**Nonlinear Processes
in Geophysics**

Parameter estimation in an atmospheric GCM using the Ensemble Kalman Filter

J. D. Annan[1]**, D. J. Lunt**[2]**, J. C. Hargreaves**[1]**, and P. J.Valdes**[2]

[1]Frontier Research Centre for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama, and
Proudman Oceanographic Laboratory, United Kingdom
[2]Bristol Research Initiative for the Dynamic Global Environment (BRIDGE), University of Bristol, United Kingdom

Abstract. We demonstrate the application of an efficient multivariate probabilistic parameter estimation method to a spectral primitive equation atmospheric GCM. The method, which is based on the Ensemble Kalman Filter, is effective at tuning the surface air temperature climatology of the model to both identical twin data and reanalysis data. When 5 parameters were simultaneously tuned to fit the model to reanalysis data, the model errors were reduced by around 35% compared to those given by the default parameter values. However, the precipitation field proved to be insensitive to these parameters and remains rather poor. The model is computationally cheap but chaotic and otherwise realistic, and the success of these experiments suggests that this method should be capable of tuning more sophisticated models, in particular for the purposes of climate hindcasting and prediction. Furthermore, the method is shown to be useful in determining structural deficiencies in the model which can not be improved by tuning, and so can be a useful tool to guide model development. The work presented here is for a limited set of parameters and data, but the scalability of the method is such that it could easily be extended to a more comprehensive parameter set given sufficient observational data to constrain them.

1 Introduction

Parameter estimation is an important part of the creation of a complex numerical model, and is especially critical for prediction of anthropogenically-forced climate change, since it is parameters (rather than initial conditions) which determine the model climate. Until recently, no practical and efficient method for automatic tuning was available, so researchers generally use a large number of trial-and-error direct perturbation sensitivity experiments in order to choose appropriate

values for model parameters (e.g. Allen, 1999; Knutti et al., 2002). However, these brute-force methods are spectacularly inefficient for even modest problems, with the cost growing exponentially with the number of parameters. Variational parameter estimation with an adjoint model does not work well for tuning the climate of chaotic models due to their sensitive dependence on initial conditions: some attempts have been made to ameliorate this problem but no wholly satisfactory method has yet been found (Köhl and Willebrand, 2002, 2003; Lea et al., 2000, 2002). Moreover, when used for climate prediction purposes, parameter estimation is not merely a search for the optimal values (which an adjoint most readily generates) but a range of parameters that represents the uncertainty in their (joint) distribution, since this is what determines the uncertainty of the climate reponse for a given scenario.

The ensemble Kalman filter or EnKF (Evensen, 1994) is an efficient Monte Carlo approximation to the Kalman filter equations (Kalman, 1960). It has been widely used in near-operational forecasting, especially for short-term numerical weather and ocean prediction. A thorough description of the theory and basic methodology together with a survey of recent applications is provided in Evensen (2003). Although the EnKF has generally been used for sequential initial state estimation, parameter estimation can readily be included in the same framework, by the means of state space augmentation (Derber, 1989; Anderson, 2001). The principle here is that the parameters can be considered to be part of the model state alongside the conventional variables, and then the covariances sampled by the ensemble members can be used directly to update parameters in exactly the same manner as for the state variables.

Although this approach to parameter estimation has been well known for some time, previous applications appear to have generated generally rather poor results. For example, Kivman (2003) found that the EnKF performed rather poorly when applied to simultaneous state and parameter estimation in the Lorenz model. He ascribed this problem to the highly

non-Gaussian probability distribution functions which arise. He found that a particle filter worked rather better, but this is generally much more computationally expensive for high-dimensional systems, and even that method will converge over time to a single set of (incorrect) parameter values unless some form of noise is added to the system.

When noise is added, and the parameter values are therefore allowed to vary through time, simultaneous parameter and state estimation can give good results via both the EnKF (Anderson, 2001) and particle filter (Losa et al., 2003). However, the need to add random noise, the amount of which is generally rather poorly determined, would greatly complicate any long-term forecasting and the model could be expected to lose skill over climatological time scales. On the other hand, estimating temporally constant parameter values by fitting a long integration of a chaotic model to a time series of data is extremely challenging even for a perfect model (Pisarenko and Sornette, 2004) and impossible in all real applications with imperfect models.

Therefore, tuning of parameters for climatological forecasting is generally treated from the standpoint of choosing temporally fixed values for which the model's climatology matches observations (Allen, 1999; Giorgi and Mearns, 2002; Murphy et al., 2004). Although the model's trajectory through state space is highly sensitive to initial conditions, the climate of a sufficiently long trajectory (for example, temporal means of particular model variables) is typically much less sensitive to initial conditions, being essentially a sample of the underlying true model climate (i.e. the limit as integration time tends to infinity) contaminated by a small (and controllable) amount of deterministic noise due to the finite integration interval. For all practical purposes, this noise can be treated as truly stochastic, and it decreases in proportion to the square root of integration time (Lea et al., 2000).

Recently, Annan et al. (2005) have presented an efficient technique for parameter estimation using the ensemble Kalman filter. This has been applied to the simultaneous estimation of 12 parameters in the low-resolution (non-chaotic) coupled atmosphere-ocean model of Edwards and Marsh (2005), and also to the chaotic 3-variable Lorenz model (Annan and Hargreaves, 2004).

These previous applications of this climatological parameter estimation method have been limited to cases where the model is either devoid of internal chaotic dynamic or chaotic but very low-dimensional. Here, we extend these results to show that the method can also work successfully when applied to a realistic intermediate complexity atmospheric GCM. Our results suggest that this method could be used for practical applications with a range of sophisticated climate models.

In Sect. 2, we describe the model and outline the estimation method. Section 3 describes an identical twin experiment, where the model is tuned towards a climate generated by a known set of parameters. Section 4 contains the results of the numerical experiments using reanalysis (i.e. based on observed) data. We conclude the paper in Sect. 5.

2   Model and methods

2.1   A simplified atmospheric GCM

The model we use is essentially that of de Forster et al. (2000), which has been used in a diverse range of studies (e.g. Rosier and Shine, 2000; Highwood and Stevenson, 2003; Joshi et al., 2003). Some modifications have been made to the original model, which will be described below. The model is an intermediate resolution (T21) spectral primitive equation atmospheric general circulation model which was originally designed to efficiently examine the mechanisms of climate change and the robustness of model behaviour under varying scenarios. To that end, it contains somewhat simplified parameterisations including a particularly efficient radiation scheme, in order to enable multiple, decadal-length integrations. However, due to the close relationship with higher complexity models, it is not unreasonable to expect that its behaviour is largely consistent with them.

Our main interest in this model is as a component of a new Earth System model of intermediate complexity which is being built as part of the GENIE project (http://www.genie.ac.uk). The ultimate goal of this project is to create a model capable of lengthy and/or ensemble simulations such as are necessary for paleoclimate (for example a glacial-interglacial cycle) and long-term climate change studies (for example the effect of a reduction in the sizes of the Greenland and Antarctic ice-sheets). An efficient 3D ocean model has already been built and coupled to a simple 2D energy and moisture balance atmosphere (Edwards and Marsh, 2005; Hargreaves et al., 2004), but we believe that the dynamic AGCM used in this paper models the atmospheric processes more realistically and therefore should be able to generate significantly improved results when coupled to the ocean model. The original version of this atmospheric model as presented by de Forster et al. (2000) had 22 vertical layers and was coupled to a slab ocean. In order to speed it up and prepare for coupling to the ocean model, the number of vertical layers in the atmospheric model has been reduced to 7, and some restructuring of the code has been carried out. The reduction in the number of vertical levels has suprisingly little effect on the convective precipitation and surface temperature, but reduces the intensity of the mid-latitude storms somewhat. The slab ocean and sea-ice layers have been temporarily replaced by user-specified sea surface and sea ice temperature fields. These will ultimately be provided by the ocean and sea-ice model components, but in the work presented here we have used climatological monthly mean values (NCEP Reanalysis data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/), linearly interpolated onto a 2-day timestep. Therefore, the model in the form described here is not suitable for generating predictions of climate change, due to the absence of feedbacks associated with interactive sea-ice (for example changes in albedo) and ocean (for example changes in surface temperature and the thermohaline

circulation). However, its speed and dynamical similarity to more complex GCMs makes it a highly suitable test-bed for investigations of parameter estimation methods which may be more widely applicable.

2.2 EnKF implementation

Previous applications of the EnKF for parameter estimation with the coupled 2D atmosphere – 3D ocean model are described in Annan et al. (2005) and Hargreaves et al. (2004). As mentioned in the Introduction, the model state of each ensemble member is augmented with parameter values and climatological diagnostics from a model run of specified duration, and time series output is not used directly.

For a steady state problem (i.e. tuning the model's climatology), the Kalman Filter can in fact be simplified to a Wiener Filter (Press et al., 1994, Sect. 13.3) and the equations can be solved in a single step. However, if this approach is attempted, the "curse of dimensionality" (Bellman, 1961) implies that the ensemble size would have to be very large in order for any of the prior sample to be close to the posterior. Moreover, if the problem is nonlinear, then this combined with the finite ensemble (and numerical approximations that are usually required for implementation) will tend to result in an inaccurate posterior estimate which does not satisfy the model equations (i.e. is unbalanced), as we show in a simple example below. We have therefore implemented an iterative approach which we now describe in more detail.

As shown in Evensen and Leeuwen (2000), data can in principle be assimilated in arbitrary order, together or separately without affecting the final estimate, so long as there are assumed to be no correlations between observational errors on data which are assimilated in different batches. We can use this result to generate sets of artificial observations, which when all are assimilated is equivalent to the original data set, but which when assimilated sequentially in batches, reduces the inaccuracies due to nonlinearity and the curse of dimensionality by virtue of replacing a single huge jump between the prior and posterior with a sequence of smaller steps.

For example, if the original data set takes the values x_o with observational error covariance matrix $\mathbf{R}$, then we can create 2 sets of artificial observations which both have the values x_o and covariance matrices $2\mathbf{R}$, with the sets of observations assumed independent of each other (the fact that they actually take the same values does not matter). These two sets of observations are exactly equivalent to the original set in terms of the posterior they generate, since they could be combined, prior to assimilation, into the values $x_o+1/2(x_o-x_o)=x_o$ with covariance matrix $2\mathbf{R}(2\mathbf{R}+2\mathbf{R})^{-1}2\mathbf{R}=\mathbf{R}$ (using the standard equations for optimal interpolation). Indeed, this is exactly how one would normally combine separate observations of the same model variable (say, duplicate independent observations taken within a specific grid box and time interval). However, these data sets can also be assimilated sequentially into the model in two steps to generate the same posterior.

For a nonlinear model with relatively diffuse prior, the single step procedure is liable to be somewhat inaccurate and generate unbalanced solutions. However, when the data are assimilated in two stages, the loss of balance and resulting inaccuracy can be reduced by reintegrating the model equations between performing the two analyses. We can generalise this approach to N sets of identical observations (for any whole number N) with the covariance matrices $N\mathbf{R}$, or even an infinite number of sets of observations, as we now show.

For convenience, we write $(x_o, \mathbf{Q})$ to denote the set of observations which take the values x_o (a vector) with covariance matrix $\mathbf{Q}$. We consider the infinite series of sets of observations

$$\{(x_o, ce^i\mathbf{R})\}\ i \in \mathbb{N},$$

where c and e are real constants.

By induction, the first N terms in this series can be combined into a single equivalent set of observations taking the same values x_o but with the covariance matrix

$$\frac{ce^{N-1}(e-1)}{e^N-1}\mathbf{R},$$

which converges to

$$\frac{c(e-1)}{e}\mathbf{R}$$

in the limit as $N\rightarrow\infty$. Therefore, if we choose $c=e/(e-1)$, the infinite series of observations is equivalent to the original set. In these equations, e and c are the squares of the "expansion" and "correction" factors described in Annan et al. (2005). $e>1$ can be chosen arbitrarily, with smaller values giving a slower convergence but more accurate final solution to the problem in the presence of model nonlinearity. We have found that values in the range $1.05\leq e\leq 1.2$ generally give good results.

We can converge towards the posterior solution defined by this infinite sequence of sets of observations by starting from an arbitrary initial guess $(i, \mathbf{S})$ and then repeating the sequence:

- Integrate the model to sample the climatology.

- Inflate the ensemble by a factor $\sqrt{e}$ about its mean (and thus increase the covariance matrix by the factor e).

- Assimilate the data set defined by $(x_o, c\mathbf{R})$.

After N iterations, the posterior is that given by interpolation of the data sets

$$\left\{(x_o, c\mathbf{R}),\ (x_o, ce\mathbf{R}),\ (x_o, ce^2\mathbf{R}), \ldots,\right.$$
$$\left.(x_o, ce^{N-1}\mathbf{R}), (i, e^N\mathbf{S})\right\}$$

(at least for a linear model) and so the ensemble converges as above to the distribution defined by the data set and model equations.
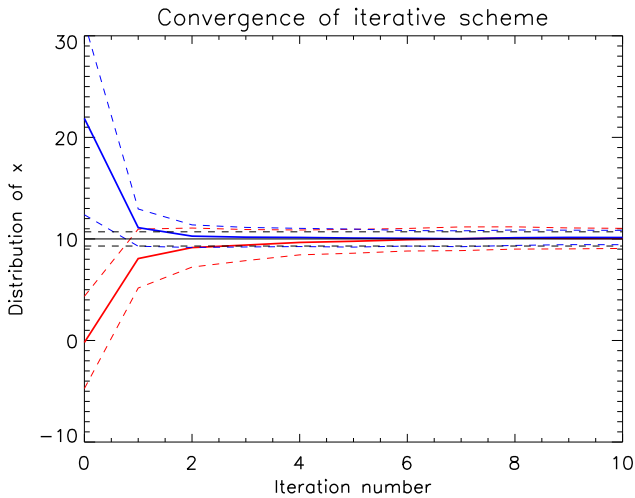
Fig. 1. Convergence of scheme for simple nonlinear problem. Thick solid lines indicate ensemble means, and thinner dashed lines show the one standard deviation widths. Red and blue lines show the results of two different experiments, with the black lines indicating the true solution. Red: initial distribution $x=1\pm5$, expansion factor $e=1.1$. Blue: initial guess $x=20\pm10$, expansion factor $e=1.44$.

So far we have ignored the prior (the influence of the initial ensemble decays with the iterative process). However, we can observe that a prior estimate is precisely equivalent to a set of observations with the same mean and covariance matrix. Therefore, the prior can also be assimilated simultaneously by the same iterative scheme and the data sets $(x_o, c\mathbf{R})$ should be viewed as consisting of the observations of model output (such as climatological values) together with prior estimates of parameter values. Convergence to the limit is linear (at least for a linear model), and the cost of each iteration barely changes with the number of parameters to be estimated (assuming that this number is substantially smaller than the total dimension of model and climatological fields) so this method has the potential to be extremely efficient compared to the alternatives which have been previously used.

We illustrate the process with the followng simple example. Our model takes a single input variable, x, for which we have a prior estimate $x=30\pm10$ (the indicated uncertainties are all one standard deviation, Gaussian, and assumed independent where possible) and generates the output y via the slightly nonlinear equation

$$y = x + 0.02x^2. \tag{1}$$

We have a single observation $y_o=12\pm1$. The posterior distributions for x and y can be easily calculated numerically, and are given by $x_a=10.0\pm0.7$, $y_a=12\pm1$ to one decimal place (the posterior distributions are not perfectly Gaussian but extremely close, and of course the uncertainties on x_a and y_a are highly correlated). Even for such a near-linear problem, a single application of the EnKF equations (using an ensemble of 10 000 members to eliminate one possible source of error) gives the rather poor solution $x_a=13\pm1.3$, $y_a=12\pm1$. The

estimate for x_a is very poor and moreover these values do not come close to satisfying the model equations ($x=13\pm1.3$ is mapped by the model to $y=16.4\pm2$). The source of this error is that the prior mean and covariance matrix cannot represent the full nonlinear distribution adequately, and this distorts the posterior even though the prior is very diffuse and should have minimal influence. Two experiments using our iterative method using an ensemble size of 100, with different ensemble expansion factors and initial guesses, are shown in Fig. 1. Both experiments converge to the correct solution and generate well-balanced (x, y) pairs, in marked contrast to the single-step procedure.

An additional benefit of this scheme is that since prior information is treated in an identical manner to observational data, it can be completely eliminated from the analysis if the observational constraints are adequate in themselves. This solves the problem of the double-counting of data through its inclusion in expert priors (Allen et al., 2002). Although it may seem rather inefficient to repeatedly integrate the ensemble members towards climatological convergence, in practice the integration interval within the iterative procedure can be kept quite short (significant stochastic noise in individual members can be tolerated due to the ensemble size) and so the total integration time is not so dissimilar from what would be required for a single integration of the ensemble to an accurate steady state. We use both 1 and 5 year iterative cycles in the work presented here, with reasonable convergence for the 1 year cycle requiring roughly 50 years in total per ensemble member, a time which is only a modest factor greater than the O(15–30) year integrations which are generally used to generate model climatologies.

In the previous applications with the coupled climate model, a parallel supercomputer was used to integrate the ensemble members simultaneously, and a domain decomposition was implemented for the analysis step following the method of Keppenne (2000). However, the domain decomposition is not necessary for relatively small models where the entire ensemble can be stored on a single processor as is the case here. The 64-member ensemble that we use here (determined by the maximum number of processors available for a single job) is integrated in parallel, one member to a processor as before, but the analysis is performed on a single processor. In order to limit the cost of the analysis (which is dominated by the inversion of the error covariance matrix), the observations are treated sequentially in blocks of 1024 values (the number of lateral grid points on a hemisphere). So long as the observational errors are uncorrelated, as is assumed here, this sequential treatment does not affect the solution. With each year of integration taking about 8 min per processor (on a Compaq Alpha SC), and the analysis requiring about 4 min (with the remaining 63 processors sitting idle), this approach is somewhat inefficient for a 1 year analysis cycle but the wastage is much less significant for a 5 year cycle. Each complete experiment of 100 years integration per ensemble member (6400 total model years) took about 20 h of real time using the 1 year cycle, with a typical load factor of about 65%. For the 5 year iterative cycle, the average
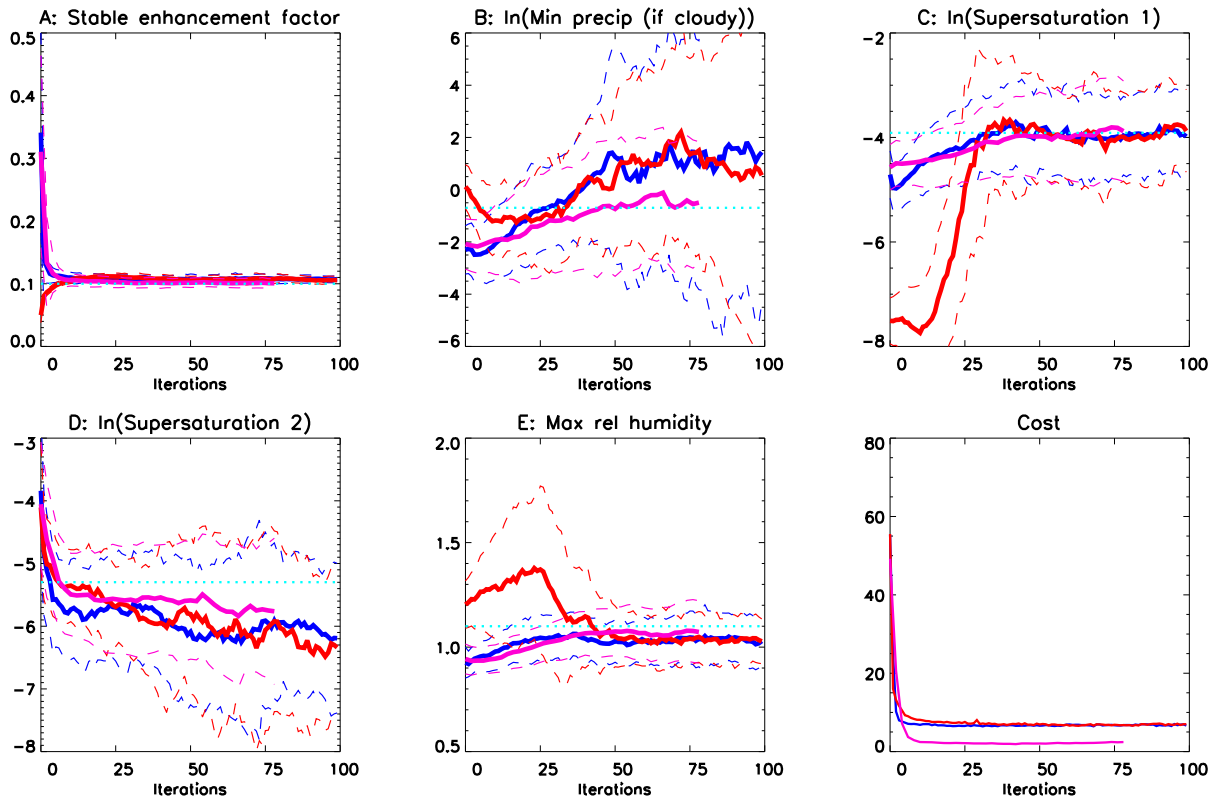
Fig. 2. Parameters converging in two identical twin tests for 1 and 5 year iteration, with different starting points. Red, dark blue and magenta lines indicate different experiments with 1, 1 and 5 year iterations respectively. Thick lines represent ensemble means, thin dashed lines indicate ensemble width (1 standard deviation). Cyan dotted lines indicate parameter values used to generate synthetic data set.

loading was improved to around 90% and the same integration length only required 15 h, however in our experiments convergence of this system required more model-years in total resulting in a longer overall integration time. Computing resources were not a limiting factor and the performance of our system may be some way from optimal. Further investigation may be worthwhile for application to more computationally demanding models.

We chose 5 parameters to tune using this system, identified via some preliminary sensitivity analyses which involved a series of 10-year integrations in which 29 tunable parameters were varied individually, within ranges believed to be physically reasonable (with the other parameters held fixed at the default values). The tunable parameters were from the radiation, convection, and surface parameterisations. The 5 variables finally selected were those which were found to have most effect on a skill score which was determined by the quality of fit to the tuning targets of December–January–February (DJF) and June–July–August (JJA) surface air temperature and precipitation (4 two-dimensional data sets in total). The parameters selected were (A) a non-dimensional linear multiplier of the sensible and latent heats, (B) the convective precipitation rate in mm/day at which convective clouds start to form, (C) the large-scale cloud supersaturation for the liquid water path calculation, (D) the convective cloud supersaturation for the liquid water path calculation,

and (E) the relative humidity at which large-scale clouds are assumed to completely cover a grid-box. Three of the parameters (B, C, and D) are constrained to be positive, but can otherwise vary by several orders of magnitude without invalidating the model. For these, we use a logarithmic transformation as in Annan et al. (2005), in order to avoid negative analysis values. The remaining two parameters (A and E) have a priori much more restricted ranges and therefore no transformation was necessary.

3 Identical twin testing

3.1 Experimental details

For the identical twin testing, a model run of 30 years was performed with a set of randomly-chosen parameters, and the output of the final 25 years analysed into synthetic "observations" of total precipitation and surface air temperature for the DJF and JJA seasons at all points on the model grid, thus making a total of 2048×4=8196 data points. Determining an appropriate estimate for observational errors of these data points is not entirely trivial, since although the typical distance of these fields from the model's true climate (given an infinite integration) can be readily estimated, it would be inappropriate to use this value for the assimilation.
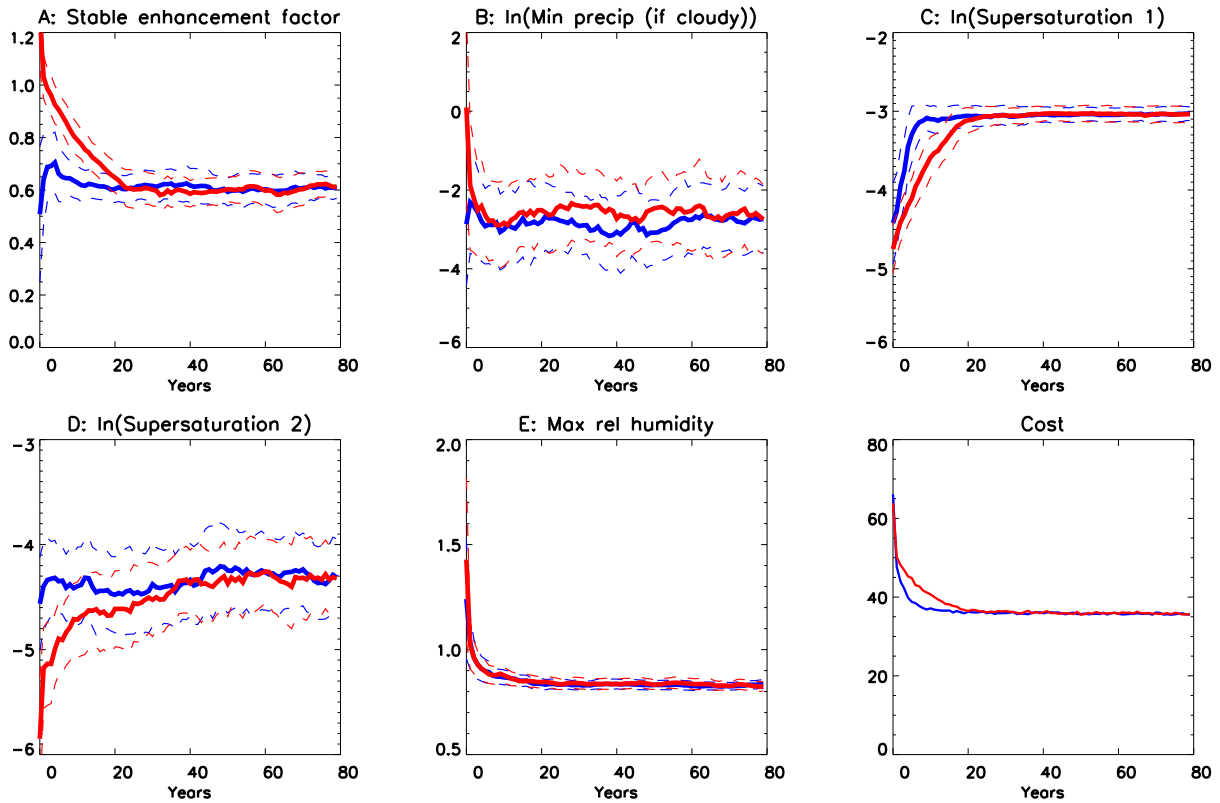
Fig. 3. Parameters converging in two experiments with reanalysis data (red and dark blue lines as for Fig. 2).

This is because the short model integrations used in the iterative scheme will have a correspondingly larger component of chaotic noise, and so it is not possible for them even in principle to match the observed values to within this observational accuracy. Attempting to fit the data more closely than is possible, forces the ensemble to collapse to a point in parameter space. Moreover, the spatial correlation of the model bias implies that the observational errors are not truly independent. Therefore, the error statistics were scaled up by a somewhat arbitrary factor of 20, the influence of which was checked by comparing the posterior model skill to the expected value for a short integration with correct parameters. Clearly this ad-hoc adjustment is not an entirely satisfactory approach and further investigations are planned. The errors were also assumed spatially invariant.

The ensemble was initialised with each member having parameters chosen from a distribution some way removed from the truth run. Since the spin-up time of the model is so fast, there is no systematic dependence of the climatology on the initial fields even for a 1 year integration. Furthermore, the model appears to have a problem when initialised with an excessively cold state in the polar regions (which can arise in some analysed model states), so rather than attempting to corect this problem here we instead decided to re-initialise from a uniform state ("cold start") rather than use the analysed model fields throughout the iterative analysis procedure. Obviously, for a model with a longer spin-up time, compu-

tational efficiency would be improved by using the analysed state which should be in reasonable balance with the analysed parameter set.

There are two primary adjustable controls on our assimilation scheme, being the length of integration between analysis steps, and the ensemble inflation factor. A longer integration interval gives more stable estimates of each ensemble member's climatology, with the noise due to deterministic chaos decreasing in inverse proportion to the square root of the run length, as if it was truly random noise (Lea et al., 2000). A larger ensemble inflation factor also gives more rapid convergence but is potentially less accurate in nonlinear situations. In practice, run lengths of 1 and 5 years, with an inflation factor of 10%, gave good results which are now described further.

3.2 Identical twin results

We wanted to investigate the value of the limited observational data used, so no prior estimates of parameter values were used in the initial experiments. Figure 2 shows the convergence of the parameter distributions and cost function for three identical twin experiments. Two of these experiments used a 1 year cycle and 100 iterations (6400 model years in total), and were identical apart from the initial ensemble. Results for a single experiment using 80 iterations of a 5 year cycle (25 600 years) are also shown. Parameters A, C, and E have converged in all experiments to the same stable
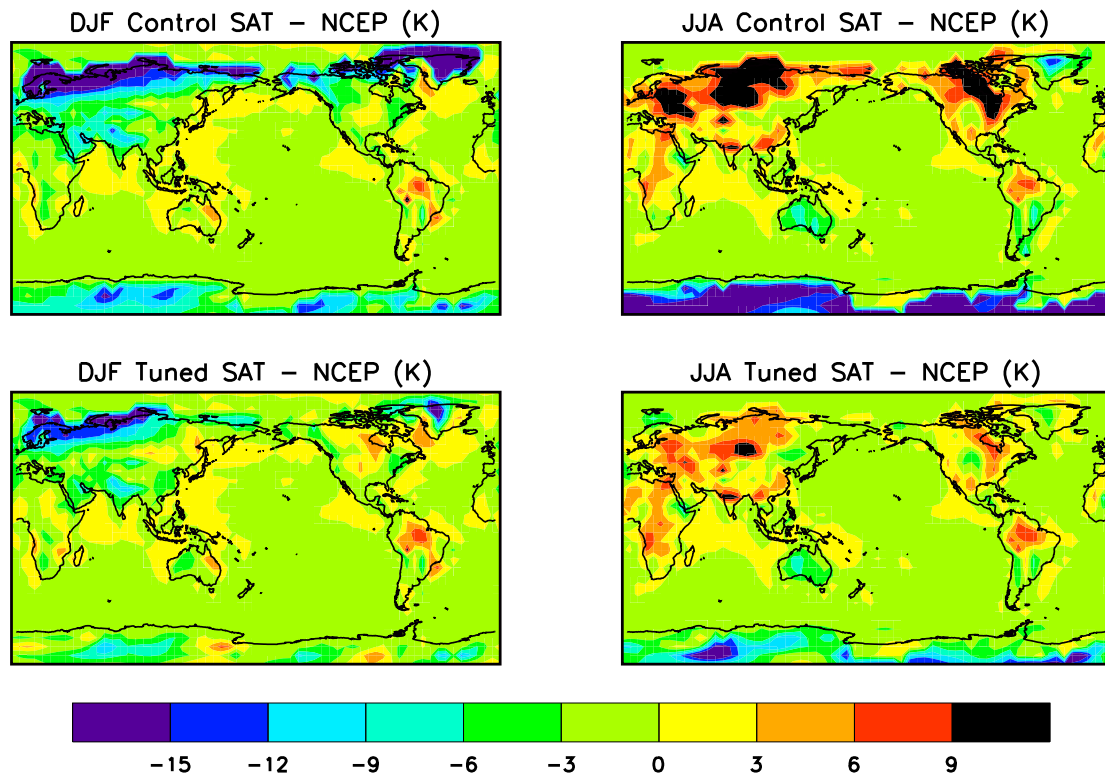
Fig. 4. Model SAT errors (degrees Kelvin), before and after tuning.

distributions, consistent with the truth. Parameter D is not so clear, with perhaps some evidence of a continued modest drift to lower values, but again it is consistent with the value used to generate the identical twin data. Parameter B, however, is clearly not constrained in the 1 year experiments although there are some signs that it is converging with the 5 year iterations. Given that these parameters were initially selected to be those for which the model was most sensitive, this suggests that the data used are barely adequate for constraining as many as 5 parameters simultaneously. Even though there are 8196 data points in all, they only represent 2 types of measurement so this is a not entirely unexpected result. It is, however, also possible that the preselection of sensitive parameters (based around the default values) may not be valid close to this new optimum.

The cost function of an ensemble member is given by the sum of the squared differences between model and synthetic data fields, normalised by the number of model grid points. The cost function line plotted in Fig. 2 is the mean of the costs of the ensemble members (rather than the cost of their mean output). A typical cost of a little over 7 in the posterior ensemble for the 1 year iterations is made up of 5.5 for the two temperature fields ($\sqrt{5.5/2}$=1.4 K RMS error at each gridpoint) and 1.6 for the precipitation (0.9 mm/d RMS precipitation error). This is essentially the same as the variability between model runs due to stochastic noise. For the 5 year integration, the cost is about 2, illustrating not a superior solution (the parameter distributions are essentially the same)

but the effect of the longer integration on reducing the magnitude of the deterministic noise. At the true parameter values and arbitrary initial conditions, stochastic noise generates a cost of 1.7, so the range of parameter values in the posterior distribution is generating a marginally worse fit to the data than that due to stochastic noise alone. All ensembles have generated similar parameter distributions despite the different experimental conditions. This suggests that multiple local minima are not a significant problem in this application, if they exist at all.

4 Application using reanalysis data

We now apply the method using observational data. The surface air temperature and the precipitation both come from climatological monthly-mean NCEP reanalysis.

In initial experiments, it quickly became apparent that it is not possible for this model to match the data closely, with any combination of values for the 5 parameters. There are significant regional biases in both temperature and precipitation which cannot all be simultaneously eliminated by parameter tuning alone. As a result of the model-data mismatch, two of the parameters (C and E) were forced towards values which were numerically unstable (and physically meaningless, in the case of parameter E) and prior distributions for them had to be provided. With this proviso, convergence was actually more rapid and consistent in these experiments than in the identical twin tests, perhaps because with two parameters
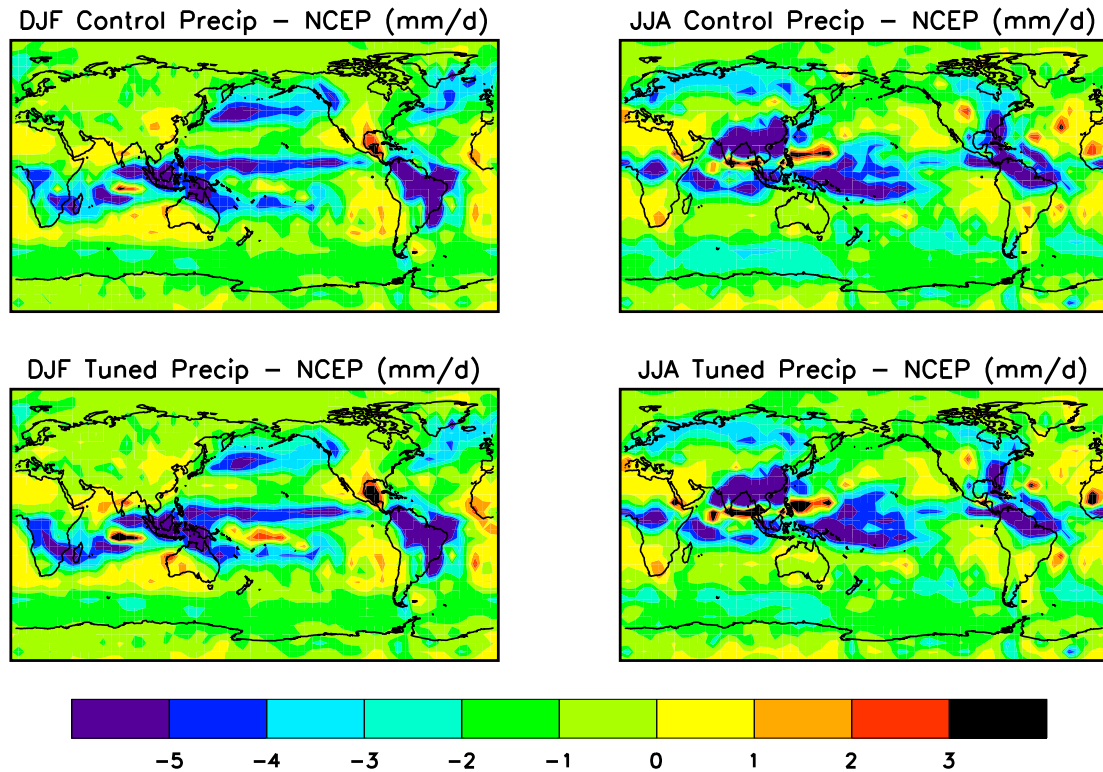
Fig. 5. Model precipitation error (mm/d), before and after tuning.

forced to the edges of their ranges, there were fewer remaining effective degrees of freedom and less nonlinearity in the parameter space. Again, repeat runs under different conditions did not find any different solutions. Figure 3 shows the evolution of the parameter distributions and cost functions for two experiments.

The ensemble mean temperature matches the data fairly well (Fig. 4), with a typical RMS error of 3K at each gridpoint. However, it should be noted that in tuning to reanalysis data, we have chosen a somewhat easier target than if we had used pure observations. Substantial cold biases over the high plateaux (especially Tibet) are present in almost all AGCMs, including that used for the NCEP reanalysis. This comparison therefore gives an optimistic impression of model skill by masking the same failing in our model. In contrast to the reasonable temperature fields, precipitation remains poor in all simulations (Fig. 5), with an RMS error as high as 3mm per day. In fact, even attempting to tune to precipitation alone, completely ignoring the fit to the temperature data, did not improve that result. The model parameters appear to have very little effect on precipitation patterns, despite several of them relating directly to hydrology. Re-examining the results from the univariate sensitivity analysis indicated that the wider range of 29 parameters tested all had a minimal effect on the precipitation. Disabling the convection scheme entirely, changed the model precipitation more substantially (and in fact led to an overall improvement). Clearly this points to a significant structural deficiency and research is

now under way to investigate alternative convection schemes. Although the model output is disappointing in this respect, the power of this multivariate tuning method is still apparent here in efficiently and rigorously identifying the limit of the parameter tuning and thereby motivating investigation of structural changes which are now under way.

The overall fit to the data, using our unweighted cost function, dropped from a value of 78 for the default parameters, to 33 for the tuned ensemble, representing an improvement of about $1-\sqrt{33/78}=35\%$ in the typical model-data mismatch.

5   Conclusions

The iterative ensemble Kalman filter for parameter estimation has been successfully applied to an intermediate complexity spectral primitive equation AGCM. The underlying similarity of this model to more complex AGCMs, for example the CCSR/NIES model (Nozawa et al., 2001), implies that they could also be tuned using this method, and further work in this direction is in progress. Identical twin testing demonstrates that the method reliably finds a unique optimal posterior pdf in parameter space, although of course this cannot be guaranteed in all applications. Tuning to reanalysis data generates realistic temperature fields. However, precipitation in this model remains rather poor and this problem appears to be due to structural deficiencies in the convection routine. Further research is now under way to improve this situation. The efficient optimal parameter tuning has already

proved its worth by showing that parameter tuning will not improve this aspect of model behaviour, and we expect it to contribute futher in the creation of the coupled atmosphere-ocean model. With more data sources, there are no obvious reasons why many more parameters could not be simultaneously tuned as the computational time appears to only scale slowly with the number of free parameters.

Although the model as presented here is not directly suited to climate prediction (being created as one component of an Earth system model), the success of the method in this application strongly suggests that there are no fundamental reasons why future applications to prediction using more complete models should not be successful.

Edited by: S. Vannitsem
Reviewed by: two referees

References

Allen, M.: Do it yourself climate prediction, Nature, 401, 642, 1999.

Allen, M., Kettleborough, J., and Stainforth, D.: Model error in weather and climate forecasting, in Proceedings of the 2002 ECMWF predictability seminar,ECMWF, Reading, United Kingdom, 275–294, 2002.

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Monthly Weather Review, 129, 2884–2902, 2001.

Annan, J. D. and Hargreaves, J. C.: Efficient parameter estimation for a highly chaotic system, Tellus, 56A, 520–526, 2004.

Annan, J. D., Hargreaves, J. C., Edwards, N. R., and Marsh, R.: Parameter estimation in an intermediate complexity Earth System Model using an ensemble Kalman filter, Ocean Modelling, 8, 135–154, 2005.

Bellman, R.: Adaptive Control Processes: A Guided Tour, Princeton University Press, 1961.

de Forster, P. M., Blackburn, M., Glover, R., and Shine, K. P.: An examination of climate sensitivity for idealised climate change experiments in an intermediate general circulation model, Climate Dynamics, 16, 833–849, 2000.

Derber, J.: A variational continuous assimilation scheme, Monthly Weather Review, 117, 2437–2446, 1989.

Edwards, N. R. and Marsh, R.: Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model, Climate Dynamics, in press, 2005.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10 143–10 162, 1994.

Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, 2003.

Evensen, G. and Leeuwen, P. J. V.: An Ensemble Kalman Smoother for nonlinear dynamics, Monthly Weather Review, 128, 1852–1867, 2000.

Giorgi, F. and Mearns, L.: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, Journal of Climate, 15, 1141–1158, 2002.

Hargreaves, J. C., Annan, J. D., Edwards, N. R., and Marsh, R.: Climate forecasting using an intermediate complexity Earth System Model and the ensemble Kalman filter, Climate Dynamics, 23, 745–760, 2004.

Highwood, E. J. and Stevenson, D. S.: Atmospheric impact of the 1783-1784 Laki eruption. Part II - Climatic effect of sulphate aerosol, Atmos. Chem. Phys., 3, 1177—1189, 2003,
SRef-ID: 1680-7324/acp/2003-3-1177.

Joshi, M., Shine, K., Ponater, M., Stuber, N., Sausen, R., and Li, L.: A comparison of climate respone to different radiative forcings in three general circulation modles: towards an improved metric of climate change, Climate Dynamics, 20, 843–854, 2003.

Kalman, R. E.: A new approach to linear filtering and prediction problems, J. Basic Engineering, 82D, 33–45, 1960.

Keppenne, C. L.: Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter, Monthly Weather Review, 128, 1971–1981, 2000.

Kivman, G. A.: Sequential parameter estimation for stochastic systems, Nonlin. Proc. Geophys., 10, 253–259, 2003,
SRef-ID: 1607-7946/npg/2003-10-253.

Knutti, R., Stocker, T. F., Joos, F., and Plattner, G.-K.: Constraints on radiative forcing and future climate change from observations and climate model ensembles, Nature, 416, 719–723, 2002.

Köhl, A. and Willebrand, J.: An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models, Tellus, 54A, 406–425, 2002.

Köhl, A. and Willebrand, J.: Variational assimilation of SSH variability from TOPEX/POSEIDON and ERS1 into an eddy-permitting model of the North Atlantic, Journal of Geophysical Research, C3, art. num. 3092, 2003.

Lea, D. J., Allen, M. R., and Haine, T. W. N.: Sensitivity analysis of the climate of a chaotic system, Tellus, 52A, 523–532, 2000.

Lea, D. J., Haine, T. W. N., Allen, M. R., and Hansen, J. A.: Sensitivity analysis of the climate of a chaotic ocean circulation model, Q. J. R. Meteorological Soc., 128, 2587–2605, 2002.

Losa, S. N., Kivman, G. A., Schröter, J., and Wenzel, M.: Sequential weak constraint parameter estimation in an ecosystem model, Journal of Marine Systems, 43, 31–49, 2003.

Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004.

Nozawa, T., Emori, S., Numaguti, A., Tsushima, Y., Takemura, T., Nakajima, T., Abe-Ouchi, A., and Kimoto, M.: Projections of Future Climate Change in the 21st Century Simulated by the CCSR/NIES CGCM under the IPCC SRES Scenarios, in Present and Future of Modeling Global Environment Change, edited by: Matsuno, T. and Hideji, K., Terra Scientific Publishing Company, 15–28, 2001.

Pisarenko, V. F. and Sornette, D.: Statistical methods of parameter estimation for deterministically chaotic time series, Physical Review E, 69, 036122, 2004.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: Numerical recipes in Fortran: the art of scientific computing, Cambridge University Press, 1994.

Rosier, S. M. and Shine, K. P.: The effect of two decades of ozone change on stratospheric temperatures as indicated by a general circulation model, Geophys. Res. Lett., 27, 2617–2620, 2000.