABLE 2. Expected posterior moments of different quadratic polynomial ensemble generation methods.

Moment	True	Stochastic QF	Stochastic DART	DART EAKF
Second	0.507	0.507	0.507	0.507
Third	0.115	0.115	0.010	0.566
Fourth	2.81	2.81	1.89	1.29

the following algorithm. First, the regular observations are associated with Gaussian random noise, $\varepsilon_{ki} = \delta_{ki}$, with variance r_i, but the pseudo-observations are associated with noise of the form

$$\varepsilon_{ki} = \delta_{ki}^2 - r_i^2 + 2(y_{ki}^0 - \bar{y}_i^0)\delta_{ki}, \qquad (29)$$

where δ_{ki} is the actual random number used to perturb the ith regular observation that is associated with the $(N_o + i)$th pseudo-observation and the y_{ki}^0 and $\bar{y}_i^0$ are the prior ensemble members and prior mean, respectively, in observation space *before* any observations have been assimilated. Note that if we calculate the variance of (29), we obtain (22) but further note that this noise is not Gaussian as it is the convolution of a χ^2 variable and the product of an ensemble perturbation and a Gaussian-distributed variable.

Irrespective of whether the present observation being assimilated is a regular observation or a pseudo-observation, the form of the observation increment equation that replaces (15) is

$$\Delta y_{ki} = \bar{y}_i^u - \bar{y}_i^p - \frac{a^2\sigma_i^p}{r_i}[\varepsilon_{ki} + (y_{ki}^p - \bar{y}_i^p)]. \qquad (30)$$

The second step of the DART algorithm is unchanged. The resulting expected posterior moments from this algorithm can be seen in Table 2 to be identical to the true *expected* moments.

The method that is used in DART (referred to as stochastic DART) however associates the regular observations with Gaussian random noise with variance r_i and the pseudo-observations with Gaussian random noise with variance equal to (22). Furthermore, the noise used with the regular observation is independent of the noise used with the pseudo-observation. Once this Gaussian random noise ε_{ki} is generated, (15) is replaced by

$$\Delta y_{ki} = a^2\sigma_i^p\left(\frac{y_{ki}^p}{\sigma_i^p} + \frac{y_i^o + \varepsilon_{ki}}{r_i}\right) - y_{ki}^p. \qquad (31)$$

Again, the second step of the DART algorithm is unchanged. The resulting expected posterior moments from this algorithm can be seen in Table 2. While the centered second moment (variance) is correct, the third and fourth moments are not. The resulting distribution

is more symmetric when using (31), as can be seen in its third moment being too small. Further evidence is provided in Fig. 1, which explicitly shows that the stochastic algorithm in DART produces the wrong posterior probability density function (PDF).

To understand the difference between (30) and (31), note that (31) can be written as

$$\Delta y_{ki} = \bar{y}_i^u - \bar{y}_i^p - \frac{a^2\sigma_i^p}{r_i}[-\varepsilon_{ki} + (y_{ki}^p - \bar{y}_i^p)], \qquad (32)$$

where we note that the difference arises only in the sign of the noise, ε_{ki}. Hence, there is technically a sign error in the observation noise term because as we have shown (30) produces the correct moments. This issue identified in (32) vanishes when the noise ε_{ki} is drawn from a symmetric distribution with mean zero because multiplying a symmetric distribution by a minus sign does not invoke a change. This explains why the present DART algorithm, when running as an EBKF with symmetric noise, does not see this issue. To be clear, this
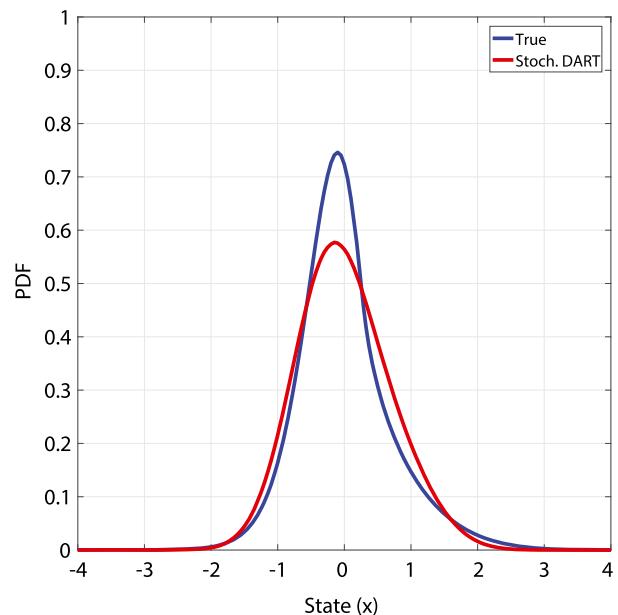

FIG. 1. The true expected posterior PDF from a stochastic algorithm is shown in blue. The resulting posterior PDF from the stochastic DART algorithm is shown in red. The stochastic QF algorithm described in (30) produces a posterior PDF identical to the blue curve.

is not an issue that is peculiar only to DART, as the original work of Burgers et al. (1998) also has this issue that it cannot properly be applied to skewed observation likelihoods. Therefore, the present DART algorithm cannot perform stochastic generation for nonsymmetric observation likelihoods, and this is true for both quadratic polynomial implementations as well as the original EBKF framework. If one does apply the present version of DART to nonsymmetric observation likelihoods, the resulting posterior distributions will have a third moment of the wrong sign. To correct this, the stochastic algorithm must be modified to be consistent with (30), and to properly produce the correct moments for the quadratic polynomial filter, the noise must be drawn from (29). Because creating the noise in (29) presents a coding challenge that we chose not to tackle, we similarly did not make the change implied by (30). As seen in Fig. 1, the differences are likely not very great at the ensemble sizes to be explored in this manuscript.

Finally, we would like to point out that for the EAKF the issue of which posterior moments it is generating can also be examined. Hodyss and Campbell (2013) showed that deterministic ensemble generation methods generally fail to produce an accurate rendition of any moment higher than the second. Testing of the EAKF algorithm in section 2b(1) with the example problem of this section reveals that it produces the exact expected second posterior moment but that its third is far too large and its fourth far too small (0.566 and 1.29, respectively). It is easy to see that the third should be far too large by noting from (15) that the centered posterior third moment in observation space for the EAKF is

$$ T_i^u = a_i^3 T_i^p , \qquad (33) $$

and whose skewness is therefore preserved by the assimilation of an observation:

$$ S_{\text{EAKF}}^u = \frac{T_i^u}{(\sigma_i^u)^{3/2}} = \frac{T_i^p}{(\sigma_i)^{3/2}} = S^p , \qquad (34) $$

where S^p is the prior skewness. This is in contrast to the skewness obtained from the stochastic algorithm, which, unlike the EAKF, is constrained to be consistent with the expected posterior moments, and whose posterior third moment is

$$ T_i^u = a_i^6 \left(T_i^p - \frac{r_{3i}}{r_i} \sigma_i^p \right) . \qquad (35) $$

There are two important differences between the updated third moment from an EAKF and that from the stochastic algorithm of (30). First, because the parameter $a_i \leq 1$, the third moment from the stochastic

algorithm is typically smaller than that from the EAKF algorithm. Second, the third moment from the EAKF algorithm is not impacted by the skewness of the observation likelihood, but as seen in (35) the third moment from the stochastic algorithm is a function of the third moment of the observation likelihood. If for the moment we assume we are dealing with a symmetric observation likelihood (i.e., $r_{3i} = 0$), which is consistent with the stochastic DART implementation, then the resulting posterior skewness from the stochastic method is

$$ S_S^u = \frac{T_i^u}{(\sigma_i^u)^{3/2}} = a_i^2 \frac{T_i^p}{(\sigma_i^p)^{3/2}} = a_i^2 S^p , \qquad (36) $$

which shows that

$$ S_S^u \leq S_{\text{EAKF}}^u = S^p \qquad (37) $$

with equality only in the trivial cases where either $r_i = 0$ or $T_i^p = 0$. Hence, data assimilation consistent with Bayes's rule will on average reduce the skewness to be less than the prior skewness, and the stochastic algorithm will obey this rule while the EAKF will not. We speculate that one reason why deterministic methods may suffer from the well-known "outlier" problem (see Lawson and Hansen 2004; Anderson 2010; Hodyss and Campbell 2013) is because of this feature that deterministic methods preserve the skewness. Nevertheless, for the small ensemble sizes used in the rest of this manuscript we will see that there are many situations where deterministic ensemble generation outperforms stochastic.

3. Understanding when quadratic polynomial filtering is useful

Quadratic polynomial filtering requires the estimation of the third and fourth moments of the prior, which are more difficult at small ensemble sizes to estimate than the second moments of the EBKF. On the other hand, when non-Gaussianity is large, the resulting curvature of the posterior mean will be missed by the EBKF. Hence, we anticipate a competition between the size of the sampling errors from small ensemble sizes and the effects of non-Gaussianity, and its implied curvature, on the relative quality of EBKF and quadratic state estimates.

To clearly illustrate this point, we construct another simple scalar data assimilation problem. In this problem we will employ a Gaussian observation likelihood with $r = 1$. The prior will be a gamma distribution with prior variance $\sigma^p = 1$ and for which we control the

skewness using the standard shape k and scale parameters θ; that is,

$$k = \left(\frac{2}{S^p}\right)^2 \tag{38}$$

$$\theta = \sqrt{\frac{\sigma^p}{k}}, \tag{39}$$

where S^p is the desired prior skewness. We again employ 10^8 trials in which truths and observational errors are resampled. Here, we will calculate only the mean squared error (MSE) of the posterior state estimates from the algorithm of section 2 (i.e., we only calculate $\overline{y}^u$) for both the traditional EBKF and the quadratic polynomial update. In addition, we will also employ the standard particle filter in which we take samples from the prior and weight them according to

$$w_k = \frac{\exp\left[-\frac{1}{2}\frac{(y^o - y_k^p)^2}{r}\right]}{\sum\limits_{k'=1}^{N_e} \exp\left[-\frac{1}{2}\frac{(y^o - y_{k'}^p)^2}{r}\right]}. \tag{40}$$

Subsequently, a posterior state estimate is produced using

$$\overline{y}^u = \sum\limits_{k=1}^{N_e} w_k y_k^p. \tag{41}$$

The results of two sets of experiments are presented in Fig. 2. In the top panel in Fig. 2 we present results for essentially an infinite ensemble size. Because the prior is Gamma we know all of the moments of the prior without sampling error. Therefore, we evaluate $\overline{y}^u$ for the EBKF and quadratic polynomial filter using the known moments of the Gamma distribution. For the standard particle filter we use an ensemble of 10^8 members.

It is interesting to see in Fig. 2 that the MSE of the EBKF is independent of skewness for infinite ensemble size. This result runs counter to the more typical philosophy that the EBKF is "worse" in non-Gaussian systems. In fact, the MSE of the EBKF for infinite ensemble size is just as good for Gaussian systems as it is for non-Gaussian systems. This result is actually already known from (16), which clearly shows that the expected squared error, σ^u (i.e., posterior variance), of the EBKF is independent of the skewness. Note that if we evaluate (16) with $r = 1$ and $\sigma^p = 1$, we obtain $\sigma^u = 1/2$, which is the value of the posterior variance for the EBKF from Fig. 2a.

What is perhaps more surprising then is that even though the prior variance is held fixed, the posterior
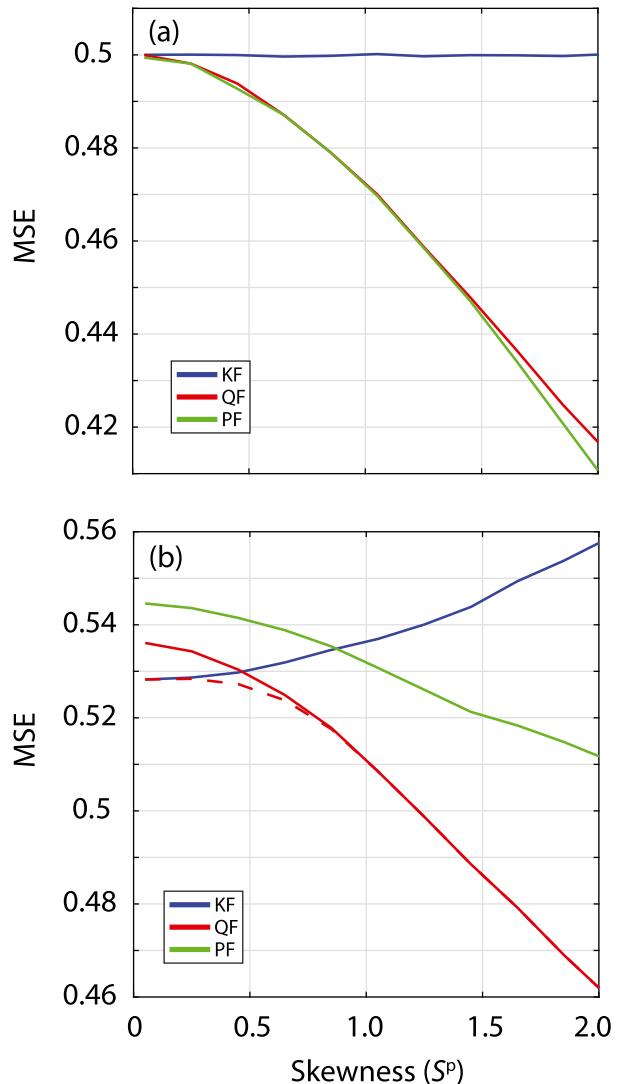

FIG. 2. Posterior MSE as a function of skewness. Results for (a) very large ensemble size and (b) for $N_e = 20$. Blue lines are for the traditional Kalman filter (KF), red lines are for the quadratic polynomial filter (QF), and green lines are for the standard particle filter (PF). The red dashed line is for the quadratic polynomial filter with third-moment damping.

MSE decreases rapidly as skewness *increases* for methods that account for the skewness (i.e., the quadratic polynomial filter and the standard particle filter). This reduction in the posterior variance by the quadratic polynomial filter was shown in Hodyss (2011) to obey

$$\frac{\sigma_i^u}{\sigma_i^p} = \frac{r_i}{\sigma_i^p + r_i}\left[1 - \frac{r_i}{\sigma_i^p + r_i}\frac{(\sigma_i^p)^2}{\Pi_i}(S^p)^2\right], \tag{42}$$

which is the posterior variance after assimilating both the ith regular observation and its corresponding pseudo-observation. Equation (42) clearly shows that

the posterior variance should decrease as the square of the prior skewness and was verified to be in agreement with Fig. 2. This shows that skewed prior distributions can actually result in a greater reduction in the MSE relative to the prior variance than nonskewed distributions and in this sense may be thought of as easier to perform data assimilation upon. By contrast, Gaussian prior distributions with the same variance as these skewed distributions will lead to larger posterior variances on average, which again in this sense suggests that they are more difficult to perform data assimilation upon. Therefore, skewed prior distributions offer something that Gaussian prior distributions cannot, which is the possibility of accounting for this skewness and realizing posterior MSE reductions.

We also report in Fig. 2 the same experiments but with an ensemble size of $N_e = 20$. In this case we now see the more intuitive result that the sampling error has a stronger negative impact on the EBKF when the skewness is large than when it is small. Nevertheless, the EBKF now has smaller MSE for very small prior skewness than either the standard quadratic polynomial filter or the standard particle filter. This result underscores the commonly held intuition that the EBKF is the best method when errors are Gaussian or nearly so.

As the skewness increases, the quadratic polynomial filter and the EBKF become comparable around a skewness of about ½ at this particular ensemble size. Beyond a skewness of ½, the non-Gaussianity is so large that it is better to account for it using the quadratic polynomial filter than to ignore it with the EBKF. The standard particle filter is not better than the EBKF until a skewness of about ¾ and is never better than the quadratic polynomial filter over the range of skewness shown here. We believe that this result—that the standard particle filter is more strongly impacted by sampling error than the quadratic polynomial filter—is because the standard particle filter is attempting to implicitly evaluate all of the terms of the polynomial expansion in (1). By contrast, the quadratic polynomial filter has truncated the polynomial to terms that are more accurately evaluated at these small ensemble sizes.

This fact that truncating the polynomial may be an advantageous method for decreasing the impact of sampling error on higher-order estimation methods lead us to attempt to damp the prior third moments for small ensemble sizes. The hypothesis is that controlling the size of the new term of the quadratic polynomial will allow for a reduction in the posterior MSE because reducing the amplitude of this term will reduce the impact of the sampling error on the state estimate from the third and fourth moments. We accomplish this by multiplying the regression coefficient by a tunable parameter α for

only the update of the regular state variables by the pseudo-observations (25) and the corresponding update of the squared state variables by the regular observation (27); namely,

$$\beta_{ji} \rightarrow \alpha\beta_{ji}, \tag{43}$$

where $0 \le \alpha \le 1$ and is tuned for minimum posterior MSE. Note that when $\alpha = 0$, the quadratic polynomial filter produces a state estimate that is identical to the EBKF state estimate, and when $\alpha = 1$, it obviously produces the quadratic state estimate. The results of tuning α for the experiments of this section are represented in Fig. 2 by the dashed red line and referred to as third moment damping. In Fig. 2 we see that the resulting MSE of the posterior is now no worse than the EBKF state estimate and can be better in regions of moderate skewness (i.e., $S^p \approx 1/2$). For skewness greater than about ¾ the minimum posterior MSE was found for $\alpha = 1$, which implies that the skewness is so large that the sampling error in estimating the third and fourth moments is minor compared to the outright neglect of the quadratic state estimate.

4. Example test problems

The goals of this section are twofold. On the one hand, we would like to describe the differences we see between the EBKF linear regression schemes available in DART and the new quadratic polynomial method. On the other hand, we would like to provide a set of test cases that potential new users of this new DART version can use for testing and validation of their setup. To these ends, we now illustrate a collection of experiments in a simple nonlinear model as well as in a simple general circulation model (GCM) that has been used previously in other studies using DART (e.g., Anderson et al. 2005).

a. Lorenz-63

The Lorenz-63 (Lorenz 1963) system of equations is a standard test case for nonlinear data assimilation and is a standard component of the DART download package. The strong non-Gaussianity in this model arises because its attractor is extremely curved and Hodyss (2011) argued that curvature of the attractor imposes curvature on the posterior mean, which is associated with large posterior third moments, and therefore this system poses a challenging data assimilation problem for the EBKF. These extremely curved posterior distributions in Lorenz-63 allow us to illustrate the full potential of the quadratic polynomial filter.

We will use four different data assimilation methods: 1) the EBKF with ensemble adjustment ensemble
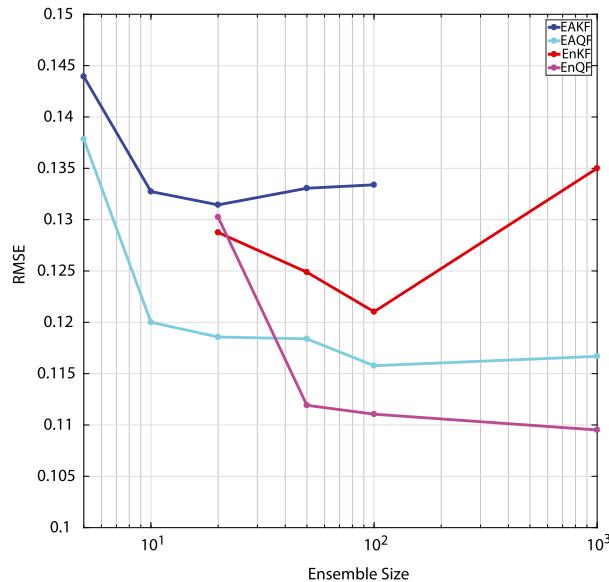
FIG. 3. Posterior RMSEs of the z variable as a function of ensemble size for four data assimilation methods. Experiments that resulted in an RMSE greater than the observational uncertainty ($\sqrt{r_i} \approx 0.33$) are not plotted and associated with filter divergence. See text for further details.

generation (EAKF), 2) the EBKF with stochastic ensemble generation (EnKF), 3) the quadratic polynomial filter with ensemble adjustment ensemble generation (EAQF), and 4) the quadratic polynomial filter with stochastic ensemble generation (EnQF). These experiments will use observations of the x and z variables with an observation error variance of 0.1. We cycle for 10 000 cycles with observations 0.12 time units apart (12 time steps of the model). We will perform experiments with ensemble sizes of 5, 10, 20, 50, 100, and 1000 members. For all methods we used the adaptive inflation technique of Anderson (2009). For both versions of the quadratic polynomial filter we tuned the third-moment damping [see (43)] to deliver the minimum posterior RMSE.

The results of these experiments in terms of the posterior RMSE of the z variable are shown in Fig. 3. The RMSEs of the x and y variables follow the same pattern as the z variables and because of this are not shown. There appear to be two regimes in Fig. 3. For ensemble sizes less than about 50 members the method with smallest RMSE is the EAQF. For ensemble sizes greater than or equal to 50 the method with smallest RMSE is the EnQF. This same two-regime pattern of behavior is also apparent between the two EBKF methods. Therefore, for all ensemble sizes, polynomial filtering is superior to the EBKF in this system. We believe that this result is explained by the fact that the average absolute value of the prior skewness found in these experiments

ranges from approximately 1 to 5 and therefore the results of this section are consistent with section 3 and Fig. 2, which showed that the benefit from polynomial filtering is large when the skewness is larger than approximately 1.

The resulting quality of the shape of the posterior ensemble distributions can also be tested. We used binned spread–skill and the continuous ranked probability score (CRPS) to examine the quality of the ensemble distributions. We found that when the adaptive inflation of Anderson (2009) was properly tuned, all four methods had approximately the same binned spread–skill score in terms of the quality of the slope of the resulting plot (not shown). Additionally, the CRPS for all four methods simply ranked the methods identically to their rankings in Fig. 2, which reveals that the quality of the posterior mean is dominating the CRPS value.

b. BGRID GCM

The model used in this section is a coarse-resolution, dry, hydrostatic general circulation model (GCM) without topography (Anderson et al. 2004). The model uses the temperature forcing of Held and Suarez (1994) to maintain a time-evolving, baroclinically unstable Rossby wave field at midlatitudes and Rayleigh damping of winds near the surface to represent friction. The model domain consists of five vertical levels with 30 grid points in latitude and 60 grid points in longitude, which is approximately 6° resolution. This model is available as a standard option in the DART download package.

A variety of experiments were performed to determine the degree of nonlinearity/non-Gaussianity in this model. It was found that at this resolution, and without physical processes related to moisture, the error growth in this model is substantially slower than in typical, full-physics atmospheric models. By trying a variety of cycling intervals and observational densities, we determined that cycling with 10 days between observations was required to overcome this very slow error growth in order to develop the small degree of nonlinear/non-Gaussian behavior that we illustrate below. Additionally, we found that for all ensemble sizes tried here the EAKF and EAQF are substantially superior to the EnKF and EnQF. While we find this result—that deterministic methods are better than the stochastic methods in BGRID, but the reverse holds in Lorenz-63—to be very interesting, we do not pursue this here as we feel a detailed analysis of the differences to be outside the scope of the present work. For this reason, we limit the presentation here to only the EAKF and EAQF.
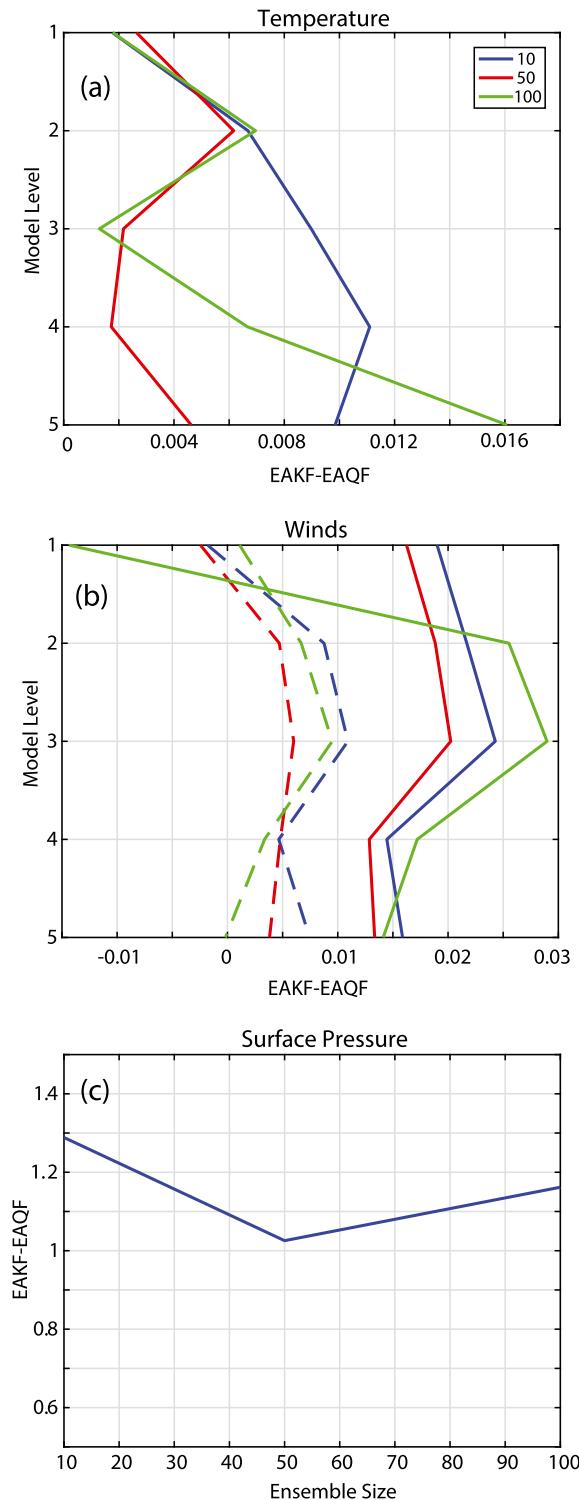
FIG. 4. Posterior RMSE differences from the BGRID experiments. The difference between the posterior RMSE from the EAKF and the EAQF is shown in all panels. Positive values denote that the EAQF had smaller RMSE. Results for (a) temperature (K); (b) winds (m s^{-1}), where solid is the zonal wind and dashed is the meridional winds; and (c) surface pressure (hPa). In (a) and (b) blue curves correspond to experiments with 10 members, red curves are for 50 members, and green curves are for 100 members.

The observational network consists of 100 randomly distributed surface pressure observations with an observational error variance of 1 hPa^2. The simulation is run for 50 yr while assimilating observations from this observational network every 10 days; this yields 1825 analysis times. Similar to the Lorenz-63 section above we employ the adaptive inflation technique of Anderson (2009), but here we also include both vertical and horizontal localization, which is tuned for minimum posterior RMSE. The RMSEs of these analyses are measured in the four prognostic variables of the model (temperature, zonal and meridional winds, and surface pressure) and will be calculated by discarding the first 100 analysis times and calculating the RMSE from the remainder.

We performed data assimilation with ensembles of size 10, 50, and 100 members. For these experiments we calculated the average absolute value of the prior skewness and found that for all variables it was approximately 0.1–0.5, which is an order of magnitude less than what was found in the Lorenz-63 experiments. Checking the impact at this level of skewness in Fig. 2 reveals that there is likely not much to be gained from quadratic polynomial filtering. Hence, we must expect from Fig. 2 that we will not see large improvements from quadratic polynomial filtering in this model setting.

The results in terms of the difference in posterior RMSEs between the EAKF and the EAQF as a function of different ensemble sizes and vertical model levels are shown in Fig. 4 and largely confirm our hypothesis that prior skewness values less than 1 lead to small improvements from quadratic polynomial filtering. The largest impact from the EAQF appears in the surface pressure, which is likely due to the observations being of that variable. The information from surface pressure observations must then be transmitted upward and to the other variables using the covariance structure as well as the higher-order comoments in the EAQF. In temperature and winds the EAQF generally provides a small improvement at all levels except for winds at the top of the model, which is inside a strong gravity-wave-damping region.

While it is tempting to anticipate that there should be greater benefit from larger ensemble sizes in Fig. 4, this line of thinking is confounded by the fact that there are generally larger errors for small ensemble sizes, which is confirmed by the fact that the posterior RMSEs for both the EAKF and the EAQF (not shown) are substantially larger for an ensemble size of 10 as compared to 100. Therefore, the non-Gaussianity is actually larger for small ensemble size, and because the EAQF responds to non-Gaussianity, the EAKF-EAQF

difference as a function of ensemble size does not follow a clear pattern. Finally, similar to the results from the Lorenz-63 experiments, binned spread–skill and CRPS were used to examine the quality of the ensemble distributions and no notable differences were found when the adaptive inflation of Anderson (2009) was properly tuned.

5. Summary and conclusions

We have shown how to perform quadratic polynomial regression entirely within the standard framework of an ensemble-based Kalman filter (EBKF) that employs serial observation processing. The key principles are to rewrite multivariate, polynomial regression in a way that is homomorphic to the EBKF and to note that high-order, multivariate moments can be thought of as the covariance between the state and a power of a state variable. Applying these principles allows high-order, multivariate regression to follow the original flow of the EBKF and makes modifying an already constructed EBKF relatively straightforward.

We have shown that the benefit from quadratic polynomial filtering is obtained through a competition between the ensemble size and the level of prior skewness in the physical system being simulated. As a general rule, and with ensemble sizes less than 100 members, we would suggest using quadratic polynomial filtering for systems with an average absolute value of the prior skewness greater than about 1. If the user can accept the computational expense of ensemble sizes greater than 100, then quadratic polynomial filtering could be beneficial for skewness values less than 1. But in this case the benefit will be small as the user must acknowledge from Fig. 2 that even with an infinite ensemble a skewness of approximately 1.5 is required for a 10% improvement. Therefore, we suggest a careful examination of the prior skewness in the system to be studied before application of any nonlinear data assimilation method.

Third-moment damping was found to be an integral component to getting quadratic polynomial filtering to outperform the EBKF at ensemble sizes substantially less than 100. In the experiments in this manuscript the same α value was used for all variables in each experiment. It is likely however that performance gains would be realized by, for example, making the α value a function of field variable, latitude, and altitude in BGRID. Because of the success of third-moment damping, work in this direction is under way.

Quadratic polynomial filtering is now an option within the Data Assimilation Research Testbed (DART) of the Data Assimilation Research Section (DARES) of the National Center for Atmospheric Research. Those interested in making use of this facility should contact the DARES team (http://www.image.ucar.edu/DAReS/DART/) for the latest version of the code.

REFERENCES

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2894–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, doi:10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2.

——, 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, **61A**, doi:10.1111/j.1600-0870.2008.00361.x.

——, 2010: A non-Gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev.*, **138**, 4186–4198, doi:10.1175/2010MWR3253.1.

——, and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, doi:10.1175/JTECH2049.1.

——, and Coauthors, 2004: The new GFDL global atmosphere and land model AM2–LM2: Evaluation with prescribed SST simulations. *J. Climate*, **17**, 4641–4673, https://doi.org/10.1175/JCLI-3223.1.

——, B. Wyman, S. Zhang, and T. Hoar, 2005: Assimilation of surface pressure observations using an ensemble filter in an idealized global atmospheric prediction system. *Bull. Amer. Meteor. Soc.*, **62**, 2925–2938, https://doi.org/10.1175/JAS3510.1.

——, T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A community facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296, doi:10.1175/2009BAMS2618.1.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, doi:10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367, doi:10.1007/s10236-003-0036-9.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, doi:10.1002/qj.49712555417.

Held, I. M., and M. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.*, **75**, 1825–1830, doi:10.1175/1520-0477(1994)075<1825:APFTIO>2.0.CO;2.

Hodyss, D., 2011: Ensemble state estimation for nonlinear systems using polynomial expansions in the innovation. *Mon. Wea. Rev.*, **139**, 3571–3588, doi:10.1175/2011MWR3558.1.

——, 2012: Accounting for skewness in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 2346–2358, doi:10.1175/MWR-D-11-00198.1.

——, and W. F. Campbell, 2013: Square root and perturbed observation ensemble generation techniques in Kalman and

quadratic ensemble filtering algorithms. *Mon. Wea. Rev.*, **141**, 2561–2573, doi:10.1175/MWR-D-12-00117.1.

——, ——, and J. S. Whitaker, 2016: Observation-dependent posterior inflation for the ensemble Kalman filter. *Mon. Wea. Rev.*, **144**, 2667–2684, doi:10.1175/MWR-D-15-0329.1.

Houtekamer, P., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, doi:10.1175/1520-0493(2001)129<0123: ASEKFF>2.0.CO;2.

Lawson, W., and J. Hansen, 2004: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.*, **132**, 1966–1981, doi:10.1175/1520-0493(2004)132<1966: IOSADF>2.0.CO;2.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, doi:10.1175/1520-0469(1963)020<0130: DNF>2.0.CO;2.

Posselt, D. J., D. Hodyss, and C. Bishop, 2014: Errors in ensemble Kalman smoother estimates of cloud microphysical parameters. *Mon. Wea. Rev.*, **142**, 1631–1654, doi:10.1175/ MWR-D-13-00290.1.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, doi:10.1175//2555.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913: EDAWPO>2.0.CO;2.