JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

Successes and Failures of Assimilating Global Integrals with an Ensemble Filter

Lisa Neef[1] and Katja Matthes[1]

Corresponding author: L.J. Neef, Düsternbrooker Weg 20, D-24105 Kiel, Germany (neef@geomar.de)

[1]Ocean Circulation and Climate Dynamics - Marine Meteorology, Helmholtz Centre for Ocean Research Kiel (GEOMAR), Kiel, Germany.

Abstract. (Type abstract here)

1. Introduction

General Issue

• Data assimilation is the process of inserting observations of the Earth system into dynamical models, by computing, at each observation time, the best estimate between what is predicted by the model on the one hand, and what is observed by measurement on the other.

• Data assimilation was first developed in the context of numerical weather prediction, where the emphasis is using data to compute the best possible model initial state, in order to improve forecasts.

• Atmospheric data assimilation is now at the stage where implementation of the algorithms is accessible even to non-experts, via tools such as the Data Assimilation Research Testbed [*Anderson et al.*, 2009, DART] and the PDAF [DEFINE and CITE]. These tools make it possible to extend the algorithms that were originally developed for NWP to climate studies. Using data assimilation, we can constrain climate simulations to reality, while still changing individual model parameters to meet the needs of an experimental design.

scientific context and State-of-the-art

• In the climate modeling context, data assimilation in some form is frequently used to capture real events (e.g. the El Nino of a given year), while other factors in the modeling environment can be changed. Assimilation can also be used to constrain one part of the Earth system, and then explore the variability that ensues in other parts [CITE EXAMPLE].

- (Say something about reanalysis.)

- But what observations should be used in such simulations?

- In general, it is not obvious which variables need to be observed in order to adequately constrain the model.

Problems

- It has been suggested to assimilation Earth rotation parameters, i.e. anomalies in the Earth's rotation rate and the orientation of its rotational pole – have been suggested as a useful variable for data assimilation because they reflect changes in the angular momentum of the Earth system. On subseasonal to interannual timescales, length-of-day anomalies (hereafter ΔLOD) is essentially a measure of the angular momentum of the atmosphere [CITE], while polar motion (hereafter PM) reflects the combined angular momentum of the atmosphere, ocean, and hydrosphere.

- *Saynisch et al.* [2011a, b] and *Saynisch and Thomas* [2012] used advanced data assimilation methods to fit the excitation of ΔLOD and PM in an ocean model to the observed variations of these paramters, by adjusting boundary parameters (atmospheric wind stresses and freshwater flux). They found that strong adjustments must be made in boundary parameters in order to close the observed discrepancy in the oceanic angular momentum budget. However, these studies did not show whether the assimilation of ΔLOD and PM actually brought the modeled ocean state closer to reality, nor whether the assimilation of these values added value to other variables that were also assimilated.

- While the prospect of assimilating non-gridpoint observations is exciting an a potential new source for information about the Earth system, the assimilation of such observations is not straightforward and has several inherent difficulties.

Solution In this study we identify some of the major difficulties associated with assimilating average or integral observations, using the example of atmospheric angular momentum. Instead, we show that the main value of global-integral observations is as a tool for evaluating the fidelity of an assimilation system consisting of more traditional, gridpoint observations. **Costs / benefits**: This result yields a handy tool for evaluating and quantifying the performance of data-constrained climate simulations.

2. Data Assimilation System

2.1. Atmospheric Model

Assimilation experiments are performed using the Community Atmosphere Model 5 [*Neale et al.*, 2011, CAM5 hereafter], within the Data Assimilation Research Testbed (DART, *Anderson et al.* [2009], *Raeder et al.* [2012]). CAM5 is run with the finite-volume dynamical core at 2.5 deg. horizontal resolution and 30 hybrid-coordinate vertical levels, with a top near 3 hPa. How is CAM damped at the top? CAM5 forms the atmospheric component of the Community Earth System Model (CESM).

2.2. Data Assimilation

DART is an open-source, community facility that makes ensemble data assimilation available for any model, using any observations. It is available online at `www.image.ucar.edu/DAReS/DART`. [Fill in details of the dart filter, etc/]

In ensemble data assimilation, covariances between components of the model state and the observations are estimated using an ensemble, which itself is updated with the observed information at each analysis time. Though the DART distribution includes several types of ensemble filter, its primary setting, and the one used in this study, is the deterministic ensemble adjustment Kalman filter [*Anderson*, 2001, EAKF] with a sequential parallel implementation [**?**].

2.3. Synthetic Angular Momentum Observations

Add some sentences justifying the use of AAM as an observation variable. The relationship between ERP variations and AAM is given by nondimensional atmospheric excitation functions (AEFs), i.e. nondimensional functions of the three components of AAM. These are derived in *Barnes et al.* [1983], and are given by the following:

$$\chi_1(t) = \frac{1.608}{\Omega\,(C - A')}\left[0.684\Omega\Delta\mathbf{I}_{13}(t) + \Delta h_1(t)\right] \quad (1)$$

$$\chi_2(t) = \frac{1.608}{\Omega\,(C - A')}\left[0.684\Omega\Delta\mathbf{I}_{23}(t) + \Delta h_2(t)\right] \quad (2)$$

$$\chi_3(t) = \frac{0.997}{\Omega C_m}\left[0.750\Omega\Delta\mathbf{I}_{33}(t) + \Delta h_3(t)\right] \quad (3)$$

where the $\mathbf{I}_{ij}$ represent components of the atmospheric inertia tensor:

$$I_{13} = -\int R^2 \cos\phi \sin\phi \cos\lambda dM \quad (4)$$

$$I_{23} = -\int R^2 \cos\phi \sin\phi \sin\lambda dM \quad (5)$$

$$I_{33} = \int R^2 \cos^2\phi dM \quad (6)$$

$$\quad (7)$$

and the h_i represent the relative angular momentum of the atmosphere in each direction:

$$h_1 = -\int R\left[u\sin\phi\cos\lambda - v\sin\lambda\right] dM \quad (8)$$

$$h_2 = -\int R\left[u\sin\phi\sin\lambda + v\cos\lambda\right] dM \quad (9)$$

$$h_3 = \int Ru\cos\phi dM \quad (10)$$

$$\quad (11)$$

Here $R = 6371.0$ km is the radius of the Earth, $\Omega = 7.292115 \times 10^{-5}$rad/s the average rotation rate, and $g = 9.81$m/s^2 is the acceleration due to gravity. $C = 8.0365 \times 10^{37}$kgm^2 and $A = 8.0101 \times 10^{37}$kgm^2 are the axial and next-largest principal moments of inertia of the solid Earth, and $C_m = 7.1236 \times 10^{37}$kgm^2 is the principal inertia tensor component of the Earth's mantle [*Gross*, 2009].

The equatorial excitation functions can be easily mapped to equivalent polar motion by a rotation of the reference frame:

$$p_1 + \frac{\dot{p}_2}{\sigma_0} = \chi_1 \tag{12}$$

$$-p_2 + \frac{\dot{p}_1}{\sigma_0} = \chi_2, \tag{13}$$

where σ_0 represents the Chandler frequency, an intrinsic wobble frequency of the solid Earth.

The axial excitation function χ_3 corresponds to unit changes in the rate of rotation of the Earth, and therefore also unit changes in the length of a day (LOD):

$$\frac{\Delta \mathrm{LOD}}{\mathrm{LOD}_0} = \Delta\chi_3, \tag{14}$$

where $\Delta\mathrm{LOD} \equiv \mathrm{UT1} - \mathrm{IAT}$ denotes anomalies in the length-of-day.

Note that in the axial AAM terms [(6) and (10)], the zonal integral is uniform, with no weighting function applied. In practice, this means that mass anomalies tend to cancel one another out in the zonal integral, with the result that the axial mass excitation (6) is usually several orders of magnitude smaller than the axial wind excitation (10) (CITE).

Nearly the opposite is due for the equatorial terms [(4)-(5) and (8)-(9)]. Here, longitudinal asymmetry determines the magnitude of the excitation. The mass excitation typically outweights the wind excitation by a few factors for the equatorial terms.

Of the polar motion angles, p_1 largely reflects surface pressure variations over the oceans, and p_2 over the continents. The effect of short-timescale surface pressure variations over the ocean tend to be evened out by corresponding displacement of the ocean surface (the so-called "inverse barometer" effect), reducing the total angular momentum transfered to the Earth. Sub-annual p_1 variations therefore depend largely on other sources of AM. — this statement makes no sense. Look at Henryk's paper and sort out what really happens with p_1. In contrast, p_2, which is strongly weighted over land, has a pronounced annual cycle due to the yearly appearance of the Siberian High (CITE) and tends to show strong negative anomalies in the two months preceding sudden stratospheric warmings [*Neef and Matthes*, in preparation].

2.4. Perfect-model experiments

In this study we perform so-called perfect-model experiments, wherein the "truth" is actually a realization of the model, and observations are generated from this realization with known error statistics. Experiments are performed with an 80-member ensemble, which itself is generated from one full year of assimilation of the uniform grid of radiosonde observations described in Section **??**.

For our experiments, observations of χ_1, χ_2, and χ_3 are generated every 24 hours. This simulates the true observations of the ERPs (p_1, p_2, and ΔLOD), (ΔLOD isn't really the observation .. it's UT1-UTC – but what are they?) which are available once daily (CITE).

Our main evaluation diagnostic is root mean square error (RMSE) between the ensemble mean the the true state. To evaluate how much information is actually gained from assimilation in each case, the RMSE for each case will be shown relative to o a case where the 80-member initial ensemble is evolved forward in time without assimilation, hereafter known as the"No-DA" run.

3. Can we constrain the state using global integral observations?

3.1. Prior and Posterior fit to Observed ERPs

Figure 1 shows the fit of the ensemble to the observed ERP excitation functions in each of the four main experiments outlined in Section 2.4. The case of no assimilation (Fig. 1(A)) starts out with the ensemble clustered closely around the true ERPs. The ensemble spreads noticeably after about a week, and by the end of January, the ensemble mean lacks any of the short-timescale features of the true ERPs.

When ERPs are assimilated (Fig. 1(B)), the ensemble members naturally agree much more in their predicted ERPs. Thus the assimilation of ERPs is successful in the sense that the wind and surface pressure fields in each ensemble member are nudged enough to give each ensemble member an AAM vector that is close to what is observed. However, the relative success in fitting the ERPs does not automatically imply that the state variables themselves are closer to the true state. This will be investigated in the following section.

An even greater contrast is seen when radiosonde observations are assimilated (Fig. 1(C)). Here the true ERPs are matched closely for all ensemble members, for the entirety of the assimilation run, even though no ERPs are assimilated. is there any point in showing this case? the fit is so perfect here that we don't really need ERPs.

3.2. Limited Error reduction in Wind and Pressure Fields

A comparison between the NODA and ERPDA experiments in terms of the model state space is given in Figures **??**-**??**. Fig. **??**(a) shows the RMSE in zonal wind (see Section

2.4), averaging over all latitudes and longitudes, as a function of height and time, in the NODA experiment. As in the corresponding observation-space plots [Fig. 1(A)] the error begins to spread noticeably after about a week, and saturates after about a month. Error growth is strongest in the tropospheric jets (around 300hPa) and near the model lid.

Fig. **??**(b) shows the same but for assimilation of ERPs. Here the error growth is similar but slightly weaker, especially during the first few weeks. The reduction of error between NODA and ERPALL is shown in Fig. **??**(c). It can be seen that assimilating ERPs reduces the error most of all at the top of the model, and visible error reduction lasts through about the end of February at which point the assimilation of ERPs actually *increases* the error.

Fig. **??**(d) shows the innovation (Prior-Posterior) at each analysis time. The innovation, like the error reduction, is also strongest near the model top. The innovation is generally strongest when the true error is largest, which means that the filter correctly chooses a strong adjustment at times when the error is largest. However, while this results in a large error reduction early on in the assimilation period, it increases error towards the end of the assimilation period.

[Insert explanation for why increments are concentrated up high.]

Figure **??** examines what happens to the wind at 300hPa as the assimilation progresses. This figure is similar to Fig. **??**, but now showing 300hPa wind as a function of latitude and time. Again, the grown of error is slower when ERPs are assimilated [Fig. **??**(b)] than with no assimilation [Fig. **??**(a)], and we see visible error reduction through January and February [Fig. **??**(c)].

Figure ?? is similar to Fig. ??, but showing surface pressure. Here the error growth, and the error reduction brought about my the assimilation of the ERPs, are largely at latitudes about 50°N. Why is that?

3.3. Evolution of the Ensemble-Estimated Covariance Field

Describle how the covariances between local wind or mass variables, and global AAM, evolve in time as estimated by the ensemble. It doesn't reach some sort of steady state but also doesn't seem to reflect reality? (How do we know this?)

4. Integral observations as an evaluation tool

4.1. Dynamical implications of each ERP component

In this section, write about how the different ERP components reflect different aspects of the model state. How can we illustrate this with a figure?

Acknowledgments. (Text here)

References

Anderson, J., T. Hoar, K. Raeder, H. Lui, N. Collins, R. Torn, and A. Avellano (2009), The data assimilation research testbed: A community facility, *Bull. Am. Met. Soc.*, pp. 1283–1296, doi:10.1175/2009BAMS2816.1.

Anderson, J. L. (2001), An ensemble adjustment filter for data assimilation, *Mon. Wea. Rev.*, *129*, 2884–2903.

Barnes, R., R. Hide, F.R.S., A. White, and C. Wilson (1983), Atmospheric angular momentum fluctuations, length-of-day changes and polar motion, *Proc. Roy. Soc. London*, *387*, 31–73.

Gross, R. S. (2009), Earth rotation variations - long period, in *Geodesy, Treatise on Geophysics*, edited by T. Herring, pp. 239–294, Elsevier.

Neale, R. B., C.-c. Chen, P. H. Lauritzen, D. L. Williamson, A. J. Conley, A. K. Smith, M. Mills, and H. Morrison (2011), Description of the NCAR Community Atmosphere Model (CAM 5.0), *National Center for Atmospheric Research Tech. Rep.*, (NCAR/TN-486+STR.), 268.

Neef, L. J., and K. Matthes (), Sudden stratospheric warmings excited Earth rotation variations, *J. Geophys. Res.*, submitted.

Raeder, K., J. L. Anderson, N. Collins, T. J. Hoar, J. E. Kay, P. H. Lauritzen, and R. Pincus (2012), DART/CAM: An ensemble data assimilation system for CESM atmospheric models, *J. Clim.*, doi:10.1175/JCLI-D-11-00395.1, in press.

Saynisch, J., and M. Thomas (2012), Ensemble Kalman-Filtering of Earth rotation observations with a global ocean model, *Journal of Geodynamics*, *62*, 24–29, doi: 10.1016/j.jog.2011.10.003.

Saynisch, J., M. Wenzel, and J. Schröter (2011a), Assimilation of earth rotation parameters into a global ocean model: excitation of polar motion, *Nonlin. Proc. Geophys.*, *18*, 581–585.

Saynisch, J., M. Wenzel, and J. Schröter (2011b), Assimilation of earth rotation parameters into a global ocean model: length of day excitation, *J. Geodesy.*, *85*, 67–73.

References

Figure 1. Fit to ERP observations with No DA and assimilating ERPs – we see that it's fairly easy to make the fit.

Figure 2. Error reduction in U300 and PS fields as the assimilation progresses – we start out with a pretty healthy error reduction, but then it goes away.

Figure 3. Ensemble versus truth in the NAO index in NoDA and ERPDA – we don't gain much as far as large-scale variability goes.

Figure 4. Evolution of the analysis increment (proportional to covariances) as the assimilation progresses. – Nothing coherent emerges, and for some reason, observations up high get all the weight.

Figure 5. Evaluation of other data assimilation experiments (with localized observations) using ERPs.