

Improving Weather Forecast Skill through Reduced-Precision Data Assimilation

SAM HATFIELD, ANEESH SUBRAMANIAN, AND TIM PALMER

Atmospheric, Oceanic and Planetary Physics, University of Oxford, Oxford, United Kingdom

PETER DÜBEN

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

(Manuscript received 11 May 2017, in final form 14 September 2017)

ABSTRACT

A new approach for improving the accuracy of data assimilation, by trading numerical precision for ensemble size, is introduced. Data assimilation is inherently uncertain because of the use of noisy observations and imperfect models. Thus, the larger rounding errors incurred from reducing precision may be within the tolerance of the system. Lower-precision arithmetic is cheaper, and so by reducing precision in ensemble data assimilation, computational resources can be redistributed toward, for example, a larger ensemble size. Because larger ensembles provide a better estimate of the underlying distribution and are less reliant on covariance inflation and localization, lowering precision could actually permit an improvement in the accuracy of weather forecasts. Here, this idea is tested on an ensemble data assimilation system comprising the Lorenz '96 toy atmospheric model and the ensemble square root filter. The system is run at double-, single-, and half-precision (the latter using an emulation tool), and the performance of each precision is measured through mean error statistics and rank histograms. The sensitivity of these results to the observation error and the length of the observation window are addressed. Then, by reinvesting the saved computational resources from reducing precision into the ensemble size, assimilation error can be reduced for (hypothetically) no extra cost. This results in increased forecasting skill, with respect to double-precision assimilation.

1. Introduction

The skill of weather forecasts has improved dramatically over the past 30 years thanks to, among other things, improved model initialization through data assimilation. Modern, ensemble-based data assimilation techniques, such as the ensemble Kalman filter (EnKF) (Houtekamer and Zhang 2016), are able to estimate the flow dependence of covariances, in spite of nonlinearities in the model, so that observations and backgrounds can be more optimally combined. Further, sophisticated inflation techniques (Whitaker and Hamill 2012; Miyoshi 2011), efficient deterministic variations on the EnKF, such as the local ensemble transform Kalman filter (LETKF) (Hunt et al. 2007), and other developments have made the EnKF feasible for operational numerical weather prediction.

In the near future, increasingly accurate and widespread observational networks of satellites and

surface instruments will allow a more accurate characterization of the atmospheric state (Bauer et al. 2015). If computing hardware keeps pace, the trends of increasing ensemble size, grid resolution, and model complexity will also continue to improve background estimates, leading to reduced initial error and further extensions of the forecast horizon (the lead time within which a forecast is still useful). However, the historical trend of increasing computing power cannot be relied on indefinitely. Moore's Law, the prediction that transistor density doubles every two years, is beginning to decline and high-performance computers are increasingly relying on parallelization, with commensurate increases in power consumption (Sutter 2005). This predicament prompts a reevaluation of current computing paradigms, and motivates this study into numerical precision.

When simulating geophysical systems, it is common practice to use "double-precision" variables, which use 64 bits of storage to accurately represent real numbers. Lowering precision will introduce larger rounding errors, which can contaminate the simulation. However, lower-precision arithmetic is computationally cheaper.

Corresponding author: Sam Hatfield, samuel.hatfield@physics.ox.ac.uk

If the simulation is tolerant to the rounding errors introduced, then by reducing precision we can free up computational resources that can be invested into other areas of the system. Palmer (2014) suggested applying such a trade-off within climate modeling, a field of considerable uncertainty because of simplified parameterizations of unresolved phenomena. If a reduction in precision can be achieved without an unacceptable deterioration of model performance, then computing resources liberated by this procedure could be reinvested toward increasing the grid resolution, thereby reducing the dependence on parameterizations.

The subject of numerical precision within climate and weather models is one of increasing publicity. Previous studies have considered both idealized models (like the one presented below) and intermediate complexity models (Düben et al. 2014; Düben and Palmer 2014). Through emulation of hypothetical inexact hardware, it was found that large portions of the code can be run at a precision significantly less than the standard, double-precision, without strongly impacting the quality of either forecasts or the climatology. The European Centre for Medium-Range Weather Forecasts (ECMWF) has also considered a reduction in precision in their forecasting model, the Integrated Forecast System (IFS) (Vaña et al. 2017), albeit without data assimilation and at a lower resolution than the operational setup. Reducing the precision of almost all floating-point variables throughout the model (physics and dynamics) from double to single precision reduced the runtime of IFS by around 40% without noticeably degrading probabilistic skill scores. Two ensemble forecasts of an equal computational cost were then performed: double-precision with 34 members and single-precision with 50 members. By reinvesting the cost savings from reducing precision into the ensemble size, skill scores could be improved.

Data assimilation is also an inherently uncertain process, because of the presence of model and observational error. The length of time between subsequent observations for the ensemble Kalman filter is an additional source of uncertainty. We might therefore expect rounding errors introduced by reducing precision to be within the bounds of uncertainty already present. If this is the case, then using reduced precision arithmetic could free up computing resources and permit an improvement in other aspects of the data assimilation system. Houtekamer et al. (2014) listed a number of parameters that can be increased in an operational ensemble Kalman filter to improve analysis quality with a commensurate increase in computational cost, including ensemble size, observation count, and model resolution. Miyoshi et al. (2014) demonstrated the benefits of

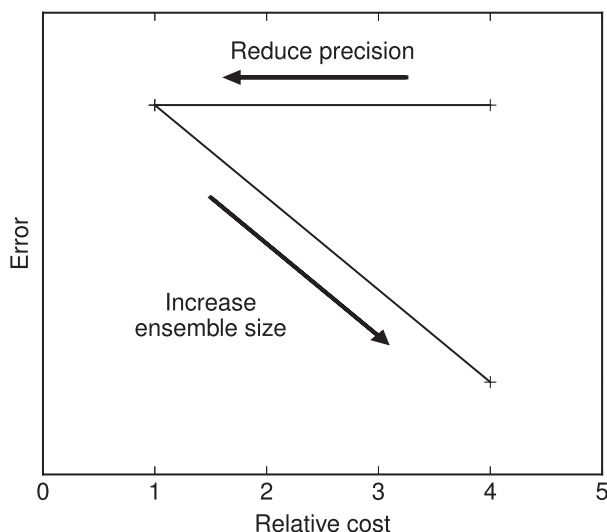


FIG. 1. The trade-off between numerical precision and ensemble size for the computation of the analysis. The relative cost refers to the total cost of data assimilation—both the forecast and the analysis computation.

increasing ensemble size, in particular, through the use of a 10 240 member ensemble Kalman filter. With such a large ensemble the localization length scale can be extended (or localization can be removed entirely) (Kondo and Miyoshi 2016), allowing more distant observations to affect the analysis at a grid point. Reducing precision could therefore improve the analysis quality, by providing the computational resources necessary to increase the ensemble size. This trade-off, illustrated in Fig. 1 as an error/cost diagram similar to that presented in Houtekamer et al. (2014), is the focus of our study.

Existing literature on numerical precision within the context of data assimilation is minimal. Miyoshi et al. (2016) compared the performance of a regional data assimilation system with double- and single-precision forecasting models. The single-precision simulation allowed a reduction in compute time of around 32%, with respect to the double-precision simulation, a figure produced by averaging the compute time reduction of the dynamics (45%) and the physics (21%).

In this paper, we assess the performance of an ensemble data assimilation system, consisting of the Lorenz '96 system and the ensemble square root filter, as precision is reduced from double-precision through to single- and half-precision. We investigate the sensitivity of these results to both the level of observation error and the length of the assimilation window. Additionally, we consider a trade-off between numerical precision and ensemble size. By reallocating the computational savings delivered by half- and single-precision arithmetic into the ensemble size, we can reduce the analysis error

for, hypothetically, no extra cost, with respect to double-precision assimilation. This in turn gives higher-quality weather forecasts. The speed-up from reducing precision is difficult to predict, so we present results for a variety of speed-up assumptions. The precision of the underlying code represents another trade-off between analysis quality and computational speed, similar to grid resolution or ensemble size. We hope that the results presented here will encourage operational weather forecasting centers to consider numerical precision more closely when designing operational data assimilation systems. Section 2 describes our experimental setup, section 3 contains the results, and section 4 gives some conclusions.

2. Data assimilation setup

Observing system simulation experiments (OSSEs) allow different assimilation algorithms to be compared and tested with complete freedom over the quantity and quality of observations. A “nature run” is first generated through a single integration of the best possible model available (the “truth model”), and observations of this nature run are then generated and fed into the assimilation system. The output from the assimilation system can then be verified against the (perfectly characterized) nature run. We employ such a system here in order to characterize the effects of a reduction of numerical precision from the double-precision standard.

a. The Lorenz ’96 system

We use the Lorenz ’96 system for both the truth and forecast models (Lorenz and Emanuel 1998; Lorenz 2006). This system provides a simple and efficient method to generate spatiotemporally chaotic time series analogous to those displayed by prognostic variables in the atmosphere. Its simplicity allows good statistics to be obtained in a short period of time, even on a desktop computer. For these reasons, it is commonly employed as a test bed in data assimilation studies (Whitaker and Hamill 2002; Ott et al. 2004; Nerger 2015).

For the truth model, we use the three-tier variant of the Lorenz ’96 system, introduced in Thornes et al. (2017). The tiers are labeled X, Y, and Z and are analogous to large-, medium- and small-scale dynamics, respectively, in the atmosphere. Each tier comprises a line of grid points with periodic boundary conditions. Each X grid point is “connected” to a number of Y grid points and each Y grid point is connected to a number of Z grid points. The forecast model used in the ensemble square root filter is a two-tier variant of the Lorenz ’96 system with a parameterization of the tendency of the missing Z tier. In this way, the difference between the forecasting

and truth models is intended to mimic the difference between a real forecasting model and the atmosphere itself. The truth and forecast models are illustrated in Fig. 2.

The three-tier system is described by the following equations:

$$\frac{dX_k}{dt} = X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F - \frac{hc}{b} \sum_{j=1}^J Y_{j,k}, \quad (1)$$

$$\begin{aligned} \frac{dY_{j,k}}{dt} = & -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} \\ & + \frac{hc}{b}X_k - \frac{he}{d} \sum_{i=1}^I Z_{i,j,k}, \end{aligned} \quad (2)$$

and

$$\frac{dZ_{i,j,k}}{dt} = edZ_{i-1,j,k}(Z_{i+1,j,k} - Z_{i-2,j,k}) - g_Z e Z_{i,j,k} + \frac{he}{d} Y_{j,k}. \quad (3)$$

The indices i , j , and k range from 1 up to I , J , and K for the Z, Y, and X tiers, respectively. The number of X variables is given by I , the number of Y variables is given by IJ , and the number of Z variables is given by IJK . In this case, I , J , and K are all equal to 8. There are therefore 584 grid points in total, as shown in Fig. 2a. The parameter F is a large-scale forcing term which determines the chaoticity of the model and is set to 20 in this case, placing the model firmly within the chaotic regime. The remaining parameters, h , c , b , e , d , and g_Z allow the frequency and amplitude of oscillation of the three tiers, as well as the coupling between tiers, to be tuned. These parameters are tuned such that the X tier oscillates slowly with a large amplitude, while the two tiers below vary more quickly at a lower amplitude. The parameter values are $h = 1$, $c = b = e = d = 10$, and $g_Z = 1$.

If Eqs. (1), (2), and (3) are defined as the “truth” then a complementary simplified forecast model can be constructed. We use a two-tier variant of the Lorenz ’96 system similar to Eqs. (1) and (2), only the Z tendency is replaced with a parameterization term, U_p :

$$\frac{dY_{j,k}}{dt} = -cbY_{j+1,k}(Y_{j+2,k} - Y_{j-1,k}) - cY_{j,k} + \frac{hc}{b}X_k - U_p. \quad (4)$$

The term U_p is tuned to match the average tendency of the missing Z tier, for which we use “additive noise,” given by

$$U_p = U_p^{\text{det}}(Y_{j,k}) + s_{j,k}, \quad (5)$$

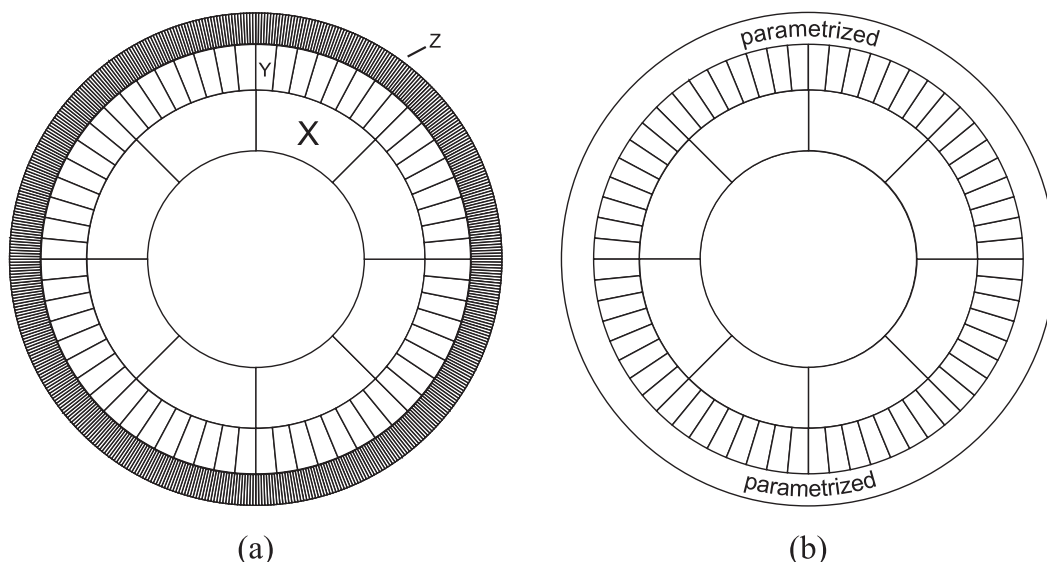


FIG. 2. (a) The three-tier Lorenz '96 system, used to generate the nature run. (b) The two-tier system, used as the forecast model. The inner tier is labeled "X," the central tier is labeled "Y," and the outer tier (which is not present in the forecast model) is labeled "Z."

where $U_p^{\text{det}}(Y_{j,k})$ is a deterministic polynomial function of the resolved variable $Y_{j,k}$ and $s_{j,k}$ is a random number generated from an autoregressive first-order stochastic process. For U_p^{det} we use a fourth-order polynomial:

$$U_p^{\text{det}}(Y_{j,k}) = c_4 Y_{j,k}^4 + c_3 Y_{j,k}^3 + c_2 Y_{j,k}^2 + c_1 Y_{j,k} + c_0. \quad (6)$$

The first-order stochastic process is generated by

$$s_{j,k} = \phi s_{j,k}^{\text{last}} + \sqrt{1 - \phi^2} e_{j,k}, \quad (7)$$

where ϕ is the lag-1 autocorrelation (characterizing the persistence of the random process), $s_{j,k}^{\text{last}}$ is the previous value of the process, and $e_{j,k}$ is a random number normally distributed with zero mean and variance σ_e^2 (characterizing the amplitude of the noise). The parameters ϕ and σ_e are tuned by matching the average behavior of the two-tier system to the three-tier system, in a similar process to that described in Arnold et al. (2013). The values of the parameters used are shown in Table 1.

Both the truth and forecast models are integrated with a fourth-order Runge–Kutta scheme of time step 0.005 nondimensional model time units (MTUs).

b. The serial ensemble square root filter

To perform data assimilation, we use the serial ensemble square root filter (Serial-EnSRF) [see Whitaker and Hamill (2002) and Tippett et al. (2003) for a detailed description of the update equations]. The Serial-EnSRF updates the ensemble mean and ensemble perturbations from the mean separately, in analogy with the original Kalman filter equations where the expectation value and covariance matrix are updated separately. As for atmospheric data assimilation, the state is only partially observed; only observations of the Y variables are assimilated. This leaves the X tier for verification.

The observation error covariance matrix \mathbf{R} is defined to be diagonal, and so observation errors for differing Y grid points are uncorrelated. Therefore, observations of each Y grid point can be processed one at a time. After the first observation is assimilated, the resulting analysis ensemble is treated as the background for assimilation of the next observation. This process is repeated until all observations have been processed. The update equations simplify greatly when processing scalar observations, and the need for matrix inversions and square roots is eliminated (Gelb 1974, Anderson 2003).

TABLE 1. The parameters defining the parameterization of the missing Z tier in the forecast model.

c_4	c_3	c_2	c_1	c_0	ϕ	σ_e^2
0.001 892	−0.066 811	0.131 826	0.242 742	0.039 970	0.997	0.015 876

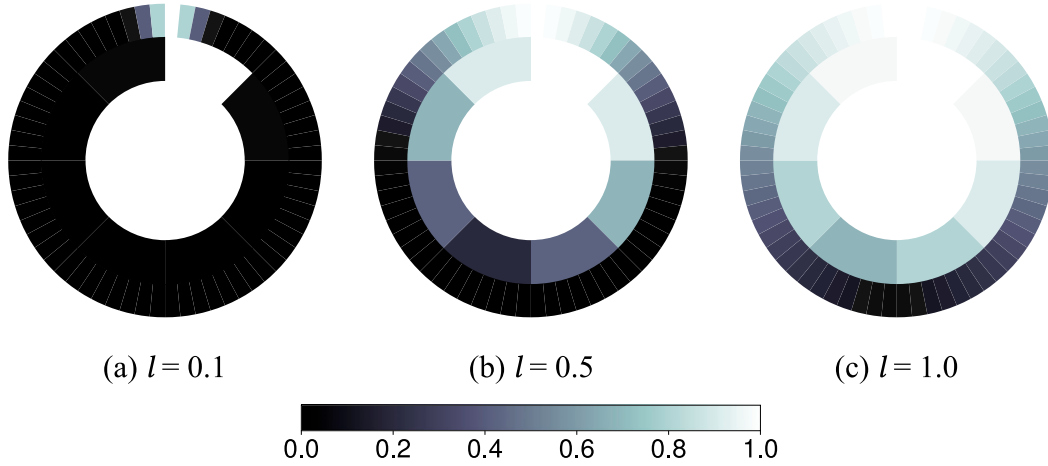


FIG. 3. The localization vector ρ applied to the Kalman gain when assimilating the first Y variable for three different length scales. The shading indicates the degree of localization.

1) COVARIANCE PREPROCESSING

At the analysis step of the Serial-EnSRF, the background covariance matrix \mathbf{P}^b is sampled from the background ensemble. The ensemble size used is typically suboptimal, and so this sample estimate suffers from sampling errors. To counter these errors, we use two covariance preprocessing techniques. The first technique is multiplicative covariance inflation, actually applied to the “scaled square root” of the background covariance matrix \mathbf{X}^b , where $\mathbf{P}^b = \mathbf{X}^b \mathbf{X}^{bT} / (N - 1)$ and N is the number of ensemble members, before the assimilation of observations:

$$\mathbf{X}_{\text{inf}}^b = r \mathbf{X}^b, \quad (8)$$

where r is a number slightly larger than 1 (e.g., 1.01). Inflation is used to counter the “inbreeding effect” in ensemble data assimilation that is present when the same ensemble is used to estimate the Kalman gain and the analysis ensemble (Houtekamer and Mitchell 1998; Anderson and Anderson 1999; Hamill et al. 2001).

The second technique is covariance localization, which is used in order to filter out spurious long-range correlations in \mathbf{P}^b that tend to contaminate the analysis (Hamill et al. 2001; Houtekamer and Mitchell 2001). The spurious correlations are a result of sampling error from an inadequate ensemble size, and so the required localization reduces as the ensemble size increases. Before observations are assimilated, a “localization vector” ρ is applied to the Kalman gain through an element-wise multiplication (a Schur product):

$$\mathbf{K}_{\text{loc}} = \rho \circ \mathbf{K}. \quad (9)$$

The localization vector has a value of 1 at the grid point where the observation currently being assimilated is

located. Moving away from this grid point, the localization vector reduces in magnitude and reaches 0 after a finite distance, known as the “localization scale.” We use the “Gaspari–Cohn” function to achieve this, which is a polynomial approximation to a Gaussian with compact support [Gaspari and Cohn (1999), their Eq. (4.19)]. This localization technique is often referred to as “B localization” (“B” refers to another notation for the background covariance matrix \mathbf{B} (Greybush et al. 2011). Note that B localization is usually applied to the full covariance matrix \mathbf{P}^b instead of to \mathbf{K} . In the case of sequential observation processing, the two techniques are equivalent.

Localization is applied not only to the covariances between Y grid points, but also to the covariances between Y grid points and X grid points. The localization between Y grid points depends simply on the distance (the number of Y grid points) between them. When localizing between a Y grid point and an X grid point, first the X grid point associated with the Y grid point is determined. Then the localization is computed based on the distance separating the two X grid points. The localization scale is assumed to be the same function of ensemble size in both cases, and so both are tuned simultaneously through the parameter l , where the Y–Y localization scale is $l_{YY} = 26l$ and the X–Y localization scale is $l_{XY} = 8l$. Figure 3 shows several examples of the localization vector ρ , when assimilating the first Y variable.

Both the optimal inflation and the optimal localization are assumed to be functions of ensemble size, assimilation frequency, and observation error, and so the inflation and localization are retuned with every change of one of these parameters. An example of the tuning process is illustrated in Fig. 4. The gray region denotes

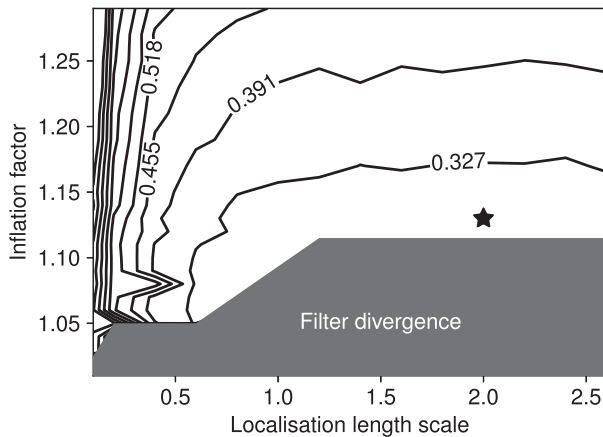


FIG. 4. Ensemble mean root-mean-square error as a function of covariance inflation and localization parameters, for 160-member assimilation. The star denotes the optimal choice of parameters. The shaded area denotes parameter choices that lead to unstable filtering, with a high likelihood of filter divergence.

parameter combinations that resulted in filter divergence. In this example, the optimal choice of inflation was 1.13 and the optimal localization length scale was 2.0, as the minimum error is located at this point, indicated by a star. On all occasions, we found that the optimal parameters for double-, single-, and half-precision assimilation were roughly the same. Therefore, we only tune the system at double-precision, and use the same parameters for single- and half-precision assimilation.

2) CHOICE OF ASSIMILATION PARAMETERS

To make a fair comparison between full- and reduced-precision results, it is important to choose a reference scenario that accurately represents an operational setup. For the operational EnKF at the Canadian Meteorological Centre, for example, assimilation is typically performed every 6 h (Houtekamer et al. 2014). Equivalently, we assimilate every 0.05 MTUs. This time scale was determined by computing the error-doubling time of the large-scale (X) tier, and comparing with the error-doubling time of the real atmosphere [approximately

1.5 days (Lorenz 2006)]. It should be noted, however, that this time scale is only approximate, and the true value could be anywhere between 0.025 and 0.075 MTUs. We use 40 ensemble members for our reference scenario and double-precision for all calculations. Additionally, observation errors for the Y variables are drawn from a zero-mean normal distribution with a standard deviation σ of 0.1, which is approximately 30% of the natural root-mean-square (RMS) variability of these variables. Figure 5 shows an excerpt of the evolution of the first Y variable from a nature run and the observations derived from this run. Limited results will be given for two other levels of observation error, $\sigma = 0.0035$ and $\sigma = 0.017$, which are approximately 1% and 5% of the natural RMS variability, respectively.

c. Reducing precision

We consider three levels of numerical precision in our investigation: double (64 bits), single (32 bits), and half (16 bits). These are described in the appendix. Lacking the hardware to implement true half-precision variables, we emulate them using the reduced-precision emulator described in Dawson and Düben (2017). Half-precision values are stored in a double-precision container, but significand bits beyond the 10th bit are zeroed manually after every operation. In some calculations, an IEEE 754 compliant rounding is performed on the resulting number to avoid biases. Furthermore, the 5-bit half-precision exponent is implicitly emulated through range checking on the results of floating-point operations: variables are forced to overflow or underflow if they are too large or small, respectively, to be stored in a half-precision floating-point variable. We also perform some limited experiments with a fixed double-precision exponent size and flexible significand width, using the same emulator.

In the experiments that follow, all calculations in the main loop (forecast/update cycle) are performed in the specified precision. The nature run is always generated using double-precision. Thornes et al. (2017) found error metrics particularly sensitive to the numerical precision of the X tier in the same Lorenz '96 system used

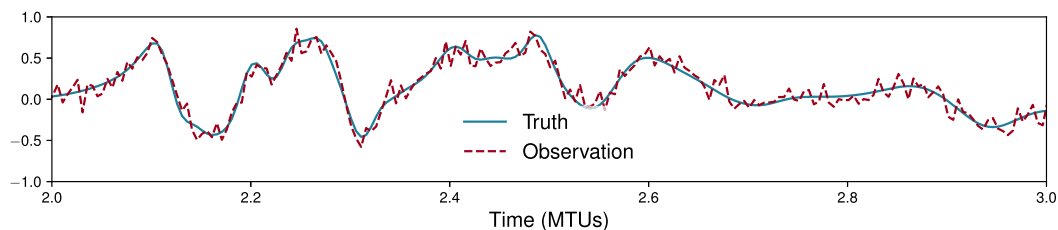


FIG. 5. A time series comparison of the signal displayed by the $Y_{1,1}$ grid point and the corresponding observational signal “seen” by the assimilation system. The observation error is 30% of the natural RMS variability.

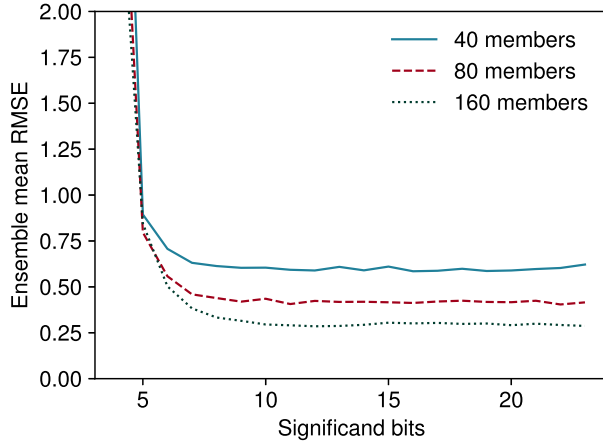


FIG. 6. Ensemble mean RMSE averaged over 250 MTUs at various levels of numerical precision, determined by the number of bits in the significand. Three sizes of ensemble are considered. The exponent is kept at the double-precision value, 11 bits.

here, reasoning that large-scale dynamics are more important than small-scale dynamics and therefore require more precise computations. Here, we do not consider such a scale-selective precision, and both tiers of the forecast model are performed at the same precision.

3. Results

a. Reduced precision analyses

Figure 6 shows the analysis ensemble mean root-mean-square error (RMSE) averaged over 5000 analyses of a data assimilation cycle for three different ensembles—40, 80, and 160 members—as a function of n_s , the number of significand bits. Assimilation is performed every 0.05 MTUs (every 10 time steps), so the

total length of simulations is 250 MTUs or around 4 atmospheric years. The ensemble mean RMSE at time k is given by

$$\text{RMSE}_k = \sqrt{\frac{1}{n_x} \sum_i^{n_x} (\bar{x}_{i,k}^a - x_{i,k}^t)^2}, \quad (10)$$

where n_x is the number of X variables (8 in this case) and $\bar{x}_{i,k}^a$ and $x_{i,k}^t$ are the i th X variables of the analysis and truth at analysis time k , respectively. The RMSE is proportional to the distance between the analysis ensemble mean and the truth, and therefore gives the simplest measure of the performance of an assimilation system. No half-precision range checking is performed in this case, so the exponent is still represented by 11 bits. The significand width can be reduced to around 8 bits with no discernible increase in the ensemble mean RMSE. Beyond this point, the error explodes and the system becomes unstable. This indicates that half-precision assimilation may be possible, as the half-precision significand width is 10 bits. To verify this, it is necessary to also emulate a 5-bit exponent, by using half-precision range checking.

Figure 7 shows the analysis ensemble mean RMSE with respect to that of the double-precision, 40-member reference, “D40”, as a function of precision for three ensemble sizes and three levels of observation error. Values are stated as a fraction of the ensemble mean RMSE of the D40 reference, for each level of observation error. The values shown were obtained by averaging over 10 experiments of 250 MTUs (5000 analyses) each, and the error bars were computed from the standard deviation of the same 10 experiments. When the observation error is 30% of the natural RMS variability, there is no noticeable degradation in the quality of the

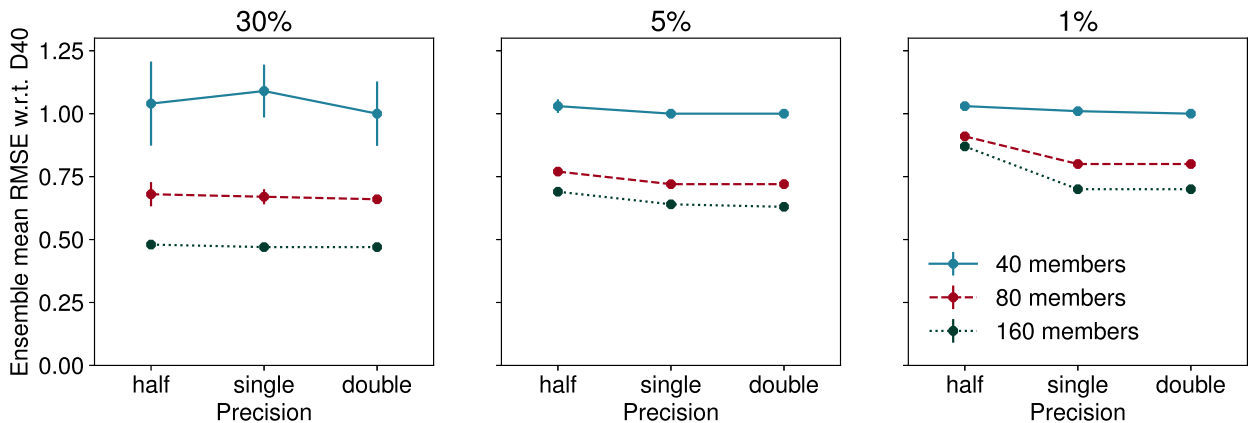


FIG. 7. RMSE as a function of numerical precision, ensemble size, and observation error. The observation error is stated as a percentage of the natural variability of the root-mean-square of the variable (from left to right) 30%, 5%, and 1%. Assimilation is performed at the “normal rate”: every 0.05 MTUs per 10 time steps.

analyses as precision is reduced from double- to half-precision, for either of the three ensemble sizes considered. When lowering the observation error to 5% of the natural RMS variability, the gains from increasing ensemble size diminish. Additionally, there is a slight, but significant, degradation in the analyses when using half-precision, which increases further as the observation error is lowered to 1% of the natural RMS variability. Our results are therefore sensitive to the observation error chosen. We consider observation errors of 5% and 1% to be unrealistic, and only consider observation errors of 30% from now on, as is done in other studies (Ott et al. 2004; Anderson et al. 2009). This may seem high compared with the error of a typical atmospheric instrument, but our study does not consider representativity error or errors in the observation operator. These two sources of error are far from negligible when doing atmospheric data assimilation (Janjić and Cohn 2006).

An assimilation system should be judged not just by mean-error statistics, but also by the reliability of the ensembles that it produces. Rank histograms allow us to measure this (Hamill et al. 2001). At each analysis time, the values of a given X variable reported by each analysis ensemble member are used to define bin limits. Thus, for an ensemble of 40 members there will be 41 bins. The truth at that time is then placed in a bin according to its value. This is repeated, in this case, for 5000 analyses (as before), and for all X variables, to produce a histogram. If this histogram is flat then the ensemble has an ideal spread—the truth is, on average, a plausible member of the ensemble. If the histogram slopes up at the edges, then the analyses are underdispersive, and if it slopes down at the edges then it is overdispersive. Figure 8 shows rank histograms of analyses generated from the same setups considered in Fig. 7 with an observation error of 30% the natural RMS variability. For the 80- and 160-member ensembles, 40 members are randomly sampled in order to determine the bin limits. Figure 8 indicates some underdispersiveness for all setups, with the degree of underdispersiveness decreasing as the ensemble size is increased, but crucially there is no apparent increase or decrease in dispersiveness as the numerical precision is altered.

b. Sensitivity to the length of the assimilation window

The results presented so far indicate that half-precision data assimilation is feasible, at least for a toy model. However, if the level of uncertainty in the system is reduced, the errors introduced by half-precision arithmetic become more noticeable. One way to reduce uncertainty is to reduce the length of the assimilation window or, in other words, to perform

assimilation more frequently. In this regime, the model will be corrected by assimilation more frequently and the average analysis error will therefore be reduced. Figure 9 shows the analysis ensemble mean RMSE when the assimilation window is made 10 times shorter. Clearly, the analyses generated using half-precision arithmetic are substantially worse than those generated using single- or double-precision. In fact, even with 160 ensemble members, half-precision is worse than single-precision with only 40 members.

As previous studies showed, some operations are affected more than others when reducing precision (Düben et al. 2017; Dawson et al. 2017). Therefore, when a degradation occurs, it is worthwhile to look closer and try to find the most sensitive operations. For example, a precision reduction can have a large impact on additions involving numbers of considerably different magnitude. Consider the operation $a = a + b$, where a and b are half-precision floating-point numbers with values 8 and 1×10^{-3} , respectively. The next representable floating-point number above 8 is 8.0078125, which is larger than the true value of the sum, 8.001. The increment b will therefore be rounded to zero, and the addition will have no effect.

Düben et al. (2017) identified a similar problem within the time-stepping scheme of a cloud-resolving model. Consider the Euler time-stepping scheme, given by

$$\psi^{t+1} = \psi^t + \Delta t \frac{\partial \psi}{\partial t}, \quad (11)$$

where ψ^t is a field at time t and Δt is the time step and all variables are at half-precision. If the increment $\Delta t(\partial \psi / \partial t)$ is smaller than the floating-point spacing then it will be rounded to zero and ψ^{t+1} will be equal to ψ^t . The field will therefore not advance in time, and significant errors will accrue.

One candidate for a precision-sensitive calculation in this particular setup is the Kalman filter update equation for the ensemble mean,

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^b + \mathbf{K}(y - \mathbf{H}\bar{\mathbf{x}}^b), \quad (12)$$

where $\bar{\mathbf{x}}^a$ and $\bar{\mathbf{x}}^b$ refer to the analysis and background ensemble mean state vectors, respectively; y is a scalar observation; \mathbf{H} is the observation operator; and \mathbf{K} is the Kalman gain. In a low-error assimilation regime (e.g., when assimilation is performed frequently) the analysis increments will be very small. If an increment is too small, it will be rounded to zero, and that observation will therefore have no effect. Figure 10 illustrates this. Over 100 forecast/assimilation cycles, the analysis increments [the individual elements of $\mathbf{K}(y - \mathbf{H}\bar{\mathbf{x}}^b)$] were

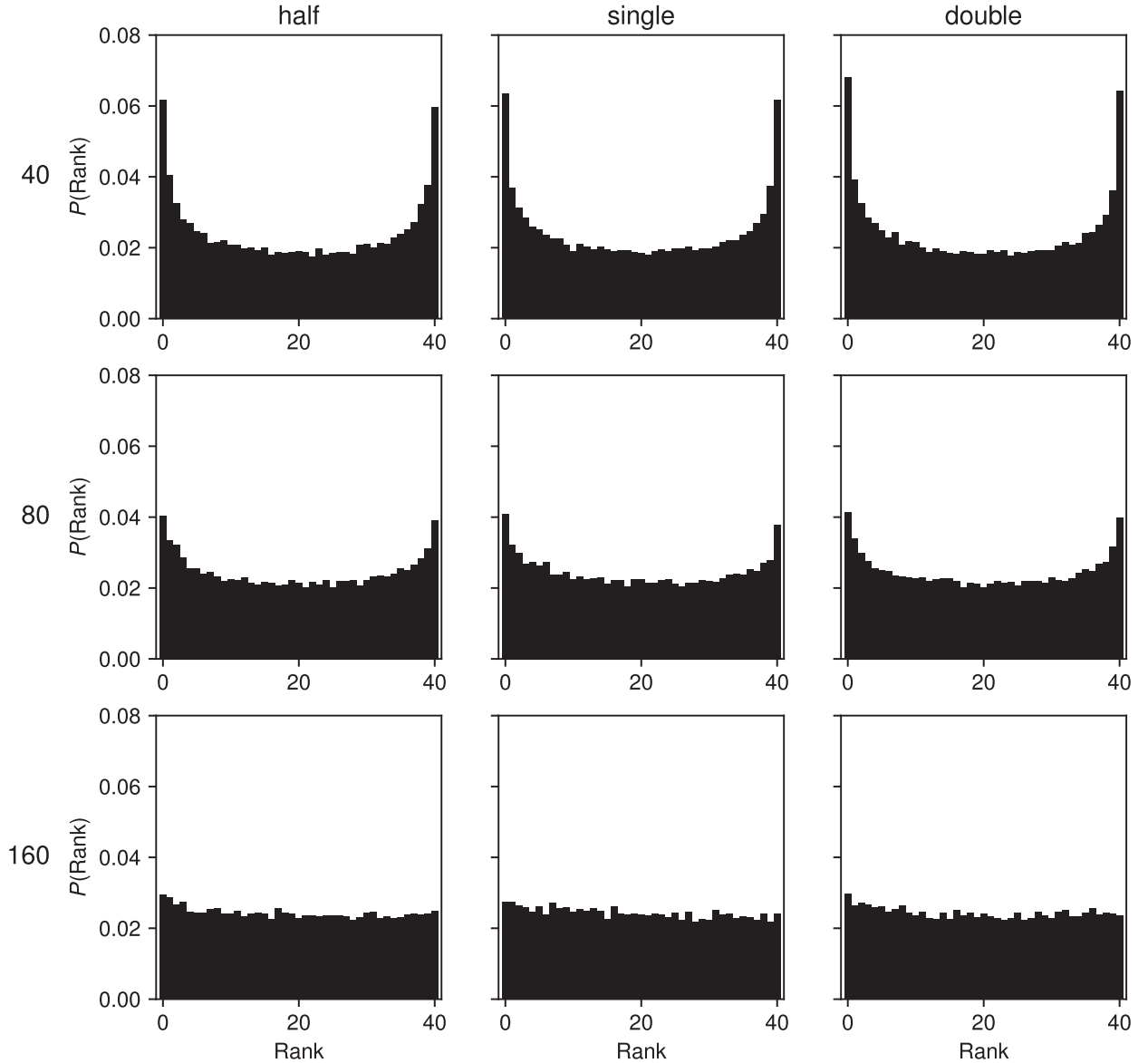


FIG. 8. Rank histograms of analyses generated using different assimilation setups: (from left to right) half-, single-, and double-precision with (from top to bottom) 40, 80, and 160 members. The observation error is 30% of the natural RMS variability.

divided based on whether they were resolved or unresolved (rounded to zero), and then plotted on a histogram (which is normalized by the total number of increments considered). Clearly a significant proportion of increments are rounded to zero. Note that some analysis increments were exactly zero already, because of the localization scheme, and are not shown on the histogram.

If the ensemble means, $\bar{\mathbf{x}}^b$ and $\bar{\mathbf{x}}^a$, are kept at higher precision, however, the addition of the analysis increment to the ensemble mean can be performed with no rounding error. When doing this, the analysis

increments are all resolved, as shown in Fig. 11. Only one variable, representing the background and analysis ensemble means, must be changed from half-precision to double-precision. Dübén et al. (2017) found a similar solution for the cloud resolving model; the contribution of advection to the $\partial\psi/\partial t$ term could be performed in half-precision, as long as the field ψ itself was kept at a higher precision.

Figure 12 gives the analysis error for frequent assimilation with the correction just mentioned. The error is considerably lower at half-precision compared to the error in Fig. 9. However, the error is still

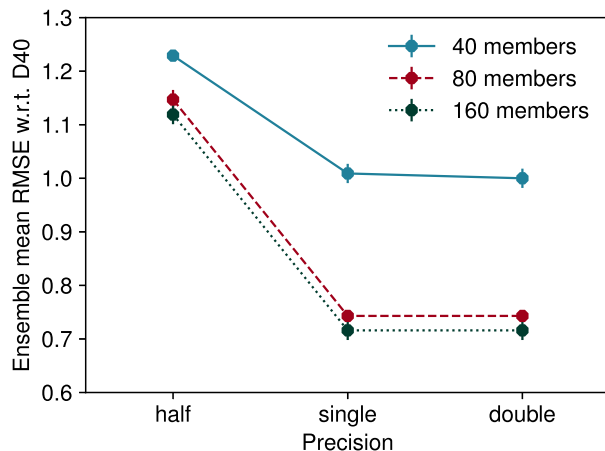


FIG. 9. Analysis ensemble mean RMSE as a function of IEEE 754 numerical precision for frequent assimilation.

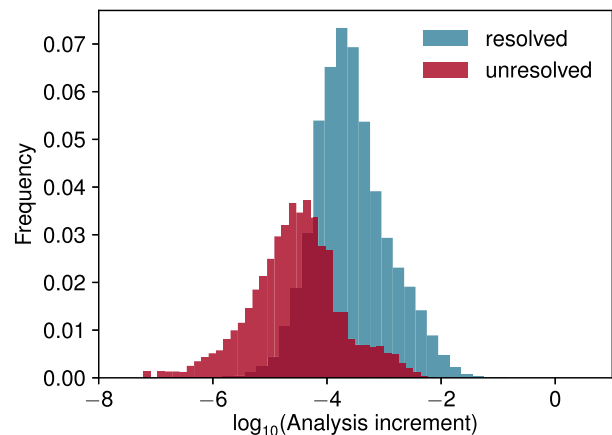


FIG. 10. Histograms of analysis increments for half-precision, frequent assimilation, both resolved and (due to rounding errors) unresolved. The histogram bars are translucent.

slightly higher than at double- and single-precision. It may be that, in this low-error assimilation regime, sensitivities in the forecast model to numerical precision become more noticeable, and that the “scale selective” approach advocated by [Thornes et al. \(2017\)](#) is warranted. This would require keeping the X variables at a higher precision. In any case, the frequent assimilation case corresponds to assimilation every 30 min, approximately, which is far more frequent than the global assimilation performed in operational centers. Therefore, the results for frequent assimilation do not undermine those presented before in [section 2c](#), as the scenario presented there is more realistic.

c. Trading numerical precision for ensemble size

The scaling of computational cost with numerical precision is difficult to predict. The speed-up obtained by moving from double- to single-precision, for example, will depend on the degree of vectorization and the arithmetic intensity of the model (the number of floating-point operations for every byte transferred from memory). To acknowledge the uncertainty in the cost estimate when reducing precision, we compare reduced precision data assimilation setups assuming a range of cost reductions. For single-precision, we assume that the speed-up will lie somewhere between 0% and 50%, and for half-precision, we consider speed-up factors between 50% and 75%.

[Figure 13a](#) shows the average analysis error, with respect to a double precision, 40-member reference experiment (“D40”), for a variety of assumptions about how the computational cost will fall when reducing precision. For each assumption, the ensemble size is

increased in proportion to the reduction in computational cost, so that each reduced precision setup has the same computational cost as the reference, D40. For all ensemble sizes, we again use an observation error standard deviation of 0.1 and perform assimilation every 0.05 MTUs. [Figure 13b](#) shows the forecast RMSE at day 3 of single-member forecasts initialized using the same assimilation setups as in [Fig. 13a](#). These forecasts were started from the ensemble mean of each analysis in an analysis cycle of 1000 MTUs (around 200 000 forecasts in total), excluding an assimilation spinup period of 10 MTUs. Note that the forecast error reaches its minimum at around 140 ensemble members, and therefore the slight increase in forecast errors for 160 members is simply a fluctuation from this saturation error.

[Figure 13](#) allows one to estimate the gains from re-investing precision into the ensemble size, for a given speed-up that is measured for the model on a given hardware architecture. For example, if we assume that single-precision computations would allow a reduction in wall-clock time of 20%, then we would be able to afford to use an additional 10 members, and the analysis error would fall by around 20% as a result. This would deliver a 5% reduction in day-3 RMSE of deterministic forecasts. Note that [Fig. 13](#) presumes a linear scaling in the cost of the data assimilation system as ensemble size is increased. This is true for the ensemble forecast, as each member is independent, and it is also true for the EnSRF algorithm, as computational complexity arguments predict ([Tippett et al. 2003](#)). The cost of the full data assimilation system therefore also scales linearly with the number of ensemble members, and we measured the speed of our configuration to confirm that this is the case.

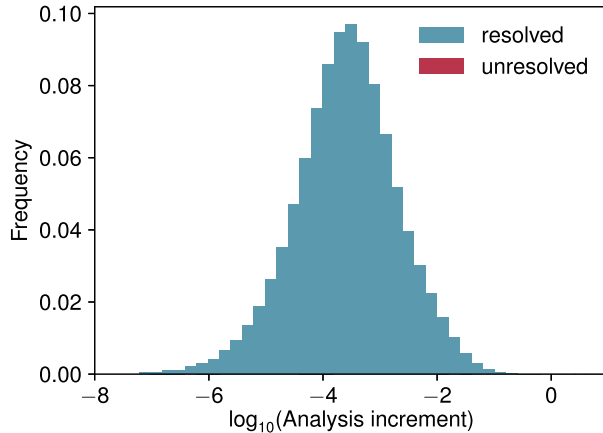


FIG. 11. As in Fig. 10, but only the analysis and background ensemble means are kept at double-precision. All increments are resolved.

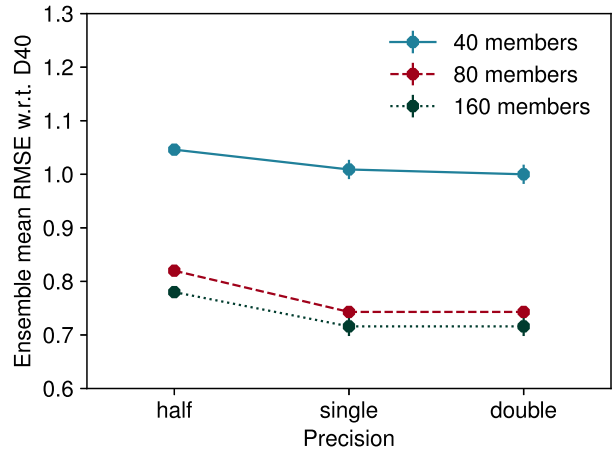


FIG. 12. Analysis ensemble mean RMSE as a function of IEEE 754 numerical precision for frequent assimilation. The ensemble mean is kept at double-precision.

What degree of cost reduction should we expect for a real atmospheric data assimilation system? It can be argued that weather and climate models, being, in general, memory-bound, will display a linear scaling of compute time with the width of the floating-point number. This would give roughly a 50% reduction in wall-clock time when reducing precision from double to single, and another 50% reduction when moving from single- to half-precision. Experiments with a single-precision version of IFS at ECMWF found a reduction in execution time of 40%, which is close to a 50% cost saving (Vána et al. 2017). Furthermore, recent developments in the graphics processing unit market suggest that another reduction of 50% in compute time from single- to half-precision can be achieved (NVIDIA 2017). However, to the best of our knowledge, no fluid dynamical model has been tested on this hardware, so naturally any results that follow from the linear scaling assumption should be considered hypothetical. Although some groups experimenting with single-precision data assimilation report a computation time reduction of as little as 10%–20% (J. Anderson 2017, personal communication), our results show that even a very small speed-up would be beneficial, if savings are reinvested. Also, note that, in our case, we do not realize any computational savings at half-precision, because we use a software emulator to emulate half-precision arithmetic. A final caveat of the discussion of cost savings is that data input and output is not considered, which can be significantly expensive in numerical weather predictions.

If a 50% reduction in wall-clock time could be realized by using single-precision instead of double-precision (in other words, if computations can be made 2 times

faster), the analysis and forecast errors would fall by around 30% and 25%, respectively, if the savings were reinvested into the ensemble size. Furthermore, if half-precision computations could deliver a 75% reduction in wall-clock time, with respect to double-precision, the analysis and forecast errors could be reduced by 55% and 40%, respectively, through another ensemble size reinvestment. Though we cannot predict the exact reduced-precision speed-up of an operational data assimilation system, we feel that these results should encourage operational centers to consider this trade-off seriously.

4. Conclusions

Using a toy model of atmospheric dynamics we have carried out an initial investigation of the effects of reducing numerical precision from double-precision to half-precision on the performance of an ensemble data assimilation routine. We found that the analysis error was largely insensitive to a significant reduction in the width of the floating-point significant, with a tolerance to as few as 8 bits. We then considered three IEEE 754 types: half, single, and double. The half-precision floating-point variables were emulated. We again found no increase in the analysis error as the precision was reduced from double to half, for a regime designed to approximate an operational data assimilation scenario (e.g., using an observation error of 30% the natural RMS variability of the prognostic variables). These results were sensitive to the level of observation error: when reducing the noise in the observations, the analysis error increased for half-precision assimilation.

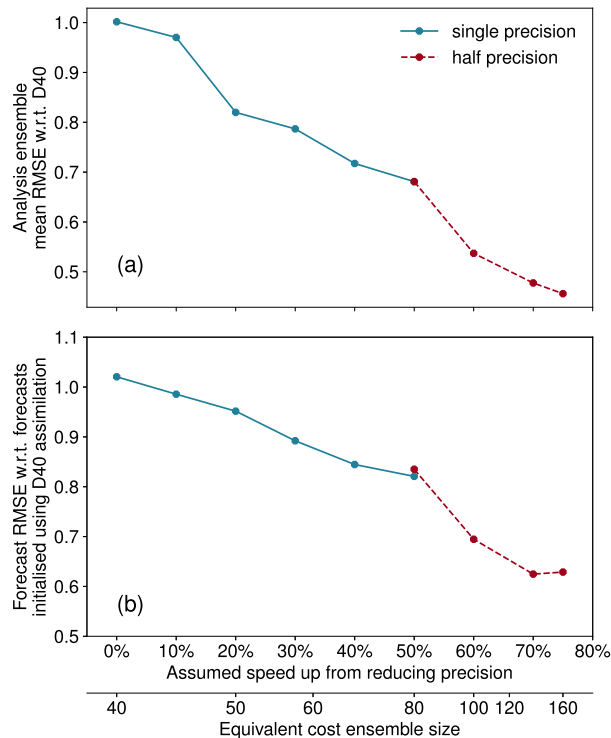


FIG. 13. (a) The analysis mean RMSE of single- and half-precision assimilation, with respect to double-precision with 40 members (“D40”), after reinvesting precision into the ensemble size. A range of possible speed-ups due to reducing precision are assumed and the equivalent cost ensemble size after reducing precision is given by the lower x axis. For example, if single-precision computations deliver a 20% reduction in the cost, then 50 members can be run at single-precision for the same cost as 40 double-precision members. This would deliver a 20% reduction in the analysis error. (b) As in (a), but the forecast RMSE at day 3 of forecasts initialized using each reduced precision assimilation setup is shown, with respect to forecasts initialized using double-precision with 40 members.

When the length of the assimilation window was reduced, half-precision assimilation gave a higher analysis error than single- and double-precision. We conjectured that this was because half-precision variables were unable to resolve the small analysis increments that result from such a low-error assimilation regime. However, the positive results from half-precision observed in the previous experiment can be partly regained by keeping just one key operation in double-precision: the Kalman filter update equation.

Finally, we considered a trade-off between numerical precision and ensemble size. By reducing precision and reinvesting the resulting cost saving into more ensemble members, the analysis and forecast errors can be reduced. The exact error reduction depends on one’s assumption about how much the cost will fall when reducing precision. For example, if

single-precision computations are 20% faster, the analysis error is reduced by 20%, but if they are 50% faster, the error falls by 30%. Additionally, if half-precision computations are 75% faster, then the analysis error can be reduced by 55%. It remains to be seen, however, whether the cost of a real atmospheric model will scale so favorably when reducing to half-precision.

These results demonstrate that forecast models are subject to different constraints when used for data assimilation instead of for, say, medium-range forecasting. [Thornes et al. \(2017\)](#) reported “deleterious results” when computing the evolution of the X variables in the two-tier Lorenz ’96 system at half-precision, so that they must be promoted to single-precision (even though the Y variables could be kept at half-precision). We experienced no such degradation, finding half-precision to be sufficient for both scales of the Lorenz ’96 system. This suggests that model error can be higher within an assimilation context, perhaps because it is partially masked by the other inherent error sources, such as observation error.

The cost of the Serial-EnSRF update algorithm scales linearly with the ensemble size ([Tippett et al. 2003](#)). Therefore, the precision-ensemble size trade-off described in [section 3c](#) can deliver a significant increase in the ensemble size. However, not all algorithms scale so favorably. For example, the scaling of the cost of the local ensemble transform Kalman filter with ensemble size is between quadratic and cubic ([Hunt et al. 2007](#)), and therefore a reduction in precision may not deliver as significant gains as with the EnSRF. Nonetheless, if the algorithm can be made faster through a precision reduction, without impacting the quality of analyses, this would be beneficial for operational data assimilation, where the analyses must be computed within an allotted time.

The results presented here suggest that precision in data storage should also be reconsidered and adjusted to the minimal level that can be justified by the level of model error. If a model can run at half-precision, then the restart files could also be reduced to half-precision, halving the file size with respect to single-precision output.

In this paper, we have demonstrated that reducing precision could provide the computational resources necessary to boost the ensemble size, and that such a reinvestment procedure would improve weather forecast skill. In the future, we plan to extend this study to a more complicated data assimilation scenario, with a general circulation model (GCM) instead of the Lorenz ’96 system. We hope the results presented here will

encourage operational forecasting centers to consider more closely the potential benefits of reducing precision and to make more variable-precision hardware available to modelers.

Acknowledgments. Many thanks are owed to Andrew Dawson, for his help in using the reduced precision emulator. The authors also thank three anonymous reviewers and Jeffrey Anderson for their very helpful reviews. Sam Hatfield is funded by NERC Grant NE/L002612/1. Peter Düben gratefully acknowledges the funding from the ESIWACE project under Grant 675191. Peter Düben, Tim Palmer, and Aneesh Subramanian received funding from an ERC grant (Towards the Prototype Probabilistic Earth-System Model for Climate Prediction, Project Reference 291406). The code used in this study is available online (Hatfield 2017).

APPENDIX

The IEEE 754 Standard for Floating-Point Numbers

Under the IEEE 754 standard, a floating-point number x has three components: a sign, a significand, and an exponent, related according to

$$x = (-1)^s \left(1 + \sum_{i=1}^{n_s} b_i 2^{-i} \right) 2^E. \quad (\text{A1})$$

The sign is determined by a single bit s for all precision levels. The significand is the second term in the product in Eq. (A1), determined by n_s bits with elements b_i . The exponent is determined by n_e bits with elements c_i and is given by

$$E = \sum_{i=1}^{n_e} c_i 2^{i-1} - 2^{n_e-1} + 1. \quad (\text{A2})$$

Double-, single-, and half-precision variables have $(n_s, n_e) = (52, 11)$, $(n_s, n_e) = (23, 8)$, and $(n_s, n_e) = (10, 5)$, respectively.

The precision of a “float” is dependent on both n_s and n_e . For example, for a fixed exponent, the distance between two consecutive floats is 2^{E-n_s} . The exponent also determines the range of permissible values, outside of which a float will over or underflow. The properties of half-, single-, and double-precision floats are summarized in Table A1. Note that values below the given ranges are possible, but these so-called subnormal values are handled differently.

TABLE A1. Summary of the floating-point types considered.

	Spacing when $2^E = 1$	Range
Half	$\sim 10^{-3}$	$(\sim 10^{-6}, 65504)$
Single	$\sim 10^{-7}$	$(\sim 10^{-38}, \sim 10^{38})$
Double	$\sim 10^{-16}$	$(\sim 10^{-308}, \sim 10^{308})$

REFERENCES

- Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, **131**, 634–642, [https://doi.org/10.1175/1520-0493\(2003\)131<0634:ALLSFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2).
- , and S. L. Anderson, 1999: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.*, **127**, 2741–2758, [https://doi.org/10.1175/1520-0493\(1999\)127<2741:AMCIOT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2).
- , 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, <https://doi.org/10.1175/2009BAMS2618.1>.
- Arnold, H. M., I. M. Moroz, and T. N. Palmer, 2013: Stochastic parametrizations and model uncertainty in the Lorenz ’96 system. *Philos. Trans. Roy. Soc. London*, **371A**, 20110479, <https://doi.org/10.1098/rsta.2011.0479>.
- Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.
- Dawson, A., and P. D. Düben, 2017: Rpe v5: An emulator for reduced floating-point precision in large numerical simulations. *Geosci. Model Dev.*, **10**, 2221–2230, <https://doi.org/10.5194/gmd-10-2221-2017>.
- , D. D. Peter, D. A. Macleod, and T. N. Palmer, 2017: Reliable low precision simulations in land surface models. *Climate Dyn.*, in press.
- Düben, P. D., and T. N. Palmer, 2014: Benchmark tests for numerical weather forecasts on inexact hardware. *Mon. Wea. Rev.*, **142**, 3809–3829, <https://doi.org/10.1175/MWR-D-14-00110.1>.
- , H. McNamara, and T. N. Palmer, 2014: The use of imprecise processing to improve accuracy in weather & climate prediction. *J. Comput. Phys.*, **271**, 2–18, <https://doi.org/10.1016/j.jcp.2013.10.042>.
- , A. Subramanian, A. Dawson, and T. N. Palmer, 2017: A study of reduced numerical precision to make superparameterization more competitive using a hardware emulator in the OpenIFS model. *J. Adv. Model. Earth Syst.*, **9**, 566–584, <https://doi.org/10.1002/2016MS000862>.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, <https://doi.org/10.1002/qj.49712555417>.
- Gelb A., Ed., 1974: *Applied Optimal Estimation*. The MIT Press, 384 pp.
- Greybush, S. J., E. Kalnay, T. Miyoshi, K. Ide, and B. R. Hunt, 2011: Balance and ensemble Kalman filter localization techniques. *Mon. Wea. Rev.*, **139**, 511–522, <https://doi.org/10.1175/2010MWR3328.1>.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129**, 2776–2790, [https://doi.org/10.1175/1520-0493\(2001\)129<2776:DDFOBE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2).

- Hatfield, S., 2017: samhatfield/lorenz96-ensrf: Publication (version v1.0.0). Zenodo, accessed 8 December 2017, <https://doi.org/10.5281/zenodo.571203>.
- Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, [https://doi.org/10.1175/1520-0493\(1998\)126<0796:DAUAEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2).
- , and —, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, [https://doi.org/10.1175/1520-0493\(2001\)129<0123:ASEKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2).
- , and F. Zhang, 2016: Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **144**, 4489–4532, <https://doi.org/10.1175/MWR-D-15-0440.1>.
- , X. Deng, H. L. Mitchell, S.-J. Baek, and N. Gagnon, 2014: Higher resolution in an operational ensemble Kalman filter. *Mon. Wea. Rev.*, **142**, 1143–1162, <https://doi.org/10.1175/MWR-D-13-00138.1>.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- Janjić, T., and S. E. Cohn, 2006: Treatment of observation error due to unresolved scales in atmospheric data assimilation. *Mon. Wea. Rev.*, **134**, 2900–2915, <https://doi.org/10.1175/MWR3229.1>.
- Kondo, K., and T. Miyoshi, 2016: Impact of removing covariance localization in an ensemble Kalman filter: Experiments with 10 240 members using an intermediate AGCM. *Mon. Wea. Rev.*, **144**, 4849–4865, <https://doi.org/10.1175/MWR-D-15-0388.1>.
- Lorenz, E. N., 2006: Predictability—A problem partly solved. *Predictability of Weather and Climate*, T. Palmer and R. Hagedorn, Eds., Cambridge University Press, 40–58, <https://doi.org/10.1017/CBO9780511617652.004>.
- , and K. A. Emanuel, 1998: Optimal sites for supplementary weather observations: Simulation with a small model. *J. Atmos. Sci.*, **55**, 399–414, [https://doi.org/10.1175/1520-0469\(1998\)055<0399:OSFSWO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<0399:OSFSWO>2.0.CO;2).
- Miyoshi, T., 2011: The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Mon. Wea. Rev.*, **139**, 1519–1535, <https://doi.org/10.1175/2010MWR3570.1>.
- , K. Kondo, and T. Imamura, 2014: The 10 240-member ensemble Kalman filtering with an intermediate AGCM. *Geophys. Res. Lett.*, **41**, 5264–5271, <https://doi.org/10.1002/2014GL060863>.
- , and Coauthors, 2016: “Big data assimilation” toward post-petascale severe weather prediction: An overview and progress. *Proc. IEEE*, **104**, 2155–2179, <https://doi.org/10.1109/JPROC.2016.2602560>.
- Nerger, L., 2015: On serial observation processing in localized ensemble Kalman filters. *Mon. Wea. Rev.*, **143**, 1554–1567, <https://doi.org/10.1175/MWR-D-14-00182.1>.
- NVIDIA, 2017: NVIDIA TESLA P100: Infinite Compute Power for the Modern Data Center. NVIDIA Corporation, accessed 8 December 2017, <http://www.nvidia.com/object/tesla-p100.html>.
- Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428, <https://doi.org/10.3402/tellusa.v56i5.14462>.
- Palmer, T. N., 2014: More reliable forecasts with less precise computations: A fast-track route to cloud-resolved weather and climate simulators? *Philos. Trans. Roy. Soc. London*, **372A**, 20130391, <https://doi.org/10.1098/rsta.2013.0391>.
- Sutter, H., 2005: A fundamental turn toward concurrency in software. *Dr. Dobbs's J.*, **30**, 16–22.
- Thornes, T., P. Düben, and T. N. Palmer, 2017: On the use of scale-dependent precision in Earth System modelling. *Quart. J. Roy. Meteor. Soc.*, **143**, 897–908, <https://doi.org/10.1002/qj.2974>.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490, [https://doi.org/10.1175/1520-0493\(2003\)131<1485:ESRF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2).
- Vána, F., P. Düben, S. Lang, T. N. Palmer, M. Leutbecher, D. Salmond, and G. Carver, 2017: Single precision in weather forecasting models: An evaluation with the IFS. *Mon. Wea. Rev.*, **145**, 495–502, <https://doi.org/10.1175/MWR-D-16-0228.1>.
- Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, [https://doi.org/10.1175/1520-0493\(2002\)130<1913:EDAWPO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2).
- , and —, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, <https://doi.org/10.1175/MWR-D-11-00276.1>.