Liu Huaran, Lu Feiyu, Liu Zhengyu, Liu Yun, and Zhang Shaoqing, 2016: Assimilating atmosphere reanalysis in coupled data assimilation. *J. Meteor. Res.*, **30**(4), doi: 10.1007/s13351-016-6014-1.(in press)

Assimilating Atmosphere Reanalysis in Coupled Data Assimilation

LIU Huaran[1, 2] (刘华然)，LU Feiyu [2]（卢飞雨）, LIU Zhengyu [2] (刘征宇)，LIU Yun[3]（刘赟）, and ZHANG Shaoqing [4]（张绍晴）

[1] *Dept. Atmospheric Sciences, Nanjing University of Information Science &Technology, China*

[2] *Dept. Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, USA*

[3] *Dept. Atmospheric and Oceanic Science, University of Maryland, USA*

[4] *Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, USA.*

(Received February 6, 2016; in final form May 5, 2016)

Supported by Nanjing University of Information Science and Technology

Corresponding author: Huaran Liu

Email: hliu287@wisc.edu.

ABSTRACT

This paper tests the idea of substituting the atmospheric observations with atmospheric reanalysis when setting up a coupled data assimilation system. The paper focuses on the quantification of the effects on the oceanic analysis resulted from this substitution and designs four different assimilation schemes for such substitution . A coupled Lorenz96 system is constructed and an ensemble Kalman filter is adopted. The atmospheric reanalysis and oceanic observations are assimilated into the system and the analysis quality is compared to a benchmark experiment where both atmospheric and oceanic observations are assimilated. Four schemes of assimilating the reanalysis are designed and they differ in the generation of the perturbed observation ensemble and the representation of the error covariance matrix. The results show that when the reanalysis is assimilated directly as independent observations, the RMSE increase of oceanic analysis relative to the benchmark is less than 16% in the perfect model framework; in the biased model case, the increase is less than 22%. This result is robust with sufficient ensemble size and reasonable atmospheric observation quality (e.g frequency, noisiness, density). If the observation is overly noisy, infrequent, sparse, or the ensemble size is insufficiently small, the analysis deterioration caused by the substitution is less severe since the analysis quality of the benchmark also deteriorates significantly due to worse observations and undersampling. The results from different assimilation schemes highlights the importance of two factors: accurate representations of the error covariance of the reanalysis and the temporal coherence along each ensemble member, which are crucial for the analysis quality of the substitution experiment.

Key words: data assimilation, reanalysis data, ensemble Kalman filter

1. Introduction

Coupled data assimilation (CDA) uses a coupled model to extract information from observations that are available in one or more media and produces continuous time series of the climate states. Compared to single component assimilation, CDA incorporates the full impact of observations across the air-sea interface and allows the covariability of the atmospheric and oceanic states, and thus it can provide consistent state estimation of the coupled system for further study of the climate variability and the initialization of coupled general circulation models (CGCM) (Kitoh and Arakawa, 1999; Arakawa and Kitoh, 2004; Luo et al., 2008; Sugiura et al., 2008; Tardif et al., 2015; Zhang et al., 2005, 2007; Zhang, 2011). Despite the huge benefits and demand for CDA, the implementation of CDA has both theoretical and technical challenges, for example, the estimation of the coupled error covariance matrix (e.g. Lu et al., 2015; Han et al., 2013) and the huge computational costs of CDA experiments in fully-coupled models. The NCEP Climate Forecast System Reanalysis (CFSR) was completed for the period 1979-2009 (Saha et al., 2010). It is a weak CDA system where the atmospheric and oceanic data assimilation are performed independently and the coupling is only through model dynamics.

In this paper, we want to explore the idea of assimilating atmospheric reanalysis data in a CDA system and its resultant consequences. This idea is motivated by the desire to find an efficient way to get ocean analysis from a CDA process that incorporates both the atmospheric and oceanic observations. The atmospheric observations include hundreds of types with different formats, coverage, frequency, etc., which makes it nearly impossible for an individual or a small group to collect and assimilate all these observations to set up the CDA system independently. Reanalysis datasets incorporate

millions of observations, which includes, but is not limited to, radiosonde, satellite, buoy, aircraft and ship reports, into a stable data assimilation system (e.g., Kalnay et al., 1996; Kistler et al., 2001; Kanamitsu et al., 2002; Uppala et al., 2005; Kobayashi et al., 2015; Saha et al., 2010; Dee et al., 2011). In addition, these datasets provide global coverage with constant spatial and temporal resolution over three or more decades for hundreds of variables (e.g., Kalnay et al., 1996; Kistler et al., 2001; Kanamitsu et al., 2002; Uppala et al., 2005; Kobayashi et al., 2015; Saha et al., 2010; Dee et al., 2011), which makes them relatively straightforward to handle from a processing standpoint. If it is feasible to substitute the atmospheric reanalysis datasets for actual observational data, we can set up CDA systems using different models much more easily and expect reasonable output of model analysis, especially oceanic analysis. Zhang et al.(2007) has assimilated atmospheric reanalysis directly as observations in a fully coupled climate model with a CDA system, without the examination of the possible effects brought out by the substitution. Yet their results are still promising and their assimilation successfully reconstructs the twentieth-century ocean heat content variability and trends in most locations. This indicates that it is feasible to substitute the atmospheric observations with reanalysis in a CDA process. However, the resultant effects on the analysis from the substitution is never carefully studied. In this paper, we will test this idea with an emphasis on the quantification of the resultant effects from the substitution, and investigates the assimilation schemes associated with this substitution.

A coupled Lorenz96 model (Lorenz, 1996) representing the atmosphere and the ocean is constructed to test the idea of assimilating atmospheric reanalysis data as observations in a CDA system. The results will be compared to the best-case scenario (benchmark)

where both the atmospheric and oceanic observations are assimilated. The paper is organized as follows: the methodology is shown in section 2; experiments and results are presented in section 3; tests on different assimilation schemes are shown in section 4; section 5 provides a concluding summary.

2. Methodology

2.1 The model

A dynamical system is set up by coupling two 40-variable (nv=40) Lorenz96 systems (Lorenz, 1996), one representing the atmosphere (Eq.1) and the other representing the ocean (Eq.2).

$$\frac{dX_j}{dt} = [X_{j+1} - X_{j-2}] * X_{j-1} - (1 - C_a) * X_j + F_a + C_a * [Y_j - X_j]. \quad (1)$$

$$M * \frac{dY_j}{dt} = [Y_{j+1} - Y_{j-2}] * Y_{j-1} - Y_j + F_O + Co * [X_j - Y_j]. \quad (2)$$

The atmosphere and ocean are coupled through the flux terms $C_a * [Y_j - X_j]$ and $C_o * [X_j - Y_j]$, where C_a =2.0 and C_o =0.1 are the coupling coefficients for the atmosphere and ocean, respectively. F_a and F_O represent the external forcing, and in this case F_a =8, F_O =0 such that the ocean is only forced by the atmosphere. The oceanic

timescale is controlled by coefficient M, which is chosen to be 20. Fig.1 shows the typical time evolution of variables X_1 and Y_1. The climatological standard deviations averaged over 40 variables for X and Y are 3.86 and 0.47, respectively. Therefore, the observational errors are arbitrarily set at 1.0 for the atmosphere and 0.1 for the ocean. In all experiments, the integration time step is 0.005, and 1 time unit is roughly 5 days. We forward the model for $2 * 10^5$ time steps (1000 time units). The first 20000 spin-up steps are discarded when we evaluate the analysis quality.

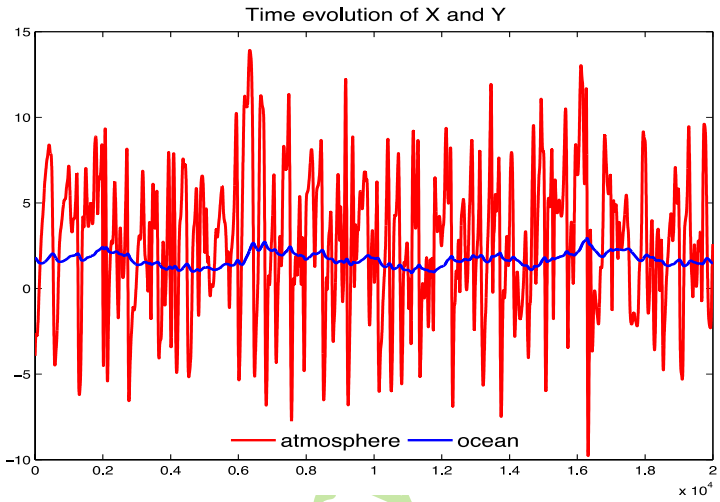


Fig. 1. Typical time evolution of X_1 (red) and Y_1 (blue).

2.2 The assimilation procedure and diagnostics

We will employ the ensemble Kalman filter (EnKF) with perturbed observation (Evensen, 1994; Burgers et al., 1998; Houtekamer and Mitchell, 1998). 80 ensemble members are used ($ens = 80$). Covariance localization is performed by applying a Schur product (Schur, 1911) to the forecast error covariance matrix (Gaspari and Cohn, 1999). Covariance inflation is also applied with the relaxation method by Zhang et al. (2004).

The root-mean-square error (RMSE) from all analysis steps is calculated to evaluate the data assimilation performance, given by

$$\mathrm{RMSE} = \sqrt{\frac{1}{nt}\frac{1}{nv}\sum_t \sum_i [X_{i,t} - X_{i,t}^T]^2}\ . \qquad (3)$$

Where X_i and X_i^T are analysis and truth, respectively, at gridpoint i and time step t. In reality, the benchmark experiment would be the best-case scenario for a CDA system and should produce the best analysis possible. Therefore, it is used to evaluate the performance of our proposed schemes. The RMSE of different experiments that assimilate the reanalysis are then normalized by that of the benchmark experiment as in Eq.4

$$Ratio = \frac{R - R_b}{R_b} \times 100\%. \qquad (4)$$

where R represents the RMSE of the experiment in which the atmospheric reanalysis is assimilated as observation, and R_b represents the RMSE of the benchmark experiment. We repeat 90 simulations for each experiment and the results are displayed via boxplots.

2.3 Model framework

In this section, we will introduce 3 models that are used to generate the true state, reanalysis and conduct the experiments respectively. The purpose of this is to allow for model bias to test the robustness of the quantified effects. If these three models are the same, then it is a perfect model framework with no model bias.

Model 1: Generate the true state and observation. The true state is a control run of this model and the observations are generated by adding a Gaussian white noise $N(0, \sigma_o)$ to the true state, where σ_o is the observational error.

Model 2: Generate reanalysis by assimilating observations from model 1. This mimic the fact that different research centers generate the reanalysis through their own GCMs (e.g. Kalnay et al., 1996; Kistler et al., 2001; Uppala et al., 2005; Kobayashi et al., 2015; Saha et al., 2010) and these GCMs are biased with regard to the the model used to generate the true state.

Model 3: Conduct the benchmark and substitution experiments. Model 3 is different from model 2 in that the model where CDA system is set up can be different from the model used to generate reanalysis.

The relationship among the three models is illustrated in Fig.2. In the perfect model framework, three models are the same and they use the default parameter values with Runge-Kutta 4 (RK4) integration scheme; In the biased mode framework, model bias is mimicked by different integration schemes and slight variations on model parameters. The detailed setup is summarized in Table.1.

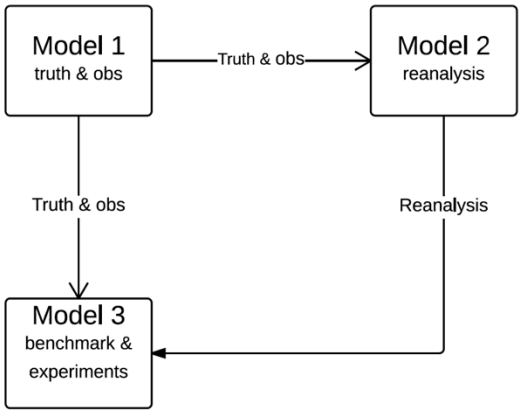


Fig. 2. Illustration of model setup. Truth and observations are generated in model 1. Reanalysis is generated in model 2. Experiments are conducted in model 3. In perfect model case the three models are the same, while in biased model case all three are different from each other.

	Integration Scheme	Model Parameters
Model 1	Runge-Kutta 2	Standard parameters
Model 2	Runge-Kutta 2	5% increase from Standard parameters
Model 3	Runge-Kutta 4	Standard parameters

Table 1. Model setup in biased model framework

2.4 The benchmark and substitution experiment

In the benchmark experiment conducted in model 3, atmospheric and oceanic observations generated in model 1 are assimilated. Variables X_1, X_3, X_5, …, X_{39} are observed every 20 integration time steps and variables Y_1, Y_3, Y_5, …, Y_{39} are observed every 40 integration time steps, unless specified otherwise. This is the best-case scenario where all observations that are available are assimilated.

The atmospheric reanalysis data used for the substitution experiments is generated in model 2. The observations are the same as that in the benchmark experiment except no oceanic observations are assimilated since the effects of the ocean data assimilation on the atmosphere are small in this coupled Lorenz96 system. The reanalysis is the ensemble mean output. We also preserve the reanalysis ensemble for the assimilation scheme design later.

In the substitution experiments, the atmospheric reanalysis and oceanic observations are assimilated into model 3. The reanalysis is assimilated with the same frequency as the benchmark experiment, while oceanic observations stay unchanged. Although the atmospheric observations are not available at every gridpoint, the reanalysis will provide additional observations at unobserved locations..

2.5 Assimilating the reanalysis

The most straightforward way to assimilate the reanalysis is to simply treat the reanalysis as independent observations. First the error covariance matrix of the reanalysis is calculated, and then the reanalysis is independently perturbed according to the diagonal values of the matrix, namely the variances of the observations. The error covariance matrix of the reanalysis can be calculated as below:

$$R_t = cov < X - X^T > \qquad (5)$$

$$= \begin{bmatrix} cov < X_1 - X_1^T, X_1 - X_1^T > & cov < X_1 - X_1^T, X_2 - X_2^T > & \dots & cov < X_1 - X_1^T, X_{40} - X_{40}^T > \\ cov < X_2 - X_2^T, X_1 - X_1^T > & cov < X_2 - X_2^T, X_2 - X_2^T > & \dots & cov < X_2 - X_2^T, X_{40} - X_{40}^T > \\ \vdots & \vdots & & \vdots \\ cov < X_{40} - X_{40}^T, X_1 - X_1^T > & cov < X_{40} - X_{40}^T, X_2 - X_2^T > & \dots & cov < X_{40} - X_{40}^T, X_{40} - X_{40}^T > \end{bmatrix}.$$

where X and X^T are the time series of the reanalysis and truth respectively; $cov <>$ calculates the covariance. $R_t$is shown in Fig.3. The reanalysis at any given location is perturbed with Gaussian noise that has the same standard deviation as the square root of the corresponding diagonal element in R_t. Consequently, the error matrix used in calculating the Kalman gain is the R_t matrix with its off-diagonal elements set to zero due to the assumption of independence among different observations. In the real world, however, X^T is unknown. We can either use observation to replace X^T in Eq.5, or we can

use the averaged sample covariance of the original reanalysis ensemble (see *R* in Eq.7), which will be introduced later in section 4.

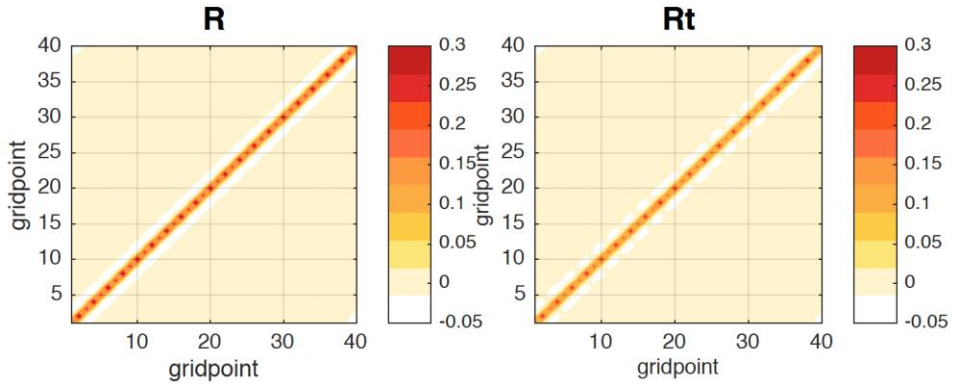


Fig. 3. (left) The time average of flow-independent ensemble covariance matrix R and (right) the temporal covariance matrix Rt. Point (i, j) indicates the covariance between the i^{th} and j^{th} atmospheric variables.

3. Experiments and results

Following the procedure in section 2d, the first type of experiment is denoted as UNCORR in Fig.4, which stands for "uncorrelated observation ensemble" since the observation ensemble is the reanalysis plus independent Gaussian white noise. Compared to the benchmark, the increase of the average RMSE of the ocean variables over 90 simulations is 11.41% in the perfect model framework and 16.93% in the biased model framework. For the 90 simulations in the perfect model framework (Fig. 4b), the maximum and minimum RMSE increase are 15.46% and 7.42%. In the biased model framework (Fig. 4d), the maximum and minimum non-outlier RMSE increase are 21.79% and 11.81%. If only ocean data assimilation is carried out in this coupled Lorenz96 model, the ocean component will not be constrained by the oceanic observations alone and the

oceanic RMSE can reach the climatological standard deviation. This is because in this simple coupled model, the ocean component is purely driven by the atmosphere and the feedback from the ocean to the atmosphere is small, hence we are not able get a reasonable oceanic analysis if the atmosphere is not well constrained. In this sense, substituting the atmospheric observations with the atmospheric reanalysis in a CDA process is better than assimilating oceanic observations alone in a coupled system. The performance of assimilating reanalysis is further tested with varied atmospheric observation frequency, atmospheric observation error, atmospheric observation density and ensemble size.

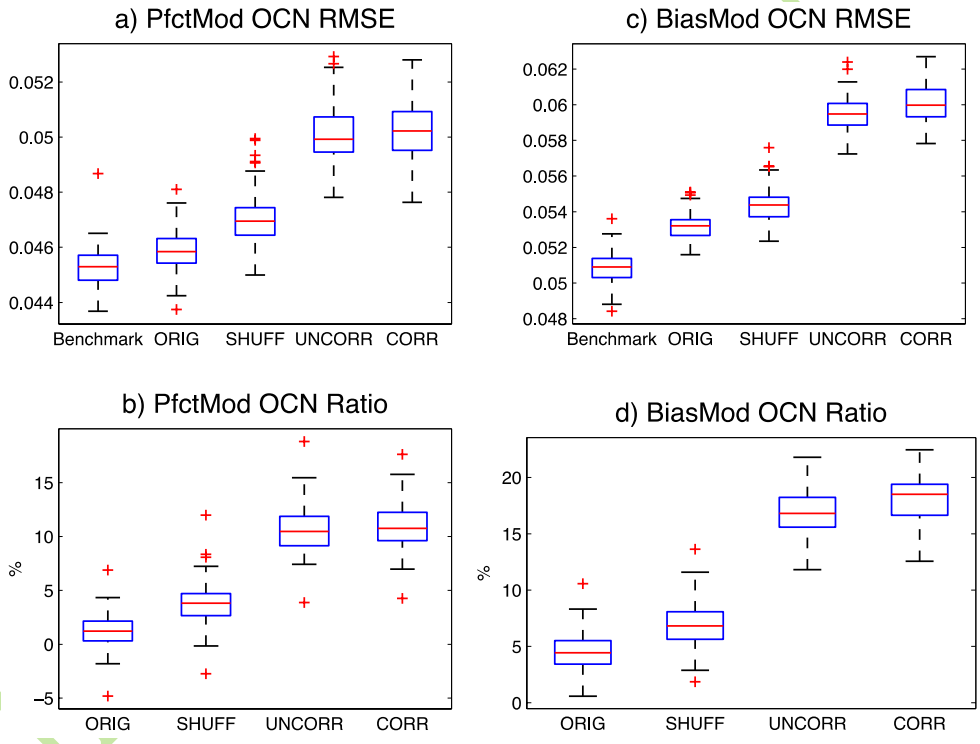


Fig. 4. Boxplots of 90 simulations for benchmark and the 4 different schemes named ORIG, SHUFF, UNCORR, and CORR. The left panel is the perfect model case and the right panel is the biased model case. The top panel shows the RMSE while the bottom

panel shows the RMSE ratio which is normalized by the benchmark experiment. The whiskers below and above the box show minimum and maximum value. The upper and lower bounds of the box are the first and third quartiles. The red line is the median and the red crosses indicate the outliers.

The performance of UNCORR is tested for different atmospheric observation frequencies and the results are shown in Fig. 5. The analysis cycle increases from 20 to 120 as the atmosphere observations becomes infrequent. the RMSE of both UNCORR and the benchmark increase due to decreased observational information (Figure not shown). The ratios in Fig.5 show no significant trend when the analysis cycle increases from 20 steps to 80 steps. However, as the atmospheric observation frequency becomes unrealistically infrequent (every 80 steps and beyond, this frequency is like less than 1 observation every 2 days), the ratios tend to decrease. This means that as the analysis quality get worse for both the benchmark and the substitution experiment due to less available observational information, the difference between the benchmark and the substitution experiment is smaller. That is to say, the resultant analysis deterioration from the substitution is less severe.

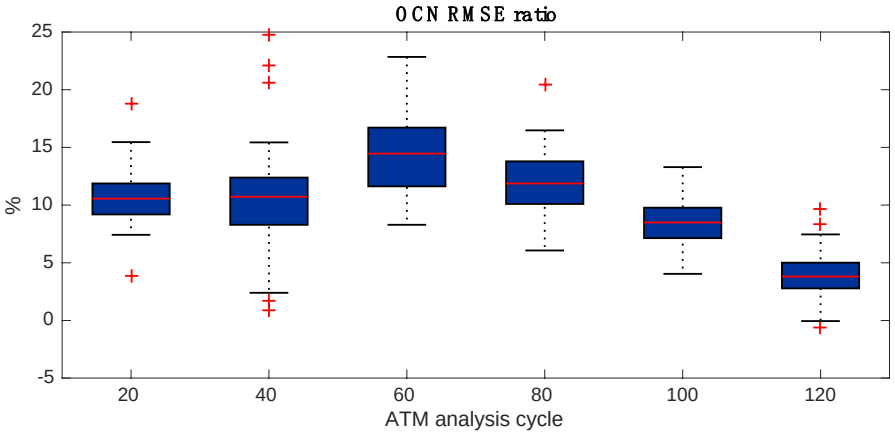
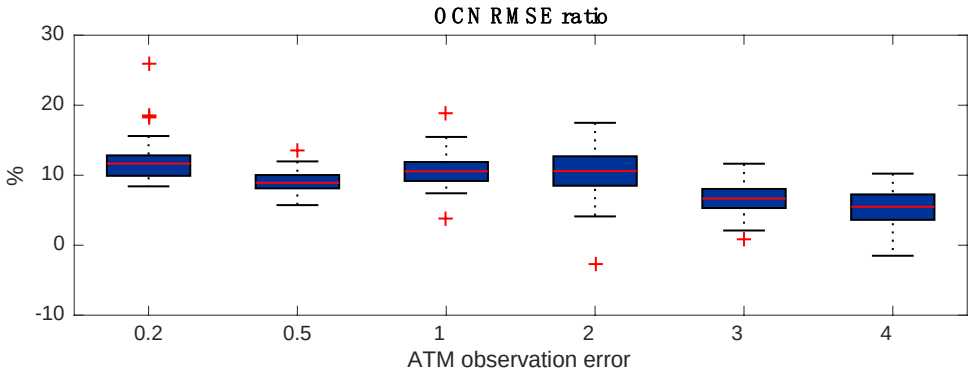


Fig. 5. Sensitivity of ocean RMSE (top) and RMSE ratio (bottom) to atmosphere observation frequency for the UNCORR scheme over 90 simulations.

The sensitivity of UNCORR to varied atmospheric observation error is shown in Fig. 6. The absolute RMSE of both UNCORR and benchmark increase as the observations become more and more noisier (Figure not shown), while the ratios in Fig. 6 show small fluctuations when the atmospheric observation error increases from 0.2 to 2.0, and eventually decrease as the error gets unrealistically large (beyond 2.0, more than half of the climatological standard deviation). This indicates that the oceanic analysis deterioration in UNCORR is fairly insensitive to the atmospheric observation error when it is in a reasonable range; In addition, the deterioration is lessened if the oceanic analyses of both the benchmark and the substitution experiment get worse due to the overly noisy observations.

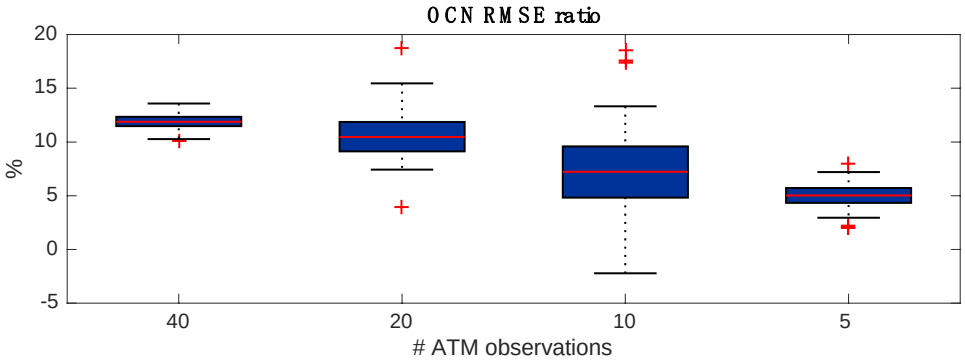


2

284 **Fig. 6.** Same as Fig.5, but for different atmopsheric observation errors.

285

286 The results of experiments with varied observation density but still evenly distributed
287 observations are shown in Fig.7. The RMSE for both experiments are larger with sparser
288 observations and smaller with denser observation (Figure not shown). The ratios in Fig. 7
289 shows a consistent decreasing trend as the observations get denser. This suggests that the
290 difference between the substitution and benchmark experiments is less significant if the
291 analyses of both experiments get worse due to less observations, which is similar to the
292 previous sensitivity tests. In contrast, if every variable is observed, the increase of RMSE
293 in the substitution experiment relative to the benchmark is the most significant. In reality,
294 the number of atmospheric observations is more or less fixed, and the density change will
295 not be as extreme as in this simple sensitivity test. It is notable that as the observations
296 get intermediately sparse, the data assimilation process gets less stable: the variance of
297 ratios among the 90 simulations is noticeably bigger when there are only 10 or 20
298 observations. This can be explained by the numerical instabilities developed in a sparse
299 observation network with finite ensemble size (Gottwald and Majda, 2013).



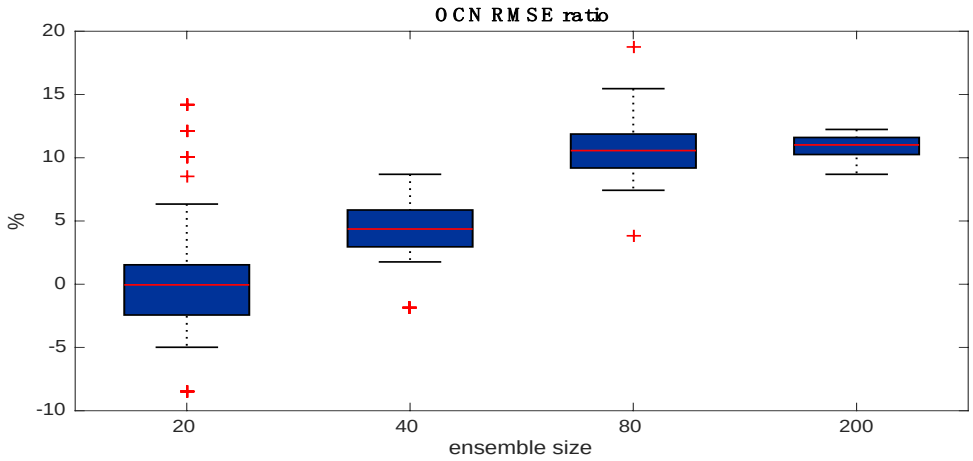
3

301 **Fig. 7.** Same as Fig.5, but for different observation densities.

302

303 The results in Figs. 5, 6 and 7 collectively indicate that the decrease in analysis
304 quality due to the substitution is insensitive to the observation quality (frequency,
305 noisiness, density) within a reasonable range. Meanwhile, when the quality of
306 observation declines too much, the decrease in analysis quality caused by substitution
307 becomes less severe compared to the benchmark because the analysis quality of the
308 benchmark also decreases significantly due to poor observational quality.

309 Ensemble size is an important factor in the estimation of error covariance and
310 correlation. Additional experiments with ensemble size 20, 40, and 200 are shown in
311 Fig.8. The RMSEs of both experiments decrease with increasing ensemble size (Figure
312 not shown). The ratios in Fig.8 increases with ensemble size and eventually levels off.
313 This suggests that with smaller sample size, the bad analysis quality for both experiments
314 will lead to smaller RMSE contrast between the substitution and the benchmark
315 experiments, thus smaller ratios. Vice versa for sufficiently large sample size.



316

317 **Fig. 8.** Same as Fig.5, but for different ensemble sizes.

318

319 **4. Tests on assimilating schemes**

320

321 In the last section, the UNCORR experiments assume that the analysis has independent
322 errors at different locations, which, however, is not the case in reality. When reanalysis is
323 generated, different model locations are connected through both model dynamics and the
324 use of localization schemes. Thus the analysis errors will be spatially correlated between
325 nearby or even far-apart locations. In addition, the analysis errors can also persist through
326 time, hence there is also temporal correlation in the time series of the analysis. The
327 previous UNCORR experiment neglects both the spatial and temporal correlations, which
328 may affect the performance of the CDA scheme. To deal with these correlations and
329 investigate how they affect the CDA, we tested three other schemes of treating the
330 reanalysis error covariance, which are named as CORR, ORIG, and SHUFF.

331 First in the CORR scheme, the off-diagonal correlation among different variables
332 of the reanalysis is taken into consideration. Instead of being set to zero as in UNCORR,

the off-diagonal elements in the reanalysis error covariance matrix R_t are retained. In correspondence, a spatially correlated observation ensemble is attained by perturbing the reanalysis with correlated noise. The spatial correlation among different variables can be calculated in a similar way as R_t in Eq.5: $C = corr < X - X^T >$. (6)

The average RMSE increase of ocean variables over 90 simulations in CORR is 11.79% for the perfect model framework and 18.13% for the biased model framework. The maximum and minimum non-outlier RMSE increase are 15.76% and 6.96% for 90 simulations in the perfect model framework (Fig.4a) and are 22.45% and 12.55% in biased model framework (Fig.4b). Although CORR includes the off-diagonal correlation among different variables, it does not outperform UNCORR and is also less stable. This is mainly caused by the additional sampling errors. Because of the chaotic nature of the model, the correlation among different locations in Lorenz96 decreases below 0.2 within five gridpoints. Thus CORR is subject to significant sampling error in two processes, firstly when the correlation matrix is calculated in Eq.6 and secondly when the correlated observation ensemble is artificially generated based on the correlation matrix. With a finite sample size, the error in the covariance or correlation estimates greatly increases when the true correlation becomes smaller. Therefore, although including the spatial correlation may improve the performance theoretically, the additional sampling error overwhelms the possible improvement.

The second ORIG scheme uses the original reanalysis ensemble as the "perturbed" observation ensemble during CDA. The original reanalysis ensemble is the byproduct of the ensemble-based DA filter during the generation of the reanalysis, therefore they could

accurately capture not only the flow-dependent correlation information among different locations, but also the temporal coherence of each ensemble member at every location. The error covariance matrix in this scheme is calculated as the time average of the error covariance matrix of the reanalysis ensemble over each analysis step.

$$R = mean < cov < X_t^{re}(nv, ens) >. \qquad (7)$$

where $cov <\cdot>$ and $mean <\cdot>$ represents sample covariance and average over time respectively. X_t^{re} represents the original reanalysis ensemble at analysis time step t. R is quantitively similar to Rt and is also diagonally dominant (Fig.3). The performance of ORIG is noticeably better than UNCORR and CORR (Fig.4) and is fairly close to the benchmark. The average oceanic RMSE increase over 90 simulations is 2.02% in the perfect model framework and 4.64% in the biased model framework. The ratio ranges from -1.81% to 4.33% in the perfect model framework over 90 simulations (Fig. 4a), and from 0.58% to 8.31% in the biased model framework.

The third SHUFF scheme is used to test the relative importance of accurate spatial correlation and temporal coherence in the improvement from CORR to ORIG. SHUFF is the same as ORIG except that the original reanalysis ensemble is shuffled at each analysis step before it is assimilated, hence the temporal coherence carried along each ensemble member is removed in SHUFF while the off-diagonal spatial correlation is still preserved. The performance of SHUFF is slightly worse than ORIG in both perfect and biased model case (Fig.4), which indicates that the temporal coherence of the reanalysis ensemble is less important for the ocean analysis. Meanwhile, SHUFF, similar to ORIG, outperforms UNCORR and CORR significantly: the highest RMSE increase in SHUFF (7.22%) almost approaches the lowest one in CORR (6.96%) and UNCORR (7.42%).

Recall that SHUFF and CORR both have the off-diagonal correlation and do not have the temporal coherence, and they primarily differ in generating the perturbations for the reanalysis (Eq.5 and Eq.7), or simply the magnitude of sampling errors for the correlation matrix. The comparisons between SHUFF and CORR and between SHUFF and ORIG therefore suggest that accurate representation of the spatial correlation is relatively more important than the temporal coherence for the ocean analysis. However, for the atmosphere component, the performance of SHUFF is closer to CORR than to ORIG (figure not shown), which suggests a relatively more important role of temporal coherence for the atmospheric analysis. All the assimilation schemes tested are summarized in Table 2.

	Covariance Matrix	**Observation Ensemble**
ORIG	Eq.7	Original reanalysis ensemble
SHUFF	Eq.7	Shuffled original reanalysis ensemble
UNCORR	Eq.5, off-diagonal elements set to zero	Perturbed ensemble (uncorrelated)
CORR	Eq.5	Perturbed ensemble based on correlation matrix

Table 2. Assimilation scheme design

5. Summary and conclusions

We substituted the atmospheric observations with reanalysis data to set up a CDA system in coupled Lorenz96 models and quantified the resultant effects on the oceanic analysis. We compared the the oceanic RMSE of the substitution experiment where atmospheric reanalysis and oceanic observations are assimilated, to a benchmark experiment where

both atmospheric and oceanic observations are assimilated. It is found that the substitution results in the deterioration of oceanic analysis quality. The magnitude of this deterioration depends on how the reanalysis is assimilated. When the reanalysis is assimilated directly as independent observations (UNCORR) as in Zhang et al. 2007, , the oceanic RMSE increase due to the substitution is about 16% in perfect model framework and about 22% in the biased model framework compared to the benchmark, or best-case scenario. Additional sensitivity tests show that this result is robust with sufficient ensemble size and reasonable atmospheric observation quality (density, frequency, noisiness). If the ensemble size is smaller, or the observation quality is worse (less frequent, sparser, noisier), the deterioration will become less severe because the analysis quality of the benchmark also decreases significantly.

In addition to the direct method, three supplementary schemes (CORR, ORIG, SHUFF) are tested with a focus on the representation of the background error covariance matrix and the generation of the perturbed observations in EnKF. We found that both the spatial correlation among the reanalysis data points and the coherence along each original reanalysis ensemble member are crucial to the analysis quality of the substitution experiments: The oceanic RMSE increase is significantly reduced when the temporal coherence along each member of the original reanalysis ensemble is preserved (ORIG); the removal of such ensemble member coherence (SHUFF, CORR) and inaccurate capture of the off-diagonal correlation (CORR, UNCORR) will result in the increase of RMSE . However, the relative importance between the off-diagonal correlation and temporal coherence on analysis quality is different for the atmosphere and ocean component. For the ocean component, the RMSE of SHUFF is closer to ORIG than

CORR, indicating a relative more important influence from the accurate representation of spatial correlation than temporal coherence, while for the atmosphere, it is the other way round.

This study has demonstrated that substituting the atmospheric observations with atmospheric reanalysis is a potentially efficient approach to implement CDA systems at the cost of moderate degradation of analysis quality. Despite the fact that this degradation can not be eliminated, the CDA products can still provide state-estimation of the coupled variability in the atmosphere-ocean system which incorporates both the observational and model information, and the dynamical balance between the atmosphere and ocean components can reduce the initial shock in the initialization of coupled GCM. There are still remaining issues regarding assimilating atmospheric reanalysis data. First, different schemes, in particular ORIG, UNCORR and CORR, should be tested on models that have higher spatial correlations and the impact on the oceanic analysis quality should be evaluated. Second, different ensemble filters such as ensemble adjustment filter (Zhang et al., 2007; Anderson, 2001), can be employed to assess robustness of the assimilating the reanalysis. Finally, this idea should be further tested in a coupled model of higher complexity.

Acknowledgment:

We gratefully thank Shan Li for helpful discussions. This research is sponsored by Nanjing University of Information Science and Technology.

References

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

Arakawa, O., and A. Kitoh, 2004: Comparison of local precipitation–SST relationship between the observation and a reanalysis dataset. *Geophys. Res. Lett.*, **31**, L12206, doi: 10.1029/2004GL020283.

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.

Chen, D., S. E. Zebiak, A. J. Busalacchi, et al., 1995: An improved procedure for EI Niño forecasting: implications for predictability. *Science*, **269**, 1699–1702, doi: 10.1126/science.269.5231.1699.

Dee, D. P., S. M. Uppala, A. J. Simmons, et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi: 10.1002/qj.828.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10143–10162, doi: 10.1029/94JC00572.

Fujii, Y., T. Nakaegawa, S. Matsumoto, et al., 2009: Coupled climate simulation by constraining ocean fields in a coupled model with ocean data. *J. Clim.*, **22**, 5541–5557, doi: 10.1175/2009JCLI2814.1.

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.

Gottwald, G. A., and A. J. Majda, 2013: A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks. *Nonlinear Process. Geophys.*, **20**, 705–712, doi: 10.5194/npg-20-705-2013.

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, doi: 10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Han, G. J., X. R. Wu, S. Q. Zhang, et al., 2013: Error covariance estimation for coupled data assimilation using a lorenz atmosphere and a simple pycnocline ocean model. *J. Clim.*, **26**, 10218–10231, doi: 10.1175/JCLI-D-13-00236.1.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.

Kalnay, E., M. Kanamitsu, R. Kistler, et al., 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–472. doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, et al., 2002: NCEP–DOE AMIP-II reanalysis (R-2). *Bull. Am. Meteor. Soc.*, **83**, 1631–1643, doi: 10.1175/BAMS-83-11-1631.

Kistler, R., W. Collins, S. Saha, et al., 2001: The NCEP-NCAR 50 -year reanalysis: monthly means CD-ROM and documentation. *Bull. Am. Meteor. Soc.*, **82**, 247–267, doi: 10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Kitoh, A., and O. Arakawa, 1999: On overestimation of tropical precipitation by an atmospheric GCM with prescribed SST. *Geophys. Res. Lett.*, **26**, 2965–2968, doi: 10.1029/1999GL900616.

Kobayashi, S., Y. Ota, Y. Harada, et al., 2015: The JRA- 55 reanalysis: general specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, doi: 10.2151/jmsj.2015-001.

Liu, Z. Y., S. Wu, S. Q. Zhang, et al., 2013: Ensemble data assimilation in a simple coupled climate model: the role of ocean-atmosphere interaction. *Adv. Atmos. Sci.*, **30**, 1235–1248, doi: 10.1007/s00376-013-2268-z.

Lorenz, E. N., 1996: Predictability—A Problem Partly Solved. *Proceedings of Seminar on Predictability*, *1*, *ECMWF*, *Reading*, Berkshire, UK, 1–18.

Lu, F. Y., Z. Y. Liu, S. Q. Zhang, et al., 2015: Strongly coupled data assimilation using leading averaged coupled covariance (LACC). Part II: CGCM Experiments. *Mon. Wea. Rev.*, **143**, 4645–4659.

Luo, J.-J., S. Masson, S. K. Behera, et al., 2008: Extended ENSO predictions using a fully coupled ocean–atmosphere model. *J. Clim.*, **21**, 84–93, doi: 10.1175/2007JCLI1412.1.

Saha, S., S. Moorthi, H. L. Pan, et al., 2010: The NCEP climate forecast system reanalysis. *Bull. Am. Meteor. Soc.*, **91**, 1015–1057, doi: 10.1175/2010BAMS3001.1.

Singleton, T., 2011: Data assimilation experiments with a simple coupled ocean- atmosphere model. Ph. D. dissertation, University of Maryland, College Park, 116 pp.

Sugiura, N., T. Awaji, S. Masuda, et al., 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res. Ocean.*, **113**(C10), C10017, doi: 10.1029/2008JC004741.

Tardif, R., G. J. Hakim, and C. Snyder, 2015: Coupled atmosphere–ocean data assimilation experiments with a low-order model and CMIP5 model data. *Climate Dyn.*, **45**, 1415–1427, doi: 10.1007/s00382-014-2390-3.

Uppala, S. M., P. W. Kållberg, A. J. Simmons, et al., 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, doi: 10.1256/qj.04.176.

Zhang, S., 2011: A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model. *J. Clim.*, **24**, 6210–6226, doi: 10.1175/JCLI-D-10-05003.1.

Zhang, F., C. Snyder, and J. Z. Sun, 2004: Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, doi: 10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

Zhang, S., M. J. Harrison, A. T. Wittenberg, et al., 2005: Initialization of an ENSO forecast system using a parallelized ensemble filter. *Mon. Wea. Rev.*, **133**, 3176–3201, doi: 10.1175/MWR3024.1.

Zhang, S., M. J. Harrison, A. Rosati, et al., 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, **135**, 3541–3564, doi: 10.1175/MWR3466.1.